INITIAL-CONDITION-ROBUST INFERENCE
IN AUTOREGRESSIVE MODELS

By

Donald W. K. Andrews, Ming Li and Yapeng Zheng

February 2026

COWLES FOUNDATION DISCUSSION PAPER NO. 2496

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Initial-Condition-Robust Inference in Autoregressive Models

Donald W. K. Andrews* Ming Li† Yapeng Zheng‡

February 22, 2026

Abstract

This paper considers confidence intervals (CIs) for the autoregressive (AR) parameter in an AR model with an AR parameter that may be close or equal to one. Existing CIs rely on the assumption of a stationary or fixed initial condition to obtain correct asymptotic coverage and good finite sample coverage. When this assumption fails, their coverage can be quite poor. In this paper, we introduce a new CI for the AR parameter whose coverage probability is completely robust to the initial condition, both asymptotically and in finite samples. This CI pays only a small price in terms of its length when the initial condition is stationary or fixed. The new CI also is robust to conditional heteroskedasticity of the errors.

Keywords: Asymptotic size, autoregressive model, confidence set, initial condition, robustness.

JEL Classification Numbers: C10, C12.

*Andrews: Department of Economics and Cowles Foundation, Yale University, donald.andrews@yale.edu.

†Li: Department of Economics and Risk Management Institute, National University of Singapore, mli@nus.edu.sg.

‡Zheng: Department of Economics, Chinese University of Hong Kong, yapengzheng@link.cuhk.edu.hk.

1. Introduction

We consider an AR model with an AR parameter that may take values close or equal to one. Existing CIs in the literature for the AR parameter assume that the initial condition is stationary, zero, or fixed, e.g., see Stock (1991), Andrews (1993), Andrews and Chen (1994), Hansen (1999), Elliott and Stock (2001), Mikusheva (2007), and Andrews and Guggenberger (2014, AG14 hereafter). If the initial-condition assumption is not satisfied, then these CIs do not have correct asymptotic coverage probabilities and their finite sample coverage probabilities can be poor. For example, simulations reported below show that the nominal 95% AG14 CI has finite sample coverage probabilities ranging from 24.1% to 93.5% across 50 cases with highly variable initial conditions and different types of conditional heteroskedasticity of the error. A quarter of these CPs are 79.0% or less.

To circumvent the sensitivity of existing CIs to the initial condition, this paper introduces a new initial-condition-robust (ICR) CI whose finite sample CPs do not depend on the initial condition. These CIs are easy to compute and do not depend on any tuning parameters. We show that the ICR CI has correct asymptotic size in a uniform sense under conditions that allow for an arbitrary initial condition and conditional heteroskedasticity. Simulations show that the finite sample CPs of the nominal 95% ICR CI are quite good. They lie between 93.5% and 95.0% in the cases considered.

Like the other CIs referenced above, the ICR CI is constructed by inverting hypothesis tests concerning the value of the AR parameter. For all of these CIs, the tests are based on a t statistic that depends on a least squares (LS) estimator of the AR parameter. In contrast to existing CIs, the LS estimator used by the ICR tests contains an additional regressor that eliminates the effect of the initial condition under the null hypothesis. (For the definition of this regressor, see Section 2 below.) In consequence, the ICR CI has a CP that is invariant to the value of the initial condition. However, the length of the ICR CI can be effected by the initial condition. Simulations show that the ICR CI is slightly shorter when the initial condition is highly variable than when it is stationary or fixed.[1]

In scenarios where existing CIs are asymptotically valid (i.e., the initial condition is stationary or fixed), the ICR CI pays a price in terms of its length compared to existing CIs. This occurs because the ICR LS estimator has a higher variance due to the additional regressor that is included in the LS regression. However, simulations show that the effect is fairly small. Across 50 scenarios with stationary or fixed initial conditions and i.i.d., GARCH,

[1]The length of the ICR CI varies mostly with the value of the AR parameter. It is shortest for AR parameters near one. This is true of existing CIs as well.

or ARCH errors, the ratio of the expected length of the ICR CI compared to that of the AG14 CI is found to vary between 1.00 and 1.11. On average, the ICR CI is 3.5% longer than the AG14 CI in these scenarios.

Based on the ICR CI, we also introduce an asymptotically median-unbiased interval estimator (MUE) of the AR parameter.

The AR model considered here has been applied in the literature to exchange rate, commodity and stock prices, and other economic time series, e.g., see Kim and Schmidt (1993).

This paper is organized as follows. Section 2 specifies the AR model and defines the ICR CI. Section 3 defines the median-unbiased interval estimator. Section 4 provides simulation results concerning CPs and average lengths of the ICR and AG14 CIs and absolute median biases of the ICR MUE. Section 5 provides asymptotic results for the ICR CI. The Supplemental Material describes how the critical values were computed, proves Theorems 1 and 2, and provides some additional simulation results.

2. Model and ICR Confidence Interval

We consider the AR(1) model with conditional-heteroskedasticity studied in AG14:

$$\begin{aligned} Y_i &= \mu + Y_i^* \quad \text{and} \\ Y_i^* &= \rho Y_{i-1}^* + U_i \quad \text{for } i = 1, 2, \ldots, n, \end{aligned} \tag{1}$$

where $\rho \in [-1+\varepsilon, 1]$ for some $0 < \varepsilon < 2$ and $\{U_i : i = 1, 2, \ldots\}$ 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 $j \le i$, conditional variance $\sigma_i^2 = E_F(U_i^2 | \mathcal{G}_{i-1})$, and unconditional variance $\sigma_U^2 \in (0, \infty)$. A detailed description of the parameter space Λ is provided in Section 5.1.

By iterative substitution in (1), we have

$$Y_i = \mu + \rho^{i-1} \cdot \rho \cdot Y_0^* + \sum_{j=1}^{n} \rho^{i-j} U_j \quad \text{for } i = 1, \ldots, n. \tag{2}$$

Let Y, U, and X_1 be n-vectors whose i-th elements are Y_i, U_i, and Y_{i-1}, respectively. Let $X_2(\rho)$ be an $n \times 2$ matrix whose i-th row is $(1, \rho^{i-1})$ when $-1 < \rho < 1$.[2] When $\rho = 1$, we define $X_2(\rho)$ to be an $n \times 2$ matrix whose i-th row is $(1, i)$. Let $X(\rho) = [X_1 : X_2(\rho)]$,

[2]Following the convention $0^0 = 1$, we use ρ^{i-1} in the regression instead of ρ^i to avoid perfect multicollinearity when $\rho = 0$.

$P_{X(\rho)} = X(\rho)(X(\rho)'X(\rho))^{-1}X(\rho)'$, and $M_{X(\rho)} = I_n - P_{X(\rho)}$. Let $\widehat{U}_i$ denote the i-th element of the residual vector $M_{X(\rho)}Y$. Let p_{ii} denote the i-th diagonal element of $P_{X(\rho)}$, and define $p^*_{ii} = \min\{p_{ii}, n^{-1/2}\}$. Let Δ be a diagonal $n \times n$ matrix whose i-th diagonal element is $\widehat{U}_i/(1 - p^*_{ii})$. Then, the ICR LS estimator $\widehat{\rho}_n(\rho)$, and the HC5 variance estimator $\widehat{\sigma}^2_n(\rho)$ are defined as

$$\begin{aligned} \widehat{\rho}_n(\rho) &= (X_1'M_{X_2(\rho)}X_1)^{-1}X_1'M_{X_2(\rho)}Y \text{ and} \\ \widehat{\sigma}^2_n(\rho) &= (n^{-1}X_1'M_{X_2(\rho)}X_1)^{-1}(n^{-1}X_1'M_{X_2(\rho)}\Delta^2 M_{X_2(\rho)}X_1)(n^{-1}X_1'M_{X_2(\rho)}X_1)^{-1}, \end{aligned} \tag{3}$$

respectively.

Let

$$T_n(\rho) = \frac{n^{1/2}(\widehat{\rho}_n(\rho) - \rho)}{\widehat{\sigma}^2_n(\rho)}. \tag{4}$$

For suitable sequences $\{\rho_n\}_{n\geq 1}$ such that $n(1-\rho_n) \to h \in [0, \infty]$, we show that

$$T_n(\rho_n) \xrightarrow{d} J_h, \tag{5}$$

where J_h is defined in Section 5.2 below. Table 1 provides the quantiles $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$ of the distribution of J_h for $\alpha = .05$ and $.1$, which are the critical values employed by the ICR CI.

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

$$CI_{\mathrm{ICR},n} := \{\rho \in [-1+\epsilon, 1] : c_h(\alpha/2) \leq T_n(\rho) \leq c_h(1-\alpha/2) \text{ for } h = n(1-\rho)\}, \tag{6}$$

which can be computed by taking a fine grid of values $\rho \in [-1+\epsilon, 1]$ and comparing $T_n(\rho)$ to $c_h(\alpha/2)$ and $c_h(1-\alpha/2)$. Given these critical values, computation of the CIs is fast.

3. ICR Median-Unbiased Interval Estimator

By definition, an estimator $\widehat{\theta}_n$ of a parameter θ is median unbiased if $P(\widehat{\theta}_n \geq \theta) \geq 1/2$ and $P(\widehat{\theta}_n \leq \theta) \geq 1/2$. In this section, we introduce an ICR MUE of ρ that satisfies an analogous condition. Also, with probability close to one, this interval estimator is a point.[3]

Let $CI^{\mathrm{up}}_{\mathrm{ICR},n}(.5)$ and $CI^{\mathrm{low}}_{\mathrm{ICR},n}(.5)$ denote level .5 one-sided upper-bound and lower-bound

[3] If the median of the asymptotic null distribution of the t-statistic is strictly decreasing in ρ, or equivalently, if $c_h(0.5)$ is strictly increasing in h, then the proposed interval estimator would be a point estimator with probability one. Because this condition fails to hold exactly, but almost holds, there is a very small probability that the estimator is a short interval rather than a point.

Table 1: Quantiles of J_h for Use with 90% and 95% Equal-Tailed Two-Sided CIs and MUEs

Values of $c_h(\alpha)$, the α^{th} Quantile of J_h, for Use with 90% and 95% Equal-Tailed Two-Sided CI's and MUE's													
h	0	.2	.4	.6	.8	1	1.4	1.8	2.2	2.6	3	3.4	3.8
c_h (.025)	-3.66	-3.63	-3.60	-3.56	-3.54	-3.52	-3.46	-3.40	-3.36	-3.31	-3.27	-3.23	-3.19
c_h (.05)	-3.41	-3.38	-3.35	-3.31	-3.28	-3.25	-3.20	-3.14	-3.08	-3.04	-3.00	-2.95	-2.91
c_h (.5)	-2.18	-2.13	-2.09	-2.04	-1.99	-1.95	-1.86	-1.78	-1.70	-1.63	-1.57	-1.50	-1.45
c_h (.95)	-.94	-.87	-.80	-.74	-.68	-.62	-.50	-.39	-.29	-.19	-.11	-.03	.05
c_h (.975)	-.65	-.59	-.52	-.45	-.38	-.32	-.21	-.08	.01	.11	.19	.28	.35
h	4.2	4.6	5	6	7	8	9	10	11	12	13	14	15
c_h (.025)	-3.16	-3.12	-3.09	-3.02	-2.97	-2.90	-2.87	-2.82	-2.79	-2.75	-2.73	-2.71	-2.69
c_h (.05)	-2.87	-2.83	-2.80	-2.72	-2.66	-2.61	-2.56	-2.52	-2.48	-2.45	-2.42	-2.40	-2.38
c_h (.5)	-1.39	-1.34	-1.30	-1.20	-1.11	-1.04	-.99	-.93	-.89	-.85	-.82	-.78	-.76
c_h (.95)	.11	.18	.24	.36	.46	.55	.61	.68	.74	.78	.81	.84	.88
c_h (.975)	.41	.48	.54	.66	.77	.86	.92	.99	1.05	1.08	1.12	1.15	1.20
h	20	25	30	40	50	60	70	80	90	100	200	300	500
c_h (.025)	-2.59	-2.53	-2.47	-2.41	-2.36	-2.32	-2.30	-2.27	-2.26	-2.25	-2.16	-2.13	-2.09
c_h (.05)	-2.28	-2.22	-2.15	-2.09	-2.05	-2.01	-1.99	-1.96	-1.94	-1.94	-1.84	-1.81	-1.78
c_h (.5)	-.65	-.58	-.52	-.45	-.41	-.37	-.34	-.32	-.30	-.28	-.20	-.16	-.13
c_h (.95)	.99	1.06	1.12	1.19	1.24	1.28	1.30	1.32	1.34	1.36	1.45	1.48	1.52
c_h (.975)	1.30	1.38	1.43	1.50	1.55	1.59	1.62	1.64	1.66	1.68	1.76	1.79	1.83

CIs for ρ, respectively. By definition,

$$\begin{aligned} CI^{\text{up}}_{\text{ICR},n}(.5) &\coloneqq \{\rho \in [-1+\epsilon, 1] : c_h(0.5) \leq T_n(\rho) \text{ for } h = n(1-\rho)\} \text{ and} \\ CI^{\text{low}}_{\text{ICR},n}(.5) &\coloneqq \{\rho \in [-1+\epsilon, 1] : T_n(\rho) \leq c_h(0.5) \text{ for } h = n(1-\rho)\}. \end{aligned} \tag{7}$$

The MUE $\widetilde{\rho}_n$ of ρ is defined by

$$\begin{aligned} \widetilde{\rho}_n &= [\widetilde{\rho}^{\text{low}}_n, \widetilde{\rho}^{\text{up}}_n], \text{ where} \\ \widetilde{\rho}^{\text{up}}_n &= \max\{\rho : \rho \in CI^{\text{up}}_{\text{ICR},n}(.5)\} \text{ and} \\ \widetilde{\rho}^{\text{low}}_n &= \min\{\rho : \rho \in CI^{\text{low}}_{\text{ICR},n}(.5)\}. \end{aligned} \tag{8}$$

By construction, we have $\widetilde{\rho}^{\text{low}}_n \leq \widetilde{\rho}^{\text{up}}_n$.[4] In addition, $\widetilde{\rho}_n$ is a singleton whenever the set $\{\rho \in [-1+\epsilon, 1] : T_n(\rho) = c_h(.5)\}$ contains a single point, in which case $\widetilde{\rho}_n$ equals this point. Table 1 provides the critical values $c_h(.5)$ for a wide range of h. Given these critical values, computation of $\widetilde{\rho}_n$ is fast.

[4]This holds because $\widetilde{\rho}^{\text{up}}_n \geq \sup\{\rho \in [-1+\epsilon, 1] : c_h(.5) = T_n(\rho)\}$ and $\widetilde{\rho}^{\text{low}}_n$ is less than or equal to the infimum of the values in the same set.

4. Monte Carlo Simulations

We compare the coverage probabilities (CPs) and average lengths (ALs) of the ICR and AG14 CIs and report the absolute median biases (ABMs) of the ICR MUE using Monte Carlo simulations. In Section D of the Supplemental Material, we compare the performance of the ICR and Mikusheva (2007, Mik07 hereafter) CIs. We focus on the nominal 95% equal-tailed two-sided CIs. For the AG14 CI, we follow the paper's constructions of the test statistic and critical values. To calculate the MUE for ρ, we follow the procedure described in Section 3 and use $\widetilde{\rho}_n$ as the MUE of ρ when $\widetilde{\rho}_n$ is a singleton. When $\widetilde{\rho}_n^{\text{up}} \neq \widetilde{\rho}_n^{\text{low}}$, we take $\widetilde{\rho}_n^{\text{up}}$ as the MUE for ρ following Andrews and Li (2025, Section 5.1).

We consider a wide range of ρ values: 0, 0.5, 0.7, 0.9, and 0.99. The innovations are of the form $U_i = \sigma_i \varepsilon_i$, where $\{\varepsilon_i : i = 0, \pm 1, \dots\}$ are identically and independently distributed (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-$(ar_1, \dots, ar_4; \psi)$ denote an ARCH(4) process with AR parameters $(ar_1, \dots, ar_4)$ and intercept ψ. Following AG14, we consider five specifications for the conditional heteroskedasticity of the innovations: (a) i.i.d. $N(0, 1)$ ("i.i.d."), (b) GARCH(1,1)-(.05,.9;.001) ("GARCH1"), (c) GARCH(1,1)-(.15,.8;.2) ("GARCH2"), (d) GARCH(1,1)-(.25,.7;.2) ("GARCH3"), and (e) ARCH(4)-(.3,.2,.2,.2;.2) ("ARCH"). For the initial conditions, we consider the following four specifications: (a) Fixed: $Y_0^* = 0$, (b) Stationary: $Y_0^* = \sum_{i=0}^{\infty} \rho^i U_{-i}$, (c) Scaled n: $Y_0^* \sim \sqrt{n} \cdot \sum_{i=0}^{\infty} \rho^i U_{-i}$, and (d) Explosive: $Y_0^* \sim n^{3/4} \cdot \sum_{i=0}^{\infty} \rho^i U_{-i}$. We consider a sample size of $n = 150$. All simulation results are based on 30,000 simulation repetitions.

Tables 2 and 3 report the CPs of AG14 and ICR CIs, respectively. The CPs of the AG14 CI in Table 2 are close to the nominal level of 95% when the initial condition is fixed at zero or stationary, but fall below 95% when the initial condition is more variable, namely in the scaled n and explosive Y_0^* cases. In particular, the AG14 CI is robust to non-i.i.d. innovations when the initial condition is fixed or stationary. On the other hand, for scaled n and explosive Y_0^* cases, the AG14 CPs lie in the interval [24.1, 93.5] with roughly a quarter of the CPs being less than 79.0. Table 3 shows that the ICR CI achieves CPs close to 95% in all of the scenarios considered and for any initial conditions. Specifically, all ICR CI CPs lie in the interval [93.5, 95.0].

Table 4 reports the ratios of the ALs of the nominal 95% ICR CI to the AG14 CI. We include only cases with fixed or stationary initial conditions because the AG14 CI is not robust to more variable initial conditions, whereas the ICR CI is. The ratios are between 1.00 and 1.11. On average, the ICR CI is 3.5% longer than the AG14 CI over the fifty cases considered. Thus, the price the ICR CI pays in terms of AL to gain robustness against the distribution of the initial condition is fairly small.

Table 2: Coverage probabilities ($\times$100) of the nominal 95% AG14 CI

Initial Conditions:	**Fixed**					**Stationary**				
$\boldsymbol{\rho}$**:**	.00	.50	.70	.90	.99	.00	.50	.70	.90	.99
i.i.d.	94.6	94.6	94.8	94.7	95.1	94.6	94.5	94.8	94.7	94.5
GARCH1	94.6	94.5	94.8	94.8	95.1	94.6	94.6	94.9	94.8	94.4
GARCH2	94.4	94.6	94.5	94.9	95.2	94.3	94.6	94.5	94.8	94.7
GARCH3	94.2	94.3	94.6	94.7	95.2	94.1	94.3	94.6	94.6	94.6
ARCH4	93.7	93.9	94.2	94.6	95.6	93.6	94.0	94.1	94.5	94.7
Initial Conditions:	**Scaled** $\boldsymbol{n}$					**Explosive**				
i.i.d.	88.4	90.5	91.8	92.7	83.4	63.7	80.5	86.6	90.8	77.7
GARCH1	59.1	79.0	86.4	91.7	83.2	24.1	69.9	82.3	90.0	77.3
GARCH2	90.2	91.8	92.4	93.0	83.9	71.0	83.3	87.8	90.9	77.4
GARCH3	90.9	92.1	92.5	93.4	84.2	74.4	84.9	88.4	91.4	77.7
ARCH4	91.1	92.3	92.8	93.5	84.4	76.7	86.3	89.4	91.8	77.9

Table 3: Coverage probabilities ($\times$100) of the nominal 95% ICR CI

Initial Conditions:	**Arbitrary**				
$\boldsymbol{\rho}$**:**	.00	.50	.70	.90	.99
i.i.d.	94.4	94.5	94.7	94.7	94.3
GARCH1	94.4	94.6	94.9	95.0	94.3
GARCH2	94.1	94.6	94.4	94.9	94.2
GARCH3	93.9	94.2	94.5	94.7	94.1
ARCH4	93.5	93.8	93.9	94.5	94.3

Table 4: Ratios of the average lengths of the nominal 95% ICR to AG14 CIs

Initial Conditions:	**Fixed**					**Stationary**				
$\boldsymbol{\rho}$**:**	.00	.50	.70	.90	.99	.00	.50	.70	.90	.99
i.i.d.	1.00	1.00	1.02	1.04	1.08	1.00	1.01	1.02	1.05	1.11
GARCH1	1.04	1.05	1.06	1.07	1.08	1.00	1.01	1.02	1.05	1.11
GARCH2	1.00	1.01	1.03	1.02	1.07	1.00	1.01	1.03	1.04	1.11
GARCH3	1.00	1.01	1.03	1.02	1.06	1.00	1.02	1.04	1.03	1.10
ARCH4	1.00	1.02	1.03	1.01	1.06	1.01	1.02	1.04	1.03	1.10

Table 5: Average lengths of the nominal 95% ICR CI

Initial Conditions:	**Fixed**					**Stationary**				
ρ:	.00	.50	.70	.90	.99	.00	.50	.70	.90	.99
i.i.d.	.32	.28	.24	.17	.08	.32	.28	.24	.17	.07
GARCH1	.33	.30	.25	.17	.08	.34	.30	.26	.17	.08
GARCH2	.38	.34	.29	.19	.08	.38	.34	.29	.19	.08
GARCH3	.43	.38	.33	.20	.09	.43	.38	.33	.20	.08
ARCH4	.49	.44	.37	.21	.09	.49	.44	.37	.21	.09
Initial Conditions:	**Scaled n**					**Explosive**				
i.i.d.	.31	.27	.23	.14	.04	.28	.23	.18	.09	.02
GARCH1	.28	.24	.19	.12	.04	.18	.15	.12	.07	.02
GARCH2	.37	.32	.26	.15	.04	.32	.27	.21	.10	.02
GARCH3	.41	.36	.29	.16	.04	.36	.30	.23	.11	.02
ARCH4	.46	.40	.32	.17	.04	.39	.32	.25	.12	.02

Table 5 presents the ALs of the ICR CI. These are substantially decreasing in ρ, as expected, longer for ARCH conditional heteroskedasticity than i.i.d., and increasing from GARCH1 to GARCH2 to GARCH3 conditional heteroskedasticity.

Table 6: Absolute Median Biases of the ICR MUE

Initial Conditions:	**Fixed**					**Stationary**				
ρ:	.00	.50	.70	.90	.99	.00	.50	.70	.90	.99
i.i.d.	.012	.011	.006	.005	.020	.012	.011	.006	.005	.015
GARCH1	.017	.011	.006	.005	.020	.017	.011	.006	.005	.015
GARCH2	.017	.011	.006	.000	.020	.017	.011	.006	.000	.015
GARCH3	.022	.011	.011	.000	.020	.022	.011	.011	.000	.015
ARCH4	.022	.016	.011	.000	.020	.022	.016	.011	.000	.015
Initial Conditions:	**Scaled n**					**Explosive**				
i.i.d.	.012	.011	.006	.005	.010	.012	.011	.006	.010	.010
GARCH1	.017	.011	.006	.005	.010	.017	.011	.011	.010	.010
GARCH2	.017	.011	.011	.005	.010	.017	.011	.011	.010	.010
GARCH3	.022	.011	.011	.005	.010	.022	.011	.011	.010	.010
ARCH4	.022	.016	.011	.005	.010	.022	.016	.011	.010	.010

Table 6 reports the AMBs of the ICR MUE. Across all cases, the AMBs range from 0.000 to 0.022, indicating that their magnitudes are generally small.

In Section D of the Supplemental Material, we compare the performance of the ICR and

Mik07 CIs via simulations. The results are similar to those reported above for the ICR and AG14 CIs except that the Mik07 CI undercovers both under conditional heteroskedasticity, and/or scaled n or explosive initial conditions, but by a lower magnitude in the latter case than the AG14 CI. Again, we find that the ICR CI is robust to more variable initial conditions while paying only a small price in terms of CI width in the situations where the Mik07 CI has correct asymptotic coverage, i.e., when the innovations are i.i.d. and the initial conditions are stationary (or fixed).

5. Asymptotic Results

This section establishes the correct uniform asymptotic size and asymptotic similarity of the ICR CI for ρ.

5.1. Parameter Space, Correct Uniform Size, and Asymptotic Similarity

Following AG14, the parameter space for (ρ, F) is given by:

$\Lambda = \Big\{ \lambda = (\rho, F)$: (i) $\rho \in [-1+\epsilon, 1]$;

(ii) $\{U_i : i = 1, 2, \ldots\}$ are stationary and strong-mixing under F with $E_F(U_i|\mathcal{G}_{i-1}) = 0$ a.s., $E_F(U_i^2|\mathcal{G}_{i-1}) = \sigma_i^2$ a.s., where $\mathcal{G}_i$ is some non-decreasing sequence of σ-fields for which $U_j \in \mathcal{G}_i$ for all $j \leq i$ for $i = 1, 2, \ldots$;

(iii) The strong-mixing numbers $\{\alpha_F(m) : m \geq 1\}$ satisfy $\alpha_F(m) \leq Cm^{-3\zeta/(\zeta-3)}$, $\forall m \geq 1$;

(iv) $\sup_{i,s,t} E_F|\prod_{a\in A} a|^\zeta \leq M$, where $0 < i, s, t < \infty$, $i \geq \max(s,t)$, and A is any non-empty subset of $\{U_{i-s}, U_{i-t}, U_{i+1}^2, U_1^2\}$;

(v) $E_F U_1^2 = 1$;

for some constants $0 < \epsilon < 2$, $\zeta > 3$, and $C < \infty \Big\}$.

We suppose $E_F U_1^2 = 1$ for simplicity because the asymptotic distribution of $T_n(\rho)$ does not depend on the unconditional variance. Parts (ii) and (iv) are slightly different from the corresponding conditions in AG14 because the random variables $\{U_{-j},\ j \geq 0\}$ do not enter into the asymptotic analysis of the ICR CI. The key innovation of this paper is that we allow arbitrary initial conditions Y_0^*, including those corresponding to explosive processes.

The main theoretical result of this paper shows that $CI_{\mathrm{ICR},n}$ 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 $1-\alpha$ ICR CI for the parameter ρ satisfies*

$$\liminf_{n\to\infty} \inf_{\lambda\in\Lambda} P_\lambda(\rho \in CI_{ICR,n}) = \limsup_{n\to\infty} \sup_{\lambda\in\Lambda} P_\lambda(\rho \in CI_{ICR,n}) = 1-\alpha.$$

The MUE $\widetilde{\rho}_n$ has the following median-unbiasedness property. This result is a corollary to the one-sided versions of Theorem 1.[5]

Corollary 1. *The MUE estimator $\widetilde{\rho}_n$ satisfies*

$$\liminf_{n\to\infty} \inf_{\lambda\in\Lambda} P_\lambda(\widetilde{\rho}_n^{up} \geq \rho) \geq 1/2 \text{ and } \liminf_{n\to\infty} \inf_{\lambda\in\Lambda} P_\lambda(\widetilde{\rho}_n^{low} \leq \rho) \geq 1/2.$$

5.2. Asymptotic Distribution

As stated in (5), J_h is the asymptotic distribution of $T_n(\rho_n)$ under suitable sequences $\{\rho_n : n \geq 1\}$. Here, we state this result formally as Theorem 2 and define J_h. This result is proved in a similar way to Theorem 1 in Andrews and Guggenberger (2012). Using Theorem 2.1 in Andrews, Cheng, and Guggenberger (2020), Theorem 2 is sufficient to establish Theorem 1, which is proved in Section B of the Supplemental Material.

Theorem 2. *For a sequence $\lambda_n = (\rho_n, F_n)$ such that $n(1-\rho_n) \to h \in [0,\infty]$, we have*

$$T_n(\rho_n) = \frac{n^{1/2}(\widehat{\rho}_n(\rho_n) - \rho_n)}{\widehat{\sigma}_n(\rho_n)} \xrightarrow{d} J_h,$$

where J_h is defined immediately below.

Remark 1. For any subsequence $\{p_n\}_{n\geq 1}$ of $\{n\}_{n\geq 1}$, Theorem 2 holds with p_n in place of n throughout and h_{p_n} in place of $h = h_n$.

For $h = \infty$, J_h is the $N(0,1)$ distribution. For $h \in [0,\infty)$,

$$J_h := \int_0^1 I_{f,h}(r)dW(r) \Big/ \left(\int_0^1 I_{f,h}(r)^2 dr\right)^{1/2}, \tag{9}$$

where $I_{f,h}(r)$ is defined as follows.

---

[5] Corollary 1 holds because $CI^{\text{up}}_{\text{ICR},n}(.5)$ and $CI^{\text{low}}_{\text{ICR},n}(.5)$ both have coverage probabilities of 1/2 or greater by the proof of Theorem 1 applied to these one-sided CIs and $(\widetilde{\rho}_n \geq \rho) \supset CI^{\text{up}}_{\text{ICR},n}(.5)$ and $(\widetilde{\rho}_n \leq \rho) \supset CI^{\text{low}}_{\text{ICR},n}(.5)$.

Let $W(\cdot)$ denote a standard Brownian motion on $[0,1]$. For $h \in (0,\infty)$, define $f_h(r) = (1, \exp(-hr))'$, and for $h = 0$, let $f_h(r) = (1, r)'$. Define

$$I_h(r) := \begin{cases} \int_0^r \exp(-(r-s)h)dW(s) & \text{for } h > 0 \\ W(r) & \text{for } h = 0 \end{cases} \quad \text{and}$$

$$I_{f,h}(r) := I_h(r) - \left(\int_0^1 I_h(s)f_h(s)ds\right)' \left(\int_0^1 f_h(s)f_h(s)'ds\right)^{-1} f_h(r), \tag{10}$$

where the subscript f indicates that we take the residual process of $I_h(\cdot)$ after projecting it onto the space spanned by $f_h(\cdot)$. When $h > 0$,

$$I_{f,h}(r) = I_h(r) - \alpha_h(r)\int_0^1 I_h(s)ds - \beta_h(r)\int_0^1 \frac{1-e^{-hs}}{h} I_h(s)ds, \tag{11}$$

where

$$\begin{aligned} \alpha_h(r) = & \ \frac{h[1-e^{-2h} - 2(1-e^{-h})e^{-hr} + 2he^{-hr} - 2(1-e^{-h})]}{h(1-e^{-2h}) - 2(1-e^{-h})^2} \quad \text{and} \\ \beta_h(r) = & \ \frac{2h^2[he^{-hr} - (1-e^{-h})]}{h(1-e^{-2h}) - 2(1-e^{-h})^2}. \end{aligned} \tag{12}$$

When $h = 0$, $I_{f,h}(r)$ reduces to detrended Brownian motion:

$$W_f(r) := W(r) - (4-6r)\int_0^1 W(s)ds - (12r-6)\int_0^1 sW(s)ds. \tag{13}$$

The following lemma establishes the connection between $I_{f,h}(r)$ for $h > 0$ and $h = 0$.

Lemma 1. *Let $h \downarrow 0$, we have*

$$\sup_{r\in[0,1]} |\alpha_h(r) - (4-6r)| \to 0 \quad \textit{and} \quad \sup_{r\in[0,1]} |\beta_h(r) - (12r-6)| \to 0.$$

Remark 2. This lemma shows that $I_{f,h}(r)$ is uniformly continuous with respect to h. Consequently, $I_{f,h}(r) \to I_{f,0}(r)$ uniformly over $r \in [0,1]$ as $h \to 0$ a.s. and the asymptotic distribution is continuous at $h = 0$.

References

Andrews, Donald W. K. (1993), "Exactly median-unbiased estimation of first order autoregressive/unit root models." *Econometrica*, 61 (1), 139–165.

Andrews, Donald W. K. and Hong-Yuan Chen (1994), "Approximately median-unbiased estimation of autoregressive models." *Journal of Business & Economic Statistics*, 12 (2), 187–204.

Andrews, Donald W. K., Xu Cheng, and Patrik Guggenberger (2020), "Generic results for establishing the asymptotic size of confidence sets and tests." *Journal of Econometrics*, 218 (2), 496–531.

Andrews, Donald W. K. and Patrik Guggenberger (2012), "Asymptotics for LS, GLS, and feasible GLS statistics in an AR(1) model with conditional heteroskedasticity." *Journal of Econometrics*, 169 (2), 196–210.

——— (2014), "A conditional-heteroskedasticity-robust confidence interval for the autoregressive parameter." *Review of Economics and Statistics*, 96 (2), 376–381.

Andrews, Donald W. K. and Ming Li (2025), "Inference in a stationary/nonstationary autoregressive time-varying-parameter model." *Quantitative Economics*, 16 (3), 823–858.

Elliott, Graham and James H. Stock (2001), "Confidence intervals for autoregressive coefficients near one." *Journal of Econometrics*, 103 (1-2), 155–181.

Hansen, Bruce E. (1999), "The grid bootstrap and the autoregressive model." *Review of Economics and Statistics*, 81 (4), 594–607.

Kim, Kiwhan and Peter Schmidt (1993), "Unit root tests with conditional heteroskedasticity." *Journal of Econometrics*, 59 (3), 287–300.

Mikusheva, Anna (2007), "Uniform inference in autoregressive models." *Econometrica*, 75 (5), 1411–1452.

Stock, James H. (1991), "Confidence intervals for the largest autoregressive root in US macroeconomic time series." *Journal of Monetary Economics*, 28 (3), 435–459.