Journal of Monetary Economics 28 (1991) 435–459. North-Holland

Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series

James H. Stock*
University of California - Berkeley, Berkeley, CA 94720, USA

Received December 1990, final version received May 1991

This paper provides asymptotic confidence intervals for the largest autoregressive root of a time series when this root is close to one. The intervals are readily constructed either graphically or using tables in the appendix. When applied to the Nelson–Plosser (1982) data set, the main conclusion is that the confidence intervals typically are wide. The conventional emphasis on testing for whether the largest root equals one fails to convey the substantial sampling variability associated with this measure of persistence.

1. Introduction

A prominent problem in empirical macroeconomics during the past decade has been the measurement of the persistence of shocks to macroeconomic time series variables. Since Nelson and Plosser (1982), much of this literature has focused on the size of the largest autoregressive root (ρ) of a time series, and tests for whether ρ is one have played a central role in the empirical analysis. This emphasis on unit root tests, which in part is attributable to the availability of appropriate statistical theory, has been criticized on several grounds. While macroeconomic theories suggest substantial serial dependence in time series data, a unit root typically is predicted only as a special case. Moreover, reporting only unit root tests and point estimates of the largest root is unsatisfying as a description of the data: this fails to convey

*The author thanks Christopher Cavanagh, Thomas Rothenberg, and Mark Watson for helpful discussions and Allan Gregory and Robin Lumsdaine for comments on an earlier draft. Financial support by the National Science Foundation through grant no. SES-89-10601 is gratefully acknowledged.

0304-3932/91/$03.50

information about the sampling uncertainty or, more precisely, the range of models (i.e., values of ρ) that are consistent with the observed data. While not new [see for example Campbell and Mankiw (1987) and Cochrane (1988)], these criticisms suggest that confidence intervals for ρ could provide a more useful summary measure of persistence than unit root tests alone.

This paper reports asymptotic confidence intervals for ρ, calculated for the fourteen historical U.S. annual macroeconomic time series studied by Nelson and Plosser (1982). The methodological contribution of the paper is to provide a set of figures and tables for use in constructing confidence intervals for ρ when ρ is large. Because the distribution of the t-statistic testing ρ is nonnormal and depends strongly on ρ when ρ is nearly one, the usual approach of constructing asymptotic confidence intervals as the point estimate ± 2 standard errors is not appropriate here. Moreover, as Cavanagh (1985), Sims (1988), and Sims and Uhlig (1988) emphasized, the first-order asymptotic theory does not provide a suitable framework for the construction of confidence intervals because it is discontinuous at $\rho = 1$. Instead, the confidence intervals reported here are constructed using the local-to-unity asymptotic theory developed by Bobkoski (1983), Cavanagh (1985), Phillips (1987), and Chan and Wei (1987). In this theory, the true value of ρ is modeled as being in a decreasing neighborhood of one, specifically $\rho = 1 + c/T$, where c is a fixed constant (the Pitman drift) and T is the sample size. This device – nesting ρ as a function of the sample size – is analogous to the usual approach used to study the asymptotic power of econometric tests against local alternatives, except that in the conventional case the alternatives are in a $1/\sqrt{T}$ rather than a $1/T$ neighborhood of the null value. Cavanagh (1985) originally described how to use this theory to construct confidence intervals based on the t-statistic testing $\rho = 1 + c/T$ for first-order autoregressions with no intercept in the regression. This paper extends his approach to the empirically more relevant case of higher-order autoregressions with an intercept, or an intercept and a time trend. Two sets of confidence intervals are studied here: one based on the augmented Dickey–Fuller (1979) (ADF) t-statistic testing $\rho = 1$, and one based on a modification of Sargan and Bhargava's (1983) uniformly most powerful test statistic (the MSB statistic).

The main new empirical result in this paper is that the confidence intervals for ρ for many of the annual Nelson–Plosser series are wide. As Nelson and Plosser emphasized, the ADF statistic rejects $\rho = 1$ against $\rho < 1$ at the 5% level only for the unemployment rate, so unemployment is the only series for which the 90% central confidence interval for ρ (based on Nelson and Plosser's ADF statistics) falls below one. But these intervals also include values of ρ substantially different from one. The 90% intervals for real GNP and real per capita GNP, based on 62 years of data, are approximately (0.6, 1.04) using the detrended ADF statistic. This provides additional empirical content to the often-voiced view, recently expressed for example by Christiano and Eichenbaum (1990), that either difference-stationary or trend-stationary models are capable of producing the autocorrelations observed in U.S. output data. Some series, however, have substantially tighter intervals than GNP. The series with the tightest interval estimates are industrial production, consumer prices, velocity, and stock prices – the series with the most observations – and the bond yield. For example, the 90% *ADF* interval for consumer prices, on which there are 111 annual observations, is (0.901, 1.037) and it is (0.873, 1.039) for stock prices (the S&P 500). These tighter intervals in part reflect the longer samples available for these series than for GNP. Reporting solely the results of unit root tests fails to convey the evident imprecision with which the largest root is estimated in many of these series, even with these long annual data.

The paper is organized as follows. The method for constructing these confidence intervals is summarized in section 2. Section 3 reports a Monte Carlo experiment that examines the finite-sample performance of the intervals. The empirical results are reported and discussed in section 4, and some conclusions are summarized in section 5. Tables of central confidence intervals as a function of the *ADF* statistics are provided in the appendix.

2. Local-to-unity asymptotic confidence intervals

2.1. The model and statistics

Let the univariate time series y_t obey

$$
\begin{aligned}
&y_t = \mu_0 + \mu_1 t + v_t, \\
&a(L)v_t = \varepsilon_t, \qquad a(L) = b(L)(1-\rho L), \qquad (1)\\
&\qquad\qquad\qquad\qquad t = 1,\dots,T,
\end{aligned}
$$

where $b(L) = \sum_{j=0}^{k} b_j L^j$, where $b_0 = 1$ and L is the lag operator [so that the lag polynomial $a(L)$ has order $k+1$], $b(1) \neq 0$, $v_0 = 0$, and ε_t is a martingale difference sequence with $\mathrm{E}\varepsilon_t^2 = \sigma^2$ and $\sup_t \mathrm{E}\varepsilon_t^4 < \infty$, with μ_0 and μ_1 nonzero in general. The factorization of $a(L)$ is used to distinguish the largest root, $\rho = 1 + c/T$, from the fixed stable roots describing short-run dynamics in $b(L)$.[1]

[1]An alternative nesting for ρ, used by Phillips (1987), is $\rho = \exp(c/T)$. Because $\exp(c/T) = 1 + c/T + \mathrm{O}(T^{-2})$, the asymptotic representations obtained in this section are the same for either nesting.

The representation (1) can be rearranged to yield the usual Dickey–Fuller regression,

$$y_t = \tilde{\mu}_0 + \tilde{\mu}_1 t + \alpha(1) y_{t-1} + \sum_{j=1}^{k} \alpha_{j-1}^{*} \Delta y_{t-j} + \varepsilon_t, \tag{2}$$

where, with $\rho = 1 + c/T$, $\alpha(L) = L^{-1}(1 - a(L))$ so $\alpha(1) = 1 + cb(1)/T$, $\tilde{\mu}_0 = -cb(1)\mu_0/T - cb^*(1)\mu_1/T + \rho b(1)\mu_1$, $\tilde{\mu}_1 = -cb(1)\mu_1/T$, $b_i^* = -\sum_{j=i+1}^{k} b_j$, and $\alpha_i^* = -\sum_{j=i+1}^{k} \alpha_j$. The *ADF* t-statistic, denoted by $\hat{\tau}^{\tau}$, is the t-statistic testing the hypothesis that $\alpha(1) = 1$ in (2).

Sargan and Bhargava (1983) proposed a different test statistic, motivated as the uniformly most powerful test statistic for testing $\rho = 1$ against the stationary alternative using an approximation to the Gaussian likelihood when $\mu_1 = 0$. Bhargava (1986) extended these results to the case of nonzero μ_0 and μ_1, showing the Sargan–Bhargava statistic to be locally most powerful invariant when computed using detrended y_t, where the detrended data are $y_t^B = y_t - (t-1)/(T-1)y_T - (T-t)/(T-1)y_1 - (\bar{y} - \frac{1}{2}(y_T + y_1))$. Although the Sargan–Bhargava statistic has these optimality properties in the first-order Gaussian case, the test is not similar when $b(1) \neq 1$. As is shown in Stock (1988), however, it is readily modified to provide an asymptotically similar test statistic. Let the spectral density of $(1 - \rho L)v_t$ at frequency zero be $\omega^2/2\pi$, so that $\omega = \sigma/b(1)$, and estimate ω^2 by $\hat{\omega}^2 = \hat{\sigma}^2/(1 - \hat{\alpha}^*(1))^2$, where σ and $\alpha^*(L)$ are estimated from the regression (2). When μ_0 and μ_1 are possibly nonzero, the modified Sargan–Bhargava statistic (in logarithms), computed using y_t^B, is

$$MSB^B = \tfrac{1}{2} \ln \left\{ \hat{\omega}^{-2} T^{-2} \sum_{t=1}^{T} \left(y_t^B \right)^2 \right\}. \tag{3}$$

The regression (2) includes t as a regressor, and the detrended series y_t^B is used to construct the *MSB* statistic in (3). This is appropriate if μ_0 and μ_1 are not restricted *a priori*, and this will be referred to as the 'detrended' case. Alternatively μ_1, but not necessarily μ_0, might be known to be zero. Then the appropriate *ADF* statistic is $\hat{\tau}^{\mu}$, the t-statistic testing $\alpha(1) = 1$ in (2) excluding the time trend, and the *MSB* statistic is computed using $y_t^{\mu} \equiv y_t - \bar{y}$ (rather than y_t^B) and is denoted MSB^{μ}. This will be referred to as the 'demeaned' case.

2.2. Asymptotic distributions

One approach to constructing confidence intervals for ρ would be to assume a distribution for ε_t and to derive the exact finite-sample confidence intervals based on an appropriate test statistic. Because any specific distributional assumption typically would not be satisfied in practice, the justification for such an approach would be that, in large samples, it might nonetheless provide a good approximation under more general conditions. This suggests instead computing confidence intervals with an explicit asymptotic justification, which is the approach taken here.

Limiting representations for the statistics at hand are obtained using the local-to-unity asymptotic distribution theory given in Bobkoski (1983), Cavanagh (1985), Chan (1988), Chan and Wei (1987), and Phillips (1987) [for a different approach, see Ahtola and Tiao (1984)]. For a technical review of this literature, see Nabeya and Tanaka (1990, sect. 1). The basic result in this literature is that the process v_t in (1) obeys a functional limit theorem, in which $V_T(\lambda) = T^{-1/2} v_{[T\lambda]}$ converges to a diffusion process as $T \to \infty$, where $[\cdot]$ is the greatest lesser integer function. Specifically, $V_T(\cdot) \Rightarrow \omega J(\cdot)$, where $J(\cdot)$ satisfies $\mathrm{d}J(s) = cJ(s)\,\mathrm{d}s + \mathrm{d}W(s)$, where $W(\cdot)$ is a standard Brownian motion and '$\Rightarrow$' denotes weak convergence in $D[0,1]$. If $c = 0$ so $\rho = 1$, this specializes to the more familiar limit, $V_T(\cdot) \Rightarrow \omega W(\cdot)$. These results are extended here to include additional regressors using the techniques of Sims, Stock, and Watson (1990).

For μ_0 and μ_1 possibly nonzero, the appropriate statistics involve detrending. It is shown in appendix A that, when $\rho = 1 + c/T$ and (2) includes a constant and a time trend,

$$T(\hat{\alpha}(1) - 1) \Rightarrow b(1)\left\{\left(\int_0^1 J^\tau(s)^2\,\mathrm{d}s\right)^{-1} \int_0^1 J^\tau(s)\,\mathrm{d}W(s) + c\right\}, \tag{4}$$

$$\hat{\tau}^\tau \Rightarrow \left(\int_0^1 J^\tau(s)^2\,\mathrm{d}s\right)^{1/2}\left\{\left(\int_0^1 J^\tau(s)^2\,\mathrm{d}s\right)^{-1} \int_0^1 J^\tau(s)\,\mathrm{d}W(s) + c\right\}, \tag{5}$$

$$MSB^B \Rightarrow \tfrac{1}{2}\ln\left\{\int_0^1 J^B(s)^2\,\mathrm{d}s\right\}, \tag{6}$$

where $J^\tau(\lambda) = J(\lambda) - \int_0^1 (4 - 6s)J(s)\,\mathrm{d}s - \lambda \int_0^1 (12s - 6)J(s)\,\mathrm{d}s$ and $J^B(\lambda) = J(\lambda) - (\lambda - \frac{1}{2})J(1) - \int_0^1 J(s)\,\mathrm{d}s$.

The limiting representations (4)–(6) are the same in the demeaned case, except that $J^\mu(\cdot)$ replaces $J^\tau(\cdot)$ and $J^B(\cdot)$, where $J^\mu(\lambda) = J(\lambda) - \int_0^1 J(s)\,\mathrm{d}s$.

2.3. Construction of asymptotic confidence intervals

The distributions corresponding to (4)–(6) are nonnormal and the dependence on c is not a simple location shift, so confidence intervals for ρ cannot be formed using a simple '± 2 standard error' rule. Still, because the representations (5) and (6) depend only on c and are continuous in c, the *ADF* and *MSB* test statistics can be used as the basis for interval estimation.

Recall that a $100(1-\alpha)\%$ confidence set for c, $S(y_1,\dots,y_T)$, is a set-valued function of the data with the property that $\Pr[c \in S(y_1,\dots,y_T)] = 1-\alpha$ for all values of c. In general, a confidence set can be constructed by 'inverting' the acceptance region of a test statistic that has a distribution which depends on c but not on the nuisance parameters. To be concrete, consider confidence sets based on $\hat{\tau}^{\tau}$. If $A_{\alpha}(c_0)$ is the (1- or 2-sided) asymptotic acceptance region for a level α test of the null of $c=c_0$, then $S(\hat{\tau}^{\tau}) = \{c\colon \hat{\tau}^{\tau} \in A_{\alpha}(c)\}$ is a $100(1-\alpha)\%$ confidence set. Because $\hat{\tau}^{\tau}$ is a scalar, a $100(1-\alpha)\%$ closed confidence set can be constructed as $S(\hat{\tau}^{\tau}) = \{c\colon f_{l;\alpha_l}(c) \le \hat{\tau}^{\tau} \le f_{u;\alpha_u}(c)\}$, where $f_{l;\alpha_l}(c)$ and $f_{u;\alpha_u}(c)$ are respectively the lower and upper α_l and $1-\alpha_u$ percentiles of $\hat{\tau}^{\tau}$ as a function of c, where $\alpha_l + \alpha_u = \alpha$. If $f_{l;\alpha_l}(c)$ and $f_{u;\alpha_u}(c)$ are strictly monotone increasing in c, the critical values can be inverted to yield the more familiar representation, $S(\hat{\tau}^{\tau}) = \{c\colon f_{u;\alpha_u}^{-1}(\hat{\tau}^{\tau}) \le c \le f_{l;\alpha_l}^{-1}(\hat{\tau}^{\tau})\}$. Here, we construct central confidence intervals, so that $\alpha_l = \alpha_u = \frac{1}{2}\alpha$.

A simple way to construct these intervals is to use the graphical device described by Kendall and Stuart (1967, ch. 20). Asymptotic local-to-unity central confidence belts (the graph of $\{f_{l;\frac{1}{2}\alpha}(c), f_{u;\frac{1}{2}\alpha}(c)\}$) are plotted in figs. 1–4 for, respectively, the demeaned *ADF* t-statistic $\hat{\tau}^{\mu}$, the detrended *ADF* t-statistic $\hat{\tau}^{\tau}$, the demeaned *MSB* statistic MSB^{μ}, and the detrended *MSB* statistic MSB^{B}. The computation of these belts by Monte Carlo simulation is described in appendix B. In each figure, the four bands describe the 95% (the widest band), 90%, 80%, and 70% confidence belts. The central line plots the median of the local-to-unity distribution of the test statistic. The $100(1-\alpha)\%$ confidence set is given by those values of c falling within the $1-\alpha$ belt for a given value of the statistic. Each $1-\alpha$ confidence belt has the property that, for a given value of c, the asymptotic probability of realizing a value of the statistic inside the belt is $1-\alpha$. For any true value of c, the confidence intervals constructed using the belt will contain c if and only if the realized statistic falls within the belt. Thus, the asymptotic probability that the confidence interval contains the true value of c is $1-\alpha$.[2]

As an example, suppose $\hat{\tau}^{\mu} = -3.0$ is calculated from a series with $T=100$. The 95% confidence interval is the c in the 95% belt in fig. 1, read

[2]The $\hat{\tau}^{\mu}$ and $\hat{\tau}^{\tau}$ statistics used here are centered around one rather than ρ (the conventional approach). The reason for centering around one is to eliminate the dependence of the distribution of the t-statistic on the nuisance parameters in the higher-order case. Cavanagh (1985), who focused on the first-order case [so that $b(1)=1$], described the construction of confidence intervals based on the t-statistic centered around $\alpha(1)=\rho$, that is, based on $\hat{\tau}(\rho) = (\hat{\alpha}(1)-\rho)/SE(\hat{\alpha}(1))$, where $\hat{\alpha}(1)$ and $SE(\hat{\alpha}(1))$ are computed using (2). In general the distribution of $\hat{\tau}(\rho)$ depends not only on c but also on $b(1)$, so its critical values cannot be inverted to obtain confidence intervals without adjusting for $b(1)$.

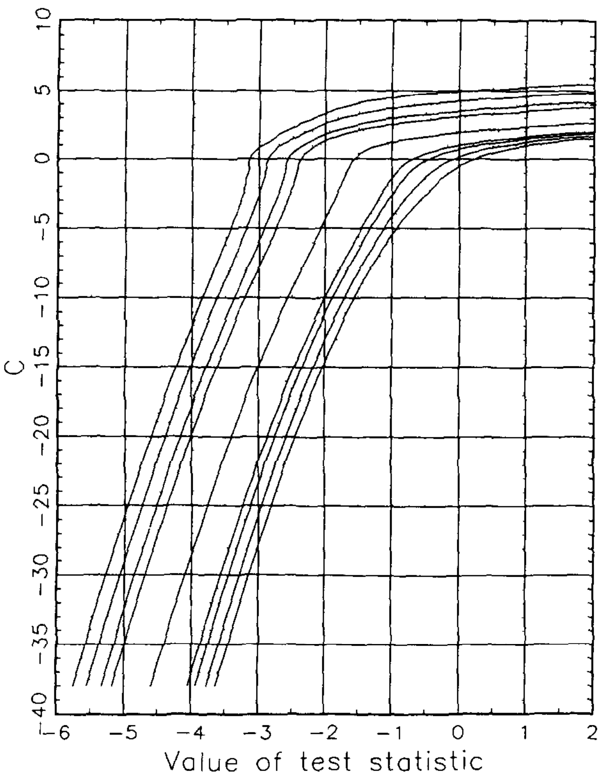

Fig. 1. Confidence belt for local-to-unity parameter c based on demeaned *ADF* t-statistic; bands in order of decreasing width: 95%, 90%, 80%, 70%, central line: median.

vertically for $\hat{\tau}^{\mu} = -3.0$ (or alternatively taken from table A.1, part A), which is $-27.9 \le c \le 0.8$. The 95% confidence interval for ρ is $(1 - 27.9/100, 1 + 0.8/100) = (0.721, 1.008)$. Because the medians in figs. 1 and 3 are monotone increasing in c, an asymptotically median-unbiased estimator is obtained using the central line in fig. 1 (or the final column of table A.1, part A). Because the median of $\hat{\tau}^{\mu}$ is -3.0 when $c = -14.9$, $\hat{c}^{\text{med}} = -14.9$ is an asymptotically median-unbiased estimate of c, corresponding to $\hat{\rho}^{\text{med}} = 1 - 14.9/100 = 0.851$. Based on the level of numerical accuracy used to produce figs. 1–4, it appears that the medians of $\hat{\tau}^{\tau}$ and MSB^{B} bend backwards for c between zero and one, so estimators thus constructed using $\hat{\tau}^{\tau}$ and MSB^{B} are not median unbiased. However, because the range of values of $\hat{\tau}^{\tau}$ and MSB^{B} over which these curves bend backwards is very small [specifically, $\hat{\tau}^{\tau} \in (-1.346, -1.334)$ and $MSB^{B} \in (-2.193, -2.184)$], the bias introduced by using this estimator with the detrended statistics appears to be negligible.

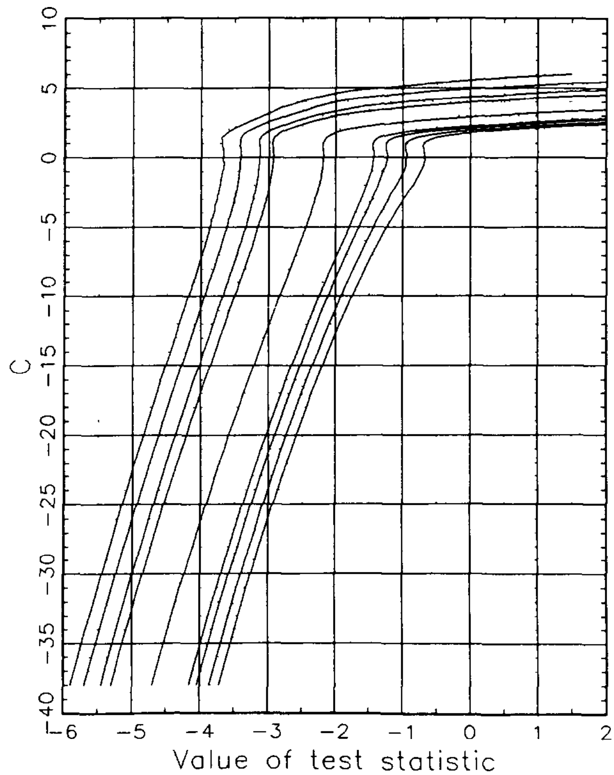

Fig. 2. Confidence belt for local-to-unity parameter c based on detrended ADF t-statistic; bands in order of decreasing width: 95%, 90%, 80%, 70%, central line: median.

2.4. Discussion

Six aspects of these results are noteworthy. First, the confidence belts are nonlinear, exhibiting a sharp bend for c just above zero. For positive values of $\hat{\tau}^{\tau}$ or $\hat{\tau}^{\mu}$ the confidence intervals are tight, for large negative values they are wide. Large positive values of $\hat{\tau}^{\tau}$ are unlikely to be realized unless c is positive, but negative values of $\hat{\tau}^{\tau}$ are likely to be realized whether c is positive or negative. A simple calculation demonstrates how different are the widths of the interval estimates for different realizations of the test statistic: if, for example, $\hat{\tau}^{\tau} = 0$, the sample must have $T = 75$ for the 95% interval to have width 0.05, but if $\hat{\tau}^{\tau} = -3.5$ is observed, T must be 725 to produce this short an interval.

Second, the detrended belts do not increase monotonically, so for some values of $\hat{\tau}^{\tau}$ the central confidence set will be disjoint. Cavanagh (1985) pointed out that disjoint sets are theoretically possible in the local-to-unity

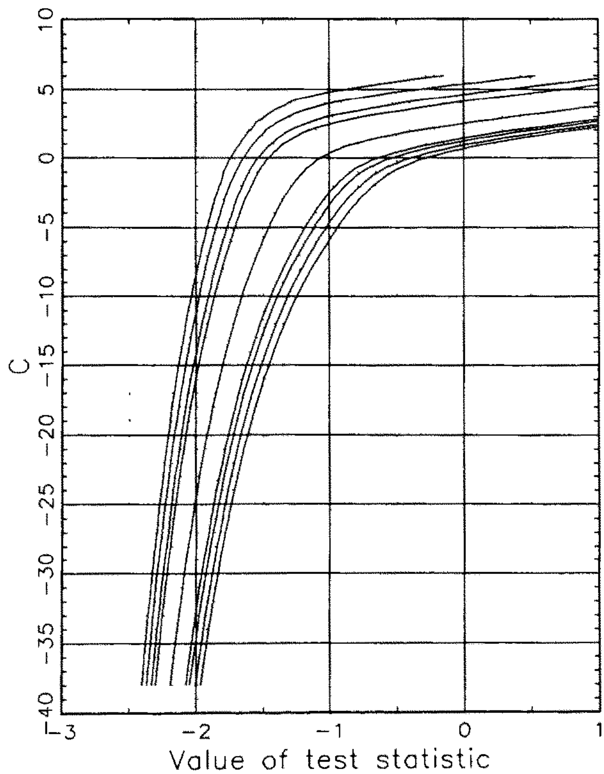

Fig. 3. Confidence belt for local-to-unity parameter c based on demeaned modified Sargan–Bhargava statistic; bands in order of decreasing width: 95%, 90%, 80%, 70%, central line: median.

setting, but his computations did not uncover any in the nondemeaned first-order case. Sims (1988) found disjoint confidence sets for ρ in the nondemeaned first-order model using first-order asymptotic theory, which is discontinuous in ρ, and conjectured that exact finite-sample distributions (which are continuous in ρ) also might result in disjoint confidence sets. Using the asymptotic local-to-unity confidence intervals, the Cavanagh–Sims conjecture is not borne out in the demeaned case, although it is in the detrended case. This is, however, of little practical importance. The largest range of discontinuities is for the 95% belt, in which case disjoint confidence sets obtain when $\hat{\tau}^{\tau}$ falls between (−3.66, −3.69) and (−0.66, −0.71). In table A.1, this issue is addressed by reporting only the outer bounds of the confidence intervals in these ranges, so that the intervals actually have asymptotic confidence coefficients slightly greater than $1-\alpha$. Note that this

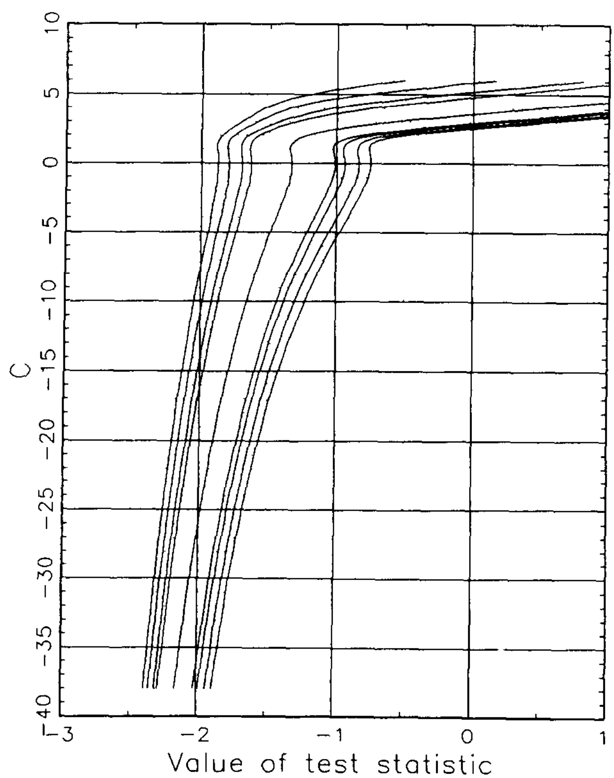

Fig. 4. Confidence belt for local-to-unity parameter c based on detrended modified Sargan–Bhargava statistic; bands in order of decreasing width: 95%, 90%, 80%, 70%, central line: median.

results in a discontinuous jump in the confidence interval (as a function of $\hat{\tau}^{\tau}$) in these regions.

Third, because the local-to-unity distribution of $T(\hat{\alpha}(1)-1)$ is skewed and moreover depends on the nuisance parameters $b(1)$, the relation between $\hat{\alpha}(1)$ and the confidence interval constructed by inverting the *ADF* t-statistic is complicated. The point estimate generally will not be at the center of the confidence interval.

Fourth, the confidence intervals based on the detrended statistics are larger than for the demeaned statistics. To be concrete, consider the median confidence interval, taken to be the confidence interval computed for the median value of the *ADF* statistic for a given value of c. For c between -20 and 2, the median 90% interval based on $\hat{\tau}^{\tau}$ is uniformly longer than the median 90% interval based on $\hat{\tau}^{\mu}$, assuming $\mu_1 = 0$. For example, for

$c=-5$, the median $\hat{\tau}^{\mu}$ is -2.06, with a 90% confidence interval for c of $(-13.6, 2.4)$, whereas the median $\hat{\tau}^{\tau}$ is -2.45, with 90% confidence interval $(-16.2, 3.4)$.

Fifth, the *MSB* confidence belts have the same general properties as the *ADF* confidence belts. Intervals are wider for large negative values of the statistic. Like the $\hat{\tau}^{\tau}$ intervals, there is a small range of MSB^{B} for which the confidence set is disjoint.

Sixth, an alternative to the asymptotic approach used here is to construct confidence intervals and median-unbiased estimators for ρ using finite-sample techniques. This approach has recently been adopted by Andrews (1990), who used exact distribution theory to construct confidence intervals and median-unbiased estimators in the Gaussian AR(1) model, and by Rudebusch (1990), who used Monte Carlo techniques to construct median-unbiased estimators in the Gaussian AR(k) model. The principal advantages of the asymptotic approach relative to the finite-sample approaches are the simplifications that arise in handling the nuisance parameters and its validity under a wide range of assumptions on the distribution of ε_t.

3. Monte Carlo analysis

The asymptotic analysis of section 2 serves two main purposes: to show that in large samples the local power functions of the *ADF* and *MSB* statistics depend only on c so that they can be used to construct asymptotic confidence intervals, and to provide large-sample approximations to the finite-sample distributions of these statistics when ρ is near one. It is well known [Schwert (1989)] that unit root tests statistics can have finite-sample distributions that differ markedly from their asymptotic approximations under the unit root null when there are nuisance parameters, specifically when there is a moving average error. A Monte Carlo analysis of the local-to-unity confidence intervals was therefore performed to assess the finite sample performance of these asymptotic approximations when ρ is near one. The probability model examined was the nearly-integrated moving average model,

$$(1-\rho L)y_t=(1+\theta L)\varepsilon_t,\quad \varepsilon_t \text{ i.i.d. } \mathrm{N}(0,1), \tag{7}$$

where $\rho=1+c/T$. The *ADF* and *MSB* statistics were computed in both the demeaned and detrended cases. For $T=100$, k in (2) was set to 4, and for $T=200$, k was set to 5. The experiment examined $c=(2,0,-2,-5,-10)$ and $\theta=(0.5,0,-0.5)$. Note that, for $\theta\neq 0$, the finite-order autoregressive approximation is misspecified so that in these cases the experiment examines both specification error and the effect of having a finite sample.

Table 1

Monte Carlo results:
Finite-sample coverage probabilities for local-to-unity asymptotic confidence intervals with asymptotic confidence coefficients 0.95, 0.90, 0.80, and 0.70.[a]

Model: $(1-\rho L)y_t = (1+\theta L)\varepsilon_t$, ε_t i.i.d. N(0, 1), $\rho = 1 + c/T$

		$\hat{\tau}^{\mu}$				$\hat{\tau}^{\tau}$				MSB^{μ}				MSB^{B}			
θ	c	0.95	0.90	0.80	0.70	0.95	0.90	0.80	0.70	0.95	0.90	0.80	0.70	0.95	0.90	0.80	0.70
Part A. $T = 100$																	
0.0	2	0.954	0.906	0.803	0.691	0.933	0.878	0.766	0.663	0.893	0.834	0.729	0.634	0.817	0.741	0.630	0.533
0.0	0	0.937	0.883	0.776	0.677	0.931	0.874	0.771	0.665	0.877	0.809	0.700	0.604	0.830	0.760	0.652	0.563
0.0	−2	0.926	0.867	0.764	0.665	0.926	0.870	0.767	0.667	0.867	0.803	0.698	0.601	0.826	0.753	0.645	0.550
0.0	−5	0.918	0.857	0.749	0.656	0.922	0.867	0.763	0.664	0.832	0.765	0.657	0.559	0.787	0.722	0.609	0.523
0.0	−10	0.896	0.834	0.722	0.625	0.913	0.849	0.734	0.628	0.775	0.698	0.586	0.502	0.731	0.655	0.551	0.461
0.5	2	0.958	0.908	0.804	0.691	0.933	0.875	0.766	0.660	0.895	0.839	0.722	0.624	0.807	0.726	0.613	0.517
0.5	0	0.929	0.868	0.764	0.664	0.926	0.869	0.770	0.662	0.876	0.805	0.694	0.599	0.832	0.765	0.655	0.555
0.5	−2	0.924	0.870	0.765	0.664	0.922	0.867	0.768	0.674	0.863	0.800	0.693	0.601	0.827	0.757	0.651	0.553
0.5	−5	0.913	0.860	0.755	0.659	0.923	0.867	0.762	0.665	0.840	0.770	0.661	0.560	0.807	0.739	0.630	0.531
0.5	−10	0.885	0.820	0.699	0.601	0.904	0.837	0.719	0.619	0.770	0.695	0.582	0.499	0.749	0.667	0.561	0.471
−0.5	2	0.948	0.899	0.805	0.703	0.934	0.873	0.765	0.656	0.922	0.869	0.764	0.671	0.852	0.779	0.670	0.572
−0.5	0	0.939	0.881	0.775	0.675	0.926	0.870	0.767	0.662	0.914	0.855	0.750	0.651	0.870	0.804	0.701	0.611
−0.5	−2	0.925	0.868	0.763	0.661	0.922	0.860	0.758	0.653	0.904	0.845	0.747	0.648	0.864	0.802	0.696	0.601
−0.5	−5	0.925	0.866	0.752	0.652	0.924	0.870	0.764	0.661	0.880	0.814	0.711	0.608	0.840	0.779	0.670	0.574
−0.5	−10	0.915	0.852	0.741	0.646	0.914	0.855	0.746	0.641	0.828	0.757	0.640	0.535	0.776	0.697	0.588	0.495

Part B. $T = 200$

0.0	2	0.960	0.912	0.810	0.701	0.944	0.887	0.783	0.677	0.922	0.869	0.762	0.662	0.878	0.808	0.700	0.606
0.0	0	0.940	0.883	0.781	0.682	0.940	0.886	0.786	0.680	0.903	0.836	0.731	0.632	0.893	0.832	0.724	0.625
0.0	−2	0.940	0.883	0.779	0.682	0.936	0.877	0.778	0.679	0.903	0.840	0.733	0.633	0.886	0.820	0.710	0.606
0.0	−5	0.927	0.872	0.765	0.664	0.930	0.876	0.775	0.673	0.885	0.821	0.707	0.609	0.866	0.798	0.682	0.583
0.0	−10	0.919	0.860	0.748	0.650	0.926	0.872	0.762	0.651	0.855	0.780	0.665	0.564	0.832	0.757	0.645	0.546
0.5	2	0.956	0.909	0.804	0.695	0.942	0.885	0.775	0.669	0.918	0.861	0.753	0.655	0.878	0.808	0.701	0.603
0.5	0	0.941	0.888	0.778	0.682	0.936	0.880	0.780	0.678	0.903	0.840	0.731	0.633	0.881	0.815	0.708	0.609
0.5	−2	0.939	0.885	0.782	0.680	0.937	0.882	0.784	0.680	0.898	0.839	0.731	0.629	0.875	0.809	0.704	0.607
0.5	−5	0.929	0.875	0.772	0.670	0.932	0.883	0.780	0.678	0.874	0.808	0.695	0.591	0.861	0.794	0.679	0.580
0.5	−10	0.915	0.855	0.743	0.645	0.925	0.866	0.754	0.650	0.831	0.762	0.657	0.565	0.822	0.749	0.634	0.540
−0.5	2	0.951	0.905	0.806	0.703	0.941	0.883	0.776	0.666	0.930	0.881	0.771	0.672	0.905	0.841	0.733	0.633
−0.5	0	0.941	0.891	0.787	0.687	0.938	0.887	0.791	0.687	0.933	0.873	0.775	0.672	0.919	0.863	0.760	0.657
−0.5	−2	0.940	0.886	0.785	0.681	0.939	0.885	0.783	0.679	0.927	0.870	0.767	0.668	0.915	0.858	0.754	0.654
−0.5	−5	0.929	0.875	0.773	0.678	0.936	0.879	0.778	0.681	0.919	0.862	0.762	0.662	0.905	0.844	0.733	0.633
−0.5	−10	0.930	0.874	0.771	0.669	0.933	0.881	0.772	0.668	0.888	0.823	0.712	0.616	0.877	0.807	0.696	0.593

[a]The entries are the fraction of Monte Carlo replications for which the asymptotic local-to-unity confidence intervals contained the true value of ρ, where $\rho = 1 + c/T$. The *ADF* statistics were computed using $k = 4$ for $T = 100$ and $k = 5$ for $T = 200$. The *MSB* statistics were computed using $\hat{\omega}^2 = \hat{\sigma}^2/(1 - \hat{\alpha}^*(1))^2$, where σ and $\alpha^*(1)$ were estimated from (2) (including t as a regressor for the MSB^B statistic, excluding t for the MSB^μ statistic), with $k = 4$ for $T = 100$ and $k = 5$ for $T = 200$. For each replication, ten initial observations were drawn. The coverage rates are based on 10,000 Monte Carlo replications for each (T, θ, c) experiment.

Table 1 reports the fraction of times that the calculated central confidence interval contains the true value of c for different experiments. Because this is just the fraction of times that the computed statistic falls within the upper and lower $\frac{1}{2}\alpha$ percentiles for that value of c, the coverage rates in table 1 were computed from the fraction of pseudo-random test statistics that reject the null hypothesis that $\rho = 1 + c/T$ in a two-sided level α test, where the critical values for each c are those used to construct figs. 1–4.

For *ADF* intervals, the asymptotic approximations perform well. For both $\hat{\tau}^{\tau}$ and $\hat{\tau}^{\mu}$ with $T = 100$, in all cases the empirical coverage rates of the asymptotic 90% confidence interval are between 82% and 91%. The performance of the *ADF* interval estimator is insensitive to θ but deteriorates somewhat as c becomes large and negative. The performance of the *MSB* intervals, particularly the MSB^{B} intervals, is less satisfactory than the *ADF* intervals. For example, for $c = -5$, $\theta = 0$, and $T = 100$, the Monte Carlo coverage rate of the asymptotic 90% MSB^{B} interval is 72.2%, compared with 86.7% for the corresponding $\hat{\tau}^{\tau}$ interval; for $T = 200$, this MSB^{B} coverage rate remains under 80%.[3] Overall, these results suggest that the *ADF* intervals will be more reliable in empirical applications.

4. Empirical results

Table 2 presents 90% and 80% confidence intervals for the largest autoregressive root ρ for the fourteen annual series studied by Nelson and Plosser (1982). Because of the poor Monte Carlo performance of the *MSB* intervals, the empirical intervals were computed using the detrended *ADF* statistic. Panel A reports estimates based on Nelson and Plosser's choice of the number of lags (k) included in (2). The 90% central confidence interval for ρ (based on $\hat{\tau}^{\tau}$) is below one for the unemployment rate and above one for the bond yield, while $\rho = 1$ is included in all the other intervals.

The calculations were repeated using $k = 5$ for each of the series; the results are reported in panel B of table 2. The primary qualitative conclusion from panel A – the striking width of the confidence intervals – remains unchanged, although several estimated intervals shift. The main differences occur for the three GNP series, the unemployment rate, real wages, velocity, and the S&P 500; each of these intervals is shifted up.

As was noted in the introduction, the series with the tightest confidence intervals are the bond yield, industrial production, consumer prices, stock prices, and velocity. The confidence intervals for the bond yield are tight because the test statistics are relatively positive, suggesting a root greater

[3]Additional Monte Carlo experiments (not reported) suggest that the poor performance of the *MSB* intervals when t is included in (2) and for large negative c arises from the imprecision of $\hat{\alpha}^*(1)$, which is used to construct $\hat{\omega}$. This suggests investigating alternative spectral density estimators in the local-to-unity model, a topic left for future research.

Table 2

Asymptotic confidence intervals for ρ for the Nelson–Plosser data set.[a]

Series	N	k	$ADF\ \hat{\tau}^{\tau}$	90% intervals	80% intervals
			Part A. Nelson–Plosser lag lengths		
Real GNP	62	1	−2.994	(0.604, 1.042)	(0.646, 1.031)
Nominal GNP	62	1	−2.321	(0.757, 1.060)	(0.793, 1.049)
Real per capita GNP	62	1	−3.045	(0.591, 1.041)	(0.634, 1.029)
Industrial production	111	5	−2.529	(0.836, 1.031)	(0.857, 1.026)
Employment	81	2	−2.655	(0.757, 1.039)	(0.787, 1.032)
Unemployment rate	81	3	−3.552	(0.577, 0.950)	(0.615, 0.893)
GNP deflator	82	1	−2.516	(0.787, 1.041)	(0.815, 1.034)
Consumer prices	111	3	−1.972	(0.901, 1.037)	(0.922, 1.031)
Wages	71	2	−2.236	(0.800, 1.054)	(0.833, 1.045)
Real wages	71	1	−3.049	(0.644, 1.035)	(0.681, 1.025)
Money stock	82	1	−3.078	(0.687, 1.030)	(0.719, 1.020)
Velocity	102	0	−1.663	(0.929, 1.042)	(0.950, 1.035)
Bond yield	71	2	0.686	(1.032, 1.075)	(1.034, 1.067)
S&P 500	100	2	−2.122	(0.873, 1.039)	(0.896, 1.033)
			Part B. Uniform lag lengths		
Real GNP	62	5	−2.123	(0.780, 1.068)	(0.820, 1.057)
Nominal GNP	62	5	−1.788	(0.847, 1.074)	(0.886, 1.062)
Real per capita GNP	62	5	−2.222	(0.760, 1.066)	(0.800, 1.055)
Industrial production	111	5	−2.529	(0.836, 1.031)	(0.857, 1.026)
Employment	81	5	−2.565	(0.764, 1.043)	(0.794, 1.035)
Unemployment rate	81	5	−2.835	(0.715, 1.037)	(0.746, 1.029)
GNP deflator	82	5	−2.466	(0.784, 1.044)	(0.813, 1.036)
Consumer prices	111	5	−2.369	(0.855, 1.033)	(0.876, 1.028)
Wages	71	5	−2.124	(0.811, 1.059)	(0.845, 1.049)
Real wages	71	5	−2.564	(0.728, 1.049)	(0.762, 1.041)
Money stock	82	5	−3.005	(0.685, 1.033)	(0.718, 1.024)
Velocity	102	5	−0.741	(1.015, 1.049)	(1.018, 1.042)
Bond yield	71	5	0.597	(1.033, 1.078)	(1.035, 1.069)
S&P 500	100	5	−1.062	(0.982, 1.048)	(1.016, 1.042)

[a]The detrended *ADF* statistic ($\hat{\tau}^{\tau}$) was obtained by estimating the regression (2), including a constant, a time tend, and k lags of Δy_t. The data are annual, with all series ending in 1970. N denotes the total number of observations on each series, including observations used for initial conditions, so that, in the notation of the paper, $T = N - k - 1$. The 90% and 80% asymptotic confidence intervals were computed using appendix table A.1, linearly interpolated, as described in section 2.

than one. As is evident from figs. 1–4, positive values of the test statistic produce much tighter intervals than do large negative values. The relatively tight intervals for industrial production, consumer prices, stock prices, and velocity arise from the greater number of observations on these series.

The width of the intervals in table 2 raises the question of whether tighter intervals could be obtained using more frequent observations over the same span of years, say quarterly rather than annual data were they available. A calculation, simplest for the *MSB* statistic, indicates that the answer is no.

Suppose that T years of data are used to compute the MSB^B statistic, first using quarterly data, then using the quarterly data aggregated to the annual level. Also suppose a sufficient number of lags are included in the quarterly and annual regressions to yield consistent estimators of ω. With a root local to unity, the quarterly and annual MSB^B statistics will be equal asymptotically, so the confidence intervals for c will be the same, say (c_0, c_1). The confidence interval for the quarterly root computed using the quarterly data is $(1 + c_0/4T,\ 1 + c_1/4T)$. But this quarterly interval converted to an annual basis is $(1 + c_0/T,\ 1 + c_1/T)$ to order $O(T^{-2})$, the same as computed using the annual data.

Two caveats should be borne in mind when interpreting the width of these intervals. First, no attempt has been made to construct optimal intervals; rather only central intervals are given. Presumably, the reported intervals overstate somewhat the sampling variability relative to optimal intervals. Second, the finite-sample properties of these intervals have been studied only in the rarefied experimental design of section 3, and further simulation experiments are in order.

5. Summary and discussion

These procedures provide asymptotic confidence intervals for the largest autoregressive root of a nearly nonstationary time series variable. Because of the nonstandard distribution theory, the relation between the observed t-statistic (or p-value) and the confidence interval for ρ is complicated. Thus substantial additional information beyond whether or not a unit root test rejects is revealed by formally constructing these interval estimates.

The classical confidence intervals developed here can be contrasted to recent Bayesian approaches to the unit root problem. In part in reaction to the discontinuity in the first-order asymptotic theory, Sims (1988) and Sims and Uhlig (1988) suggested computing Bayesian interval estimates for ρ. This has been implemented empirically by DeJong and Whiteman (1989), Schotman and van Dijk (1989), and Phillips (1990) using various priors.

The classical analysis here has several advantages over these Bayesian approaches. First, it sidesteps the debate over priors. As Phillips (1990) emphasizes, the flat and Jeffreys priors differ most in their treatment of roots near and greater than one, and not surprisingly the posteriors – and thus inferences about ρ – differ sharply depending on the choice of prior. The choice of prior is further complicated in this problem because it must be specified over the nuisance parameters as well as over the parameter of interest, ρ. Second, the classical approach does not require the additional conceptual device of treating the unknown parameters as random. Third, and most important, the classical confidence intervals are precise expressions of a common form of reasoning in the 'unit roots' debate in empirical macroeconomics: if a computed test statistic is a likely realization from some hypothesized model (value of ρ), then that model ought to be treated as possibly true. Christiano and Eichenbaum (1990) can be interpreted as using this logic to argue that specific models of interest are within classical confidence sets of some reasonable but unspecified confidence coefficient. In contrast, the Bayesian posteriors can be interpreted only with reference to the priors, the appropriateness of which are inherently difficult to judge. Were econometricians able to agree on the best priors for reporting results to a general scientific readership, or were inferences on ρ framed as an explicit decision problem, then the Bayesian approach would have more appeal; but neither condition is satisfied here.

The main empirical message of table 2 is that the confidence intervals for ρ are wide. For all series except unemployment and the bond yield, the intervals contain one, but they also contain values that could be substantially different from one in terms of their implications for quantities of interest to macroeconomists. This sampling uncertainty is large despite having more than a century of observations on several of the series. A next step in this research is to calculate confidence intervals for the several-year-ahead impulse response function analyzed by Campbell and Mankiw (1987) and subsequent researchers, allowing for a root that is nearly, but not exactly, one. Although that calculation is beyond the scope of this paper, the findings here [and those in Christiano and Eichenbaum (1990) and Rudebusch (1990)] suggest that the resulting confidence intervals would be wide relative to ones calculated under a maintained unit root assumption.

Appendix A: Derivation of eqs. (4), (5), and (6)

The results provided here apply to the 'detrended' case. The results for the 'demeaned' statistics are obtained by dropping the deterministic time trend terms in these derivations. The approach used to derive (4)–(6) is to rewrite the regression (1) in 'canonical form' as defined by Sims, Stock, and Watson (1990), in which the regressors are transformed so that their limiting moment matrix is nonsingular. The distribution of statistics from the canonical regression is then used to obtain the distribution of the statistics of interest in the original regression (2).

Write the regression (2) as

$$y_t = \beta' X_{t-1} + \varepsilon_t, \tag{A.1}$$

where $X_{t-1} = (\Delta y_{t-1}, \ldots, \Delta y_{t-k}, 1, y_{t-1}, t)$ and $\beta = (\beta_1', \beta_2, \beta_3, \beta_4)'$, where $\beta_1 = (\alpha_1^*, \ldots, \alpha_k^*)'$, $\beta_2 = \tilde{\mu}_0$, $\beta_3 = \alpha(1)$, and $\beta_4 = \tilde{\mu}_1$. Because a constant and t are included in the regression, without loss of generality set $\mu_0 = \mu_1 = 0$, so $\tilde{\mu}_0 = \tilde{\mu}_1 = 0$. Also set $X_0 = 0$.

The canonical regression is obtained by rewriting (A.1) so that all but three of the regressors have mean zero and are stationary. Let $\tilde{\Delta} = (1-\rho L)$ and let $u_t = b(L)^{-1}\varepsilon_t$, so that $\tilde{\Delta} y_t = u_t$ and $\mathrm{E}\,\tilde{\Delta} y_t = 0$. The canonical regression is

$$y_t = \tilde{\mu}_0 + \tilde{\mu}_1 t - \sum_{j=1}^{k} b_j \tilde{\Delta} y_{t-j} + \rho y_{t-1} + \varepsilon_t, \tag{A.2}$$

where $b(L)$ is defined in (1). Written more compactly, this is

$$y_t = \delta' Z_{t-1} + \varepsilon_t, \tag{A.3}$$

where $Z_{t-1} = (Z_{t-1}^{1\prime}, Z_{t-1}^2, Z_{t-1}^3, Z_{t-1}^4)'$, where $Z_{t-1}^1 = (\tilde{\Delta} y_{t-1}, \dots, \tilde{\Delta} y_{t-k})'$, $Z_{t-1}^2 = 1$, $Z_{t-1}^3 = y_{t-1}$, and $Z_{t-1}^4 = t$.

The transformation from the original regressors X_{t-1} to the canonical regressors Z_{t-1} is

$$Z_{t-1} = \begin{bmatrix} \rho & 0 & \cdot & \cdot & \cdot & 0 & 0 & 1-\rho & 0 \\ -(1-\rho) & \rho & \cdot & \cdot & \cdot & 0 & 0 & 1-\rho & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -(1-\rho) & -(1-\rho) & \cdot & \cdot & \cdot & \rho & 0 & 1-\rho & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 1 & 0 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 1 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \Delta y_{t-2} \\ \cdot \\ \cdot \\ \Delta y_{t-k} \\ 1 \\ y_{t-1} \\ t \end{bmatrix}$$

$$= DX_{t-1}. \tag{A.4}$$

Let $\Upsilon_T = \mathrm{diag}(T^{1/2} I_k, T^{1/2}, T, T^{3/2})$, where I_k is the $k \times k$ identity matrix, and let $\hat{\delta}$ be the OLS estimator of δ, $\hat{\delta} = (\sum_{t=1}^{T} Z_{t-1} Z_{t-1}')^{-1} (\sum_{t=1}^{T} Z_{t-1} y_t)$. The results of Bobkoski (1983), Chan and Wei (1987, lemma 2.1), and Phillips (1987, lemma 1(a)) show that $V_T(\cdot) \Rightarrow \omega J(\cdot)$, where $\omega^2 = \sigma^2/b(1)^2 = 2\pi$ times the spectral density of u_t at frequency zero, $V_T(\lambda) = T^{-1/2} v_{[T\lambda]}$ and $J(\cdot)$ satisfies $\mathrm{d}J(\lambda) = cJ(\lambda) + \mathrm{d}W(\lambda)$, where $W(\lambda)$ is a standard Brownian motion and $J(0) = 0$. This result, combined with Sims, Stock, and Watson (1990, theorem 1), yields

$$\Upsilon_T(\hat{\delta} - \delta) \Rightarrow \{(\Gamma_{11}^{-1}\phi_1)', (\Gamma_{22}^{-1}\phi_2)'\}', \tag{A.5}$$

where $\Gamma_{11} = \mathrm{E} Z_t^1 Z_t^{1\prime}$ [so that $(\Gamma_{11})_{ij} = \mathrm{E} u_t u_{t-j+i}$], ϕ_1 is distributed $\mathrm{N}(0, \Gamma_{11}\sigma^2)$, ϕ_1 is independent of (Γ_{22}, ϕ_2), Γ_{22} is a symmetric 3×3 matrix with elements $(\Gamma_{22})_{11} = 1$, $(\Gamma_{22})_{12} = \omega \int_0^1 J(s)\,\mathrm{d}s$, $(\Gamma_{22})_{13} = \frac{1}{2}$, $(\Gamma_{22})_{22} = \omega^2 \int_0^1 J(s)^2\,\mathrm{d}s$, $(\Gamma_{22})_{23} = \omega \int_0^1 sJ(s)\,\mathrm{d}s$, and $(\Gamma_{22})_{33} = \frac{1}{3}$, and where $\phi_2 = \sigma\{W(1), \omega\int_0^1 J(s)\,\mathrm{d}W(s), \int_0^1 s\,\mathrm{d}W(s)\}'$. Note that $\hat{\beta} = D'\hat{\delta}$, so (A.5) implies that $\hat{\beta} - \beta \xrightarrow{p} 0$.

A.1. Derivation of (4)

From $\hat{\beta} = D'\hat{\delta}$ and the definition of D in (A.4), $\hat{\beta}_3 = (1-\rho)\sum_{j=1}^{k}\hat{\delta}_{1,j} + \hat{\delta}_3$. Because $\hat{\beta}_3 = \hat{\alpha}(1)$, $T(\hat{\alpha}(1)-1) = T(1-\rho)\sum_{j=1}^{k}\hat{\delta}_{1,j} + T(\hat{\delta}_3 - 1)$. Direct calculation shows that $T(\hat{\delta}_3 - \delta_3) \Rightarrow (\sigma/\omega)(\int_0^1 J^\tau(s)^2\,\mathrm{d}s)^{-1}(\int_0^1 J^\tau(s)\,\mathrm{d}W(s))$, where J^τ is defined in section 2. From (A.2) and (A.3), $\delta_3 = \rho$ and $\sum_{j=1}^{k}\delta_{1,j} = -\sum_{j=1}^{k}b_j = 1 - b(1)$. Thus,

$$T(\hat{\alpha}(1)-1) = -c\sum_{j=1}^{k}\hat{\delta}_{1,j} + T(\hat{\delta}_3 - \delta_3) + T(\rho - 1)$$

$$\Rightarrow b(1)\left\{\left(\int_0^1 J^\tau(s)^2\,\mathrm{d}s\right)^{-1}\left(\int_0^1 J^\tau(s)\,\mathrm{d}W(s)\right) + c\right\}.$$

A.2. Derivation of (5)

Using the device in Sims, Stock, and Watson (1990, theorem 2), one obtains

$$\hat{\tau}^\tau = \left\{\hat{\sigma}^2\left(T^{-2}\sum_{t=1}^{T}(y_{t-1}^\tau)^2\right)^{-1}\right\}^{-1/2} T(\hat{\alpha}(1)-1) + o_p(1), \qquad \text{(A.6)}$$

where y_t^τ is the residual from regressing y_t on $(1,t)$. Because $\mathrm{E}\varepsilon_t^4 < \infty$ and $\hat{\beta} - \beta \xrightarrow{p} 0$, $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$. The result (5) follows from (4), (A.6), and $V_T^\tau(\cdot) \Rightarrow \omega J^\tau(\cdot)$, where $V_T^\tau(\lambda) = T^{-1/2}y_{[T\lambda]}^\tau$.

A.3. Derivation of (6)

From $\hat{\beta} = D'\hat{\delta}$ and (A.4), $\sum_{j=1}^{k}\hat{\beta}_{1,j} = \rho\sum_{j=1}^{k}\hat{\delta}_{1,j} - (1-\rho)\sum_{j=1}^{k}(j-1)\hat{\delta}_{1,j} \xrightarrow{p} 1 - b(1)$. Thus,

$$\hat{\omega}^2 = \hat{\sigma}^2 \Bigg/ \left(1 - \sum_{j=1}^{k}\hat{\beta}_{1,j}\right)^2 \xrightarrow{p} \sigma^2/b(1)^2 = \omega^2. \qquad \text{(A.7)}$$

Let $V_T^B(\lambda) = T^{-1/2}y_{[T\lambda]}^B$. Then $V_T^B(\cdot) \Rightarrow \omega J^B(\cdot)$, where $J^B(\lambda) = J(\lambda) - (\lambda - \frac{1}{2})J(1) - \int_0^1 J(s)\,\mathrm{d}s$ [this follows from $V_T(\cdot) \Rightarrow \omega J(\cdot)$ and by straightforward calculations]. The desired expression (6) obtains from these results and the definition of the MSB^B statistic in (3).

Appendix B: Numerical issues in the tabulation and computation of the confidence belts

Table A.1 summarizes the central confidence intervals obtained by inverting the *ADF* statistic as discussed in section 2. The tables report the minimal and maximal limits of the confidence set, so in the small ranges for which the confidence set is disjoint, the tabulated interval joins the outer limits of the set. Confidence intervals for observed values of the statistics can be obtained by linear interpolation. (This introduces some numerical inaccuracy in the neighborhood of $\hat{\tau}^{\tau}$'s or MSB^{B}'s for which the confidence sets are disjoint.) In some cases bounds are omitted because the lower bound falls outside the range of c used for the calculations. If so, the confidence interval constructed from the table is a $1-\frac{1}{2}\alpha$ open interval.

Various procedures are available for the evaluation of the limiting distribution of the *ADF* $\hat{\tau}$ and $\hat{\rho}$ statistics. The literature has focused on the case with no deterministic regressors (with $\mu_1=\mu_2=0$). Dickey and Fuller (1979) provide representations in terms of infinite sums of independent normal variates when $c=0$; Cavanagh (1985) and Chan (1988) generalize these to nonzero c. Bobkoski (1983) and Perron (1989) numerically invert moment generating functions, and Nabeya and Tanaka (1990) compute limiting distributions using the theory of Fredholm determinants. Results in Chan (1988), Nabeya and Tanaka (1990), and Perron (1989) suggest that the asymptotic approximations work well for Gaussian AR(1) models in finite samples, even for $T=50$ and certainly for $T=500$. These latter results imply that suitable approximations to the limiting distribution can be obtained by Monte Carlo simulation with $T=500$, where the number of replications is sufficiently large to provide the desired numerical accuracy. Indeed, Chan's (1988) comparison of several numerical procedures in the nondemeaned $\mu_1=\mu_2=0$ case led him to conclude that direct Monte Carlo simulation with T large produced the most reliable approximations.

Chan's (1988) recommendation is adopted here, and the limiting distributions were evaluated by Monte Carlo simulation for $T=500$ with 20,000 replications. The pseudo-data were generated according to $y_t=\rho y_{t-1}+\varepsilon_t$, ε_t i.i.d. N(0, 1), with $\rho=1+c/T$ and $y_0=0$. The distributions were evaluated on a grid of 87 values of c for $-38\leq c\leq 6$, with the grid most dense on $(-5,6)$. For each c, the percentiles of $\hat{\tau}^{\mu}$, $\hat{\tau}^{\tau}$, MSB^{μ}, and MSB^{B} (computed with $k=0$) were recorded. The 0.025, 0.05, 0.10, 0.15, 0.50, 0.85, 0.90, 0.95, and 0.975 percentiles are plotted in figs. 1–4. The intervals in table A.1 were computed by linear interpolation of the *ADF*-based confidence belt as a function of the statistic, using the outer bounds of the belt in the disjoint cases. Computer procedures in RATS and GAUSS to calculate these intervals are available from the author on request.

Table A.1

Confidence belts for ρ based on *ADF* statistics.

Stat	95%		90%		80%		70%		Median
	c_0	c_1	c_0	c_1	c_0	c_1	c_0	c_1	
Part A. Based on demeaned *ADF* t-statistic $\hat{\tau}^{\mu}$									
−5.70	—	−37.12	—	—	—	—	—	—	—
−5.60	—	−35.41	—	—	—	—	—	—	—
−5.50	—	−33.70	—	−37.17	—	—	—	—	—
−5.40	—	−32.04	—	−35.50	—	—	—	—	—
−5.30	—	−30.47	—	−33.84	—	−37.57	—	—	—
−5.20	—	−28.88	—	−32.17	—	−35.87	—	—	—
−5.10	—	−27.30	—	−30.60	—	−34.17	—	−36.73	—
−5.00	—	−25.81	—	−28.98	—	−32.52	—	−35.05	—
−4.90	—	−24.32	—	−27.41	—	−30.88	—	−33.42	—
−4.80	—	−22.87	—	−25.94	—	−29.26	—	−31.77	—
−4.70	—	−21.35	—	−24.46	—	−27.70	—	−30.15	—
−4.60	—	−19.93	—	−22.97	—	−26.19	—	−28.58	—
−4.50	—	−18.56	—	−21.53	—	−24.71	—	−27.02	−36.49
−4.40	—	−17.21	—	−20.11	—	−23.26	—	−25.55	−34.82
−4.30	—	−15.88	—	−18.72	—	−21.84	—	−24.08	−33.21
−4.20	—	−14.51	—	−17.35	—	−20.44	—	−22.62	−31.63
−4.10	—	−13.23	—	−15.98	—	−19.09	—	−21.19	−30.05
−4.00	—	−11.96	—	−14.68	—	−17.78	−37.07	−19.83	−28.52
−3.90	—	−10.70	—	−13.39	−37.43	−16.48	−35.38	−18.49	−27.02
−3.80	—	−9.45	—	−12.10	−35.72	−15.19	−33.74	−17.17	−25.53
−3.70	—	−8.19	−36.79	−10.89	−34.06	−13.97	−32.13	−15.89	−24.09
−3.60	−37.46	−7.00	−35.11	−9.74	−32.44	−12.77	−30.55	−14.63	−22.67
−3.50	−35.79	−5.87	−33.47	−8.57	−30.86	−11.58	−28.99	−13.43	−21.28
−3.40	−34.13	−4.74	−31.86	−7.45	−29.31	−10.42	−27.48	−12.25	−19.93
−3.30	−32.51	−3.48	−30.27	−6.31	−27.78	−9.30	−26.01	−11.06	−18.63
−3.20	−30.93	−1.57	−28.72	−5.19	−26.29	−8.15	−24.56	−9.96	−17.35
−3.10	−29.39	0.32	−27.20	−3.98	−24.83	−7.03	−23.15	−8.86	−16.11
−3.00	−27.86	0.80	−25.71	−2.55	−23.39	−5.91	−21.77	−7.78	−14.90
−2.90	−26.39	1.09	−24.25	−0.71	−22.00	−4.77	−20.45	−6.67	−13.72
−2.80	−24.86	1.39	−22.85	0.43	−20.64	−3.59	−19.13	−5.61	−12.57
−2.70	−23.41	1.74	−21.54	0.88	−19.32	−2.39	−17.85	−4.54	−11.46
−2.60	−22.03	1.97	−20.19	1.16	−18.03	−0.56	−16.61	−3.36	−10.39
−2.50	−20.69	2.26	−18.86	1.44	−16.79	0.40	−15.41	−2.10	−9.33
−2.40	−19.41	2.55	−17.60	1.69	−15.60	0.81	−14.26	−0.30	−8.29
−2.30	−18.13	2.75	−16.38	1.94	−14.43	1.12	−13.12	0.43	−7.29
−2.20	−16.93	2.91	−15.23	2.16	−13.30	1.35	−12.05	0.79	−6.32
−2.10	−15.76	3.08	−14.08	2.35	−12.21	1.55	−11.00	1.06	−5.35
−2.00	−14.65	3.28	−13.00	2.56	−11.15	1.74	−9.99	1.30	−4.38
−1.90	−13.52	3.51	−11.96	2.72	−10.15	1.92	−9.02	1.49	−3.37
−1.80	−12.43	3.67	−10.91	2.87	−9.19	2.09	−8.06	1.64	−2.29
−1.70	−11.37	3.80	−9.91	3.02	−8.24	2.24	−7.14	1.80	−1.15
−1.60	−10.36	3.93	−8.96	3.13	−7.33	2.38	−6.25	1.92	−0.24
−1.50	−9.41	4.06	−8.05	3.25	−6.46	2.49	−5.38	2.05	0.26
−1.40	−8.53	4.18	−7.21	3.36	−5.60	2.61	−4.55	2.17	0.59
−1.30	−7.71	4.28	−6.38	3.45	−4.82	2.71	−3.71	2.26	0.80
−1.20	−6.91	4.35	−5.61	3.54	−4.03	2.80	−2.86	2.35	0.96
−1.10	−6.13	4.42	−4.90	3.62	−3.30	2.87	−2.06	2.43	1.09
−1.00	−5.43	4.49	−4.17	3.69	−2.61	2.94	−1.37	2.51	1.21

Table A.1 (continued)

Stat	95% c_0	95% c_1	90% c_0	90% c_1	80% c_0	80% c_1	70% c_0	70% c_1	Median
−0.90	−4.81	4.55	−3.52	3.75	−1.93	3.00	−0.78	2.58	1.29
−0.80	−4.22	4.61	−2.90	3.81	−1.31	3.06	−0.35	2.64	1.38
−0.70	−3.66	4.68	−2.35	3.87	−0.78	3.12	−0.05	2.70	1.46
−0.60	−3.12	4.74	−1.81	3.92	−0.42	3.18	0.19	2.76	1.53
−0.50	−2.61	4.77	−1.32	3.98	−0.13	3.24	0.38	2.82	1.61
−0.40	−2.16	4.80	−0.89	4.03	0.10	3.28	0.54	2.87	1.67
−0.30	−1.59	4.83	−0.55	4.07	0.29	3.33	0.68	2.92	1.73
−0.20	−1.18	4.86	−0.28	4.12	0.44	3.38	0.80	2.98	1.79
−0.10	−0.85	4.89	−0.05	4.17	0.58	3.43	0.89	3.02	1.84
0.00	−0.54	4.92	0.15	4.21	0.70	3.47	0.98	3.07	1.89
0.10	−0.28	4.95	0.32	4.26	0.81	3.51	1.06	3.11	1.94
0.20	−0.06	4.98	0.46	4.30	0.90	3.56	1.14	3.16	1.99
0.30	0.12	5.01	0.59	4.34	1.00	3.60	1.22	3.20	2.03
0.40	0.28	5.04	0.71	4.38	1.07	3.64	1.28	3.24	2.07
0.50	0.42	5.07	0.82	4.41	1.15	3.68	1.34	3.27	2.12
0.60	0.56	5.10	0.91	4.45	1.22	3.72	1.40	3.31	2.16
0.70	0.68	5.13	1.00	4.49	1.28	3.76	1.45	3.35	2.20
0.80	0.79	5.16	1.07	4.52	1.34	3.80	1.50	3.38	2.24
0.90	0.87	5.19	1.14	4.55	1.40	3.84	1.56	3.42	2.27
1.00	0.96	5.22	1.21	4.58	1.44	3.87	1.61	3.45	2.31
1.10	1.03	5.25	1.27	4.61	1.49	3.91	1.65	3.49	2.35
1.20	1.10	5.28	1.32	4.64	1.54	3.94	1.69	3.52	2.38
1.30	1.16	5.31	1.38	4.67	1.59	3.98	1.72	3.56	2.42
1.40	1.22	5.34	1.43	4.70	1.63	4.01	1.76	3.59	2.45
1.50	1.28	5.36	1.47	4.73	1.66	4.04	1.80	3.62	2.48
1.60	1.33	5.39	1.51	4.76	1.70	4.07	1.84	3.65	2.51
1.70	1.38	5.42	1.56	4.78	1.74	4.09	1.87	3.69	2.54
1.80	1.43	5.45	1.60	4.81	1.78	4.12	1.90	3.72	2.57
1.90	1.47	5.48	1.63	4.83	1.81	4.15	1.93	3.75	2.60
2.00	1.51	5.50	1.67	4.86	1.84	4.18	1.97	3.78	2.63
Part B. Based on detrended *ADF* *t*-statistic $\hat{\tau}^{\tau}$									
−5.90	—	−37.94	—	—	—	—	—	—	—
−5.80	—	−36.20	—	—	—	—	—	—	—
−5.70	—	−34.42	—	—	—	—	—	—	—
−5.60	—	−32.70	—	−36.30	—	—	—	—	—
−5.50	—	−31.00	—	−34.56	—	—	—	—	—
−5.40	—	−29.25	—	−32.80	—	−37.01	—	—	—
−5.30	—	−27.55	—	−31.08	—	−35.25	—	−37.89	—
−5.20	—	−25.86	—	−29.36	—	−33.51	—	−36.09	—
−5.10	—	−24.24	—	−27.72	—	−31.78	—	−34.36	—
−5.00	—	−22.62	—	−26.07	—	−30.10	—	−32.64	—
−4.90	—	−21.02	—	−24.53	—	−28.45	—	−30.91	—
−4.80	—	−19.44	—	−22.95	—	−26.77	—	−29.24	—
−4.70	—	−17.91	—	−21.38	—	−25.15	—	−27.57	−37.84
−4.60	—	−16.34	—	−19.81	—	−23.59	—	−25.99	−36.07
−4.50	—	−14.87	—	−18.29	—	−22.02	—	−24.41	−34.39
−4.40	—	−13.31	—	−16.75	—	−20.49	—	−22.89	−32.70
−4.30	—	−11.73	—	−15.22	—	−18.97	—	−21.37	−31.05
−4.20	—	−10.25	—	−13.75	—	−17.51	—	−19.86	−29.41
−4.10	—	−8.81	—	−12.26	—	−16.07	−36.82	−18.41	−27.81

Table A.1 (continued)

Stat	95%		90%		80%		70%		Median
	c_0	c_1	c_0	c_1	c_0	c_1	c_0	c_1	
−4.00	—	−7.17	—	−10.81	−37.08	−14.62	−35.10	−16.97	−26.25
−3.90	—	−5.57	—	−9.29	−35.36	−13.18	−33.40	−15.56	−24.72
−3.80	—	−3.60	−36.83	−7.79	−33.67	−11.73	−31.74	−14.12	−23.23
−3.70	−37.63	1.42	−35.11	−6.31	−32.03	−10.32	−30.11	−12.72	−21.73
−3.60	−35.87	1.80	−33.39	−4.65	−30.40	−8.86	−28.49	−11.36	−20.27
−3.50	−34.13	2.07	−31.71	−2.85	−28.81	−7.39	−26.90	−9.99	−18.87
−3.40	−32.44	2.32	−30.08	1.49	−27.25	−5.89	−25.37	−8.58	−17.50
−3.30	−30.76	2.53	−28.48	1.87	−25.75	−4.34	−23.88	−7.17	−16.14
−3.20	−29.15	2.72	−26.90	2.11	−24.26	−2.37	−22.44	−5.70	−14.78
−3.10	−27.55	2.91	−25.35	2.34	−22.77	1.51	−21.00	−4.11	−13.44
−3.00	−26.05	3.11	−23.86	2.53	−21.34	1.84	−19.60	−2.28	−12.13
−2.90	−24.54	3.36	−22.35	2.69	−19.94	2.07	−18.24	1.50	−10.83
−2.80	−23.04	3.57	−20.90	2.85	−18.59	2.26	−16.90	1.83	−9.56
−2.70	−21.64	3.73	−19.53	3.01	−17.24	2.44	−15.56	2.03	−8.30
−2.60	−20.29	3.87	−18.18	3.16	−15.93	2.59	−14.30	2.20	−7.01
−2.50	−18.93	4.01	−16.87	3.31	−14.62	2.73	−13.04	2.36	−5.70
−2.40	−17.58	4.15	−15.59	3.46	−13.38	2.86	−11.78	2.50	−4.35
−2.30	−16.26	4.27	−14.36	3.61	−12.15	2.99	−10.60	2.63	−2.87
−2.20	−15.01	4.38	−13.18	3.73	−10.97	3.10	−9.43	2.74	−0.97
−2.10	−13.79	4.49	−12.03	3.85	−9.82	3.20	−8.29	2.84	1.53
−2.00	−12.61	4.57	−10.91	3.96	−8.67	3.30	−7.17	2.94	1.75
−1.90	−11.49	4.65	−9.77	4.05	−7.56	3.40	−6.07	3.03	1.89
−1.80	−10.44	4.74	−8.67	4.13	−6.51	3.47	−5.02	3.11	2.01
−1.70	−9.41	4.80	−7.61	4.20	−5.48	3.55	−3.86	3.19	2.09
−1.60	−8.38	4.85	−6.58	4.27	−4.44	3.62	−2.73	3.25	2.18
−1.50	−7.38	4.90	−5.61	4.32	−3.44	3.68	−1.47	3.30	2.24
−1.40	−6.42	4.96	−4.66	4.37	−2.39	3.74	1.27	3.36	2.30
−1.30	−5.50	5.01	−3.78	4.42	−1.14	3.80	1.47	3.41	2.36
−1.20	−4.68	5.05	−2.85	4.47	1.27	3.85	1.61	3.46	2.42
−1.10	−3.85	5.09	−2.00	4.51	1.44	3.90	1.69	3.51	2.46
−1.00	−3.10	5.14	−1.18	4.56	1.55	3.95	1.77	3.55	2.51
−0.90	−2.36	5.18	1.22	4.60	1.64	3.99	1.83	3.60	2.55
−0.80	−1.60	5.22	1.40	4.64	1.71	4.03	1.88	3.64	2.60
−0.70	−0.64	5.26	1.50	4.68	1.78	4.07	1.94	3.68	2.63
−0.60	1.26	5.30	1.60	4.72	1.83	4.11	1.99	3.72	2.67
−0.50	1.41	5.34	1.66	4.76	1.88	4.14	2.03	3.76	2.71
−0.40	1.50	5.38	1.71	4.79	1.93	4.18	2.07	3.80	2.75
−0.30	1.58	5.43	1.77	4.83	1.98	4.22	2.12	3.83	2.79
−0.20	1.64	5.47	1.82	4.86	2.02	4.25	2.16	3.87	2.82
−0.10	1.69	5.51	1.86	4.89	2.06	4.29	2.20	3.90	2.85
0.00	1.74	5.54	1.90	4.92	2.10	4.32	2.23	3.94	2.88
0.10	1.79	5.58	1.95	4.96	2.13	4.35	2.27	3.97	2.91
0.20	1.83	5.61	1.99	4.99	2.17	4.38	2.30	4.01	2.95
0.30	1.87	5.65	2.02	5.02	2.21	4.41	2.33	4.04	2.98
0.40	1.91	5.68	2.06	5.04	2.24	4.44	2.37	4.07	3.01
0.50	1.95	5.72	2.09	5.07	2.27	4.47	2.40	4.10	3.04
0.60	1.99	5.75	2.13	5.09	2.30	4.50	2.43	4.13	3.07
0.70	2.02	5.78	2.16	5.12	2.34	4.53	2.46	4.16	3.09
0.80	2.05	5.81	2.20	5.15	2.37	4.56	2.49	4.19	3.12
0.90	2.08	5.84	2.23	5.17	2.40	4.58	2.52	4.22	3.15
1.00	2.12	5.86	2.26	5.20	2.43	4.61	2.55	4.25	3.18

Table A.1 (continued)

Stat	95%		90%		80%		70%		Median
	c_0	c_1	c_0	c_1	c_0	c_1	c_0	c_1	
1.10	2.15	5.89	2.28	5.23	2.46	4.64	2.57	4.27	3.21
1.20	2.18	5.92	2.31	5.25	2.48	4.66	2.60	4.30	3.23
1.30	2.21	5.95	2.34	5.28	2.51	4.69	2.63	4.32	3.25
1.40	2.24	5.98	2.37	5.30	2.54	4.72	2.65	4.35	3.28
1.50	2.27	—	2.40	5.32	2.57	4.74	2.68	4.37	3.30
1.60	2.29	—	2.43	5.34	2.59	4.77	2.70	4.39	3.33
1.70	2.32	—	2.45	5.37	2.62	4.79	2.73	4.42	3.35
1.80	2.35	—	2.48	5.39	2.64	4.81	2.75	4.44	3.38
1.90	2.38	—	2.50	5.41	2.66	4.83	2.78	4.47	3.40
2.00	2.40	—	2.53	5.43	2.69	4.85	2.80	4.49	3.42

References

Andrews, D.W.K., 1990, Exact unbiased estimation of first order autoregressive/unit root models, Manuscript (Cowles Foundation for Economic Research, Yale University, New Haven, CT).

Ahtola, J. and G.C. Tiao, 1984, Parameter inference for a nearly nonstationary first order autoregressive model, Biometrika 71, 263–272.

Bhargava, A., 1986, On the theory of testing for unit roots in observed time series, Review of Economic Studies 53, 369–384.

Bobkoski, M.J., 1983, Hypothesis testing in nonstationary time series, Unpublished Ph.D. thesis (Department of Statistics, University of Wisconsin, Madison, WI).

Campbell, J.C. and N.G. Mankiw, 1987, Are output fluctuations transitory?, Quarterly Journal of Economics 102, 857–880.

Cavanagh, C., 1985, Roots local to unity, Manuscript (Department of Economics, Harvard University, Cambridge, MA).

Chan, N.H., 1988, The parameter inference for nearly nonstationary time series, Journal of the American Statistical Association 83, 857–862.

Chan, N.H. and C.Z. Wei, 1987, Asymptotic inference for nearly nonstationary AR(1) processes, Annals of Statistics 15, 1050–1063.

Christiano, L.C. and M. Eichenbaum, 1990, Unit roots in GNP: Do we know and do we care?, Carnegie-Rochester Conference Series on Public Policy 32, 7–62.

Cochrane, J., 1988, How big is the random walk component in GNP?, Journal of Political Economy 96, 893–920.

DeJong, D.N. and C.H. Whiteman, 1989, Trends and random walks in macroeconomic time series: A reconsideration based on the likelihood principle, Working paper no. 89-4 (Department of Economics, University of Iowa, Iowa City, IA).

Dickey, D.A. and W.A. Fuller, 1979, Distribution of the estimates for autoregressive time series with a unit root, Journal of the American Statistical Association 74, 427–431.

Kendall, M.G. and A. Stuart, 1967, The advanced theory of statistics, Vol. 2, 2nd ed. (Charles Griffin, London).

Nabeya, S. and K. Tanaka, 1990, A general approach to the limiting distribution for estimators in time series regression with nonstable autoregressive errors, Econometrica 58, 145–163.

Nelson, C.R. and C.I. Plosser, 1982, Trends and random walks in macroeconomic time series: Some evidence and implications, Journal of Monetary Economics 10, 129–162.

Perron, P., 1989, The calculation of the limiting distribution of the least-squares estimator in a near-integrated model, Econometric Theory 5, 241–255.

Phillips, P.C.B., 1987, Toward a unified asymptotic theory for autoregression, Biometrika 74, 535–547.

Phillips, P.C.B., 1990, To criticize the critics: An objective Bayesian analysis of stochastic trends, Cowles Foundation discussion paper no. 950 (Yale University, New Haven, CT).

Rudebusch, G., 1990, Trends and random walks in macroeconomic time series: A reexamination, FEDS discussion paper no. 139 (Federal Reserve Board, Washington, DC).

Sargan, D. and A. Bhargava, 1983, Testing for residuals from least squares regression for being generated by the Gaussian random walk, Econometrica 51, 153–174.

Schotman, P. and H.K. van Dijk, 1989, A Bayesian analysis of the unit root hypothesis, Manuscript (Erasmus University, Rotterdam).

Schwert, G.W., 1989, Tests for unit roots: A Monte Carlo investigation, Journal of Business and Economic Statistics 7, 147–160.

Sims, C.A., 1988, Bayesian skepticism on unit root econometrics, Journal of Economic Dynamics and Control 12, 463–474.

Sims, C.A., J.H. Stock, and M.W. Watson, 1990, Inference in linear time series models with some unit roots, Econometrica 58, 113–144.

Sims, C.A. and H. Uhlig, 1988, Understanding unit rooters: A helicopter tour, Discussion paper no. 4 (Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, MN), forthcoming in: Econometrica.

Stock, J.H., 1988, A class of tests for integration and cointegration, Manuscript (Kennedy School of Government, Harvard University, Cambridge, MA).