ON CONFIDENCE INTERVALS FOR AUTOREGRESSIVE ROOTS AND PREDICTIVE REGRESSION

BY

Peter C. B. Phillips

COWLES FOUNDATION PAPER NO. 1422

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

2014

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

Econometrica, Vol. 82, No. 3 (May, 2014), 1177–1195

ON CONFIDENCE INTERVALS FOR AUTOREGRESSIVE ROOTS AND PREDICTIVE REGRESSION

BY PETER C. B. PHILLIPS[1]

Local to unity limit theory is used in applications to construct confidence intervals (CIs) for autoregressive roots through inversion of a unit root test (Stock (1991)). Such CIs are asymptotically valid when the true model has an autoregressive root that is local to unity ($\rho = 1 + \frac{c}{n}$), but are shown here to be invalid at the limits of the domain of definition of the localizing coefficient c because of a failure in tightness and the escape of probability mass. Failure at the boundary implies that these CIs have zero asymptotic coverage probability in the stationary case and vicinities of unity that are wider than $O(n^{-1/3})$. The inversion methods of Hansen (1999) and Mikusheva (2007) are asymptotically valid in such cases. Implications of these results for predictive regression tests are explored. When the predictive regressor is stationary, the popular Campbell and Yogo (2006) CIs for the regression coefficient have zero coverage probability asymptotically, and their predictive test statistic Q erroneously indicates predictability with probability approaching unity when the null of no predictability holds. These results have obvious cautionary implications for the use of the procedures in empirical practice.

KEYWORDS: Autoregressive root, confidence belt, confidence interval, coverage probability, local to unity, localizing coefficient, predictive regression, tightness.

1. INTRODUCTION

A PRIMARY REASON FOR THE INTRODUCTION of local to unity limit theory was to develop asymptotic power functions for unit root test procedures. This theory facilitated comparisons between different test procedures. The limit theory also provided convenient approximations to the distributions of estimators and tests for models with an autoregressive parameter in the vicinity of unity of the form $\rho = 1 + \frac{c}{n}$, giving approximations that depend on the value of the localizing coefficient c. A prominent application of this theory in empirical work is the construction of confidence intervals (CIs) for autoregressive roots through the inversion of unit root test statistics. The approach was suggested in Stock (1991). It has been recommended and used in later work on CI construction for autoregressive roots (Elliott and Stock (2001)) and in predictive regression tests with persistent regressors (Cavanagh, Elliott, and Stock (1995), Campbell and Yogo (2006)).

Simulations in Hansen (1999) revealed that this inversion procedure performed well for ρ in the immediate vicinity of unity but poorly for stationary ρ values distant from unity. The limit theory in Phillips (1987) shows that appropriately centered statistics have limits as $c \to -\infty$ that correspond to the stationary limit theory for fixed $|\rho| < 1$, which suggests that inversion of appropriately centered test statistics should lead to CIs that correspond to those

[1]My thanks to three referees and the co-editor for helpful comments on earlier versions of this paper. Support is acknowledged from the NSF under Grants SES-0956687 and SES-1258285.

 DOI: 10.3982/ECTA11094

that apply for the stationary region and are based on stationary asymptotics. Mikusheva (2007) recently confirmed this supposition by demonstrating that CIs obtained in this way are asymptotically valid uniformly for ρ over local to unity and stationary regions. The methods of Mikusheva (2007) and Hansen (1999) are based on correctly centered statistics and have uniform validity.

On the other hand, inversion procedures based on unit root tests are not valid uniformly for $|\rho| \leq 1$. Since these procedures form the basis of much empirical work, it is important to understand the reasons for their failure and their properties when they are applied to data with stationary regressors, as they may be in predictive regressions of the type considered in Campbell and Yogo (2006).

The present paper contributes to this literature by providing an asymptotic analysis of the properties of CIs obtained by the inversion procedure applied to unit root tests. It is shown that such CIs are invalid at the limits of the domain of definition of the localizing coefficient as $c \to -\infty$. Consideration of the boundary case shows that these CIs manifest severe locational bias and width distortion. The asymptotic coverage probability of the intervals is zero in the stationary case even though the intervals are wider than those constructed in the usual way under stationarity. The reason for these distortions is a failure of tightness in the unit root test statistic sequence as $c \to -\infty$. Tightness is usually viewed as a purely technical requirement in the development of asymptotic theory. The present example shows how failure of tightness can have damaging implications for applied work when testing procedures based on localizing sequences are improperly constructed.

Similar consequences are shown to follow when these procedures are used in predictive regression tests of the type considered in Campbell and Yogo (2006). In particular, the commonly used Q test is biased toward accepting predictability and associated CIs for the regressor coefficient asymptotically have zero coverage probability in the stationary regressor case. These results have important empirical consequences in practice given that degrees of persistence in predictive regressors are determined imprecisely. Tests are needed that are robust for a wide range of such regressors. Alternative approaches that do achieve robustness are discussed in the paper.

The remainder of the paper is organized as follows. Section 2 develops new analytic approximations for CI constructions based on unit root tests and studies the limit properties of these intervals in the stationary case showing their zero coverage probability. Constructions involving correctly centered statistics are discussed in Section 3. The implications of these results for predictive regressions, including both CI construction and predictability tests, are explored in Section 4. Section 5 gives some extensions and Section 6 concludes. Proofs are in the Appendix.

2. UNIT ROOT TEST INVERSION

It will be sufficient for our purpose to consider the simple AR(1) model initialized at $x_0 = 0$:

$$(1) \qquad x_t = \rho x_{t-1} + u_t,$$

where $u_t \sim$ i.i.d. $(0, \sigma^2)$ with finite fourth cumulant κ_4. The least squares regression estimate of ρ is $\hat\rho$ and the associated unit root t test is $t_{\hat\rho} = \frac{\hat\rho - 1}{\sigma_{\hat\rho}}$, with $\sigma_{\hat\rho}^2 = \hat\sigma^2 / \sum_{t=1}^n x_{t-1}^2$ and $\hat\sigma^2 = n^{-1} \sum_{t=1}^n (x_t - \hat\rho x_{t-1})^2$. In the local to unity case $\rho = 1 + \frac{c}{n}$, we have the following limit theory as $n \to \infty$ for any fixed localizing coefficient c (Phillips (1987)):

$$(2) \qquad t_{\hat\rho} = \frac{n(\hat\rho - \rho) + n(\rho - 1)}{\left\{\hat\sigma^2 \Big/ \left(n^{-2} \sum_{t=1}^n x_{t-1}^2\right)\right\}^{1/2}} \Rightarrow \frac{\int J_c \, dW}{\left(\int J_c(r)^2 \, dr\right)^{1/2}} + c\left(\int J_c(r)^2 \, dr\right)^{1/2} =: \tau_c,$$

where $J_c(r) = \int_0^r e^{c(r-s)} \, dW(s)$ is a linear diffusion and W is standard Brownian motion. All integrals in (2) and elsewhere in the paper when the limits are unspecified are taken over the interval $[0, 1]$. The limit representation for τ_c holds for all $c \in \mathbb{R}$. In what follows, we concentrate on the half line $(-\infty, 0)$, the domain of definition relevant for stationary alternatives.

2.1. *Confidence Belt Asymptotics*

The method of confidence belts suggested in Stock (1991) proceeds as follows: compute a unit root test statistic such as $t_{\hat\rho}$ and use the known asymptotic distribution of that statistic under the alternative, as given in (2) above, to construct a CI for c by inversion of the test. Confidence belts provide a graphical method of performing this operation and have been tabulated and interpolated for implementation in practice.

Since (2) is the appropriate asymptotic distribution of $t_{\hat\rho}$ under the local alternative $\rho = 1 + \frac{c}{n}$ to a unit root for all c, it may not be immediately obvious why the procedure fails to deliver CIs with good properties for stationary ρ. To explain the failure, we develop the asymptotics in (2) further, focusing on the limit of the domain of definition, viz. $c \to -\infty$, corresponding to the stationary case. Phillips (1987) proved that the two components of (2) satisfy

$$\lambda_c := \frac{\int J_c(r)\,dW(r)}{\left(\int J_c(r)^2\,dr\right)^{1/2}} \Rightarrow \xi \equiv N(0,1) \quad \text{and} \quad (-2c)\int J_c(r)^2\,dr \to_p 1, \tag{3}$$

as $c \to -\infty$. Importantly, (2)–(3) imply that the sequence $\tau_c = O_p(|c|^{1/2})$ is not tight as $c \to -\infty$, unlike the sequence λ_c. This failure of tightness in τ_c leads to an escape of probability mass in the limit which has material consequences for induced CIs and tests of predictability that are founded on unit root test statistics like (2).

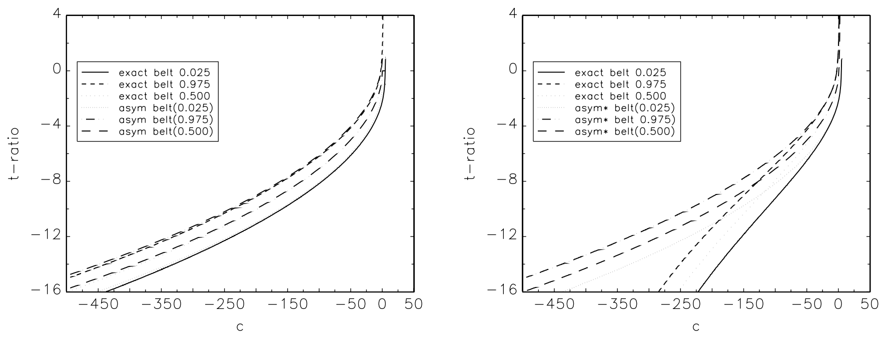
THEOREM 1: *As $c \to -\infty$, the asymptotic form of τ_c in a suitably expanded probability space is*

$$\tau_c = -\frac{|c|^{1/2}}{2^{1/2}} + \frac{1}{2}\xi + O_p\big(|c|^{-1/2}\big) \sim N\left(-\frac{|c|^{1/2}}{2^{1/2}}, \frac{1}{4}\right) + O_p\big(|c|^{-1/2}\big). \tag{4}$$

The leading term of τ_c is divergent because the test statistic $t_{\hat\rho} = \frac{\hat\rho-1}{\sigma_{\hat\rho}}$ is miscentered on unity. The distribution $N(-\frac{|c|^{1/2}}{2^{1/2}}, \frac{1}{4})$ delivers approximate percentiles of the limit variate τ_c based on the local to unity limit theory when $|c|$ is large. The percentile functions in the inversion of the test produce the confidence belts used in constructing CIs for c.

To illustrate, the $2\frac{1}{2}\%$, 50% (median), and $97\frac{1}{2}\%$ confidence belts produced from τ_c are shown in Figure 1(a) over the region $-500 < c < 0$. The confidence belts were computed with 100,000 replications and a grid of 2,000 values of c using the model (1) with Gaussian errors, $\rho = 1 + \frac{c}{n}$, and a sample size of $n = 10{,}000$. These belts can be used to execute the inversion of the t-ratio to produce an induced CI for c for any given value of the test statistic $t_{\hat\rho}$. Being "asymptotic" (for $n \to \infty$), there is no (apparent) dependence on n and the belts are used to compute first CIs for c and then implied CIs for $\rho = 1 + c/n$ given some sample of size n. The practical facility of this method is that an investigator need only compute the unit root test statistic $t_{\hat\rho}$, plug its value into the belts to extract a CI for c, and then deduce the implied CI for $\rho = 1 + \frac{c}{n}$ for the specific value of n in the sample.

In addition to the three confidence belts computed by simulation, Figure 1(a) also shows the approximate belts based on the limit theory (4) for the $2\frac{1}{2}\%$, $97\frac{1}{2}\%$, and 50% (median) percentiles, viz., $\{-\frac{|c|^{1/2}}{2^{1/2}} \pm \frac{1.96}{2}; -\frac{|c|^{1/2}}{2^{1/2}}\}$, which show broad conformity of these large $|c|$ asymptotic functional forms to the $2\frac{1}{2}\%$, $97\frac{1}{2}\%$, and 50% belts over the region $-500 < c < 0$ computed exactly



(a) (b)

FIGURE 1.—Exact confidence belts (at levels 2.5%, 50%, and 97.5%) computed by simulation for $n = 10{,}000$ for the t-ratio $t_{\hat\rho} = \frac{\hat\rho - 1}{\sigma_{\hat\rho}}$ are shown together with the asymptotic form (6) (denoted "asym") in (a) ($\rho = 1 + \frac{c}{n} \in (0.95, 1.0)$). The same (simulation based) confidence belts (denoted "asym*") are shown against the exact confidence belts for $n = 250$. The deviations between the belts in (b) reveal the escaping probability mass as $|c|$ increases ($\rho = 1 + \frac{c}{n} \in (-1, 1)$).

for $n = 10{,}000$. The difference between the curves is small because the implied interval $\rho \in (0.95, 1)$ is close to unity. But since $\tau_c \sim N(-\frac{|c|^{1/2}}{2^{1/2}}, \frac{1}{4})$, probability mass begins to escape to infinity as $|c|$ increases.

Figure 1(b) dramatically illustrates the implications of this loss of mass when the intervals are applied in smaller sample sizes with stationary values of ρ. This figure shows the two belts over the same interval $c \in (-500, 0)$ but for sample size $n = 250$. The "exact" belts in this figure are computed by simulation with $n = 250$ using $\rho = 1 + c/n$ over the grid $c \in (-500, 0)$. The "asym*" belts in the figure are identical to those obtained for Figure 1(a) from simulations with $n = 10{,}000$ (these belts correspond to the "asymptotic" distribution of τ_c). For $n = 250$, the given range of values of c span the full stationary interval $\rho \in (-1, 1)$. The escape of probability mass (as $|c|$ increases) is obvious in Figure 1(b) as the belts diverge and become disjoint for large $|c|$, documenting the increasing failure of the local to unity asymptotics to produce satisfactory CIs for stationary ρ. This implementation corroborates the failure of uniformity as well as the quantitative finding in Hansen (1999) that when $\rho = 0.6$ (read $c = -100$ for $n = 250$), the true value of ρ lies completely to the right of the local to unity 90% CI for ρ (implying that the CI based on local to unity asymptotics has zero coverage probability). In Figure 1(b), the true value $\rho = 0.6$ lies close to the point where the exact ($n = 250$) median belt crosses the asymptotic local to unity $2\frac{1}{2}\%$ belt for ρ, indicating very low coverage, on average, of the 95% CI for ρ.

2.2. *Induced Confidence Intervals for c and ρ*

Using asymptotic confidence belts obtained numerically in the manner just described, Stock (1991) suggested that a $100(1-\alpha)\%$ confidence set can be constructed as $S(\tau)=\{c: f_{L,\alpha/2}(c)\leq \tau \leq f_{U,\alpha/2}(c)\}$, where $f_{L,\alpha/2}(c)$ and $f_{U,\alpha/2}(c)$ are the lower $\alpha/2$ and upper $1-\alpha/2$ percentiles of τ_c as a function of c. Inverting the test critical values in the belts yields the CI

$$\{c: f^{-1}_{U,\alpha/2}(\hat{\tau})\leq c\leq f^{-1}_{L,\alpha/2}(\hat{\tau})\} \tag{5}$$

for c, where (5) is calculated for some given observed t ratio $\hat{\tau}$. The implied CI for ρ is $\{\rho=1+\frac{c}{n}: f^{-1}_{U,\alpha/2}(\hat{\tau})\leq c\leq f^{-1}_{L,\alpha/2}(\hat{\tau})\}$. The limits $f^{-1}_{U,\alpha/2}(\hat{\tau})$ and $f^{-1}_{L,\alpha/2}(\hat{\tau})$ in (5) are obtained by reading off the values of c that correspond to the observed $\hat{\tau}$ in the calculated confidence belts (computed using the local to unity limit theory, as in Figure 1). This process amounts to solving an equation for c based on the form of the confidence belts, for some given $\hat{\tau}$. As (4) shows, the distribution that produces the confidence belts is approximately normal with mean $-\frac{|c|^{1/2}}{2^{1/2}}$ and variance $\frac{1}{4}$. Thus, the confidence belts have the explicit asymptotic form

$$\{f_{L,\alpha/2}(c), f_{U,\alpha/2}(c)\}=\left\{-\frac{|c|^{1/2}}{2^{1/2}}-\frac{z_{\alpha/2}}{2}, -\frac{|c|^{1/2}}{2^{1/2}}+\frac{z_{\alpha/2}}{2}\right\}, \tag{6}$$

where $z_{\alpha/2}$ is the $1-\alpha/2$ percentile of the standard normal distribution $N(0,1)$. Note that this interval has length $2\times\frac{z_{\alpha/2}}{2}=1.96$ for a 95% interval and this length conforms with the vertical distance between the belts shown in Figure 1. We can proceed to derive the length of the induced CIs of c (and for ρ) and the coverage probability $\mathbb{P}\{f^{-1}_{U,\alpha/2}(\hat{\tau})\leq c\leq f^{-1}_{L,\alpha/2}(\hat{\tau})\}$ of the interval when the true ρ is fixed and stationary, that is, $|\rho|<1$, by considering the corresponding limits of these quantities at the boundary of the local to unity limit theory when $c\to-\infty$.

Inverting equation (4) for c, we have $|c|^{1/2}=-2^{1/2}\{\tau_c+\frac{\xi}{2}\}+O_p(\frac{1}{|c|^{1/2}})$, giving the following analytic expressions for the lower and upper limits $c_L=f^{-1}_{U,\alpha/2}(\hat{\tau})$ and $c_U=f^{-1}_{L,\alpha/2}(\hat{\tau})$:

$$|c_L|^{1/2}=2^{1/2}\left\{-\hat{\tau}-\frac{z_{\alpha/2}}{2}\right\}+O_p\left(\frac{1}{|c|^{1/2}}\right),$$

$$|c_U|^{1/2}=2^{1/2}\left\{-\hat{\tau}+\frac{z_{\alpha/2}}{2}\right\}+O_p\left(\frac{1}{|c|^{1/2}}\right),$$

and the CI for c

$$[c_L, c_U]=\left[-2\left(\hat{\tau}^2-z_{\alpha/2}\hat{\tau}\right), -2\left(\hat{\tau}^2+z_{\alpha/2}\hat{\tau}\right)\right], \tag{7}$$

whose length is $|c_L - c_U| = 4z_{\alpha/2}|\hat{\tau}|$, both holding up to an error of $O_p(1)$ as $c \to -\infty$.

2.3. *Properties in the Stationary Case*

When $n \to \infty$, the centered statistics have the following limits:

$$\frac{\sqrt{n}(\hat{\rho} - \rho)}{\sqrt{1-\rho^2}} \Rightarrow N(0,1), \quad t_{\hat{\rho},\rho} = \frac{\hat{\rho} - \rho}{\sigma_{\hat{\rho}}} \Rightarrow N(0,1),$$

for all fixed $|\rho| < 1$, as well as uniformly over all $|\rho| < 1$ for which $n(\rho - 1) \to 0$, as shown in Giraitis and Phillips (2006)—see also Phillips and Magdalinos (2007). It is convenient to take a probability space for which the convergences apply in probability, so that, for fixed $|\rho| < 1$,

$$t_{\hat{\rho},\rho} = \frac{\hat{\rho} - \rho}{\sigma_{\hat{\rho}}} = \zeta_\rho + o_p(1), \quad \zeta_\rho \equiv N(0,1). \tag{8}$$

For an assumed local to unity model $\rho = 1 + \frac{c}{n}$, the interval (7) for c implies the following CI for ρ:

$$[\rho_L, \rho_U] = \left[1 - \frac{2}{n}\big(\hat{\tau}^2 - z_{\alpha/2}\hat{\tau}\big), 1 - \frac{2}{n}\big(\hat{\tau}^2 + z_{\alpha/2}\hat{\tau}\big)\right]. \tag{9}$$

The coverage probability of this interval and its behavior in the stationary case can be evaluated by using the limit behavior of the unit root statistic $\hat{\tau}$ under stationarity. We have the following result.

THEOREM 2: *The nominal* $100(1-\alpha)\%$ *confidence interval* $[\rho_L, \rho_U]$ *for* $|\rho| < 1$ *based on inversion of the unit root test statistic* $\hat{\tau}$ *has the following form up to* $O_p(n^{-1/2})$:

$$[\rho_L, \rho_U] = \left[1 - 2A_\rho + 2\frac{2\zeta - z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}, 1 - 2A_\rho + 2\frac{2\zeta + z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\right], \tag{10}$$

where $A_\rho = \frac{1-\rho}{1+\rho}$, $\zeta = N(0, v_\zeta)$, $v_\zeta = 1 + \frac{A_\rho^2}{2\sigma^4}(\kappa_4 + \frac{2\sigma^4}{1-\rho^2})$, *and* $\pm z_{\alpha/2}$ *are the lower and upper symmetric* $\alpha/2$ *percentiles of the standard normal distribution. The coverage probability as* $n \to \infty$ *is*

$$\begin{aligned} &\mathbb{P}\big\{\rho \in [\rho_L, \rho_U]\big\} \\ &\quad = \mathbb{P}\left\{\frac{\sqrt{n}}{4}(1-\rho)A_\rho^{1/2} - \frac{z_{\alpha/2}}{2} \le \zeta \le \frac{\sqrt{n}}{4}(1-\rho)A_\rho^{1/2} + \frac{z_{\alpha/2}}{2}\right\} \end{aligned} \tag{11}$$

and $\mathbb{P}\{\rho \in [\rho_L, \rho_U]\} \to 0$ *if* $\sqrt{n}(1-\rho)A_\rho^{1/2} \to \infty$. *The average length of the interval is* $\frac{4}{\sqrt{n}}(\frac{1-\rho}{1+\rho})^{1/2}z_{\alpha/2}$.

It follows that the CI based on inverting a unit root test using local to unity limit theory has zero coverage probability for ρ asymptotically whenever $\sqrt{n}(1-\rho)A_\rho^{1/2} \to \infty$. This includes all fixed $|\rho| < 1$ and all ρ for which $n^{1/3}(1-\rho) \to \infty$, that is, all ρ whose distance from unity exceeds $O(n^{-1/3})$ as $n \to \infty$. The induced CI (10) is centered on $\bar{\rho} = 1 - 2A_\rho = \frac{3\rho-1}{\rho+1}$, and the interval shrinks to the pseudo true value $\bar{\rho}$ as $n \to \infty$. Observe that $\bar{\rho} < \rho$ for all $|\rho| < 1$ and $\bar{\rho} = \rho$ iff $\rho = 1$. When the true $\rho = 0$, $\bar{\rho} = -1$ and when $\rho = 0.5$, $\bar{\rho} = 1/3$. So the bias in $\bar{\rho}$ is substantial for much of the stationary region, with corresponding distortions in the CIs for ρ. These asymptotics explain the poor performance of these CIs in the stationary case in Hansen's (1999) simulations, as noted in the discussion of Figure 1(b).

We can compare these results with the standard CI based on $t_{\hat{\rho},\rho}$ under stationarity

$$(12) \qquad \{\hat{\rho} \pm z_{\alpha/2}\sigma_{\hat{\rho}}\} = \left\{\hat{\rho} \pm z_{\alpha/2}\left(\frac{1-\rho^2}{n} + O_p\left(\frac{1}{n^{3/2}}\right)\right)^{1/2}\right\},$$

whose average length is $2z_{\alpha/2}(\frac{1-\rho^2}{n})^{1/2} \leq \frac{4}{\sqrt{n}}(\frac{1-\rho}{1+\rho})^{1/2} z_{\alpha/2}$ up to an error of $O_p(n^{-1})$. Thus, the implied CI from inversion of local to unity limit theory has length greater than that of the standard interval for all $|\rho| < 1$. The length exceeds that of the standard interval by the factor $2/(1+\rho)$ and so is more than twice as large when $\rho < 0$.

3. HANSEN AND MIKUSHEVA CONSTRUCTIONS

Hansen (1999) and Mikusheva (2007) suggested to construct CIs by performing test inversion with a properly centered t-ratio statistic (Hansen) or a general test function involving a centered numerator and separate denominator components (Mikusheva). In place of (2), these approaches effectively amount to working with the statistic

$$(13) \qquad t_{\hat{\rho},\rho} = \frac{n(\hat{\rho}-\rho)}{\left\{\hat{\sigma}^2 \Big/ \left(n^{-2}\sum_{t=1}^{n} x_{t-1}^2\right)\right\}^{1/2}} \Rightarrow \frac{\int J_c\, dW}{\left(\int J(r)^2\, dr\right)^{1/2}} = \lambda_c,$$

or a coefficient-based version of the test instead.

Proceeding as in the above analysis for large $|c|$, we find that $\lambda_c \sim N(0,1) + O_p(\frac{1}{|c|^{1/2}})$ in place of (4). The induced CI for c upon inversion of $t_{\hat{\rho},\rho}$ is, up to an error of $O_p(|c|^{-1/2})$,

$$(14) \qquad \{f_{L,\alpha/2}(t_{\hat{\rho},\rho}), f_{U,\alpha/2}(t_{\hat{\rho},\rho})\} \sim \{-z_{\alpha/2}, +z_{\alpha/2}\} \quad \text{as} \quad c \to -\infty.$$

From the centered form of the t-ratio $t_{\hat{\rho},\rho}$ in (13), the corresponding interval for ρ when $c \to -\infty$ is simply $\{\hat{\rho} - z_{\alpha/2}\sigma_{\hat{\rho}}, \hat{\rho} + z_{\alpha/2}\sigma_{\hat{\rho}}\}$, giving the classical stationary interval based on normal asymptotics and confirming uniformity over stationary ρ (cf. Giraitis and Phillips (2006), Phillips and Magdalinos (2007)). Mikusheva (2014) showed that the nonparametric grid bootstrap procedure achieves a second order refinement in the local to unity region relative to the approach based on centered local to unity asymptotics.

4. PREDICTIVE REGRESSION TESTS

In predictive regressions, it is now common empirical practice to allow for unknown persistence in the predictors. Such predictors complicate testing procedures by introducing nonstandard limit theory and dependence on nuisance parameters like the localizing coefficient in a near integrated regressor. Various approaches are now available to deal with these complications. A recent overview was given in Phillips and Lee (2013).

A popular procedure was implemented in Campbell and Yogo (2006, hereafter CY), following an earlier suggestion by Cavanagh, Elliott, and Stock (1995, hereafter CSE). These procedures both use a Bonferroni method in conjunction with Stock's (1991) CI construction for the autoregressive coefficient of the regressor-predictor to produce tests of predictability that are intended to be robust to persistence.

Both the Q test of CY and the t-ratio test in CSE involve t ratios computed by simple regression in combination with Stock's (1991) CI construction. To fix ideas, consider the standard predictive regression model without an intercept:

$$y_t = \beta x_{t-1} + u_{0t}, \tag{15}$$

$$x_t = \rho x_{t-1} + u_{xt}, \tag{16}$$

where $\rho = 1 + \frac{c}{n}$ for $c \leq 0$, and the innovations $u_t = (u_{0t}, u_{xt}) \sim$ i.i.d. $(0, \Sigma)$ with $\text{vech}(\Sigma) = \{\sigma_{00}, \sigma_{0x}, \sigma_{xx}\}$ and finite fourth moments. Standardized partial sums of u_{xt} satisfy a functional law with limit Brownian motion $\sigma_{xx}^{1/2} W(\cdot)$.

Upon fitting (15) by least squares, the usual t-ratio test on β has a limit theory that can be decomposed as (e.g., CSE)

$$t_{\hat{\beta}} \Rightarrow \left(1 - \delta^2\right)^{1/2} Z + \delta\lambda_c, \tag{17}$$

where $\delta = \frac{\sigma_{0x}}{(\sigma_{xx}\sigma_{00})^{1/2}}$, $Z \equiv N(0, 1)$, and λ_c is defined as in (3). The limit variates Z and λ_c are independent. The corresponding unit root t-ratio statistic $t_{\hat{\rho}} = \frac{\hat{\rho}-1}{\sigma_{\hat{\rho}}} \Rightarrow \tau_c$, as in (2). CSE and CY addressed inference when $\delta \neq 0$ by using a Bonferroni method, finding possible values for c (or ρ) and using the most conservative ones to produce a robust test.

4.1. *The t-Ratio Test* $t_{\hat{\beta}}$

The t-ratio statistic $t_{\hat{\beta}}$ was considered in both CSE and CY, although CY recommended for implementation a different t-ratio test called the Q test, discussed later. The mixture limit theory (17) means that tests and CIs need to allow for the unknown value of c. This is achieved by finding a nominal $100(1-\alpha_1)\%$ CI for c using the inversion process of Stock (1991). Then, for each c in this CI, a $100(1-\alpha_2)\%$ CI is constructed for β, denoted as $\mathrm{CI}_{\beta|c}(\alpha_2)$. A CI for β that is free of c is obtained as the union $\mathrm{CI}_\beta(\alpha_1,\alpha_2)=\bigcup_{c\in\mathrm{CI}_c(\alpha_1)}\mathrm{CI}_{\beta|c}(\alpha_2)$. More specifically, the estimate $\hat{\rho}$ is used to find a CI $\mathrm{CI}_c(\alpha_1)=[c_L(\alpha_1),c_U(\alpha_1)]$ for c by inverting a t-ratio ADF unit root test statistic for ρ or, in the case of CY, a version of this statistic that improves efficiency in trend removal by quasi-differencing (so-called GLS detrending), which is unnecessary in the present case. Then (given δ or a consistent estimate of δ), the authors used the critical value $d_{t_{\hat{\beta}},c,\alpha_2/2}$ of $t_{\hat{\beta}}\sim\delta\lambda_c+(1-\delta^2)^{1/2}Z$ at the percentile $\alpha_2/2$ to find the following CI for β:

$$\mathrm{CI}_\beta(\alpha_1,\alpha_2):=\Big[\min_{c_L(\alpha_1)\le c\le c_U(\alpha_1)} d_{t_{\hat{\beta}},c,\alpha_2/2},\ \max_{c_L(\alpha_1)\le c\le c_U(\alpha_1)} d_{t_{\hat{\beta}},c,1-\alpha_2/2}\Big]. \tag{18}$$

Then, by Bonferroni as $n\to\infty$,

$$\begin{aligned}\mathbb{P}\{t_{\hat{\beta}}\notin\mathrm{CI}_\beta(\alpha_1,\alpha_2)\}&\to\mathbb{P}\{(1-\delta^2)^{1/2}Z+\delta\lambda_c\notin\mathrm{CI}_\beta(\alpha_1,\alpha_2)\}\\&\le\alpha_1+\alpha_2.\end{aligned}$$

As noted, the interval $\mathrm{CI}_c(\alpha_1)=[c_L(\alpha_1),c_U(\alpha_1)]$ advances progressively toward $-\infty$ as x_t approaches stationarity. Although the implied CI for ρ has zero coverage probability asymptotically, it nonetheless still holds that $c\to-\infty$ for all $c\in[c_L(\alpha_1),c_U(\alpha_1)]$ when ρ is stationary, as is immediately evident from (7) and (29). The width of this interval $|c_U(\alpha_1)-c_L(\alpha_1)|=4z_{\alpha/2}|\hat{\tau}|\sim4z_{\alpha/2}\sqrt{n}A_\rho^{1/2}$ also tends to infinity. Hence, the actual coverage probability associated with the interval $\mathrm{CI}_c(\alpha_1)$ tends to unity rather than $1-\alpha_1$ as n tends to infinity in the stationary case. In effect, the restriction $c\in[c_L(\alpha_1),c_U(\alpha_1)]$ is vacuous in the stationary case because both limits $c_L(\alpha_1)$ and $c_U(\alpha_1)\to-\infty$. The probability mass $\alpha_1=\mathbb{P}\{c\notin[c_L(\alpha_1),c_U(\alpha_1)]\}$ therefore escapes to zero due to the failure of tightness in the sequence of unit root test statistics from which this interval is constructed.

Moreover, when $c\to-\infty$, we have $\lambda_c\to_d\xi\equiv N(0,1)$ and the limit variate ξ is independent of Z, as remarked above. Hence, for all $c\in[c_L(\alpha_1),c_U(\alpha_1)]$, we actually have

$$(1-\delta^2)^{1/2}Z+\delta\lambda_c\to_d(1-\delta^2)^{1/2}Z+\delta\xi\equiv N(0,1) \tag{19}$$

when ρ is stationary and $n\to\infty$. The CI in the limit for the stationary case then has coverage probability determined only by the secondary controlled level α_2

of the test. The CSE test is, therefore, undersized in the limit (here by the full primary probability α_1 which is lost in the limit by test inversion and failure of tightness), just as it is also (partially) undersized for finite values of c (because of the Bonferroni bounds). Hence, the CSE CI has excess coverage probability for stationary ρ and has longer length than the usual stationary regression interval. So the CSE CIs are not uniform with the stationary case in the limit and they remain wider than the usual stationary intervals with their nominal coverage level understating the actual coverage probability.

The limit statistic λ_c used in the CSE construction of the CI is properly centered. The improperly centered unit root test statistic used in the test inversion to create the induced CI for c is effective in properly revealing that $c \to -\infty$ when ρ is stationary. But the critical values of the CSE test are not based on the correct limit theory (19) in this case, so the test is conservative because of the use of the Bonferroni bounds, and the induced CI correspondingly has incorrect coverage probability in the stationary limit. These asymptotic results have been confirmed in simulations (not reported here) which show that the coverage probability for CSE intervals is close to $(1-\alpha_2)100\%$ for stationary ρ, as predicted by the asymptotic theory.

4.2. *The Q Test*

CY (2006) recommended a different t-ratio test, called the Q test, that is based on the augmented regression equation (cf. Phillips and Hansen (1990)) $y_t = \beta x_{t-1} + \frac{\sigma_{0x}}{\sigma_{xx}}(x_t - \rho x_{t-1}) + u_{0.xt}$. Specifically, the Q test employs the following coefficient estimator (conditional on ρ and with no need to fit an intercept here):

$$\hat{\beta}(\rho) = \frac{\sum_{t=1}^{n} x_{t-1}\left[y_t - \frac{\hat{\sigma}_{x0}}{\hat{\sigma}_{xx}}(x_t - \rho x_{t-1})\right]}{\sum_{t=1}^{n} x_{t-1}^2}, \tag{20}$$

where $\hat{\sigma}_{x0}$ and $\hat{\sigma}_{xx}$ are obtained in the usual way from the least squares residuals in regressions of (15) and (16). The induced CI for β is based on the t statistic

$$t_{\hat{\beta}(\rho)} = \frac{\hat{\beta}(\rho) - \beta}{\sigma_{\hat{\beta}(\rho)}}, \quad \sigma^2_{\hat{\beta}(\rho)} = \hat{\sigma}^2_{00.x} \Big/ \sum_{t=1}^{n} x_{t-1}^2, \tag{21}$$

and an asymptotic normal distribution for $t_{\hat{\beta}(\rho)}$, together with the CI $[\rho_L, \rho_U]$ for ρ that is calculated using the unit root test inversion process (based on the statistic $t_{\hat{\rho}}$ calculated from the autoregression (16)) which produces an induced CI $[c_L, c_U]$ for c. In the numerical implementation of their test, CY bound the interval $[c_L, c_U]$ to lie within $[-50, 5]$, which arbitrarily restricts the allowable range of c (and hence ρ), inducing bias if the true value lies outside these bounds. This restriction is relaxed for the purpose of the following discussion, which explores the properties of the CY procedure for large $|c|$ and stationary ρ.

The asymptotic form of the induced interval $[\rho_L, \rho_U]$ is given by (10) in the stationary ρ case. This interval is asymptotically centered on $\bar{\rho} = \frac{3\rho-1}{\rho+1}$ and shrinks to this pseudo value as $n \to \infty$ when $|\rho| < 1$. It follows that the induced Bonferroni CI for β is from CY (equations (15)–(17), pp. 38–39) and conditional[2] on $\hat{\sigma}_{x0} < 0$ and $\rho_U, \rho_L > 0$:

$$[\beta_L(\rho_U), \beta_U(\rho_L)] = [\hat{\beta}(\rho_U) - z_{\alpha_2/2}\sigma_{\hat{\beta}(\rho)}, \hat{\beta}(\rho_L) + z_{\alpha_2/2}\sigma_{\hat{\beta}(\rho)}], \tag{22}$$

where $\sigma^2_{\hat{\beta}(\rho)} = \hat{\sigma}^2_{00.x}/\sum_{t=1}^n x^2_{t-1} = \hat{\sigma}^2_{00}(1-\hat{\delta}^2)/\sum_{t=1}^n x^2_{t-1} \to_p 0$. The interval (22) correspondingly shrinks as $n \to \infty$ to

$$\begin{aligned}[\beta_L(\rho_U), \beta_U(\rho_L)] \to \bar{\beta} &:= \beta + \frac{\sigma_{x0}}{\sigma_{xx}}(\rho - \bar{\rho}) \\ &= \beta + \frac{\sigma_{x0}}{\sigma_{xx}}\frac{(\rho-1)^2}{\rho+1} \neq \beta\end{aligned}$$

for all $|\rho| < 1$ whenever $\sigma_{x0} \neq 0$. It follows that the CY CI based on the Q test has zero coverage probability in the limit for all stationary $|\rho| < 1$ whenever there is regressor endogeneity ($\sigma_{x0} \neq 0$). Observe that the pseudo true value $\bar{\beta} \lesseqgtr \beta$ according as $\sigma_{x0} \lesseqgtr 0$ and the bias is greater the greater is $|\frac{\sigma_{x0}}{\sigma_{xx}}|$ and the smaller is ρ. Moreover, the Q test is biased and, when the true $\beta = 0$, the (two sided) test will erroneously indicate predictability with probability approaching unity as $n \to \infty$.[3]

Figure 2 shows actual coverage probabilities at a nominal asymptotic level of 90% of the CY CIs for the predictive regression coefficient β for regressors x_t with AR coefficient $\rho \in \{0.01, 0.03, \ldots, 0.99\}$, $n = 200$, and endogeneity coefficient $r_{0x} \in \{-0.99, -0.9, -0.6, -0.04\}$ where $r_{0x} = \delta = \sigma_{x0}/(\sigma_{xx}\sigma_{xx})^{1/2}$. The

[2]The estimate $\hat{\beta}(\rho)$ of β declines linearly with ρ when $\hat{\sigma}_{x0} < 0$, so the lower bound of the confidence interval for β is $\beta_L(\rho_U)$. If $\hat{\sigma}_{x0} > 0$ and $\rho_U, \rho_L > 0$, the corresponding confidence interval for β should be $[\beta_L(\rho_L), \beta_U(\rho_U)]$. This dependence of the interval on the signs of $\hat{\sigma}_{x0}$, ρ_L, and ρ_U does not appear to be mentioned in CY (2006), so their stated interval only applies when $\hat{\sigma}_{x0} < 0$ and $\rho_U, \rho_L > 0$. CY did assume that the true covariance $\sigma_{x0} < 0$ and, for roots ρ local to unity, seemed to presume that $\rho_U, \rho_L > 0$. Of course, $\hat{\sigma}_{x0} > 0$ with probability greater than zero even when the true covariance $\sigma_{x0} < 0$.

[3]One sided tests correspondingly have size unity or zero depending on the direction of the test. For instance, if $\sigma_{x0} < 0$ so that the probability limit $\bar{\beta} < \beta$, we would reject the null $\beta = 0$ in a left sided test against $\beta < 0$ with probability unity in the limit or in a right sided test against $\beta > 0$ with probability zero.

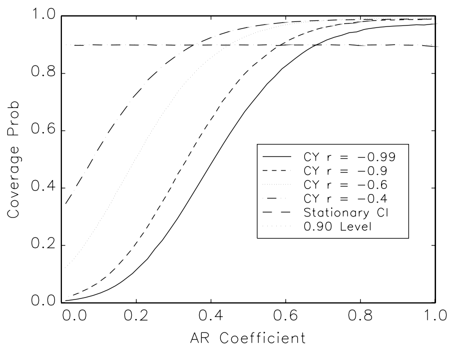


FIGURE 2.—Coverage probabilities of Campbell–Yogo and stationary confidence intervals for the predictive regression coefficient β plotted against the autoregressive coefficient ρ of x_t, shown for various values of the endogeneity coefficient r_{0x}. The nominal asymptotic level is 90%, sample size is $n = 200$, and the number of replications is 50,000.

results are based on 50,000 replications and use model (15) and the confidence belts shown in Figure 1 for the inversion of the unit root t-statistic $t_{\hat{\rho}}$. Evidently, the coverage probability monotonically declines with ρ, with sharper declines that approach zero when there is stronger endogeneity in the predictive regression (higher $|\delta_{0x}|$), thereby corroborating the limit theory. The graphs reveal that the CY Q test is typically undersized for ρ close to unity and seriously oversized when ρ is distant from unity. Also shown in Figure 2 is the coverage probability of the standard regression CI based on stationary x_t with strong endogeneity $r_{0x} = -0.99$. The stationary interval has close to nominal 90% coverage for all values of $\rho \leq 0.99$. Note that with $n = 200$, $\rho = 0.99$ corresponds to $c = -2.0$, so that this value of ρ may be regarded as being in the local to unity range.[4]

4.3. *Modifying the Q Test*

The Q test can be modified by using a CI for ρ that is based on a centered statistic such as (13), as in Mikusheva (2007), rather than a unit root test statistic. In this event, as shown above, the centered statistic $t_{\hat{\rho},\rho} \Rightarrow \lambda_c \sim N(0, 1) + O_p(|c|^{-1/2})$ as $c \to -\infty$, and under these conditions, the induced CI for ρ is approximately $[\rho_L, \rho_U] = \{\hat{\rho} - z_{\alpha_1/2}\sigma_{\hat{\rho}}, \hat{\rho} + z_{\alpha_1/2}\sigma_{\hat{\rho}}\}$ for a nominal level α_1 test, which is asymptotically valid for $|\rho| < 1$. The corresponding CI for β is

[4]Autoregressive bias is well known to be greater in the fitted intercept case. Correspondingly, the stationary test and confidence interval have more distortion, particularly for the strong endogenous case in the immediate vicinity of unity (see Figure 2 of Campbell and Yogo (2006)).

obtained using $t_{\hat{\beta}(\rho)}$ in (21) and leads, as before and under the same conditioning as (22), to

$$\text{(23)}\qquad \big[\beta_L(\rho_U), \beta_U(\rho_L)\big] = \big[\hat{\beta}(\rho_U) - z_{\alpha_2/2}\sigma_{\hat{\beta}(\rho)}, \hat{\beta}(\rho_L) + z_{\alpha_2/2}\sigma_{\hat{\beta}(\rho)}\big],$$

where $\sigma^2_{\hat{\beta}(\rho)} = \hat{\sigma}^2_{00.x}/\sum_{t=1}^n x^2_{t-1}$. But since the interval $[\rho_L, \rho_U]$ is now asymptotically valid for stationary ρ (as well as local to unity ρ), the coverage probability of (23) is at least $100(1-\alpha_2-\alpha_1)\%$ by Bonferroni. Hence, use of the centered test statistic $t_{\hat{\rho},\rho}$ for ρ leads to a robust interval for which the Bonferroni bound holds, and this construction of the CI avoids the zero coverage probability in the stationary case of the CY interval based on the Q test. Computation of this modified interval requires the use of confidence belts for ρ based on the centered statistic $t_{\hat{\rho},\rho}$.

5. SIMPLE EXTENSIONS

The analysis above used models with no fitted intercept. The same results on zero coverage probability and distended length of the CIs apply when demeaned and detrended data are used in the unit root t-tests on which the CIs are based. For example, in the demeaned case, we note that

$$\begin{aligned}(-2c)^{1/2}\tilde{J}_c(r) &= (-2c)^{1/2}J_c(r) - (-2c)^{1/2}\int J_c(s)\,ds \\ &= (-2c)^{1/2}J_c(r) + O_p\left(\frac{1}{|c|^{1/2}}\right),\end{aligned}$$

because $(-2c)^{1/2}\int J_c(s)\,ds$ has zero mean and variance $\frac{1}{-c} + O(\frac{1}{|c|^2}) \to 0$ as $c \to -\infty$. It follows that

$$\begin{aligned}(-2c)\int \tilde{J}(r)^2\,dr &= (-2c)\int J(r)^2\,dr - \left((-2c)^{1/2}\int J_c(s)\,ds\right)^2 \\ &= (-2c)\int J(r)^2\,dr + O_p\left(\frac{1}{|c|}\right),\end{aligned}$$

and

$$\begin{aligned}&(-2c)^{1/2}\int \tilde{J}_c(r)\,dW(r) \\ &\quad= (-2c)^{1/2}\int J_c(r)\,dW(r) - (-2c)^{1/2}\int J_c(s)\,ds W(1) \\ &\quad= \xi + O_p\big(|c|^{-1/2}\big),\end{aligned}$$

so the earlier arguments continue to apply with the same error order as $c \to -\infty$.

The results here also hold when we use test statistics based on a local alternative $\bar{\rho} = 1 + \frac{\bar{c}}{n}$ for some fixed $\bar{c} < 0$. The findings therefore apply to the procedure in Elliott and Stock (2001) involving inversion of a sequence of point optimal tests based on some fixed local alternative.

6. CONCLUSION

The method of Campbell and Yogo (2006) has been particularly popular among empirical researchers, is an industry standard in finance, and provides an econometric benchmark for competitors. But it suffers from the bias and zero coverage probability problems pointed out here and is only available in the simple regression case. Jansson and Moreira (2006) developed a conditional likelihood approach with certain optimal asymptotic properties that, in principle, extends to multiple regressors. They found in simulations that the Campbell–Yogo test had superior power against most alternatives. The Jansson–Moreira test has not yet been used in empirical research and simulations in Kasparis, Andreou, and Phillips (2012) indicate that the test encounters difficulties in numerical implementation and has size distortion in cases of strong endogeneity. In other recent work, Elliott, Müller, and Watson (2012) developed a "nearly optimal" test, treating the localizing coefficient c as a nuisance parameter and using a likelihood ratio test that optimizes weighted average power over $c \in [-40, 5]$ and switches to the standard t test when the maximum likelihood estimate $\hat{c} < -35$. The switching mechanism of this test ensures good behavior as $c \to -\infty$, so the procedure does not suffer from the problems inherent in the Campbell–Yogo test. Computational and design considerations make this method most useful in the scalar regressor case. Instrumental variable techniques are also available for use in predictive regression, such as the IVX method of Magdalinos and Phillips (2009) which applies to stationary, mildly integrated, and local to unity regressors (Kostakis, Magdalinos, and Stamatogiannis (2012)). That method has good size and power properties in simulations (Kostakis, Magdalinos, and Stamatogiannis (2012)), accommodates multiple regressors easily, and allows for varying degrees of persistence (as often occurs in empirical work) as well as mildly explosive roots. Implementation is by straightforward linear regression and the use of standard statistical tables, but the method requires user input on the parameter choice for instrument construction.

The applied predictive regression literature is large and continues to grow rapidly in empirical finance and macroeconomics. Against this background of applied research, the ongoing interest in econometrics in uniform procedures of inference, the challenges presented by multiple predictors, and the pitfalls pointed out in the current contribution, there is substantial need for continuing econometric research on methods of inference that can cope with potential nonstationarities in many regressors, control size, and deliver good discriminatory power in detecting predictability. The credibility of applied research on this subject ultimately depends on the reliability of the inferential machinery available in econometrics. There is still much to do.

APPENDIX

LEMMA A: *In a suitably expanded probability space as* $c \to -\infty$, *we have*

$$(24)\qquad (-2c)^{1/2}\int J_c(r)\,dW(r) = \xi + O_p\big(|c|^{-1/2}\big),$$

$$(25)\qquad (-2c)^{1/2}J_c(1) = \eta + O_p\big(e^{2c}\big),$$

$$(26)\qquad (-2c)\int J_c(r)^2 dr = 1 + 2\frac{\xi}{(-2c)^{1/2}}\big\{1 + O_p\big(|c|^{-1/2}\big)\big\} - \frac{\eta^2}{-2c}\big\{1 + O_p\big(e^{2c}\big)\big\},$$

where ξ *and* η *are independent* $N(0, 1)$.

PROOF: Result (24) follows by expanding the moment generating function of $\int J_c(r)\,dW(r)$ given in Phillips (1987) and Phillips, Magdalinos, and Giraitis (2010), result (25) follows directly from $(-2c)^{1/2}J_c(1) \equiv N(0, 1 - e^{2c})$, and (26) follows from (24)–(25) and the formula $(-2c)\int J_c(r)^2\,dr = 1 + 2\int J_c(r)\,dW(r) - J_c(1)^2$. Independence of ξ and η follows from $\mathbb{E}(\xi\eta) = 0$. *Q.E.D.*

LEMMA B: *For* $|\rho| < 1$, $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_t \Rightarrow N(0, V_w)$, *where* $w_t = (x_{t-1}u_t, u_t^2 - \sigma^2, x_{t-1}^2 - \frac{\sigma^2}{1-\rho^2})'$ *and* $V_w = \mathrm{diag}(\frac{\sigma^4}{1-\rho^2}, \kappa_4 + 2\sigma^4, \frac{1}{(1-\rho^2)^2}\{\kappa_4 + 2\sigma^4\frac{1+\rho^2}{1-\rho^2}\})$.

PROOF: The proof follows by standard time series CLT arguments (Phillips and Solo (1992)). *Q.E.D.*

PROOF OF THEOREM 1: Using (24) and (26), the stated result follows from

$$(27)\qquad \tau_c = \frac{(-2c)^{1/2}\int J_c(r)\,dW(r)}{\left((-2c)\int J_c(r)^2\,dr\right)^{1/2}} - \frac{|c|^{1/2}}{2^{1/2}}\left((-2c)\int J_c(r)^2\,dr\right)^{1/2}$$

$$= \frac{\xi\{1 + O_p(|c|^{-1/2})\}}{\left\{1 + \left(2\frac{\xi}{(-2c)^{1/2}} - \frac{\eta^2}{-2c}\right)\{1 + O_p(|c|^{-1/2})\}\right\}^{1/2}} - \frac{|c|^{1/2}}{2^{1/2}}\left\{1 + \left(2\frac{\xi}{(-2c)^{1/2}} - \frac{\eta^2}{-2c}\right)\big\{1 + O_p\big(|c|^{-1/2}\big)\big\}\right\}^{1/2}$$

$$
\begin{aligned}
&= \xi\left\{1 - \frac{\xi}{(-2c)^{1/2}} + O_p\left(\frac{1}{c}\right)\right\}\left\{1 + O_p\left(|c|^{-1/2}\right)\right\} \\
&\quad - \frac{|c|^{1/2}}{2^{1/2}}\left\{1 + \frac{\xi}{(-2c)^{1/2}} + O_p\left(\frac{1}{c}\right)\right\} \\
&= -\frac{|c|^{1/2}}{2^{1/2}} + \frac{1}{2}\xi + O_p\left(|c|^{-1/2}\right),
\end{aligned}
$$

and the fact that $\xi \equiv N(0,1)$ from Lemma A. *Q.E.D.*

PROOF OF THEOREM 2: By expansion of the components of $\hat{\tau}$ for $|\rho| < 1$, we have

$$
\begin{aligned}
\hat{\tau} &= t_{\hat{\rho},\rho} + \frac{\sqrt{n}(\rho - 1)}{\left\{\hat{\sigma}^2 \Big/ \left(\frac{1}{n}\sum_{t=1}^{n} x_{t-1}^2\right)\right\}^{1/2}} \\
&= t_{\hat{\rho},\rho} + \frac{\sqrt{n}(\rho - 1)}{(1-\rho^2)^{1/2}}\left(1 + \frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2}\right)^{-1/2} \\
&\quad \times \left\{1 + \frac{1}{n}\sum_{t=1}^{n}\left(x_{t-1}^2 - \frac{\sigma^2}{1-\rho^2}\right)\frac{(1-\rho^2)}{\sigma^2}\right\}^{1/2} \\
&= \frac{n^{-1/2}\sum_{t=1}^{n} x_{t-1}u_t}{(\sigma^4/(1-\rho^2))^{1/2}} - \sqrt{n}A_\rho^{1/2}\left\{1 - \frac{1}{2\sigma^2}\frac{1}{n}\sum_{t=1}^{n}\left(u_t^2 - \sigma^2\right)\right. \\
&\quad \left. + \frac{(1-\rho^2)}{2\sigma^2}\frac{1}{n}\sum_{t=1}^{n}\left(x_{t-1}^2 - \frac{\sigma^2}{1-\rho^2}\right)\right\} \\
&\quad + O_p\left(n^{-1/2}\right),
\end{aligned}
$$

where $A_\rho = \frac{1-\rho}{1+\rho}$. Then, setting $a' = (\frac{(1-\rho^2)^{1/2}}{\sigma^2}, \frac{A_\rho^{1/2}}{2\sigma^2}, -A_\rho^{1/2}\frac{(1-\rho^2)}{2\sigma^2})$ and using Lemma B, we obtain

$$
\begin{aligned}
\hat{\tau} &= -\sqrt{n}A_\rho^{1/2} + a'\frac{1}{\sqrt{n}}\sum_{t=1}^{n} w_t + O_p\left(n^{-1/2}\right) \qquad (28) \\
&= -\sqrt{n}A_\rho^{1/2} + \zeta + O_p\left(n^{-1/2}\right),
\end{aligned}
$$

where $\zeta = N(0, v_\zeta)$ with $v_\zeta = a'V_w a = 1 + \frac{A_\rho}{2\sigma^4}(\kappa_4 + \frac{2\sigma^4}{1-\rho^2})$. Hence

$$\hat{\tau}^2 \pm z_{\alpha/2}\hat{\tau} = \{-\sqrt{n}A_\rho^{1/2} + \zeta + o_p(1)\}^2 \pm z_{\alpha/2}\{-\sqrt{n}A_\rho^{1/2} + \zeta + o_p(1)\} \\ = nA_\rho - (2\zeta \pm z_{\alpha/2})\sqrt{n}A_\rho^{1/2}\{1 + o_p(1)\}. \tag{29}$$

Then, up to $O_p(n^{-1/2})$, the asymptotic CI (9) for ρ is

$$[\rho_L, \rho_U] = \left[1 - 2A_\rho + 2\frac{2\zeta - z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}, 1 - 2A_\rho + 2\frac{2\zeta + z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\right], \tag{30}$$

and the coverage probability can be computed as

$$\begin{aligned} &\mathbb{P}\{\rho \in [\rho_L, \rho_U]\} \\ &= \mathbb{P}\left\{1 - 2A_\rho + 2\frac{2\zeta - z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2} \le \rho \le 1 - 2A_\rho + 2\frac{2\zeta + z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\right\} \\ &= \mathbb{P}\left\{2\frac{2\zeta - z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2} \le \rho + 2A_\rho - 1 \le 2\frac{2\zeta + z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\right\}. \end{aligned} \tag{31}$$

Now $\rho + 2A_\rho - 1 = (1-\rho)A_\rho$. So for $|\rho| < 1$, (31) is

$$\begin{aligned} &\mathbb{P}\left\{2\frac{2\zeta - z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2} \le (1-\rho)A_\rho \le 2\frac{2\zeta + z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\right\} \\ &= \mathbb{P}\left\{\frac{\sqrt{n}}{4}(1-\rho)A_\rho^{1/2} - \frac{z_{\alpha/2}}{2} \le \zeta \le \frac{\sqrt{n}}{4}(1-\rho)A_\rho^{1/2} + \frac{z_{\alpha/2}}{2}\right\} \\ &\to 0, \quad \text{if} \quad \sqrt{n}(1-\rho)A_\rho^{1/2} \to \infty \text{ as } n \to \infty, \end{aligned}$$

giving the stated result (11) since $\sqrt{n}(1-\rho)A_\rho^{1/2} \to \infty$ iff $n^{1/3}(1-\rho) \to \infty$. To calculate the length of the CI when $|\rho| < 1$, we deduce from (30) that $|\rho_L - \rho_U| = \frac{4z_{\alpha/2}}{\sqrt{n}}A_\rho^{1/2}\{1 + O_p(\frac{1}{\sqrt{n}})\}$, as required. *Q.E.D.*

REFERENCES

CAMPBELL, J., AND M. YOGO (2006): "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81 (1), 27–60. [1177,1178,1185,1187-1189,1191]

CAVANAGH, C., G. ELLIOTT, AND J. H. STOCK (1995): "Inference in Models With Nearly Integrated Regressors," *Econometric Theory*, 11 (05), 1131–1147. [1177,1185]

ELLIOTT, G., AND J. H. STOCK (2001): "Confidence Intervals for Autoregressive Coefficients Near One," *Journal of Econometrics*, 103, 155–181. [1177,1191]

ELLIOTT, G., U. K. MÜLLER, AND M. W. WATSON (2012): "Nearly Optimal Tests When a Nuisance Parameter Is Present Under the Null Hypothesis," Working Paper, UCSD. [1191]

GIRAITIS, L., AND P. C. B. PHILLIPS (2006): "Uniform Limit Theory for Stationary Autoregression," *Journal of Time Series Analysis*, 27, 51–60. [1183,1185]

HANSEN, B. E. (1999): "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics*, 81, 594–607. [1177,1178,1181,1184]

JANSSON, M., AND M. J. MOREIRA (2006): "Optimal Inference in Regression Models With Nearly Integrated Time Series," *Econometrica*, 74, 681–714. [1191]

KASPARIS, I., E. ANDREOU, AND P. C. B. PHILLIPS (2012): "Nonparametric Predictive Regression," Cowles Foundation Discussion Paper 1878, Yale University. [1191]

KOSTAKIS, A., A. MAGDALINOS, AND M. STAMATOGIANNIS (2012): "Robust Econometric Inference for Stock Return Predictability," Unpublished Manuscript, University of Nottingham. [1191]

MAGDALINOS, T., AND P. C. B. PHILLIPS (2009): "Econometric Inference in the Vicinity of Unity," Working Paper, Yale University. [1191]

MIKUSHEVA, A. (2007): "Uniform Inference in Autoregressive Models," *Econometrica*, 75, 1411–1452. [1177,1178,1184,1189]

——— (2014): "Second Order Expansion of the t-Statistic in AR(1) Models," *Econometric Theory* (forthcoming). [1185]

PHILLIPS, P. C. B. (1987): "Towards a Unified Asymptotic Theory for Autoregression," *Biometrika*, 74, 535–547. [1177,1179,1180,1192]

PHILLIPS, P. C. B., AND B. E. HANSEN (1990): "Statistical Inference in Instrumental Variables Regression With I(1) Processes," *Review of Economic Studies*, 57, 99–125. [1187]

PHILLIPS, P. C. B., AND J. H. LEE (2013): "Predictive Regression Under Various Degrees of Persistence and Robust Long-Horizon Regression," *Journal of Econometrics*, 177, 250–264. [1185]

PHILLIPS, P. C. B., AND T. MAGDALINOS (2007): "Limit Theory for Moderate Deviations From a Unit Root," *Journal of Econometrics*, 136, 115–130. [1183,1185]

PHILLIPS, P. C. B., AND V. SOLO (1992): "Asymptotics for Linear Processes," *The Annals of Statistics*, 20, 971–1001. [1192]

PHILLIPS, P. C. B., T. MAGDALINOS, AND L. GIRAITIS (2010): "Smoothing Local-to-Moderate Unit Root Theory," *Journal of Econometrics*, 158, 274–279. [1192]

STOCK, J. H. (1991): "Confidence Intervals for the Largest Autoregressive Root in US Macroeconomic Time Series," *Journal of Monetary Economics*, 28 (3), 435–459. [1177,1179,1182,1185, 1186]

Cowles Foundation for Research in Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06511, U.S.A., University of Auckland, Singapore Management University, and University of Southampton; peter.phillips@yale.edu.

Manuscript received September, 2012; final revision received August, 2013.