A Conditional-Heteroskedasticity-Robust Confidence Interval for the Autoregressive Parameter

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

Patrik Guggenberger
Department of Economics
UCSD

December 2010
Revised: July 2011

* Andrews and Guggenberger gratefully acknowledge the research support of the National Science Foundation via grant numbers SES-0751517 and SES-1058376, and SES-1021101, respectively.

Abstract

This paper introduces a new confidence interval (CI) for the autoregressive parameter (AR) in an AR(1) model that allows for conditional heteroskedasticity of general form and AR parameters that are less than or equal to unity. The CI is a modification of Mikusheva's (2007a) modification of Stock's (1991) CI that employs the least squares estimator and a heteroskedasticity-robust variance estimator. The CI is shown to have correct asymptotic size and to be asymptotically similar (in a uniform sense). It does not require any tuning parameters. No existing procedures have these properties. Monte Carlo simulations show that the CI performs well in finite samples in terms of coverage probability and average length, for innovations with and without conditional heteroskedasticity.

Keywords: Asymptotically similar, asymptotic size, autoregressive model, conditional heteroskedasticity, confidence interval, hybrid test, subsampling test, unit root.

JEL Classification Numbers: C12, C15, C22.

1 Introduction

We consider confidence intervals (CI's) for the autoregressive parameter (AR) ρ in a conditionally heteroskedastic AR(1) model in which ρ may be close to, or equal to, one. The observed time series $\{Y_i : i = 0, ..., n\}$ is based on a latent no-intercept AR(1) time series $\{Y_i^* : i = 0, ..., n\}$:

$$\begin{aligned} Y_i &= \mu + Y_i^*, \\ Y_i^* &= \rho Y_{i-1}^* + U_i \text{ for } i = 1, ..., n, \end{aligned} \tag{1.1}$$

where $\rho \in [-1 + \varepsilon, 1]$ for some $0 < \varepsilon < 2$, $\{U_i : i = ..., 0, 1, ...\}$ are stationary and ergodic under the distribution F, with conditional mean 0 given a σ-field $\mathcal{G}_{i-1}$ for which $U_j \in \mathcal{G}_i$ for all $j \leq i$, conditional variance $\sigma_i^2 = E_F(U_i^2 | \mathcal{G}_{i-1})$, and unconditional variance $\sigma_U^2 \in (0, \infty)$. The distribution of Y_0^* is the distribution that yields strict stationarity for $\{Y_i^* : i \leq n\}$ when $\rho < 1$. That is, $Y_0^* = \sum_{j=0}^{\infty} \rho^j U_{-j}$when $\rho < 1$. When $\rho = 1$, Y_0^* is arbitrary.

Models of this sort are applicable to exchange rate and commodity and stock prices, e.g., see Kim and Schmidt (1993). Simulations in Mikusheva (2007b, Table II) show that CI's not designed to handle conditional heteroskedasticity may perform poorly in terms of coverage probabilities when conditional heteroskedasticity is present. In fact, most have incorrect asymptotic size in this case.[1]

For the case of conditional homoskedasticity, several CI's with correct asymptotic size have been introduced, including those in Stock (1991), Andrews (1993), Andrews and Chen (1994), Nankervitz and Savin (1996), Hansen (1999), Elliot and Stock (2001), Romano and Wolf (2001), Chen and Deo (2007), and Mikusheva (2007a).[2] Of these CI's the only one that has correct asymptotic size in the presence of conditional heteroskedasticity is the symmetric two-sided subsampling CI of Romano and Wolf (2001).[3] The latter CI has the disadvantages that it is not asymptotically similar, requires a tuning parameter (the subsample size), and is far from being equal-tailed when ρ is near one.[4]

The first CI's that were shown to have correct asymptotic size under conditional heteroskedasticity and an AR parameter close to unity were introduced in Andrews

[1] Throughout this paper we use the term "asymptotic size" to mean the limit as $n \to \infty$ of the finite-sample size. Uniformity in the asymptotics is built into this definition because finite-sample size is a uniform concept. By the definition of asymptotic size, the infimum of the coverage probability over different values of ρ and different innovation distributions is taken before the limit as $n \to \infty$ is taken.

[2] The CI of Stock (1991) needs to be modified as in Mikusheva (2007a) to have correct asymptotic size.

[3] The correct asymptotic size of this CI is established in the Supplemental Appendix. The equal-tailed subsampling CI of Romano and Wolf (2001) does not have correct asymptotic size under homoskedasticity or heteroskedasticity, see Mikusheva (2007a) and Andrews and Guggenberger (2009).

[4] Lack of asymptotic similarity implies that the CI over-covers asymptotically for some sequences of ρ values. This may yield a longer CI than is possible.

and Guggenberger (2009) (AG09).[5] These CI's are based on inverting a t statistic constructed using a feasible quasi-generalized least squares (FQGLS) estimator of ρ. AG09 shows that equal-tailed and symmetric two-sided CI's based on hybrid (fixed/subsampling) critical values have correct asymptotic size.[6] These CI's are robust to misspecification of the form of the conditional heteroskedasticity. However, they are not asymptotically similar and require the specification of a tuning parameter—the subsample size.

The contribution of this note is to introduce a CI that (i) has correct asymptotic size for a parameter space that allows for general forms of conditional heteroskedasticity and for an AR parameter close to, or equal to, unity, (ii) is asymptotically similar, and (iii) does not require any tuning parameters.

The CI is constructed by inverting tests constructed using a t statistic based on the LS estimator of ρ and a heteroskedasticity consistent (HC) variance matrix estimator. For the latter, we use a variant of the HC3 version defined in MacKinnon and White (1985), which we call HC4. It employs an adjustment that improves the finite-sample coverage probabilities. This t statistic is asymptotically nuisance parameter-free under the null hypothesis under drifting sequences of null parameters ρ, whether or not these parameters are local to unity. In consequence, critical values can be obtained by matching the given null value of ρ and sample size n with a local-to-unity parameter $h = n(1-\rho)$. Then, one uses the quantile(s) from the corresponding local-to-unity asymptotic distribution which depends on h. This method is employed by Stock (1991), Andrews and Chen (1994, Sec. 4), and Mikusheva (2007a) (in her modification of Stock's CI).[7] The resulting CI is the same as Mikusheva's (2007a) modification of Stock's (1991) CI applied to the LS estimator of ρ, except that we use the HC4 variance estimator in place of the homoskedastic variance estimator and we use a stationary initial condition rather than a zero initial condition.[8,9] We refer to the new CI as the CHR CI (which abbreviates "conditional-heteroskedasticity-robust").

The use of the LS estimator, rather than the FQGLS estimator, is important because the latter has an asymptotic distribution in the local-to-unity case that is

[5] Gonçalves and Kilian (2007) also consider inference in autoregressive models with conditional heteroskedasticity but do not allow for unit roots or roots near unity.

[6] AG09 also introduces several other CI's that have correct asymptotic size under conditional heteroskedasticity using size-corrected fixed critical values and size-corrected subsampling critical values (for equal-tailed CI's). The performance of these CI's is not as good as that of the FQGLS-based hybrid CI, so we do not discuss these CI's further here.

[7] As in Mikusheva's (2007a) modification of Stock's CI, we invert the t statistic that is designed for a given value of ρ, not the t statistic for testing $H_0 : \rho = 1$ which is employed in Stock (1991). This is necessary to obtain correct asymptotic coverage when ρ is not $O(n^{-1})$ local to unity.

[8] Mikusheva's (2007a) results do not cover the new CI because (i) she does not consider innovations that have conditional heteroskedasticity and (ii) even in the i.i.d. innovation case the t statistic considered here does not lie in the class of test statistics that she considers.

[9] The use of a stationary intial condition when $\rho < 1$, rather than a zero initial condition, is not crucial to obtaining robustness to conditional heteroskedasticity. Our results also apply to the case of a zero initial condition, in which case the second component of $I_h^*(r)$ in (2.4) below is deleted.

a convex combination of a random variable with a unit-root distribution and an independent standard normal random variable with coefficients that depend on the strength of the conditional heteroskedasticity, see Seo (1999), Guo and Phillips (2001), and Andrews and Guggenberger (2011). Hence, a nuisance parameter appears in the asymptotic distribution of the FQGLS estimator that does not appear with the LS estimator. This yields a trade-off when constructing a CI between using a more efficient estimator (FQGLS) combined with critical values that do not lead to an asymptotically similar CI and using a less efficient estimator (LS) with critical values that yield an asymptotically similar CI.

The use of an HC variance matrix estimator with the new CHR CI is important to obtain a (nuisance-parameter free) standard normal asymptotic distribution of the t statistic when the sequence of true ρ parameters converges to a value less than one as $n \to \infty$ and conditional heteroskedasticity is present. It is not needed to yield a nuisance parameter-free asymptotic distribution when ρ converges to unity (either at a $O(n^{-1})$ rate or more slowly).[10] This follows from results in Giraitis and Phillips (2006) and Andrews and Guggenberger (2011).

Simulations indicate that the CHR CI has good finite-sample coverage probabilities and has shorter average lengths—often noticeably shorter—than the hybrid CI of AG09 (based on the FQGLS estimator) for a variety of GARCH$(1,1)$ processes whose parametrizations are empirically relevant. When no conditional heterskedasticity is present, the CHR CI performs very well in finite samples relative to CI's that are designed for the i.i.d. innovation case. Hence, there is little cost to achieving robustness to conditional heteroskedasticity.

The asymptotic size and similarity results for the new CI are obtained rather easily by employing the asymptotic results of Andrews and Guggenberger (2011) for FQGLS estimators under drifting sequence of distributions, which include LS estimators as a special case, combined with the generic uniformity results in Andrews, Cheng, and Guggenberger (2009).

The CHR CI yields a unit root test that is robust to conditional heteroskedasticity. One rejects a unit root if the CI does not include unity. Seo (1999) and Guo and Phillips (2001) also provide unit root tests that are robust to conditional heteroskedasticity.

The CHR CI for ρ can be extended to give a CI for the sum of the AR coefficients in an AR(k) model when all but one root is bounded away from the unit circle, e.g., as in Andrews and Chen (1994, Sec. 4) and Mikusheva (2007a), and to models with a linear time trend. In the former case, the asymptotic distributions (and hence the CHR critical values) are unchanged and in the latter case the asymptotic distributions are given in (7.7) of Andrews and Guggenberger (2009) with $h_{2,7} = 1$. Extending the proof of Theorem 1 below for these cases requires additional detailed analysis, e.g., as in Mikusheva (2007a, Sec. 7). For brevity, we do not provide such proofs here.

[10] That is, when ρ converges to unity, one obtains the same asymptotic distribution whether an HC or a homoskedastic variance estimator is employed.

The note is structured as follows. Section 2 defines the new CI and establishes its large sample properties. Section 3 contains a Monte Carlo study. The Supplemental Appendix provides: (i) tables of critical values, (ii) the local asymptotic false coverage probabilities of the CHR CI, (iii) asymptotic and finite-sample assessments of the price the CHR CI pays in the i.i.d. case for obtaining robustness to conditional heteroskedasticity, (iv) probabilities of obtaining disconnected CHR CI's, (v) simulation results for several symmetric two-sided CI's, (vi) details concerning the simulations, (vii) proofs of the asymptotic results for the CHR CI, and (viii) a proof that the symmetric two-sided subsampling CI of Romano and Wolf (2001) has correct asymptotic size under conditional heteroskedasticity.

2 The CHR CI for the AR parameter

For the exposition of the theory, we focus on equal-tailed two-sided CI's for ρ.[11,12] The CI is obtained by inverting a test of the null hypothesis that the true value is ρ. The model (1.1) can be rewritten as $Y_i = \widetilde{\mu} + \rho Y_{i-1} + U_i$, where $\widetilde{\mu} = \mu(1-\rho)$ for $i = 1, ..., n$. We use the t statistic

$$T_n(\rho) = \frac{n^{1/2}(\widehat{\rho}_n - \rho)}{\widehat{\sigma}_n}, \tag{2.1}$$

where $\widehat{\rho}_n$ is the LS estimator from the regression of Y_i on Y_{i-1} and 1 and $\widehat{\sigma}_n^2$ is the $(1,1)$ element of the HC4 heteroskedasticity-robust variance estimator, defined below, for the LS estimator in the preceding regression. More explicitly, let Y, U, X_1, and X_2 be n-vectors with ith elements given by Y_i, U_i, Y_{i-1}, and 1, respectively. Let $X = [X_1 : X_2]$, $P_X = X(X'X)^{-1}X'$, and $M_X = I_n - P_X$. Let $\widehat{U}_i$ denote the ith element of the residual vector $M_X Y$. Let p_{ii} denote the ith diagonal element of P_X. Let $p_{ii}^* = \min\{p_{ii}, n^{-1/2}\}$. Let Δ be the diagonal $n \times n$ matrix with ith diagonal element given by $\widehat{U}_i/(1-p_{ii}^*)$.[13] Then, the LS estimator of ρ and the HC4 estimator of its variance are

$$\begin{aligned} \widehat{\rho}_n &= \left(X_1' M_{X_2} X_1\right)^{-1} X_1' M_{X_2} Y, \text{ and} \\ \widehat{\sigma}_n^2 &= \left(n^{-1} X_1' M_{X_2} X_1\right)^{-1} \left(n^{-1} X_1' M_{X_2} \Delta^2 M_{X_2} X_1\right) \left(n^{-1} X_1' M_{X_2} X_1\right)^{-1}. \end{aligned} \tag{2.2}$$

[11] Symmetric two-sided and one-sided CI's can be handled in a similar fashion, see the Supplemental Appendix for details.

[12] We prefer equal-tailed CI's over symmetric CI's in the AR(1) context because the latter can have quite unequal coverage probabilities for missing the true value above and below when ρ is near unity, which is a form of biasedness, due to the lack of symmetry of the near-unit root distributions.

[13] The quantity p_{ii}^* used in HC4 is a finite-sample adjustment to the standard HC variance estimator. In contrast, the HC3 variance estimator uses p_{ii} in the definition of Δ. The use of p_{ii}^* guarantees that the finite-sample adjustment does not affect the asymptotics. When $n(1-\rho_n) \to h < \infty$, it is straightforward to show that the use of p_{ii} is valid asymptotically. In other cases, it is more difficult to do so. However, the finite-sample results reported below are essentially the same whether p_{ii}^* or p_{ii} is used. Note that the asymptotic results given in the paper hold if one sets $p_{ii}^* = 0$, which yields the standard HC variance estimator.

The parameter space for (ρ, F) is given by

$$\Lambda = \{\lambda = (\rho, F) : \rho \in [-1+\varepsilon, 1],\ \{U_i : i = 0, \pm1, \pm2, ...\} \text{ are stationary and strong mixing under } F \text{ with } E_F(U_i|\mathcal{G}_{i-1}) = 0 \text{ a.s.}, E_F(U_i^2|\mathcal{G}_{i-1}) = \sigma_i^2 \text{ a.s., where } \mathcal{G}_i \text{ is some non-decreasing sequence of } \sigma\text{-fields for } i = ..., 1, 2, ... \text{ for which } U_j \in \mathcal{G}_i \text{ for all } j \leq i, \text{ the strong-mixing numbers } \{\alpha_F(m) : m \geq 1\} \text{ satisfy } \alpha_F(m) \leq Cm^{-3\zeta/(\zeta-3)}\ \forall m \geq 1,\ \sup_{i,s,t,u,v,A} E_F|\prod_{a \in A} a|^\zeta \leq M, \text{ where } 0 \leq i, s, t, u, v < \infty, \text{ and } A \text{ is any non-empty subset of } \{U_{i-s}, U_{i-t}, U_{i+1}^2, U_{-u}, U_{-v}, U_1^2\}, \text{ and } E_F U_1^2 \geq \delta\}, \tag{2.3}$$

for some constants $0 < \varepsilon < 2$, $\zeta > 3$, $C < \infty$, and $\delta > 0$.

Next, we define the critical values used in the construction of the CI. They are based on the asymptotic distributions of the test statistic under drifting sequences $\{\lambda_n = (\rho_n, F_n) : n \geq 1\}$ of AR parameters ρ_n and distributions F_n, when $n(1-\rho_n) \to h \in [0, \infty)$. When F_n depends on n, $\{U_i : i \leq n\}$ for $n \geq 1$ form a triangular array of random variables and $U_i = U_{n,i}$. To describe the asymptotic distribution, let $W(\cdot)$ be a standard Brownian motion on $[0, 1]$. Let Z_1 be a standard normal random variable that is independent of $W(\cdot)$. Define

$$I_h(r) = \int_0^r \exp(-(r-s)h)dW(s),\ I_h^*(r) = I_h(r) + \frac{1}{\sqrt{2h}}\exp(-hr)Z_1 \text{ for } h > 0,$$
$$I_h^*(r) = W(r) \text{ for } h = 0, \text{ and } I_{D,h}^*(r) = I_h^*(r) - \int_0^1 I_h^*(s)ds. \tag{2.4}$$

Andrews and Guggenberger (2011, Theorem 1) (with a minor adjustment for the p_{ii}^* term in Δ) shows that, under any sequence $\lambda_n = (\rho_n, F_n) \in \Lambda$ such that $n(1-\rho_n) \to h \in [0, \infty]$,

$$T_n(\rho_n) \to_d J_h, \tag{2.5}$$

where J_h is defined as follows. For $h = \infty$, J_h is the $N(0, 1)$ distribution, and for $h \in [0, \infty)$, J_h is the distribution of

$$\int_0^1 I_{D,h}^*(r)dW(r) / \left(\int_0^1 I_{D,h}^*(r)^2 dr\right)^{1/2}. \tag{2.6}$$

For $\alpha \in (0, 1)$, let $c_h(1-\alpha)$ denote the $(1-\alpha)$-quantile of J_h. The second component of $I_h^*(r)$ in (2.4) is due to the stationary start-up of the AR(1) process when $\rho < 1$, as in Elliott (1999), Elliott and Stock (2001), Müller and Elliott (2003), and Andrews and Guggenberger (2009, 2011). Giraitis and Phillips (2006) provide similar results for the LS estimator for the case $h = \infty$.

The new nominal $1-\alpha$ equal-tailed two-sided CHR CI for ρ is

$$CI_{CHR,n} = \{\rho \in [-1+\varepsilon, 1] : c_h(\alpha/2) \leq T_n(\rho) \leq c_h(1-\alpha/2) \text{ for } h = n(1-\rho)\}. \tag{2.7}$$

The CI $CI_{CHR,n}$ can be calculated by taking a fine grid of values $\rho \in [-1+\varepsilon, 1]$ and comparing $T_n(\rho)$ to $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$, where $h = n(1-\rho)$. Tables of values of $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$ are given in the Supplemental Appendix. Given these values, calculation of $CI_{CHR,n}$ is simple and fast.[14,15]

The main theoretical result of this note shows that $CI_{CHR,n}$ has correct asymptotic size for the parameter space Λ and is asymptotically similar. Let P_λ denote probability under $\lambda = (\rho, F) \in \Lambda$.

Theorem 1 *Let $\alpha \in (0,1)$. For the parameter space Λ, the nominal $1-\alpha$ confidence interval $CI_{CHR,n}$ for the AR parameter ρ satisfies*

$$AsySz \equiv \liminf_{n\to\infty} \inf_{\lambda=(\rho,F)\in\Lambda} P_\lambda(\rho \in CI_{CHR,n}) = 1-\alpha.$$

Furthermore, $CI_{CHR,n}$ is asymptotically similar, that is,

$$\liminf_{n\to\infty} \inf_{\lambda=(\rho,F)\in\Lambda} P_\lambda(\rho \in CI_{CHR,n}) = \limsup_{n\to\infty} \sup_{\lambda=(\rho,F)\in\Lambda} P_\lambda(\rho \in CI_{CHR,n}).$$

Theorem 2 in the Supplemental Appendix establishes the local asymptotic false coverage probabilities of the CHR CI, which are directly related to their length.

3 Finite-Sample Simulation Results

Here we compare the finite-sample coverage probabilities (CP's) and average lengths of the new CHR CI and the hybrid CI of AG09.[16,17] For brevity, we focus on nominal 95% equal-tailed two-sided CI's. Results for symmetric CI's, including the symmetric subsampling CI of Romano and Wolf (2001), are provided in the Supplemental Appendix.

We consider a wide range of ρ values: $.99, .9, .5, .0, -.9$. The innovations are of the form $U_i = \sigma_i \varepsilon_i$, where $\{\varepsilon_i : i \geq 1\}$ are i.i.d. standard normal and σ_i is the multiplicative conditional heteroskedasticity. Let GARCH-$(ma, ar; \psi)$ denote a GARCH$(1,1)$ process with MA, AR, and intercept parameters $(ma, ar; \psi)$ and let ARCH-$(ar_1, ..., ar_4; \psi)$ denote an ARCH(4) process with AR parameters $(ar_1, ..., ar_4)$ and intercept ψ. We consider five specifications for the conditional heteroskedasticity of the innovations: (i) GARCH-$(.05, .9; .001)$, (ii) GARCH-$(.15, .8; .2)$, (iii) i.i.d.,

[14] The Supplemental Appendix also gives tables of critical values for symmetric two-sided and one-sided CHR CI's.

[15] Note that it is possible for $CI_{CHR,n}$ to consist of two disconnected intervals of the form $[a,b] \cup [c,1]$, where $-1+\varepsilon \leq a < b < c \leq 1$. This occurs with very low probability in most cases, and low probability in all cases, see the Supplemental Appendix for details.

[16] See MacKinnon and White (1985) and Long and Ervin (2000) for simulation results concerning the properties of the HC3 estimator in the standard linear regression model with i.i.d. observations.

[17] The hybrid CI is defined as in AG09 using the standard HC variance estimator with $p_{ii}^* = 0$, not the HC4 estimator.

(iv) GARCH-$(.25, .7; .2)$, and (v) ARCH-$(.3, .2, .2, .2; .2)$. Specifications (i)-(iii) are the most relevant ones empirically.[18] Specifications (iv) and (v) are included for purposes of robustness. They exhibit stronger conditional heteroskedasticity than in cases (i)-(iii). In cases (i)-(iv), the hybrid CI has an unfair advantage over the CHR CI, because it uses a GARCH$(1, 1)$ model which is correctly specified in these cases. The results are invariant to the choice of μ.

We consider a sample size of $n = 130$. The hybrid CI is based on a GARCH$(1, 1)$ specification.[19] The hybrid critical values use subsamples of size $b = 12$, as in Andrews and Guggenberger (2009).

We report average lengths of "CP-corrected" CI's. A CP-corrected CI equals the actual nominal 95% CI if its CP is at least .95 (for the given data-generating process), but otherwise equals the CI implemented at a nominal CP that makes the finite-sample CP equal to .95.[20] All simulation results are based on $30,000$ simulation repetitions.

Table I reports the results. CHR denotes the CI in (2.7). Hyb denotes the hybrid CI of AG09. The new CHR CI has very good finite-sample coverage probabilities. Specifically, its CP's are in the range $[94.1, 94.8]$ for all values of ρ in cases (i)-(iii). For cases (iv) and (v), the range is $[93.2, 94.5]$. The hybrid CI has CP's in the range $[94.2, 98.5]$ for cases (i)-(iii) and $[93.9, 98.5]$ for cases (iv) and (v). These CP's reflect the fact that the hybrid CI is not asymptotically similar due to its reliance on subsampling.

The average length results of Table I (CP-corrected) show that the CHR CI is shorter than the hybrid CI for all values of ρ in cases (i)-(iv). The greatest length reductions are for $\rho = .5, .0$, where the CHR CI is from .69 to .83 times the length of the hybrid CI in cases (i)-(iii). For $\rho = .99, .9$, it is from .86 to .91 times the length of the hybrid CI in cases (i)-(iii). In cases (iv) and (v), the CHR and hybrid CI's have similar lengths for $\rho = .99, .9$. In cases (iv) and (v), for $\rho = .5, .0$, the CHR CI is from .82 to .98 times the length of the hybrid CI. In conclusion, in an overall sense, the CHR CI out-performs the hybrid CI in terms of average length by a noticeable margin in the cases considered.[21]

---

[18] For example, see Bollerslev (1987), Engle, Ng, and Rothschild (1990), and, for more references, Ma, Nelson, and Startz (2007).

[19] See the Supplemental Appendix for more details concerning the definition and computation of the hybrid CI. Note that the hybrid CI has correct asymptotic size whether or not the GARCH(1, 1) specification is correct.

[20] When calculating the average length of a CI, we restrict the CI to the interval $[-1, 1]$. The search to find the nominal size such that the actual finite-sample CP ($\times 100$) equals 95.0 is done with stepsize .025. In the case of a disconnected CI, the "gap" in the CI is not included in its length.

[21] The CHR CI also out-performs the hybrid CI based on the infeasible QGLS estimator, see the Supplemental Appendix. The CP ($\times 100$) results of Table I using p_{ii}, rather than p_{ii}^*, are the same in all cases except case (i) $\rho = .99$, case (iv) $\rho = .5, .0$, and case (v) $\rho = .99$, where the differences are .1% (e.g., 94.2% versus 94.3%), and case (v) $\rho = .5, .0$, where the differences are .2% and .3%, respectively. There are no differences in the average lengths. For the symmetric two-sided CHR CI, the CP results and the average length results compared to the hybrid CI are similar to those in

Simulations for the symmetric two-sided subsampling CI of Romano and Wolf (2001) given in the Supplemental Appendix show that the latter CI under-covers substantially in some cases (e.g., its CP ($\times 100$) is 88.9, 88.3, 86.7 for $b = 8$, 12, 16 in case (ii) with $\rho = .0$). It is longer than the symmetric and equal-tailed CHR CI's when $\rho = .99$ in cases (i)-(v) and has similar average length (CP-corrected) in other cases. Hence, the CHR CI's out-perform the Romano and Wolf (2001) CI in the finite-sample cases considered.

Results reported in the Supplemental Appendix compare the CHR CI in the i.i.d. case with the analogous CI that employs the homoskedastic variance estimator.[22] The use of the HC4 variance matrix estimator increases the deviations of the CP's from 95.0 compared to the homoskedastic variance estimator somewhat, but even so, the deviations for the equal-tailed CI's are only equal to .3 on average over the five ρ values. It has no impact on the average lengths except when $\rho = .99$, in which case the impact is very small (8.3 for the equal-tailed CHR CI versus 8.1 for the equal-tailed homoskedastic variance CI). Hence, the CHR CI pays a very small price in the i.i.d. case for its robustness to conditional heteroskedasticity.

Table I. Coverage Probabilities and (CP-Corrected) Average Lengths of Nominal 95% Equal-Tailed Two-Sided CI's: CHR and Hybrid

			Coverage Probabilities ($\times 100$)					Average Lengths ($\times 100$) (CP-Corrected)				
Innovations	CI	ρ:	.99	.9	.5	.0	-.9	.99	.9	.5	.0	-.9
(i) GARCH(1,1)-	CHR		94.2	94.7	94.8	94.5	94.4	8.5	19	33	37	17
(.05,.9;.001)	Hyb		98.5	98.3	96.5	95.2	95.6	9.6	21	46	47	19
(ii) GARCH(1,1)-	CHR		94.2	94.6	94.7	94.1	94.2	8.8	20	37	43	18
(.15,.8;.2)	Hyb		98.0	97.9	96.0	94.3	95.0	9.8	22	49	52	21
(iii) I.i.d.	CHR		94.5	94.7	94.8	94.7	94.6	8.3	18	31	35	16
	Hyb		97.7	97.6	95.7	94.2	94.8	9.6	21	45	48	19
(iv) GARCH(1,1)-	CHR		94.3	94.5	94.4	93.7	94.1	9.2	21	42	49	20
(.25,.7;.2)	Hyb		98.4	98.3	95.9	94.8	95.1	9.5	22	51	54	21
(v) ARCH(4)-	CHR		94.5	94.3	93.9	93.2	94.0	9.6	23	48	56	22
(.3,.2,.2,.2;.2)	Hyb		98.5	98.2	95.9	94.2	95.4	9.1	22	53	57	21

Table I, although slightly better in both dimensions, see the Supplemental Appendix.

[22] The latter CI is Mikusheva's (2007a) modification of Stock's (1991) CI applied to the LS estimator of ρ, but with a stationary initial condition when $\rho < 1$, rather than a zero initial condition.

References

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

Andrews, D. W. K. and H.-Y. Chen (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12, 187-204.

Andrews, D. W. K., X. Cheng, and P. Guggenberger (2009): "Generic Results for Establishing the Asymptotic Size of Confidence Sets and Tests," Cowles Foundation Discussion Paper No. 1813, Yale University.

Andrews, D. W. K. and P. Guggenberger (2009): "Hybrid and Size-Corrected Subsampling Methods," *Econometrica,* 77, 721-762.

——— (2011): "Asymptotics for LS, GLS, and Feasible GLS Statistics in an AR(1) Model with Conditional Heteroskedasticity," *Journal of Econometrics,* forthcoming.

Bollerslev, T. (1987): "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, 542-547.

Chen, W. and R. Deo (2007): "A Smooth Transition to the Unit Root Distribution Via the Chi-square Distribution with Interval Estimation for Nearly Integrated Autoregressive Processes," unpublished manuscript, Stern School of Business, New York University.

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.

Engle, R., V. K. Ng, and M. Rothschild (1990): "Asset Pricing with a Factor-ARCH Covariance Structure-Empirical Estimates for Treasury Bills," *Journal of Econometrics,* 45, 213-237.

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

Gonçalves, S. and L. Kilian (2007): "Asymptotic and Bootstrap Inference for AR(∞) processes with Conditional Heteroskedasticity," *Econometric Reviews,* 26, 609-641.

Guo, B. and P. C. B. Phillips (2001): "Testing for Autocorrelation and Unit Roots in the Presence of Conditional Heteroskedasticity of Unknown Form," unpublished working paper, Department of Economics, UC Santa Cruz.

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

Kim, K. and P. Schmidt (1993): "Unit Root Tests with Conditional Heteroskedasticity," *Journal of Econometrics,* 59, 287-300.

Long, J. S. and L. H. Ervin (2000): "Using Heteroskedasticity Consistent Standard Errors in the Linear Regression Model," *The American Statistician,* 54, 217-224.

Ma, J., C. R. Nelson, and R. Startz (2007): "Spurious Inference in the GARCH(1, 1) Model when It Is Weakly Identified," *Studies in Nonlinear Dynamics and Econometrics,* 11, 1-27.

MacKinnon, J. G. and H. White (1985): "Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics,* 29, 53-57.

Mikusheva, A. (2007a): "Uniform Inferences in Autoregressive Models," *Econometrica,* 75, 1411-1452. Corrigendum, 2010, 78, 1773.

——— (2007b): "Supplement to 'Uniform Inferences in Autoregressive Models'," *Econometrica,* 75, 1411-1452. Available on the Econometric Society website.

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

Nankervis, J. C. and N. E. Savin (1996): "The Level and Power of the Bootstrap t-Test in the Trend Model with AR(1) Errors," *Journal of Business and Economic Statistics,* 14, 161-168.

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

Seo, B. (1999): "Distribution Theory for Unit Root Tests with Conditional Heteroskedasticity," *Journal of Econometrics,* 91, 113-144.

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

Supplemental Appendix

for

A Conditional-Heteroskedasticity-Robust Confidence Interval for the Autoregressive Parameter

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

Patrik Guggenberger
Department of Economics
UCSD

December 2010
Revised: July 2011

4 Outline

The Supplemental Appendix is organized as follows: Section 5 provides tables with the quantiles $c_h(\alpha)$ used to compute equal-tailed, symmetric, and one-sided CHR confidence intervals (CI's). Section 6 establishes the asymptotic false coverage probabilities (FCP's) of the equal-tailed CHR CI under local alternatives to the true values. Section 7 assesses the asymptotic and finite-sample price the CHR CI pays in the i.i.d. innovation case for its robustness to conditional heteroskedasticity. Section 8 provides the probabilities of obtaining disconnected CHR CI's. Section 9 defines symmetric two-sided CHR CI's and provides simulation results for symmetric two-sided confidence intervals, including the CHR CI, the hybrid and subsampling CI's of Andrews and Guggenberger (2009a), and the subsampling CI of Romano and Wolf (2001). Section 10 provides details regarding the implementation of the Monte Carlo simulations in the paper and the Supplement. Section 11 provides the proof of Theorem 1. Section 12 gives the proof of the asymptotic FCP result that is stated in Section 6. Section 13 provides a proof of the correct asymptotic size in the presence of conditional heteroskedasticity of Romano and Wolf's (2001) symmetric two-sided subsampling CI, which is based on a least squares (LS) based t statistic with a homoskedastic variance estimator (designed for the i.i.d. innovations case).

5 Tables of Critical Values

Table S-I reports the quantiles $c_h(.025)$ and $c_h(.975)$ (for a broad range of values of h) which are used to calculate 95% equal-tailed CHR CI's. Table S-II reports analogous quantiles used to calculate 90% equal-tailed CHR CI's. These tables also can be used for 97.5% and 95% lower and upper one-sided CI's. Tables S-III to S-V report $c_h^{sym}(1-\alpha)$ for $\alpha = .05, .01, .1$, respectively, and a range of h values, which are used to calculate 95%, 99%, and 90% symmetric two-sided CHR CI's (defined in (9.8) below).

For given α, $c_h(\alpha)$ (the α-quantile of J_h in (2.5)) is simulated by simulating the asymptotic distribution J_h. To do so, $300,000$ independent AR(1) sequences are generated from the model in (1.1) with innovations $U_i \sim iid\ N(0,1)$, $\mu = 0$, stationary startup, $n = 25,000$, and $\rho_h = 1 - h/n$. For each sequence, the test statistic $T_n(\rho_h)$ (defined in (2.1) but using the homoskedastic variance estimator) is calculated. Then, the simulated estimate of $c_h(\alpha)$ is the α-quantile of the empirical distribution of the $300,000$ realizations of the test statistic.

In Table S-I, the critical values do not reach the $h = \infty$ values of -1.96 and 1.96 for $h = 500$. Larger values of h, which would be needed only in very large samples, yield the following: $c_{1,000}(.025) = -2.02$, $c_{5,000}(.025) = -1.98$, $c_{10,000}(.025) = -1.97$, $c_{1,000}(.975) = 1.90$, $c_{5,000}(.975) = 1.93$, $c_{10,000}(.975) = 1.94$.

Table S-I(a). Values of $c_h(.025)$, the .025 Quantile of J_h, for Use with 95% Equal-Tailed Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
c_h	-3.13	-3.09	-3.06	-3.03	-3.00	-2.98	-2.93	-2.89	-2.85	-2.83	-2.80	-2.77	-2.75
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
c_h	-2.73	-2.71	-2.69	-2.65	-2.62	-2.59	-2.56	-2.54	-2.52	-2.50	-2.48	-2.47	-2.45
h	20	25	30	40	50	60	70	80	90	100	200	300	500
c_h	-2.39	-2.35	-2.32	-2.28	-2.24	-2.23	-2.21	-2.19	-2.18	-2.17	-2.11	-2.08	-2.05

Table S-I(b). Values of $c_h(.975)$, the .975 Quantile of J_h, for Use with 95% Equal-Tailed Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
$c_h(.975)$	.24	.31	.36	.41	.45	.50	.57	.64	.69	.74	.79	.84	.88
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
$c_h(.975)$	.92	.95	.99	1.06	1.12	1.17	1.21	1.25	1.29	1.32	1.34	1.37	1.39
h	20	25	30	40	50	60	70	80	90	100	200	300	500
$c_h(.975)$	1.47	1.51	1.55	1.61	1.65	1.67	1.69	1.71	1.73	1.74	1.81	1.83	1.86

Table S-II(a). Values of $c_h(.05)$, the .05 Quantile of J_h, for Use with 90% Equal-Tailed Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
c_h	-2.87	-2.83	-2.79	-2.76	-2.73	-2.70	-2.65	-2.61	-2.57	-2.54	-2.51	-2.48	-2.46
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
c_h	-2.44	-2.42	-2.39	-2.35	-2.32	-2.29	-2.26	-2.23	-2.21	-2.19	-2.18	-2.16	-2.14
h	20	25	30	40	50	60	70	80	90	100	200	300	500
c_h	-2.09	-2.05	-2.01	-1.97	-1.93	-1.91	-1.89	-1.87	-1.86	-1.85	-1.79	-1.76	-1.74

Table S-II(b). Values of $c_h(.95)$, the .95 Quantile of J_h, for Use with 90% Equal-Tailed Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
$c_h(.95)$	-.07	-.02	.04	.08	.13	.17	.25	.31	.37	.43	.48	.52	.57
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
$c_h(.95)$	.61	.64	.68	.75	.81	.87	.91	.95	.98	1.01	1.03	1.05	1.08
h	20	25	30	40	50	60	70	80	90	100	200	300	500
$c_h(.95)$	1.15	1.20	1.24	1.30	1.34	1.36	1.39	1.40	1.42	1.43	1.49	1.52	1.55

Table S-III. Values of $c_h^{sym}(.95)$, the .95 Quantile of $|J_h|$, for Use with 95% Symmetric Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
$c_h^{sym}(.95)$	2.87	2.83	2.79	2.76	2.73	2.70	2.65	2.61	2.57	2.54	2.51	2.49	2.46
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
$c_h^{sym}(.95)$	2.44	2.42	2.40	2.36	2.32	2.30	2.27	2.25	2.23	2.21	2.20	2.19	2.17
h	20	25	30	40	50	60	70	80	90	100	200	300	500
$c_h^{sym}(.95)$	2.13	2.10	2.08	2.05	2.03	2.02	2.01	2.01	2.00	2.00	1.98	1.97	1.96

Table S-IV. Values of $c_h^{sym}(.99)$, the .99 Quantile of $|J_h|$, for Use with 99% Symmetric Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
$c_h^{sym}(.99)$	3.44	3.40	3.38	3.35	3.32	3.30	3.26	3.22	3.19	3.16	3.14	3.11	3.09
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
$c_h^{sym}(.99)$	3.07	3.05	3.03	3.00	2.97	2.94	2.92	2.90	2.88	2.87	2.85	2.84	2.82
h	20	25	30	40	50	60	70	80	90	100	200	300	500
$c_h^{sym}(.99)$	2.77	2.74	2.72	2.68	2.66	2.65	2.64	2.63	2.62	2.61	2.59	2.58	2.57

Table S-V. Values of $c_h^{sym}(.90)$, the .90 Quantile of $|J_h|$, for Use with 90% Symmetric Two-Sided Confidence Intervals

h	0	.2	.4	.6	.8	1.0	1.4	1.8	2.2	2.6	3.0	3.4	3.8
$c_h^{sym}(.9)$	2.57	2.53	2.49	2.45	2.42	2.39	2.34	2.30	2.26	2.22	2.19	2.16	2.13
h	4.2	4.6	5.0	6.0	7.0	8.0	9.0	10	11	12	13	14	15
$c_h^{sym}(.9)$	2.11	2.09	2.07	2.02	1.98	1.95	1.93	1.91	1.89	1.87	1.86	1.85	1.83
h	20	25	30	40	50	60	70	80	90	100	200	300	500
$c_h^{sym}(.9)$	1.79	1.77	1.75	1.72	1.70	1.69	1.69	1.68	1.68	1.67	1.66	1.65	1.65

6 Local Asymptotic False Coverage Probabilities of the CHR CI

In this section, we determine the asymptotic FCP's of the CHR CI for sequences of local alternatives $\{\rho_n^* : n \geq 1\}$ to the true parameters $\{\rho_n : n \leq 1\}$. We provide results for the full spectrum of cases in which $n(1-\rho_n) \to h$ for (i) $0 \leq h < \infty$, (ii) $h = \infty$ and $\rho_n \to 1$, and (iii) $h = \infty$ and $\rho_n \to \rho_\infty < 1$. Asymptotic results of this sort are not available currently for any of the CI's in the literature under conditional homoskedasticity or conditional heteroskedasticity.

Theorem 1 of Andrews and Guggenberger (2011) shows that under sequences $\{(\rho_n, F_n) \in \Lambda : n \geq 1\}$ for which $n(1-\rho_n) \to h \in [0,\infty]$ the LS estimator $\widehat{\rho}_n$, the HC4 variance estimator $\widehat{\sigma}_n^2$, and the corresponding t statistic $T_n(\rho) = n^{1/2}(\widehat{\rho}_n - \rho)/\widehat{\sigma}_n$ satisfy:

$$(n^{1/2} d_n(\widehat{\rho}_n - \rho_n),\ d_n \widehat{\sigma}_n,\ T_n(\rho_n))' \to_d (N_h, S_h, T_h). \tag{6.1}$$

The result of Andrews and Guggenberger (2011) actually is for the HC variance estimator with 0 in place of p^*_{ii}. Because $\max_{i \le n} p^*_{ii} \le n^{-1/2} \to 0$, the proofs go through with $\widehat{\sigma}^2_n$ being the HC4 variance estimator.

In (6.1), (i) $\{d_n : n \ge 1\}$ is a sequence of constants defined below, (ii) $T_h = N_h/S_h$, (iii) when $h = \infty$, $N_h \sim N(0,1)$ and $S_h = 1$, and (iv) when $0 \le h < \infty$,

$$N_h = \int_0^1 I^*_{D,h}(r)dW(r) / \int_0^1 I^*_{D,h}(r)^2 dr \text{ and } S_h = \left(\int_0^1 I^*_{D,h}(r)^2 dr \right)^{-1/2}, \tag{6.2}$$

where $W(\cdot)$ and $I^*_{D,h}(\cdot)$ are defined in (2.4) of the paper.[23]

The constants $\{d_n : n \ge 1\}$ depend on (ρ_n, F_n), although their order of magnitude only depends on ρ_n. When $0 \le h < \infty$, $d_n = n^{1/2}$. When $h = \infty$,

$$d_n = E_{F_n} Y^{*2}_{n,i-1} / (E_{F_n} Y^{*2}_{n,i-1} U^2_{n,i})^{1/2}, \tag{6.3}$$

where by stationarity d_n does not depend on i.

The result in (6.1) shows that the local alternatives for which the CHR CI has non-trivial asymptotic FCP's (i.e., asymptotic FCP's less than $1 - \alpha$) are of the form:

$$\rho^*_n = \rho_n - \frac{\delta_n}{n^{1/2} d_n} \tag{6.4}$$

for any sequence of constants $\{\delta_n : n \ge 1\}$ such that $\delta_n \to \delta \in R$ and $h + \delta \ge 0$ if $0 \le h < \infty$.

If $0 \le h < \infty$, then the local alternatives $\{\rho^*_n : n \ge 1\}$ are n^{-1}-local alternatives from the true values $\{\rho_n : n \ge 1\}$ because $n^{1/2} d_n = n$.

If $h = \infty$, $\rho_n \to \rho_\infty < 1$, and $\lim_{n \to \infty} d_n = d_\infty \in (0, \infty)$, then

$$\rho^*_n = \rho_n - (\delta d^{-1}_\infty / n^{1/2})(1 + o(1)). \tag{6.5}$$

In this case, $\{\rho^*_n : n \ge 1\}$ are $n^{-1/2}$-local alternatives from $\{\rho_n : n \ge 1\}$. Note that the condition on d_n is not stringent. For example, it holds if (ρ_n, F_n) does not depend on n and $E_F(Y^*_i - E_F Y^*_i)^2 U^2_i \in (0, \infty)$.

If $h = \infty$ and $\rho_n \to 1$, then

$$d_n = (1 - \rho^2_n)^{-1/2}(1 + o(1)) \tag{6.6}$$

by equation (11) of Andrews and Guggenberger (2011) and the first two results of Lemma 6 in Andrews and Guggenberger (2011). The $n^{1/2}(1 - \rho^2_n)^{-1/2}$ rate of convergence of the LS estimator in this case was first obtained by Giraitis and Phillips (2006). In the present case, the local alternatives are of the form

$$\rho^*_n = \rho_n - \frac{\delta(1 - \rho^2_n)^{1/2}}{n^{1/2}}(1 + o(1)), \tag{6.7}$$

[23] When $\widehat{\rho}_n$ is the LS estimator, then (in the notation of Andrews and Guggenberger (2011)) $\widehat{\phi}_{n,i} = \phi_{n,i} = 1 \ \forall i \le n$ and the constants in Thm. 1 of Andrews and Guggenberger (2011) simplify: $h_{2,1} = h_{2,2} = \lim_{n \to \infty} E_{F_n} U^2_i$ and $h_{2,5} = h_{2,7} = 1$. This yields the form of the asymptotic distributions in (6.1) and (6.2).

which constitute deviations that are smaller than $O(n^{-1/2})$ and larger than $O(n^{-1})$.

By definition, let $h + \delta = \infty$ if $h = \infty$ and $\delta \in R$.

We have the following asymptotic FCP result.

Theorem 2 *Let* $\alpha \in (0, 1)$. *Let* $\{(\rho_n, F_n) \in \Lambda : n \geq 1\}$ *be any sequence of true parameters and distributions for which* $n(1 - \rho_n) \to h \in [0, \infty]$, $\rho_n \to \rho_\infty \in (-1, 1]$, *and* $n^{1/2} d_n \to \infty$ *if* $\rho_\infty < 1$. *Let* $\{\rho_n^* : n \geq 1\}$ *be any sequence of alternative parameters that satisfies* $\rho_n^* = \rho_n - \delta_n/(n^{1/2} d_n)$, *where* $\delta_n \to \delta \in R$ *and* $h + \delta \geq 0$ *if* $0 \leq h < \infty$. *Then, the equal-tailed nominal* $1-\alpha$ *confidence interval* $CI_{CHR,n}$ *for the AR parameter* ρ *satisfies*

$$\lim_{n \to \infty} P_{\lambda_n}(\rho_n^* \in CI_{CHR,n}) = P(c_{h+\delta}(\alpha/2) \leq T_h + \delta/S_h \leq c_{h+\delta}(1 - \alpha/2)).$$

Comments. 1. If $h = \infty$, then $h + \delta = \infty$, $c_{h+\delta}(\alpha/2) = -z_{1-\alpha/2}$, $c_{h+\delta}(1 - \alpha/2) = z_{1-\alpha/2}$, $T_h \sim N(0, 1)$, $S_h = 1$, and the limit FCP in the result of Theorem 2 equals $P(|T_h + \delta| \leq z_{1-\alpha/2})$, which is less than $1 - \alpha$ for all $\delta \neq 0$. This result holds even if $\rho_n \to 1$ provided $n(1 - \rho_n) \to \infty$. In this case, the distance of the local alternatives from the true values depends on how fast ρ_n goes to one via $1 - \rho_n^2$, as in (6.7), but the form of the asymptotic FCP is the same as when $\rho_n \to \rho_\infty < 1$. If $0 \leq h < \infty$, then the limit FCP in the result of Theorem 2 is not a standard quantity, but it can be simulated quite easily.

2. Using the results of Andrews and Guggenberger (2011), one could establish analogous results to those in Theorem 2 for the hybrid CI of Andrews and Guggenberger (2009a) and the symmetric subsampling CI of Romano and Wolf (2001). For brevity, we do not do so here.

7 Price for Robustness of the CHR CI

This section assesses the price the CHR CI pays, in terms of asymptotic FCP's and finite-sample average lengths, in the i.i.d. innovation case to obtain robustness to conditional heteroskedasticity.

First, we show that the asymptotic price (to first order) is zero. Consider the CHR CI and the same CI but constructed with the homoskedastic variance estimator in place of the heteroskedasticity-consistent variance estimator. The latter CI is referred to as the MSS CI, because it is the same as Mikusheva's (2007) modification of Stock's (1991) CI applied to the LS estimator of ρ, but with a stationary initial condition when $\rho < 1$, rather than a zero initial condition. The latter affects the asymptotic distribution of the t statistic and hence the definition of the critical values. Under i.i.d. innovations, these two CI's have the same asymptotic FCP's (and CP's).[24] The asymptotic FCP's of the CHR CI are given in Theorem 2. Identical results for the

[24] Note that the MSS CI does not have correct asymptotic size when conditionally heteroskedastic innovation distributions are included in the parameter space. This is because the MSS t statistic

MSS CI under i.i.d. innovations hold because the two test statistics have the same asymptotic distributions in this case both when $\rho_n \to 1$ and when $\rho_n \to \rho_\infty < 1$. When $\rho_n \to 1$ this holds by the proof given in Section 13 below of the correct asymptotic size of Romano and Wolf's (2001) symmetric subsampling CI. When $\rho_n \to \rho_\infty < 1$ and the innovations are i.i.d., it holds by standard results because the two variance estimators have the same probability limit.

Next, because the asymptotic price is zero, we use simulations to assess the finite-sample price. The data generating process considered is the i.i.d. standard normal innovation case (i.e., case (iii) in the paper). The sample size is $n = 130$.

Table S-VI reports CP and (CP-corrected) average lengths for the CHR and MSS CI's. Results are given for both equal-tailed and symmetric two-sided CI's.

Table S-VI shows that the use of the HC4 variance matrix estimator increases the deviations of the CP's ($\times 100$) from 95.0 slightly compared to the homoskedastic variance estimator. It has no impact on average length except when $\rho = .99$, in which case the impact is very small. Thus, the CHR CI pays a very small price in the i.i.d. case for its robustness to conditional heteroskedasticity.

Note that CI's analogous to the CHR CI's (equal-tailed, symmetric, and one-sided) based on the HC4 variance estimator can be defined with other versions of the HC variance matrix, such as the HC, HC1, HC2, and HC3 estimators defined in MacKinnon and White (1985), but we find that the HC4 variance estimator gives the best finite-sample CP's and the choice has very little effect on the (CP-corrected) average lengths. As noted in a footnote in Section 3 of the paper, the HC4 estimator and the HC3 (without the $(n-1)/n$ term) estimator have almost the same finite-sample properties for the cases in Table I.

Next, we briefly discuss simulation comparisons between the (equal-tailed) CHR CI and the infeasible hybrid CI which is based on the infeasible QGLS estimator. By definition, the infeasible QGLS estimator takes the GARCH$(1,1)$ (or ARCH(4)) specification as known and its parameter values as known. Simulations for the hybrid CI based on the infeasible QGLS estimator (not reported) show that it over-covers in many cases and its CP's exceed those of the FQGLS hybrid CI in almost all cases. In consequence, its average lengths are the same or slightly longer than those of the FQGLS hybrid CI in cases (i)-(iv) and only slightly shorter in case (v), reported in Table I of the paper. Hence, the CHR CI out-performs the infeasible QGLS hybrid CI, as well as the FQGLS hybrid CI in the cases considered.

has an asymptotic distribution that depends on the form of the conditional heteroskedasticity and is not standard normal when $\rho_n \to \rho_\infty < 1$, but the critical value is taken from the i.i.d. innovation case which is a standard normal quantile when $\rho_n \to \rho_\infty < 1$.

Table S-VI. Coverage Probabilities and (CP-Corrected) Average Lengths of Nominal 95% Two-Sided CI's with I.i.d. Innovations: CHR and MSS

			Coverage Probabilities (×100)					Average Lengths (×100) (CP-Corrected)				
CI-Type	CI	ρ:	.99	.9	.5	.0	-.9	.99	.9	.5	.0	-.9
Equal-tailed	CHR		94.5	94.7	94.8	94.7	94.6	8.3	18	31	35	16
	MSS		95.0	94.9	95.1	94.9	94.6	8.1	18	31	35	16
Symmetric	CHR		94.5	95.2	95.1	94.8	95.6	10.6	19	31	35	16
	MSS		95.1	95.2	95.4	95.2	95.4	10.4	19	31	35	16

8 Disconnected CHR Confidence Intervals

It is possible for CHR CI's to consist of two disjoint intervals of the form $[a,b]\cup[c,1]$ for $-1+\varepsilon \leq a < b < c \leq 1$.[25] To see why, consider the nominal 95% equal-tailed CHR CI. The critical values $c_h(.025)$ and $c_h(.975)$ are increasing and concave as functions of h, see Tables S-I and S-II above. Thus, $c_{n(1-\rho)}(.025)$ and $c_{n(1-\rho)}(.975)$ are decreasing concave functions of ρ. Viewed as a function of ρ on $[-1,1]$, the critical value functions are essentially flat and take values close to -1.96 and 1.96, respectively, for most of $[-1,1]$ and dip down to the values -3.13 and $.24$ as ρ approaches one.

The test statistic is a linear function of ρ with negative slope. The test statistic takes the value 0 for the value of ρ equal to $\widehat{\rho}_n$. The CI consists of all values ρ where the linear test statistic function lies between the mostly horizontal critical value curves. Drawing the corresponding picture, one can see that disconnected CI's can occur if the linear test statistic line cuts across the lower critical value curve as it dips near one and intersects with it at two places. This only occurs for a very small range of values of $n^{1/2}$, $\widehat{\rho}_n$, and $\widehat{\sigma}_n$.

No such disconnected CI feature occurs if the test statistic line intersects the upper critical value curve in two places because ρ values between the two intersection points are in the CI, not excluded from it, in this case.

Table S-VII provides simulated values of the probability that the CHR equal-tailed and symmetric CI's are disconnected for the five cases considered in Table I of the paper.[26] The sample size is $n = 130$. In cases (i)-(iii) except for $\rho = .9$ and in cases (iv) and (v) except for $\rho = .9, .5$, the probability of a disconnected CI is essentially zero. For $\rho = .9$ in cases (i)-(iii) the probability is still quite small ($\leq 5/1000$). For $\rho = .9, .5$ in cases (iv) and (v), which are the stronger and less realistic forms of conditional heteroskedasticity, the probabilities are larger, but still small ($\leq 19/1000$).

[25] The same is true of Mikusheva's (2007) modification of Stock's (1991) CI.

[26] These results are based on 30,000 simulation repetitions with the asymptotic critical values computed using 100,000 repetitions and $n = 30,000$.

Table S-VII. Probabilities of Obtaining Disconnected CHR CI's

Innovations	CI	ρ: .99	.9	.5	.0	-.9
		Probability of Disconnected CHR CI				
(i) GARCH(1, 1)-	eq-tail	.0001	.0008	.0000	.0000	.0000
(.05,.9;.001)	sym	.0001	.0028	.0000	.0000	.0000
(ii) GARCH(1, 1)-	eq-tail	.0001	.0023	.0020	.0000	.0000
(.15,.8;.2)	sym	.0002	.0050	.0011	.0000	.0000
(iii) I.i.d.	eq-tail	.0000	.0008	.0000	.0000	.0000
	sym	.0002	.0014	.0000	.0000	.0000
(iv) GARCH(1, 1)-	eq-tail	.0001	.0035	.0089	.0002	.0000
(.25,.7;.2)	sym	.0010	.0074	.0054	.0002	.0000
(v) ARCH(4)-	eq-tail	.0003	.0054	.0190	.0020	.0000
(.3,.2,.2,.2;.2)	sym	.0008	.0102	.0170	.0016	.0000

9 Simulation Results for Symmetric Two-Sided Confidence Intervals

Let $CI_{CHR,n}^{sym}$ denote the symmetric two-sided nominal $1-\alpha$ CHR CI that is analogous to the equal-tailed CHR CI introduced in the paper. It is defined as follows:

$$CI_{CHR,n}^{sym} = \{\rho \in [-1+\varepsilon, 1] : |T_n(\rho)| \leq c_h^{sym}(1-\alpha) \text{ for } h = n(1-\rho)\}, \tag{9.8}$$

where $c_h^{sym}(1-\alpha)$ is the $1-\alpha$ quantile of the asymptotic distribution $|J_h|$ of $|T_n(\rho)|$.

Under the conditions of Theorem 1, the symmetric two-sided nominal $1-\alpha$ CHR CI has asymptotic size equal to $1-\alpha$ and is asymptotically similar. Under the conditions of Theorem 2, the FCP's of the symmetric two-sided CHR CI satisfy

$$\lim_{n\to\infty} P_{\lambda_n}(\rho_n^* \in CI_{CHR,n}) = P(|T_h + \delta/S_h| \leq c_{h+\delta}^{sym}(1-\alpha)). \tag{9.9}$$

(The proofs of these results are analogous to those given for the equal-tailed CHR CI in Sections 11 and 12 below.)

Table S-VIII reports simulation results analogous to those in Table I except for symmetric two-sided CI's, rather than equal-tailed CI's. It reports results for the symmetric CHR, hybrid, and FQGLS subsampling (Sub_{GLS}) CI's. The hybrid and Sub_{GLS} CI's are proposed in AG09a. They are based on the FQGLS estimator with

Table S-VIII. Coverage Probabilities and (CP-Corrected) Average Lengths of Nominal 95% Symmetric Two-Sided CI's: CHR, Hybrid, and Sub$_{GLS}$

			Coverage Probabilities (×100)					Average Lengths (×100) (CP-Corrected)				
Innovations	CI	ρ:	.99	.9	.5	.0	-.9	.99	.9	.5	.0	-.9
(i) GARCH(1, 1)-	CHR		94.6	95.2	95.1	94.8	95.4	11	20	33	37	16
(.05,.9;.001)	Hyb		95.9	98.0	97.2	96.1	95.8	15	28	46	45	19
	Sub$_{GLS}$		95.9	98.0	97.0	95.6	95.0	15	28	45	44	18
(ii) GARCH(1, 1)-	CHR		94.9	95.3	95.0	94.5	95.3	11	21	37	42	18
(.15,.8;.2)	Hyb		96.0	97.8	97.0	96.0	95.7	15	29	49	49	20
	Sub$_{GLS}$		96.0	97.8	96.8	95.4	95.1	15	29	49	48	20
(iii) I.i.d.	CHR		94.5	95.2	95.1	94.8	95.6	11	19	31	35	16
	Hyb		95.3	97.4	96.8	95.6	95.2	15	27	45	45	19
	Sub$_{GLS}$		95.3	97.4	96.6	95.1	94.5	15	27	45	44	19
(iv) GARCH(1, 1)-	CHR		95.3	95.4	94.8	94.1	95.2	11	23	41	48	19
(.25,.7;.2)	Hyb		96.5	98.0	97.2	96.2	95.9	15	30	51	51	20
	Sub$_{GLS}$		96.5	98.0	97.0	95.6	95.4	15	30	51	50	20
(v) ARCH(4)-	CHR		95.8	95.5	94.4	93.6	95.0	12	24	47	55	21
(.3,.2,.2,.2;.2)	Hyb		96.9	98.1	97.2	96.0	96.4	15	30	53	53	21
	Sub$_{GLS}$		96.9	98.1	96.9	95.5	95.9	15	30	53	52	21

standard heteroskedasticity-consistent variance estimator (as in Table I), coupled with hybrid (fixed/subsampling) and subsampling critical values, respectively, see Section 10 for more details. The sample size is $n = 130$ and the subsample size is $b = 12$.

The results in Table S-VIII are similar to those in Table I, but the CHR CI performs slightly better in terms coverage probabilities (CP's) and in terms of average length compared to the hybrid CI. The hybrid and Sub$_{GLS}$ CI's have very similar finite-sample properties (which is expected because they have the same asymptotic properties in terms of CP's and FCP's).

Next, Table S-IX reports results analogous to those of Table I but for the symmetric two-sided CHR and symmetric two-sided subsampling CI of Romano and Wolf (2001) (Sub$_{RW}$). The Sub$_{RW}$ CI is based on the LS estimator and the homoskedastic variance matrix estimator. For the Sub$_{RW}$ CI, we compute results for subsample sizes $b = 8, 12, 16, 20$. In Table S-IX, we report results for $b = 8, 12$, because they provide the best results in terms of CP's and average lengths.

We discuss the results for Sub$_{RW}$ CI with $b = 8$ because they are better than those for $b = 12$ in terms of CP's. Table S-IX shows that the Sub$_{RW}$ CI exhibits problems

Table S-IX. Coverage Probabilities and (CP-Corrected) Average Lengths of Nominal 95% Symmetric Two-Sided CI's: CHR and the Romano and Wolf (2001) Subsampling CI (Sub_{RW}) with Subsample Sizes $b = 8$ and $b = 12$

			Coverage Probabilities (×100)					Average Lengths (×100) (CP-Corrected)				
Innovations	CI	ρ:	.99	.9	.5	.0	-.9	.99	.9	.5	.0	-.9
(i) GARCH(1,1)-(.05,.9;.001)	CHR		94.6	95.2	95.1	94.8	95.4	11	20	33	37	16
	Sub_{RW} $b=8$		95.3	97.3	95.2	91.8	93.4	14	23	34	39	14
	Sub_{RW} $b=12$		93.4	96.1	93.6	90.8	91.8	15	22	35	40	15
(ii) GARCH(1,1)-(.15,.8;.2)	CHR		94.9	95.3	95.0	94.5	95.3	11	21	37	42	18
	Sub_{RW} $b=8$		94.7	96.5	93.0	88.9	92.2	15	23	38	45	15
	Sub_{RW} $b=12$		93.0	95.2	91.5	88.3	90.6	15	22	40	45	17
(iii) I.i.d.	CHR		94.5	95.2	95.1	94.8	95.6	11	19	31	35	16
	Sub_{RW} $b=8$		95.6	97.6	95.8	92.8	93.8	14	23	34	36	13
	Sub_{RW} $b=12$		93.7	96.4	94.2	91.6	92.3	15	22	34	37	15
(iv) GARCH(1,1)-(.25,.7;.2)	CHR		95.3	95.4	94.8	94.1	95.2	11	23	41	48	19
	Sub_{RW} $b=8$		94.2	95.4	90.5	85.9	91.3	15	23	43	52	16
	Sub_{RW} $b=12$		92.6	94.1	89.2	85.6	90.0	15	23	45	47	17
(v) ARCH(4)-(.3,.2,.2,.2;.2)	CHR		95.8	95.5	94.4	93.6	95.0	12	24	47	55	21
	Sub_{RW} $b=8$		93.8	93.8	87.4	82.5	90.5	16	25	52	54	17
	Sub_{RW} $b=12$		92.4	92.5	86.2	82.9	89.5	17	26	46	49	18

with under-coverage in some cases. For example, when $\rho = .0$, its CP's (×100) in the five cases considered lie in the interval $[82.5, 92.8]$, whereas those of the CHR CI lie in $[93.6, 94.8]$. When $\rho = .5$, the CP's of the Sub_{RW} CI in the five cases lie in $[87.4, 95.8]$, whereas those of the CHR CI lie in $[94.4, 95.1]$. The average (CP-corrected) lengths of the Sub_{RW} CI's are noticeably longer than those of the CHR CI for $\rho = .99$ (for all five specifications of the conditional heteroskedasticity), but similar for most other parameter values.

In sum, the symmetric two-sided subsampling CI of Romano and Wolf (2001) does not perform as well as the symmetric CHR CI due its noticeable under-coverage in some cases.

Comparing the results of the Sub_{GLS} and Sub_{RW} CI's in Tables S-VIII and S-IX, it is clear that the use of the feasible FQGLS estimator of ρ combined with a HC variance matrix estimator, compared to the LS estimator of ρ combined with the homoskedastic variance estimator, improves the finite-sample coverage probabilities of the subsampling CI's noticeably.

Additional simulations show that much of the difference is due to the use of the HC variance matrix estimator (even though the latter is not necessary to obtain correct asymptotic size). Specifically, we computed CP's for the symmetric subsampling CI based on the LS estimator combined with the standard heteroskedasticity consistent variance estimator, denoted $\mathrm{Sub}_{LS,Het}$, which differs from Sub_{RW} only in the choice of the variance matrix estimator. For i.i.d. innovations, $n = 130$, $b = 12$, and $\rho = .99, .9, .5, .0, -.9$, the Sub_{RW} and $\mathrm{Sub}_{LS,Het}$ CI's have CP's ($\times 100$): $(93.7, 98.7)$, $(96.4, 99.1)$, $(94.2, 98.4)$, $(91.6, 97.3)$, $(92.3, 96.6)$. Hence, even in the i.i.d. innovations case, there are substantial differences between the $n = 130$ finite-sample CP's of the Sub_{RW} and $\mathrm{Sub}_{LS,Het}$ CI's.

10 Monte Carlo Details

The hybrid and Sub_{GLS} CI's reported in Tables I and S-VIII are based on a t statistic constructed using a FQGLS estimator of ρ that employs estimators $\{\widehat{\phi}^2_{n,i} : i \leq n\}$ of the conditional variances $\{\sigma_i^2 : i \leq n\}$. The studentized t statistic is $T_{GLS,n}(\rho) = n^{1/2}(\widehat{\rho}_{GLS,n} - \rho)/\widehat{\sigma}_{GLS,n}$, where $\widehat{\rho}_{GLS,n}$ is the LS estimator from the regression of $Y_i/\widehat{\phi}_i$ on $Y_{i-1}/\widehat{\phi}_i$ and $1/\widehat{\phi}_i$ for $i = 1, ..., n$ and $\widehat{\sigma}^2_{GLS,n}$ is the $(1,1)$ element of the standard heteroskedasticity-robust variance estimator for the LS estimator in the preceding regression (which does not employ the HC4 adjustment factor $1/(1 - p^*_{ii})$).

The estimators $\{\widehat{\phi}^2_{n,i} : i \leq n\}$ are based on a GARCH$(1,1)$ parametric specification of the conditional heteroskedasticity. The GARCH$(1,1)$ model is estimated using the closed-form estimator of Kristensen and Linton (2006) applied to the LS residuals. This estimator is employed in the simulations because it is very quick to compute. More precisely, we use two Newton-Raphson iterations (see Kristensen and Linton's (2006) equation (17)), and we initialize the iteration using their closed-form estimator (see their equation (10) on p. 326) implemented with $w_1 = w_2 = w_3 = 1/3$ and with their $\widehat{\phi}$ Winsorized to the interval $[.001, .999]$. In each iteration step, we initialize the $\widehat{\sigma}^2_{k,t}$ (see their p. 329, line 5 from the bottom) by setting it equal to the squared first data observation.[27]

In the simulations using an ARCH(4) data generating process, the GARCH$(1,1)$ specification is incorrect. Nevertheless, the hybrid and Sub_{GLS} CI's still have correct asymptotic size, see Andrews and Guggenberger (2009a).

The asymptotic distribution of the FQGLS estimator in the n^{-1}-local to unity case depends on the parameter $h_{2,7} = Corr(U_i, U_i/\phi_i^2)$, where ϕ_i^2 is the conditional variance of the innovations based on the GARCH$(1,1)$ specification (which may or may not be correctly specified), with GARCH$(1,1)$ parameter values evaluated at the probability limit of the GARCH parameter estimators, see Andrews and Guggenberger (2011).

[27] For simplicity, this estimator is not discretized and the GARCH(1, 1) process is not truncated to conform to the theoretical results given in the Section 3.4 of Andrews and Guggenberger (2011) for the asymptotic equivalence of feasible and infeasible QGLS statistics.

For the five processes considered in the simulations, $h_{2,7}$ equals .98, .86, 1.00, .74, and .54, respectively.

For the equal-tailed two-sided nominal $1-\alpha$ hybrid CI, the upper critical value is the maximum of the subsampling critical value for $1-\alpha/2$ and the standard normal quantile $z_{1-\alpha/2}$. The lower critical value is the minimum of the subsampling critical value for $\alpha/2$ and $z_{\alpha/2}$. For the symmetric two-sided nominal $1-\alpha$ subsampling CI, i.e., Sub_{GLS}, the test statistic is the absolute value of $T_{GLS,n}(\rho)$ and the critical value is the $1-\alpha$ sample quantile of the absolute values of the subsample t statistics. For the symmetric hybrid CI, the test statistic is the same, but the critical value is the maximum of the latter subsampling critical value and $z_{1-\alpha/2}$.

The subsample FQGLS t statistics use the full-sample estimator of the conditional heteroskedasticity $\{\widehat{\phi}_{n,i} : i \leq n\}$, which is justified because the feasible QGLS and infeasible QGLS t statistics are asymptotically equivalent in the full sample and in subsamples. In addition, the subsample FQGLS t statistics are defined with the full-sample FQGLS estimators $\widehat{\rho}_{GLS,n}$ in place of the null value ρ in the expression for $T_{GLS,n}(\rho)$. That is, the subsample t statistic is of the form: $b^{1/2}(\widehat{\rho}_{GLS,b}-\widehat{\rho}_{GLS,n})/\widehat{\sigma}_{GLS,b}$, where $\widehat{\rho}_{GLS,b}$ and $\widehat{\sigma}_{GLS,b}$ are the estimators based on the subsample of size b.

The Sub_{RW} CI reported in Table S-IX is based on the LS estimator and the homoskedastic variance estimator, as in Romano and Wolf (2001). The Sub_{RW} critical values are based on subsample statistics that are defined analogously to those for Sub_{GLS} except that $\widehat{\rho}_{GLS,n}$, $\widehat{\rho}_{GLS,b}$, and $\widehat{\sigma}_{GLS,b}$ are replaced by the full-sample and subsample LS estimators and the subsample homoskedastic standard error estimator, respectively.

The sample size, subsample size, and number of subsamples (for the subsampling and hybrid CI's) employed are 130, 12, and 119. For the Sub_{RW} CI we also consider results for subsample sizes $b = 8, 16, 20$ (with $n - b + 1 = 131 - b$ subsamples in each case).

To mitigate the effect of the initialization on the (G)ARCH processes, we simulate time series of innovations of length 1130 and eliminate the first 1000 observations.

The CHR CP and (CP-corrected) average length results in Tables I, S-VI, S-VIII, and S-IX are computed in two steps. First, we simulate the asymptotic critical values using $30,000$ repetitions, $n = 25,000$, and standard normal innovations. Then, using these critical values, we simulate the CP's and (CP-corrected) average lengths using $30,000$ repetitions and $n = 130$. To compute CP's, all we need to consider are the true ρ values of interest: $.99, .9, .5, .0, -.9$ and one or two quantiles, such as $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$, where $h = n(1-\rho)$, for equal-tailed CHR CI's. However, to compute the average lengths we need to determine which values of ρ are in the CI. To do this, we consider 401 equally spaced grid points for ρ in $[-1, 1]$ and we determine whether each of these points is in the CI or not. This requires computing the appropriate quantiles for each of the 401 ρ values, such as $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$ for $h = n(1-\rho)$ and $n = 130$. Furthermore, to carry out CP-correction of the average lengths, we need to determine the value α' such that the nominal $1-\alpha'$ CI has finite-sample CP equal to the desired value $1-\alpha$ for the data generating process being considered. To do this, we need to compute the asymptotic critical values not only for one or two quantiles, but rather, for a broad range of potential values of $1-\alpha'$. Hence, when computing the asymptotic critical values in the first step, we consider a grid of $3,200$ values of $100 \times \beta$ (taking values in $[20, 99.999]$ with a step size of .025) and 401 values of ρ and we compute $c_h(\beta)$ for $h = n(1-\rho)$ for all of these values. Given this 3200 by 401 dimensional matrix of $c_h(\beta)$ values, we compute the CP's and (CP-corrected) average lengths of the CHR CI in the second step.

For the subsampling and hybrid CI's, we use the same grid of 401 ρ values and $3,200$ values of $100 \times \beta$ when computing the CP-corrected average lengths.

11 Proof of Theorem 1

The proof of Theorem 1 relies heavily on Theorem 1 of Andrews and Guggenberger (2011), which provides the asymptotic distribution of the t statistic under certain drifting sequences of distribution, as specified in (2.5) of the paper. As noted above, the proofs in Andrews and Guggenberger (2011) need to be adjusted slightly because of the p_{ii}^* term in the HC4 variance estimator, which does not appear in the variance estimator in Andrews and Guggenberger (2011). Because $\max_{i \leq n} p_{ii}^* \leq n^{-1/2} \to 0$, the adjustment is simple. Theorem 1 of Andrews and Guggenberger (2011) applies because the restrictions imposed in the definition of Λ include those imposed in Assumption INNOV in Andrews and Guggenberger (2011) simplified to the case where $\phi_{n,i} = 1$ in that paper.[28]

The asymptotic results in (2.5) are sufficient to determine the asymptotic size of the CHR CI and to show that it is asymptotically similar using Theorem 2.1 of Andrews, Cheng, and Guggenberger (2009) (ACG).

To describe the result in that paper, using general terminology, let $\{CS_n : n \geq 1\}$ be a sequence of confidence sets for a parameter $r(\lambda)$, where λ indexes the true distribution of the observations. The parameter space for λ is denoted by Λ. Let $CP_n(\lambda)$ denote the coverage probability of CS_n under λ. The *asymptotic size* of CS_n is defined as

$$AsySz = \liminf_{n\to\infty} \inf_{\lambda \in \Lambda} CP_n(\lambda). \tag{11.10}$$

[28] Note that Assumption INNOV(v) in Andrews and Guggenberger (2011) is only needed for the asymptotic results in the case where $\rho_n \to 1$. Assumption INNOV(v) in the case where $\phi_{n,i} = 1$ reduces to the smaller eigenvalue of the 2×2 matrix with diagonal elements $E_{F_n} Y_0^{*2} U_1^2$ and $E_{F_n} U_1^2$ and off-diagonal element equal to $E_{F_n} Y_0^* U_1^2$ being larger than δ for all n sufficiently large. By Lemma 6 in Andrews and Guggenberger (2011) $E_{F_n} Y_0^{*2} U_1^2 \to \infty$ and $E_{F_n} Y_0^* U_1^2 = O(1)$ when $\rho_n \to 1$. Assuming $E_{F_n} U_1^2 \geq \delta$ then ensures that the smaller eigenvalue of the matrix above is not smaller than δ for all large enough n. This can be seen by straightforward calculations using l'Hôpital's rule.

We say a sequence $\{CS_n : n \geq 1\}$ is *asymptotically similar* (in a uniform sense) if

$$\liminf_{n\to\infty} \inf_{\lambda\in\Lambda} CP_n(\lambda) = \limsup_{n\to\infty} \sup_{\lambda\in\Lambda} CP_n(\lambda). \tag{11.11}$$

Corollary 2.1(c) of ACG shows that under Assumptions B1* and B2, stated below, $\{CS_n : n \geq 1\}$ is asymptotically similar and satisfies $AsySz = CP$. Let $\{h_n(\lambda) : n \geq 1\}$ be a sequence of functions on Λ, where $h_n(\lambda) = (h_{n,1}(\lambda), ..., h_{n,J}(\lambda), h_{n,J+1}(\lambda))'$, $h_{n,j}(\lambda) \in R\ \forall j \leq J$, and $h_{n,J+1}(\lambda) \in \mathcal{T}$ for some compact pseudo-metric space $\mathcal{T}$.

Assumption B1*. For any sequence $\{\lambda_n \in \Lambda : n \geq 1\}$ for which $h_n(\lambda_n) \to h \in H$, $CP_n(\lambda_n) \to CP$ for some constant $CP \in [0,1]$ and some index set H.

Assumption B2. For any subsequence $\{p_n\}$ of $\{n\}$ and any sequence $\{\lambda_{p_n} \in \Lambda : n \geq 1\}$ for which $h_{p_n}(\lambda_{p_n}) \to h \in H$, there exists a sequence $\{\lambda_n^* \in \Lambda : n \geq 1\}$ such that $h_n(\lambda_n^*) \to h \in H$ and $\lambda_{p_n}^* = \lambda_{p_n}\ \forall n \geq 1$.

To prove Theorem 1, it is sufficient to verify Assumptions B1* and B2 for $CS_n = CI_{CHR,n}$. In the present case $\lambda = (\rho, F)$, $r(\lambda) = \rho$, $h_n(\lambda) = n(1-\rho) \in R$, $H = [0, \infty]$, and the parameter space Λ is defined in (2.3). Thus, $J = 1$ and there is no $(J+1)$-st component in $h_n(\lambda)$. For Assumption B1*, consider a sequence $\{\lambda_n = (\rho_n, F_n) \in \Lambda : n \geq 1\}$ for which $h_n(\lambda_n) \to h \in H$, i.e., $\rho_n = 1 - h_n/n$ and $h_n \to h \in [0, \infty]$. We have $CP_n(\lambda_n) = P_{\lambda_n}(\rho_n \in CI_{CHR,n}) = P_{\lambda_n}(c_{h_n}(\alpha/2) \leq T_n(\rho_n) \leq c_{h_n}(1-\alpha/2))$. By (2.5) of the paper, we have $T_n(\rho_n) \to_d J_h$ under $\{\lambda_n \in \Lambda : n \geq 1\}$. In addition, $c_{h_n}(\beta) \to c_h(\beta)$ for $\beta = \alpha/2, 1-\alpha/2$. To obtain the latter result, we apply Lemma 5(a) in Andrews and Guggenberger (2010) noting that J_h is increasing at its β-quantile $c_h(\beta)$ for $\beta \in (0,1)$. By the definition of convergence in distribution and continuity of J_h, it follows that $CP_n(\lambda_n) \to 1-\alpha$. Assumption B1* therefore holds with $CP = 1-\alpha$ for all $h \in H$.

For Assumption B2, assume we are given $\{\lambda_{p_n} \in \Lambda : n \geq 1\}$ for a subsequence $\{p_n\}$ of $\{n\}$ such that $h_{p_n}(\lambda_{p_n}) \to h \in H$. Define $\{\lambda_n^* : n \geq 1\}$ by (i) $\lambda_{p_n}^* = \lambda_{p_n}$ $\forall n \geq 1$, (ii) when $h < \infty$ and $m \neq p_n$, define $\lambda_m^* = (1 - h/m, F^*)$, and (iii) when $h = \infty$ and $m \neq p_n$, define $\lambda_m^* = (0, F^*)$, where F^* is the distribution such that $\{U_i : i = 0, \pm 1, \pm 2, ...\}$ are i.i.d., standard normal. Then, $\lambda_n^* \in \Lambda$ for all $n \geq 1$ and by construction $h_n(\lambda_n^*) \to h \in H$. This verifies Assumption B2 and completes the proof. □

12 Proof of Theorem 2

By the result of Theorem 1 of Andrews and Guggenberger (2011) stated in (6.1), we have

$$T_n(\rho_n^*) = T_n(\rho_n) + \delta_n/(d_n\widehat{\sigma}_n) \to_d T_h + \delta/S_h \tag{12.12}$$

under sequences $\{(\rho_n, F_n) \in \Lambda : n \geq 1\}$ for which $n(1-\rho_n) \to h$.

We show now that $\{\rho_n^* : n \geq 1\}$ satisfies

$$n(1-\rho_n^*) \rightarrow h + \delta \in [0, \infty]. \tag{12.13}$$

First, suppose $0 \leq h < \infty$. Then, $d_n = n^{1/2}$, $\rho_n^* = \rho_n - \delta_n/n$, and $n(1-\rho_n^*) = n(1-\rho_n) + \delta_n \rightarrow h + \delta \in [0, \infty)$. Second, suppose $h = \infty$ and $\rho_n \rightarrow \rho_\infty < 1$. Then, $\rho_n^* = \rho_n - \delta_n/(n^{1/2}d_n) \rightarrow \rho_\infty < 1$ and $n(1-\rho_n^*) \rightarrow \infty = h + \delta$. Third, suppose $h = \infty$ and $\rho_n \rightarrow 1$. Then, $d_n = (1-\rho_n^2)^{-1/2}(1+o(1))$ by (6.6). Note that $n(1-\rho_n^2) = (1+\rho_n)n(1-\rho_n) = 2n(1-\rho_n)(1+o(1))$. Hence, we have

$$\begin{aligned} n(1-\rho_n^*) &= n(1-\rho_n) + n^{1/2}\delta_n/d_n \\ &= n(1-\rho_n) + \delta(n(1-\rho_n^2))^{1/2}(1+o(1)) \\ &= n(1-\rho_n) + \delta 2^{1/2}(n(1-\rho_n))^{1/2}(1+o(1)) \\ &= n(1-\rho_n)[1 + \delta 2^{1/2}(n(1-\rho_n))^{-1/2}(1+o(1))] \\ &\rightarrow \infty = h + \delta. \end{aligned} \tag{12.14}$$

Hence, (12.13) is established.

Given (12.13), we have $c_{n(1-\rho_n^*)}(\beta) \rightarrow c_{h+\delta}(\beta)$ for all $\beta \in (0,1)$ by the same argument as in the proof of Theorem 1 with $(\rho_n^*, h+\delta)$ in place of (ρ_n, h). Using these results and the definition of the CHR CI, we obtain

$$\begin{aligned} & P_{\lambda_n}(\rho_n^* \in CI_{CHR,n}) \\ =\ & P_{\lambda_n}(c_{n(1-\rho_n^*)}(\alpha/2) \leq T_n(\rho_n^*) \leq c_{n(1-\rho_n^*)}(1-\alpha/2)) \\ \rightarrow\ & P(c_{h+\delta}(\alpha/2) \leq T_h + \delta/S_h \leq c_{h+\delta}(1-\alpha/2)), \end{aligned} \tag{12.15}$$

where the convergence holds by the definition of convergence in distribution and the continuity of the distribution of $T_h + \delta/S_h$. □

13 Asymptotic Validity of Romano and Wolf's (2001) Symmetric Subsampling CI

In this section, we show that the symmetric two-sided subsampling CI of Romano and Wolf (2001) (RW), denoted Sub_{RW} above, has asymptotic size equal to its nominal size for the parameter space Λ defined in (2.3), which allows for conditional heteroskedasticity. The derivations below also imply that the lower one-sided version of this CI has correct asymptotic size. The CI in RW is based on a t statistic that employs the LS estimator of ρ, a homoskedastic standard error estimator, and subsampling critical values.

RW demonstrate that this CI is pointwise asymptotically valid, while Andrews and Guggenberger (2007, Sections 9, 15) show that it has correct asymptotic size for a parameter space that imposes conditional homoskedasticity. (However, the equal-tailed two-sided and upper one-sided versions of this CI do not have asymptotically correct size under homoskedasticity or conditional heteroskedasticity, see Mikusheva (2007) and Andrews and Guggenberger (2007).)

Note that Andrews and Guggenberger (2009a) (AG09a) also analyze a CI based on a t statistic and subsampling critical values, denoted by Sub_{GLS} above. They consider a different test statistic than RW. Specifically, they consider a t statistic based on a FQGLS estimator that employs estimators $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ of the conditional variances $\{\sigma_i^2 : i \leq n\}$, combined with a heteroskedasticity-consistent standard error estimator, see Section 10 above for more details. AG09a prove that the resulting symmetric two-sided subsampling CI has asymptotic size equal to its nominal size for a parameter space that is comparable to Λ but with some additional restrictions on the quantities $\{\phi_{n,i}^2 : i \leq n\}$ that $\{\widehat{\phi}_{n,i}^2 : i \leq n\}$ estimate.

The CI in RW is based on a studentized t statistic

$$|T_{Hom,n}(\rho)| = |\frac{n^{1/2}(\widehat{\rho}_n - \rho)}{\widehat{\sigma}_{Hom,n}}|, \tag{13.16}$$

where $\widehat{\rho}_n$ is the LS estimator defined in (2.2) and $\widehat{\sigma}_{Hom,n}^2$ is the $(1,1)$ element of the standard variance estimator for the LS estimator under the assumption of homoskedasticity. More explicitly,

$$\widehat{\sigma}_{Hom,n}^2 = \left(n^{-1}X_1' M_{X_2} X_1\right)^{-1} \left(n^{-1}Y' M_X Y\right). \tag{13.17}$$

RW use subsampling critical values, denoted here by $c_{n,b}(1-\alpha)$, where b denotes the subsample size that satisfies $b \to \infty$ and $b/n \to 0$, and $1-\alpha$ is the nominal size. The critical value is the $(1-\alpha)$-quantile of the empirical distribution of the subsample test statistics over the $q = n - b + 1$ subsamples of data consisting of b consecutive observations from the original data set. The subsample test statistics $|T_{Hom,n,b,s}(\widehat{\rho}_n)|$ for $s = 1, ..., q$ are defined in the same way as the full-sample statistic $|T_{Hom,n}(\rho)|$ except that only the b observations in the s-th subsample are used and the hypothesized parameter ρ is replaced by the full-sample LS estimator $\widehat{\rho}_n$.

The symmetric two-sided CI in RW is given by the collection of all ρ ($\in [-1+\varepsilon, 1]$ for some $\varepsilon > 0$) that satisfy

$$|T_{Hom,n}(\rho)| \leq c_{n,b}(1-\alpha). \tag{13.18}$$

Equivalently, the RW CI can be written as

$$[\widehat{\rho}_n - n^{-1/2} c_{n,b}(1-\alpha)\widehat{\sigma}_{Hom,n}, \widehat{\rho}_n + n^{-1/2} c_{n,b}(1-\alpha)\widehat{\sigma}_{Hom,n}]. \tag{13.19}$$

We now show that this subsampling CI has correct asymptotic size. The proof is quite similar to that for the symmetric subsampling CI based on the FQGLS estimator in Sec. 7 of AG09a and Sec. S10 of Andrews and Guggenberger (2009b) (AG09b). A special case of the FQGLS estimator obtained by taking $\widehat{\phi}_{n,i} = 1 \ \forall i \leq n$ is the LS estimator. In this case, the only difference between the test statistics considered in

RW and AG09a,b is that the former uses the homoskedasticity variance estimator, whereas the latter uses the standard heteroskedasticity-consistent variance estimator.

The asymptotic size calculations given in AG09a,b depend on the limit as $n \to \infty$ of CP's of the CI under sequences $\{\lambda_n = (\rho_n, F_n) \in \Lambda : n \geq 1\}$ for which $\rho_n = 1 - h_n/n$, $h_n \to h \in [0, \infty]$, $\rho_n \to \rho_\infty$ for some $-1+\varepsilon \leq \rho_\infty \leq 1$, $E_{F_n} U_{n,i}^2 \to \sigma_{U,\infty}^2 > 0$, and $b_n(1-\rho_n) \to g \in [0, \infty]$ for $g \leq h$ (where $b = b_n$ is the subsample size).[29] Provided we show that the limit of the CP's of the RW symmetric two-sided CI is greater than or equal to the nominal size $1-\alpha$ for all such sequences, the remainder of the proof of the correct asymptotic size for the RW CI is almost the same as that given in AG09a,b.

When $\rho_n \to \rho_\infty < 1$, the subsample and the full-sample t statistic $|T_{Hom,n}(\rho_n)|$ have the same limiting distribution, a zero mean normal with a sandwich variance expression, and no asymptotic discontinuity arises. Hence, by standard arguments, e.g., see AG09b, the limit of the CP of the RW CI in this case equals the nominal size $1-\alpha$.

Below we show: when $\rho_n \to 1$ the asymptotic distribution of the RW statistic $|T_{Hom,n}(\rho_n)|$ is the same as the asymptotic distribution of the AG09a,b t statistic $|T_n(\rho_n)| = \left|n^{1/2}(\widehat{\rho}_n - \rho_n)/\widehat{\sigma}_n\right|$ (defined in (2.1) of the paper) based on the LS estimator and a heteroskedasticity-consistent variance estimator. Given this, by the arguments in AG09b, the limit of the CP's of the RW CI when $\rho_n \to 1$ equals that of the AG09a,b CI, which is greater than or equal to the nominal size $1-\alpha$. Hence, the RW CI has correct asymptotic size.

It remains to show the result stated in the previous paragraph. Without loss of generality, we can assume that $\mu = 0$, because both $\widehat{\rho}_n - \rho_n$ and $\widehat{\sigma}_{Hom,n}$ are invariant to the choice of μ. All limits below are taken as $n \to \infty$.

First, suppose $\rho_n \to 1$ and $h < \infty$. By Theorem 1 of Andrews and Guggenberger (2011) (AG11), we have

$$n(\widehat{\rho}_n - \rho_n) \to_d \frac{\int_0^1 I_{D,h}^*(r) dW(r)}{\int_0^1 I_{D,h}^*(r)^2 dr}, \tag{13.20}$$

where the right-hand side expression uses the fact that the quantities $h_{2,1}, h_{2,2}$, $h_{2,5}$, and $h_{2,7}$ in AG11 equal $h_{2,1} = h_{2,2} = \sigma_{U,\infty}^2$ $(= \lim_{n\to\infty} E_{F_n} U_i^2)$, $h_{2,5} = 1$, and $h_{2,7} = 1$ when $\widehat{\rho}_n$ is the LS estimator (which corresponds to $\widehat{\phi}_{n,i} = \phi_{n,i} = 1$). From eqn. (28) in AG11, it follows that jointly with (13.20) we have

$$n^{-2} X_1' M_{X_2} X_1 \to_d \sigma_{U,\infty}^2 \Big(\int_0^1 I_h^*(r)^2 dr - \big(\int_0^1 I_h^*(r) dr\big)^2\Big) = \sigma_{U,\infty}^2 \int_0^1 I_{D,h}^*(r)^2 dr. \tag{13.21}$$

Next, Lemma 5(c)-(d) and 5(f)-(h) in AG11 implies that when $0 < h < \infty$,

$$\frac{n^{-1} Y' M_X Y}{\sigma_{U,\infty}^2} = \frac{n^{-1} U' M_X U}{\sigma_{U,\infty}^2} = \frac{n^{-1} U' U}{\sigma_{U,\infty}^2} - \frac{n^{-1} U' P_X U}{\sigma_{U,\infty}^2} \to_p 1. \tag{13.22}$$

[29] In AG09a,b, h and g are denoted by h_1 and g_1.

When $h = 0$, the same result holds by Lemma 5(l) and the arguments in (35) and (36) of AG11 by writing the projection matrix P_X equivalently as the projection matrix P_{X^*}, where $X^* = [X_1 - \overline{Y}^*_{-1,n} 1_n : 1_n]$, where $1_n = (1, ..., 1)' \in R^n$ and $\overline{Y}^*_{-1,n} = n^{-1} \sum_{i=1}^n Y^*_{n,i-1}$ $(= n^{-1} 1_n' X_1)$.

Combining (13.20), (13.21), and (13.22), it follows that the asymptotic distribution of $|T_{Hom,n}(\rho_n)|$ is

$$|\int_0^1 I^*_{D,h}(r) dW(r)| / (\int_0^1 I^*_{D,h}(r)^2 dr)^{1/2}, \tag{13.23}$$

which is the same as that of $|T_n(\rho_n)|$, see Theorem 1(a) in AG11 with $h_{2,7} = 1$ (or (2.6) of the present paper).

Next, suppose $\rho_n \to 1$ and $h = \infty$. By Theorem 1(b) and the definition of a_n in eqn. (11) in AG11, we have

$$n^{1/2} \frac{E_{F_n} Y^{*2}_{n,0}}{(E_{F_n} Y^{*2}_{n,0} U^2_{n,1})^{1/2}} (\widehat{\rho}_n - \rho_n) \to_d N(0, 1). \tag{13.24}$$

By (40) of AG11,

$$\frac{n^{-1} X_1' M_{X_2} X_1}{E_{F_n} Y^{*2}_{n,0}} \to_p 1. \tag{13.25}$$

We also have

$$\frac{n^{-1} Y' M_X Y}{\sigma^2_{U,\infty}} = \frac{n^{-1} U' M_X U}{\sigma^2_{U,\infty}} = \frac{n^{-1} U' U}{\sigma^2_{U,\infty}} - \frac{n^{-1} U' P_X U}{\sigma^2_{U,\infty}} \to_p 1 \tag{13.26}$$

by a law of large numbers and

$$n^{-1} U' P_X U \to_p 0. \tag{13.27}$$

When $\rho_n \to 1$ and $h = \infty$, (13.27) holds by the following calculations:

$$n^{-1} U' P_X U = U' X (X'X)^{-1} (n^{-1} X'X)(X'X)^{-1} X'U = v_n' (n^{-1} X'X) v_n = o_p(1), \tag{13.28}$$

where $v_n = (X'X)^{-1} X'U = (O_p((1 - \rho_n)^{1/2} n^{-1/2}), O_p(n^{-1/2}))'$ by Lemma 8(d) in AG11, the first equality holds because $P_X = P_X P_X$, the second equality holds by the definition of v_n, and the third equality holds by the properties of v_n and the result that the $(1, 1)$, $(1, 2)$, and $(2, 2)$ components of $n^{-1} X'X$ are $O_p((1 - \rho_n)^{-1})$, $O_p((1 - \rho_n)^{-1/2})$, and $O(1)$, respectively, by the first result of Lemma 6 and Lemma 8(a) and (b) in AG11. Hence, (13.27) and (13.26) both hold.

Note that when $\rho_n \to 1$ and $h_n \to \infty$, then from the first two results in Lemma 6 of AG11, we have

$$\begin{aligned}
\frac{E_{F_n} Y_{n,0}^{*2} U_{n,1}^2}{E_{F_n} Y_{n,0}^{*2} E_{F_n} U_{n,1}^2} &= \frac{(1-\rho_n^2)^{-1}(E_{F_n} U_{n,1}^2)^2 + O(1)}{((1-\rho_n^2)^{-1}(E_{F_n} U_{n,1}^2) + O(1))(E_{F_n} U_{n,1}^2)} \\
&= \frac{(E_{F_n} U_{n,1}^2)^2 + o(1)}{((E_{F_n} U_{n,1}^2) + o(1))(E_{F_n} U_{n,1}^2)} \\
&= 1 + o(1). \qquad (13.29)
\end{aligned}$$

Combining (13.24), (13.25), (13.26), and (13.29), it follows that the asymptotic distribution $|N(0,1)|$ of $|T_{Hom,n}(\rho_n)|$ is the same as that of $|T_n(\rho_n)|$, see Theorem 1(b) in AG11. This completes the proof.

References for the Supplemental Appendix

Andrews, D. W. K., X. Cheng, and P. Guggenberger (2009): "Generic Results for Establishing the Asymptotic Size of Confidence Sets and Tests," Cowles Foundation Discussion Paper No. 1813, Yale University.

Andrews, D. W. K. and P. Guggenberger (2007): "Hybrid and Size-Corrected Subsample Methods," Cowles Foundation Discussion Paper No. 1606, Yale University.

——— (2009a): "Hybrid and Size-Corrected Subsampling Methods," *Econometrica,* 77, 721-762.

——— (2009b): "Supplemental Material for 'Hybrid and Size-Corrected Subsampling Methods'," *Econometrica,* 77, 721-762. Available on the Econometric Society website.

——— (2010): "Asymptotic Size and a Problem with Subsampling and the m out of n Bootstrap," *Econometric Theory,* 26, 426-468.

——— (2011): "Asymptotics for LS, GLS, and Feasible GLS Statistics in an AR(1) Model with Conditional Heteroskedasticity," *Journal of Econometrics,* forthcoming.

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

Kristensen, D. and O. Linton (2006): "A Closed-Form Estimator for the GARCH(1, 1) Model," *Econometric Theory,* 22, 323-337.

MacKinnon, J. G. and H. White (1985): "Some Heteroskedasticity Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics,* 29, 53-57.

Mikusheva, A. (2007): "Uniform Inferences in Autoregressive Models," *Econometrica,* 75, 1411-1452. Corrigendum, 2010, 78, 1773.

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

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