Econometric Theory, **25**, 2009, 857–872. Printed in the United States of America.
doi:10.1017/S0266466608090690

NOTES AND PROBLEMS

BARTLETT CORRECTION IN THE STABLE AR(1) MODEL WITH INTERCEPT AND TREND

NOUD P.A. VAN GIERSBERGEN
Universiteit van Amsterdam

Bartlett corrections are derived for testing hypotheses about the autoregressive parameter ρ in the stable (a) AR(1) model, (b) AR(1) model with intercept, (c) AR(1) model with intercept and linear trend. The correction is found explicitly as a function of ρ. In the models with deterministic terms, the correction factor is asymmetric in ρ. Furthermore, the Bartlett correction is monotonically increasing in ρ and tends to infinity when ρ approaches the stability boundary of $+1$. Simulation results indicate that the Bartlett corrections are useful in controlling the size of the likelihood ratio statistic in small samples, although these corrections are not the ultimate panacea.

1. INTRODUCTION

Several decades ago, Lawley (1956) showed in the independent and identically distributed (i.i.d.) setup that the finite-sample distribution of a Barlett-corrected (Bartlett, 1937) likelihood ratio (LR) test statistic is closer to the χ^2-distribution than the original LR statistic; see Cribari-Neto and Cordeiro (1996) for an econometric-oriented review. Recently, Bartlett-*type* corrections in *un*stable autoregressive models have attracted much attention (see, among others, Bravo, 1999; Nielsen, 1997; Larsson, 1998; Johansen, 2004), although Jensen and Wood (1997) have shown that the usual conditions for a Bartlett correction are not fulfilled in the unstable first-order autoregressive—AR(1)—model.

In this paper, we analyze the stable AR(1) model, which without deterministic terms was also considered by Taniguchi (1988, 1991), Omtzigt (2003), and Lagos and Morettin (2004). However, in contrast to earlier work, we also consider the AR(1) model with intercept (and linear trend). Anticipating the results, we can say

The author thanks a co-editor for encouraging remarks and two referees for critical comments that have led to a complete revision of this paper. A third referee is appreciated for drawing attention to the impact of the starting value. Furthermore, helpful comments from Peter Boswijk and participants of the ESEM 2004 meeting (Madrid, Spain) and the UvA-Econometrics seminar (Amsterdam, The Netherlands) are gratefully acknowledged. All errors remain my responsibility. Address correspondence to Noud P.A. van Giersbergen, Department of Quantitative Economics, Amsterdam School of Economics, Universiteit van Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, e-mail: N.P.A.vanGiersbergen@uva.nl.

© 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

that the correction factor in models with deterministic components is considerably different from the factor obtained in the pure AR(1) model.

Although a number of exact inference techniques are available for the AR(1) model (see, among others, Andrews, 1993; Kiviet and Dufour, 1997), these methods are based on numerical (Monte Carlo) procedures that only give limited analytical insight into the structure of the finite-sample problems. Furthermore, analytical techniques such as Edgeworth expansions (see, e.g., Phillips, 1977) and saddlepoint approximations often lead to very complicated formulas, which are difficult to interpret.

Suppose that the likelihood function, whose logarithm will be denoted by L, depends upon $p+1$ population parameters $\theta_1,\ldots,\theta_p,\theta_{p+1}$. In our case, $\theta_1,\ldots,\theta_p$ will denote the nuisance parameters such as the error variance, intercept, and trend coefficient, whereas θ_{p+1} is the parameter of interest, namely, the first-order autocorrelation coefficient ρ. For testing the null hypothesis $H_0:\rho=\rho_0$, the LR statistic can be written as $2(L^{(p+1)}-L^{(p)})$, where $L^{(k)}$ denotes the result of maximizing L with respect to $\theta_1,\ldots,\theta_k$ and substituting true values under the null for the remaining parameters. Under some continuity assumptions on the likelihood and its derivatives, together with the assumption that the second-order derivatives of L with respect to the parameters are of order T (the sample size), Lawley (1956) has shown that under H_0 the expectation of the LR statistic may be written as

$$\mathrm{E}[LR]=1+\xi_{p+1}-\xi_p+O(T^{-2}),$$

where

$$\begin{aligned}\xi_k=&\sum_{r,s,t,u=1}^{k}\lambda^{rs}\lambda^{tu}\left\{\tfrac{1}{4}\lambda_{rstu}-(\lambda_{rst})_u+(\lambda_{rt})_{su}\right\}-\sum_{r,s,t,u,v,w=1}^{k}\lambda^{rs}\lambda^{tu}\lambda^{vw}\\&\times\left\{\tfrac{1}{6}\lambda_{rtv}\lambda_{suw}+\tfrac{1}{4}\lambda_{rtu}\lambda_{svw}-\lambda_{rtv}(\lambda_{sw})_u-\lambda_{rtu}(\lambda_{sw})_v+(\lambda_{rt})_v(\lambda_{sw})_u\right.\\&\left.+(\lambda_{rt})_u(\lambda_{sw})_v\right\},\end{aligned}\tag{1}$$

with

$$\begin{aligned}L_{rs}&=\partial^2L/\partial\theta_r\partial\theta_s, & L_{rst}&=\partial^3L/\partial\theta_r\partial\theta_s\partial\theta_t, \quad \text{etc.},\\ \lambda_{rs}&=\mathrm{E}[L_{rs}], & \lambda_{rst}&=\mathrm{E}[L_{rst}], \quad \text{etc.},\\ (\lambda_{rs})_t&=\partial\lambda_{rs}/\partial\theta_t, & (\lambda_{rst})_u&=\partial\lambda_{rst}/\partial\theta_u, \quad \text{etc.},\\ (\lambda_{rs})_{tu}&=\partial^2\lambda_{rs}/\partial\theta_t\partial\theta_u,\end{aligned}$$

and $[\lambda^{rs}]$ is the inverse matrix of $[\lambda_{rs}]$. Because the calculations required for a Bartlett correction are similar to the calculations required for an Edgeworth expansion (see, e.g., Hayakawa, 1977, Rmk. 1, p. 367), related work for the AR(1) model with intercept can be found in Sul, Phillips, and Choi (2005).

In the class of stationary Gaussian autoregressive moving average models, Taniguchi (1988) has shown that the LR test is the only Bartlett-correctable test among the LR, Wald, modified Wald, and Lagrange multiplier tests, i.e.,

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

$$\mathbb{P}[LR^* \le x] = P\left[\chi_1^2 \le x\right] + o(T^{-1}), \tag{2}$$

where $LR^* = LR/(1+\xi_{p+1}-\xi_p)$ denotes the Bartlett-corrected LR statistic. Because all Bartlett corrections are calculated according to the general formula shown in (1) and the dynamics in our models are of the autoregressive type, the higher order result stated in (2) should also apply to our Bartlett-corrected LR statistics.

The paper is organized as follows. In Sections 2–5, the Bartlett corrections are derived in the AR(1) models with and without intercept and linear trend. Section 6 presents some simulation results to shed some light on the small-sample properties of the Bartlett-corrected tests, and Section 7 concludes.

A word on notation. Throughout this paper, the symbol $\stackrel{1}{=}$ will indicate that we have kept terms of order T^{-1}; i.e., if the stochastic expansion is of the form $V = V_0 + T^{-1}V_1 + T^{-2}V_2 + \ldots$, then $V \stackrel{1}{=} V_0 + T^{-1}V_1$, where $V_i \in O_p(1)$ are random variables. Furthermore, we use $\sum$ to indicate summation over $t = 1, \ldots, T$.

2. THE AR(1) MODEL

First, consider the asymptotic stationary, mean zero, Gaussian AR(1) model

$$y_t = \rho y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, \sigma^2), \qquad |\rho| < 1, \tag{3}$$

for $t = 1, \ldots, T$. Inference is made conditional on the starting value y_0 so that ρ can be efficiently estimated by ordinary least squares (OLS). The LR statistic for testing $H_0 : \rho = \rho_0$ against $H_1 : \rho \neq \rho_0$ in this model is equal to

$$LR = -T \log\left(1 - r_{y\varepsilon}^2\right), \tag{4}$$

where

$$r_{y\varepsilon} \equiv \frac{\sum y_{t-1}\varepsilon_t}{\sqrt{\sum y_{t-1}^2 \sum \varepsilon_t^2}}$$

denotes the sample correlation coefficient between y_{t-1} and ε_t assuming zero means; see, e.g., formula (4) of Nielsen (1997). Note that (y_t/σ) is an AR(1) process with (ε_t/σ) as innovations and starting value (y_0/σ). Therefore, $r_{y\varepsilon}$ (and hence the LR statistic and its Bartlett correction) depends on σ only through the scaled starting value (y_0/σ); see also Evans and Savin (1981). The Bartlett correction in the pure AR(1) model is stated in Theorem 1, which follows.

THEOREM 1. *In the stable Gaussian AR(1) model, given in (3), the expectation of the LR test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\text{E}[LR] \stackrel{1}{=} 1 - \frac{1}{6T} - \frac{1}{3T} = 1 - \frac{1}{2T}. \tag{5}$$

Proof of Theorem 1. We have to determine ξ_1 and ξ_2 in formula (1) for $(\theta_1, \theta_2) = (\sigma^2, \rho)$. Because all the λ's in the pure AR(1) model are a subset of

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

the λ's in the AR(1) model with intercept and trend, the λ's are only shown in the proof of Theorem 3. Because the index of ρ changes in the various models (for this particular model it is 2), the subindex of this parameter is replaced by ρ. The following λ's are nonzero: $\lambda_{11}, \lambda_{\rho\rho}, \lambda_{111}, \lambda_{(1\rho\rho)}, \lambda_{1111}, \lambda_{(11\rho\rho)}$. Here, the subindex $(1\rho\rho)$ means all permutations of $(1\rho\rho)$, so that $\lambda_{(1\rho\rho)} \in \{\lambda_{1\rho\rho}, \lambda_{\rho 1\rho}, \lambda_{\rho\rho 1}\}$. Carrying out the summation shown in (1) for $k = 2$ and $k = 1$ gives $\xi_2 = -1/(6T)$ and $\xi_1 = 1/(3T)$, which completes the proof. ■

Somewhat surprisingly, the Bartlett correction turns out to be independent of the AR(1) parameter ρ. Furthermore, the factor is always smaller than 1. Hence, the Bartlett correction predicts that the uncorrected LR statistic tends to underreject a correct null hypothesis, i.e., its rejection probability tends to be lower than the nominal level.

Note that the result in (5) is different from Taniguchi (1988), which does not treat σ^2 as a nuisance parameter. In fact, if σ^2 is assumed to be known, we have $\lambda_{\rho\rho} = -(1-\rho^2)^{-1}T$ and $\lambda_{\rho\rho\rho} = \lambda_{\rho\rho\rho\rho} = 0$, so that formula (1) for $k = 1$ reduces to

$$\xi_1 = \frac{(\lambda_{\rho\rho})_{\rho\rho}}{\{\lambda_{\rho\rho}\}^2} - 2\frac{\{(\lambda_{\rho\rho})_{\rho}\}^2}{\{\lambda_{\rho\rho}\}^3} = -\frac{2}{T}.$$

This leads to a Bartlett factor of $1 - 2/T\,(= 1 + \xi_1 - \xi_0)$, which is the correction term reported in Taniguchi (1988, p. 504).

3. AR(1) MODEL WITH INTERCEPT

In this section, the AR(1) process is (asymptotically) stationary around a nonzero mean. Consider the AR(1) model with an intercept

$$(y_t - \mu) = \rho(y_{t-1} - \mu) + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, \sigma^2), \qquad |\rho| < 1, \tag{6}$$

conditional on the starting value y_0. To examine the invariance properties, define $z_t = (y_t - \mu)/\sigma$, so that

$$z_t = \rho z_{t-1} + \eta_t, \qquad \eta_t = \varepsilon_t/\sigma, \qquad |\rho| < 1, \tag{7}$$

and $z_0 = (y_0 - \mu)/\sigma$ conditional on y_0. Let variables with a dot denote the residuals from regression on an intercept, so that $\dot{y}_t(\dot{z}_t)$ denotes $y_t(z_t)$ in deviation from its mean. Because of the relationship between y_t and z_t, we have $\dot{y}_t = \sigma \dot{z}_t$ and $\dot{\varepsilon}_t = \sigma \dot{\eta}_t$. Analogously to (4), the LR statistic is given by

$$LR^c = -T \log\left(1 - \left(r^c_{y\varepsilon}\right)^2\right),$$

where

$$r^c_{y\varepsilon} \equiv \frac{\sum \dot{y}_{t-1}\dot{\varepsilon}_t}{\sqrt{\sum \dot{y}^2_{t-1} \sum \dot{\varepsilon}^2_t}} = \frac{\sum \dot{z}_{t-1}\dot{\eta}_t}{\sqrt{\sum \dot{z}^2_{t-1} \sum \dot{\eta}^2_t}} \tag{8}$$

denotes the sample correlation coefficient between $\dot{y}_{t-1}$ and $\dot{\varepsilon}_t$. From the second equality in (8), we observe that the test statistic LR^c computed from the demeaned

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

y_t process is the same as when computed from the demeaned z_t process. Thus, for the purpose of finding the conditional distribution of LR^c in model (6) it suffices to consider a process z_t satisfying (7) using z_0 as initial value. Therefore, the distribution of LR^c only depends on (μ, σ^2) through the starting value z_0. However, the Bartlett correction as reported in Theorem 2, which follows, shows that the effect of z_0 on the first-order Bartlett correction is negligible.

THEOREM 2. *In the stable Gaussian AR(1) model with intercept, given in (6), the expectation of the LR test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\mathrm{E}[LR^c] \stackrel{1}{=} 1 + \frac{7+5\rho}{3(1-\rho)T} - \frac{11}{6T} = 1 + \frac{1+7\rho}{2(1-\rho)T}. \tag{9}$$

Proof of Theorem 2. In this setup, we have $(\theta_1, \theta_2, \theta_3) = (\sigma^2, \mu, \rho)$. The following λ's are nonzero: λ_{11}, λ_{22}, $\lambda_{\rho\rho}$, λ_{111}, $\lambda_{(122)}$, $\lambda_{(1\rho\rho)}$, $\lambda_{(22\rho)}$, λ_{1111}, $\lambda_{(1122)}$, $\lambda_{(11\rho\rho)}$, $\lambda_{(122\rho)}$, $\lambda_{(22\rho\rho)}$; see Theorem 3 in Section 4 for their values. As before, a subindex between parentheses means all permutations of that subindex. Carrying out the summation shown in (1) for $k = 3$ and $k = 2$ gives $\xi_3 = (7+5\rho)/(3(1-\rho)T)$ and $\xi_2 = 11/(6T)$. Calculating $1 + \xi_3 - \xi_2$ completes the proof. ■

This Bartlett correction agrees with the factor calculated by van Giersbergen (2004, Thm. 2), where the expectation is determined in a more laborious but direct way. Contrary to the case when there is no intercept, i.e., formula (5), we now see that the factor depends on the AR(1) parameter ρ. For $\rho > -\frac{1}{3}$, the Bartlett correction in the model with intercept is larger than the Bartlett correction in the model without intercept. Furthermore, the factor is increasing in ρ, and hence it is asymmetric with respect to the origin; it even has an asymptote for $\rho \uparrow 1$. From the proof of Theorem 2, we see that the explosive nature of the factor is due to ξ_3 that refers to the unrestricted model in which both μ and ρ are estimated. In contrast to the pure AR(1) model, where we only had the nonexplosive terms $\lambda^{\sigma^2\sigma^2} = -2\sigma^4 T^{-1}$ and $\lambda^{\rho\rho} = -(1-\rho^2)T^{-1}$, the intercept introduces an explosive term $\lambda^{\mu\mu} = -\sigma^2(1-\rho)^{-2}T^{-1}$ in Lawley's formula shown in (1). In the unrestricted AR(1) model with intercept, the interaction between terms such as $(1-\rho)^{-2}$ and $(1-\rho^2)$ leads to the explosive factor $(1-\rho)^{-1}$. Note that $\lambda^{\mu\mu}$, which equals $1/\lambda_{\mu\mu}$ in model (6), is related to the reciprocal of the information with respect to μ, i.e., $-\lambda_{\mu\mu} = \sigma^{-2}(1-\rho)^2 T$, which tends to zero as ρ tends to 1. Overall, the explosiveness of the Bartlett correction seems to be caused by the fact that μ is poorly identified as ρ approaches 1.

4. AR(1) MODEL WITH INTERCEPT AND TREND

Consider the AR(1) model with intercept and linear trend

$$(y_t - \mu - \beta t) = \rho(y_{t-1} - \mu - \beta(t-1)) + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{NID}(0, \sigma^2), \qquad |\rho| < 1. \tag{10}$$

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

Inference is again conditional on the starting value y_0. For the invariance properties, define $z_t = (y_t - \mu - \beta t)/\sigma$. As in the previous section, we have that $z_t = \rho z_{t-1} + \eta_t$ with $\eta_t = \varepsilon_t/\sigma$ and $z_0 = (y_0 - \mu)/\sigma$. Let variables with a double dot denote the residuals from regression on an intercept and trend, so that $\ddot{y}_t(\ddot{z}_t)$ denotes y_t (z_t) in deviation from a constant and trend. Because of the relationship between y_t and z_t, we have $\ddot{y}_t = \sigma \ddot{z}_t$ and $\ddot{\varepsilon}_t = \sigma \ddot{\eta}_t$. The LR statistic is given by

$$\begin{aligned} LR^\tau &= -T \log\left(1 - \left(r^\tau_{y\varepsilon}\right)^2\right) \\ &= -T \log\left(1 - \left(r^\tau_{z\varepsilon}\right)^2\right), \end{aligned}$$

where $r^\tau_{y\varepsilon}$ $\left(r^\tau_{z\varepsilon}\right)$ denotes the sample correlation coefficient between $\ddot{y}_{t-1}$ $(\ddot{z}_{t-1})$ and $\ddot{\varepsilon}_t$. Analogously to the AR(1) model with intercept, the test statistic LR^τ computed from the detrended y_t process is the same as when computed from the detrended z_t process. Therefore, the conditional distribution of LR^τ only depends on (μ, β, σ^2) through the starting value z_0. Theorem 3, which follows, gives an expression for the Bartlett factor in the AR(1) model with intercept and trend.

THEOREM 3. *In the stable Gaussian AR(1) model with intercept and linear trend, given in (10), the expectation of the LR test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\mathrm{E}[LR^\tau] \stackrel{1}{=} 1 + \frac{47 + 25\rho}{6(1-\rho)T} - \frac{13}{3T} = 1 + \frac{7 + 17\rho}{2(1-\rho)T}.$$

Proof of Theorem 3. First, we divide the linear trend by T, so that the second-order derivatives are of order T as required by Lawley (1956). This can be done without loss of generality because the LR^τ statistic is invariant to this transformation. So, we have $z_t = (y_t - \mu - \beta t/T)/\sigma$ and $z_t = \sum_{i=0}^{t-1} \rho^i \eta_{t-i} + \rho^t z_0$ with $\eta_t \sim \mathrm{NID}(0, 1)$. The log-likelihood (conditional on z_0) based on z_t is given by

$$L(\sigma^2, \mu, \beta, \rho) \propto -\frac{T}{2} \log(\sigma^2) - \frac{1}{2} \sum (z_t - \rho z_{t-1})^2, \tag{11}$$

where $z_t = z_t(\mu, \beta, \sigma^2)$ is seen as a function of μ, β, and σ^2. Note that the first- and higher order derivatives of $\sum (z_t - \rho z_{t-1})^2$ with respect to μ, β, σ^2, and ρ can be written in terms of z_t, z_{t-1}, σ^2, and ρ only; e.g.,

$$\begin{aligned} \frac{\partial}{\partial \beta} \sum (z_t - \rho z_{t-1})^2 &= 2 \sum (z_t - \rho z_{t-1}) \left(\frac{\partial z_t}{\partial \beta} - \rho \frac{\partial z_{t-1}}{\partial \beta} \right) \\ &= 2 \sum (z_t - \rho z_{t-1}) \frac{-(t + \rho(t-1))}{\sigma T}. \end{aligned}$$

Hence, it is sufficient to consider the pure AR(1) z_t process for evaluating the expectations. The log-likelihood in (11) is characterized by $(\theta_1, \theta_2, \theta_3, \theta_4)$

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

$= (\sigma^2, \mu, \beta, \rho)$. All derivatives up to the fourth order were determined using Mathematica 6.0 (see Wolfram, 1991). Conditional on z_0, we have

$$\mathrm{E}\left[\sum z_t\right] = \sum \rho^t z_0 = \rho \frac{1-\rho^T}{1-\rho} z_0 = O(1).$$

Analogously, we find

$$\mathrm{E}\left[\sum z_{t-1}\right] = O(1), \qquad \mathrm{E}\left[\sum (t/T) z_t\right] = O(1), \qquad \text{and} \quad \mathrm{E}\left[\sum (t/T) z_{t-1}\right] = O(1).$$

Furthermore,

$$\begin{aligned}\mathrm{E}\left[\sum z_t^2\right] &= \mathrm{E}\left[\sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\rho^i \eta_{t-i} + \rho^t z_0\right)^2\right] = \sum_{t=1}^{T}\left(\sum_{i=0}^{t-1}\rho^{2i} + \rho^{2t} z_0^2\right) \\ &= \sum_{t=1}^{T}\left(\frac{1-\rho^{2t}}{1-\rho^2} + \rho^{2t} z_0^2\right) \\ &= \frac{(1-\rho^2)T - \rho^2(1-\rho^{2T})}{(1-\rho^2)^2} + \rho^2 \frac{1-\rho^{2T}}{1-\rho^2} z_0^2 = \frac{1}{(1-\rho^2)} T + O(1),\end{aligned} \tag{12}$$

so that

$$\mathrm{E}\left[\sum z_{t-1}^2\right] = \frac{1}{(1-\rho^2)} T + O(1) \quad \text{and} \quad \mathrm{E}\left[\sum z_t z_{t-1}\right] = \frac{\rho}{(1-\rho^2)} T + O(1).$$

Because only $O(T)$ terms contribute to the Bartlett correction, the contribution of z_0 is negligible for our purpose. Last, we can use the following two approximations:

$$\sum (t/T) = \frac{1}{2} T + O(1) \quad \text{and} \quad \sum (t/T)^2 = \frac{1}{3} T + O(1).$$

The expectations of the second-order derivatives are given by (after replacing all subindexes 4 by ρ and omitting $O(1)$ terms)

$$\begin{aligned} \lambda_{11} &= -\tfrac{1}{2\sigma^4} T, & \lambda_{33} &= -\tfrac{(1-\rho)^2}{3\sigma^2} T, \\ \lambda_{22} &= -\tfrac{(1-\rho)^2}{\sigma^2} T, & \lambda_{\rho\rho} &= -\tfrac{1}{(1-\rho^2)} T, \\ \lambda_{(23)} &= -\tfrac{(1-\rho)^2}{2\sigma^2} T. \end{aligned}$$

The nonzero expectations of the third-order derivatives are given by (omitting $O(1)$ terms)

$$\begin{aligned} \lambda_{111} &= \tfrac{2}{\sigma^6} T, & \lambda_{(1\rho\rho)} &= \tfrac{1}{(1-\rho^2)\sigma^2} T, \\ \lambda_{(122)} &= \tfrac{(1-\rho)^2}{\sigma^4} T, & \lambda_{(22\rho)} &= \tfrac{2(1-\rho)}{\sigma^2} T, \\ \lambda_{(123)} &= \tfrac{(1-\rho)^2}{2\sigma^4} T, & \lambda_{(23\rho)} &= \tfrac{(1-\rho)}{\sigma^2} T, \\ \lambda_{(133)} &= \tfrac{(1-\rho)^2}{3\sigma^4} T, & \lambda_{(33\rho)} &= \tfrac{2(1-\rho)}{3\sigma^2} T. \end{aligned}$$

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

The nonzero expectations of the fourth-order derivatives are given by (omitting $O(1)$ terms)

$$\lambda_{1111} = -\frac{9}{\sigma^8}T, \qquad \lambda_{(122\rho)} = -\frac{2(1-\rho)}{\sigma^4}T,$$
$$\lambda_{(1122)} = -\frac{2(1-\rho)^2}{\sigma^6}T, \qquad \lambda_{(123\rho)} = -\frac{(1-\rho)}{\sigma^4}T,$$
$$\lambda_{(1123)} = -\frac{(1-\rho)^2}{\sigma^6}T, \qquad \lambda_{(133\rho)} = -\frac{2(1-\rho)}{3\sigma^4}T,$$
$$\lambda_{(1133)} = -\frac{2(1-\rho)^2}{3\sigma^6}T, \qquad \lambda_{(22\rho\rho)} = -\frac{2}{\sigma^2}T = 2\lambda_{(23\rho\rho)},$$
$$\lambda_{(11\rho\rho)} = -\frac{2}{(1-\rho^2)\sigma^4}T, \qquad \lambda_{(33\rho\rho)} = -\frac{2}{3\sigma^2}T.$$

Using Mathematica to carry out the summation shown in (1) for $k = 4$ and $k = 3$ gives $\xi_4 = (47 + 25\rho)/(6(1-\rho)T)$ and $\xi_3 = 13/(3T)$. Calculating $1 + \xi_4 - \xi_3$ completes the proof. ■

The Bartlett correction in the AR(1) model with intercept and trend has the same functional form as in the model with only an intercept. As before, the factor is increasing in ρ and goes to infinity for $\rho \uparrow 1$. The latter property is related to the fact that both μ and β are poorly identified in the unrestricted model as ρ approaches 1. For $\rho > -3/5$, the correction factor in the model with trend is larger than the factor in the model without a trend. Hence, for this range, the finite-sample problems in the model with trend are expected to be more severe than in the model without trend.

5. SOME FURTHER RESULTS

In this section, we shall consider three issues that were brought up by a referee. First, we investigate the effect of using the unconditional likelihood. Second, we indicate which asymptotically negligible terms might be important in finite samples. Third, we investigate the effect of a large starting value.

Consider the stationary AR(1) model

$$y_t^* = \rho y_{t-1}^* + \varepsilon_t, \qquad |\rho| < 1, \qquad y_0^* \sim N(0, \sigma^2/(1-\rho^2)).$$

The (unconditional) log-likelihood is given by

$$L^* \propto -\frac{1}{2}\log\left(\frac{\sigma^2}{(1-\rho^2)}\right) - \frac{(1-\rho^2)}{2\sigma^2}(y_0^*)^2 - \frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum\left(y_t^* - \rho y_{t-1}^*\right)^2,$$

where the first two terms are due to y_0^*. The second-order derivatives L_{11}^*, $L_{(1\rho)}^*$, and $L_{\rho\rho}^*$ are easy to derive. In this model specification, we have the following exact results:

$$\mathrm{E}\left[\sum y_t^*\right] = \mathrm{E}\left[\sum y_{t-1}^*\right] = 0 \quad \text{and} \quad \mathrm{E}[y_0^*] = 0,$$
$$\mathrm{E}\left[\sum (y_t^*)^2\right] = \mathrm{E}\left[\sum (y_{t-1}^*)^2\right] = \frac{\sigma^2}{1-\rho^2}T \quad \text{and} \quad \mathrm{E}\left[\sum y_t^* y_{t-1}^*\right] = \rho\frac{\sigma^2}{1-\rho^2}T,$$

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

and the expectations of the derivatives under the stationary distribution are given by

$$\lambda^*_{11} = -\frac{1}{2\sigma^4}T + \frac{1}{2\sigma^4} = -\frac{1}{2\sigma^4}T + O(1), \qquad \lambda^*_{(1\rho)} = 0,$$
$$\lambda^*_{(\rho\rho)} = -\frac{1}{(1-\rho^2)}T - \frac{1+\rho^2}{(1-\rho^2)^2} = -\frac{1}{(1-\rho^2)}T + O(1).$$

Because only $O(T)$ terms contribute to the Bartlett correction, the expectations of the derivatives are the same as in the previous section. This also holds for the higher order derivatives and in the models with deterministic components. Therefore, the Bartlett correction based on the unconditional likelihood is the same as the one based on the conditional likelihood.

Next, we take a closer look at the various terms that are neglected in the first-order Bartlett expansions. For simplicity, consider the pure AR(1) model shown in (3) with $y_0 = 0$. Substitution of $z_t = y_t/\sigma$ in formula (12) leads to

$$\mathrm{E}\left[\sum y_t^2\right] = \frac{\sigma^2}{(1-\rho^2)}T - \rho^2\frac{(1-\rho^{2T})}{(1-\rho^2)^2}\sigma^2.$$

When ρ is close to 1, say, $1-\rho^2 = c/T$ for $c > 0$, we have that the first term, which is $O(T^2)$, is just as large as the second term, which is also $O(T^2)$. Note that this second term was neglected in the analysis so far but is expected to show up in higher order expansions. This issue is explored by Nielsen (1999) in the bivariate setup, although this route will not be pursued in this paper. When $\rho = 1$, Bartlett corrections removing T-order terms do not exist in general; see also Section 1.

Last, we analyze the effect of a large starting value on the Bartlett corrections. Because the effect of y_0 is only $O(1)$ on $\mathrm{E}[\sum y_{t-i}y_{t-j}]$ for $i, j \in \{0, 1\}$ (cf. (12) for $z_t = y_t/\sigma$), we shall assume

$$y_0^2 = \gamma\sigma^2 T. \tag{13}$$

Under assumption (13), the Bartlett corrections for the three AR(1) models ((14a), without deterministic terms; (14b), with intercept; and (14c), with intercept and trend) are given by

$$\mathrm{E}[LR] \stackrel{1}{=} 1 + \frac{3\gamma - 1}{2(1+\gamma)T}, \tag{14a}$$

$$\mathrm{E}[LR^c] \stackrel{1}{=} 1 + \frac{5\gamma}{2(1+\gamma)T} + \frac{1+7\rho}{2(1+\gamma)(1-\rho)T}, \tag{14b}$$

$$\mathrm{E}[LR^\tau] \stackrel{1}{=} 1 + \frac{7\gamma}{2(1+\gamma)T} + \frac{7+17\rho}{2(1+\gamma)(1-\rho)T}. \tag{14c}$$

When $\gamma = 0$, these formulas reduce to the Bartlett factors as shown in Theorems 1–3. Although unrealistic, if γ becomes very large with respect to T, we have that the Bartlett factors converge to $1+3/(2T)$, $1+5/(2T)$, and $1+7/(2T)$. In the

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

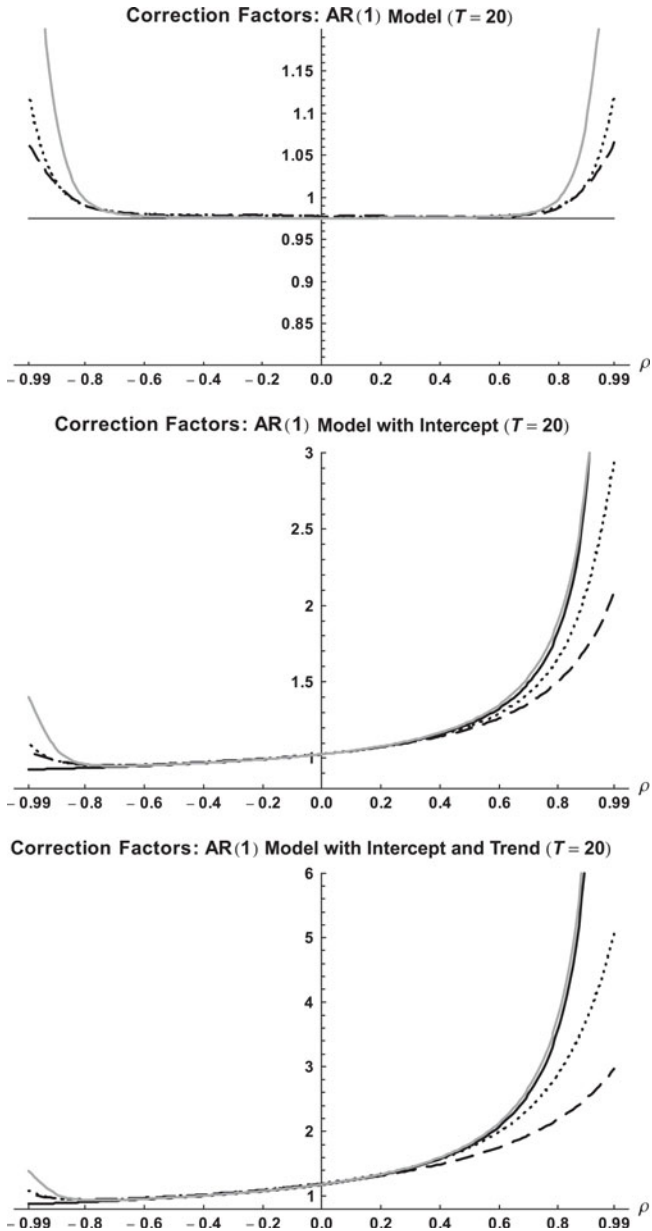


FIGURE 1. Various correction factors in the AR(1) model with possible intercept and trend. The solid line is the Bartlett factor as shown in Theorems 1–3. The dotted line is based on simulation of E[LR] (top), E[LR^c] (middle), or E[LR^τ] (bottom), and the dashed line is based on a simulated correction factor that leads to a LR test with 5% size.

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

next section, the usefulness of the corrections shown in (14a)–(14c) is investigated by simulation.

Figure 1 shows the various correction factors. Beside the first-order Bartlett factors, it also shows the expectation of the LR statistics, i.e., E[LR], E[LR^c], and E[LR^τ], approximated by simulation (using $y_0 = 0$) and simulation-based correction factors that lead to LR statistics having a 5% size. Note that the latter factors are not exact for other significance levels (or $y_0 \neq 0$). In the pure AR(1) model, it seems that the simulated values of E[LR] are too large for $|\rho|$ near to 1. In the AR(1) model with intercept (and also trend), the simulated values of E[LR^c] (and E[LR^τ]) substantially deviate from the adjustment that leads to a correct size. This shows that the Bartlett approach based on the expectation of the LR statistic is not very adequate in the model containing deterministic terms when ρ is close to 1. In these models, inference based on the restricted likelihood seems promising; see Chen and Deo (2006).

6. MONTE CARLO RESULTS

To assess the quality of the asymptotic expansions, a small simulation study has been carried out. All the simulations were done on a PC using Gauss 8. Observations were generated according to an AR(1) process. All reported results are based on 10^6 replications, and two sample sizes were considered: $T \in \{20, 80\}$. The AR(1) parameter was taken as $\rho \in \{-0.99, -0.98, \ldots, 0.99\}$. Because the distribution of the LR statistic only depends on σ (and μ in the model with intercept) through the starting value, we set $\sigma = 1$ (and $\mu = 0$). In addition, $\beta = 0$ in the model with a trend. The results shown in Figure 2 are for $y_0 = 0$. The simulations were also carried out for $y_0 \sim N(0, 1/(1-\rho^2))$, which did not change the conclusions significantly. The nominal significance level was taken to be 5%. Results remain qualitatively the same for the 1% and 10% levels.

Figure 2 shows the rejection frequencies for the three AR(1) models considered in this paper. In the pure AR(1) model, inference based on the uncorrected LR test and the χ^2-distribution performs remarkably well. Chen and Deo (2006) argue that this is due to the facts that (a) there is little estimation bias in the absence of deterministic terms and (b) the LR test statistic is invariant to monotonic transformations, so that the LR approach automatically makes a normalizing transformation (like Fisher's z-transformation in the bivariate setup). Only when $|\rho|$ is very close to 1 do the rejection frequencies rise slightly above 6%. Note that the rejection error rate seems to be an even function of ρ. Because the Bartlett correction does not depend on ρ, the curve of rejection frequencies is shifted toward the nominal significance level for $|\rho| < 0.8$. In line with Figure 1, inference based on simulated values of E[LR] as a correction factor becomes conservative for $|\rho| > 0.9$.

The middle two graphs of Figure 2 show the rejection frequencies when an intercept is added to the AR(1) model. As expected, the discrepancy between the

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

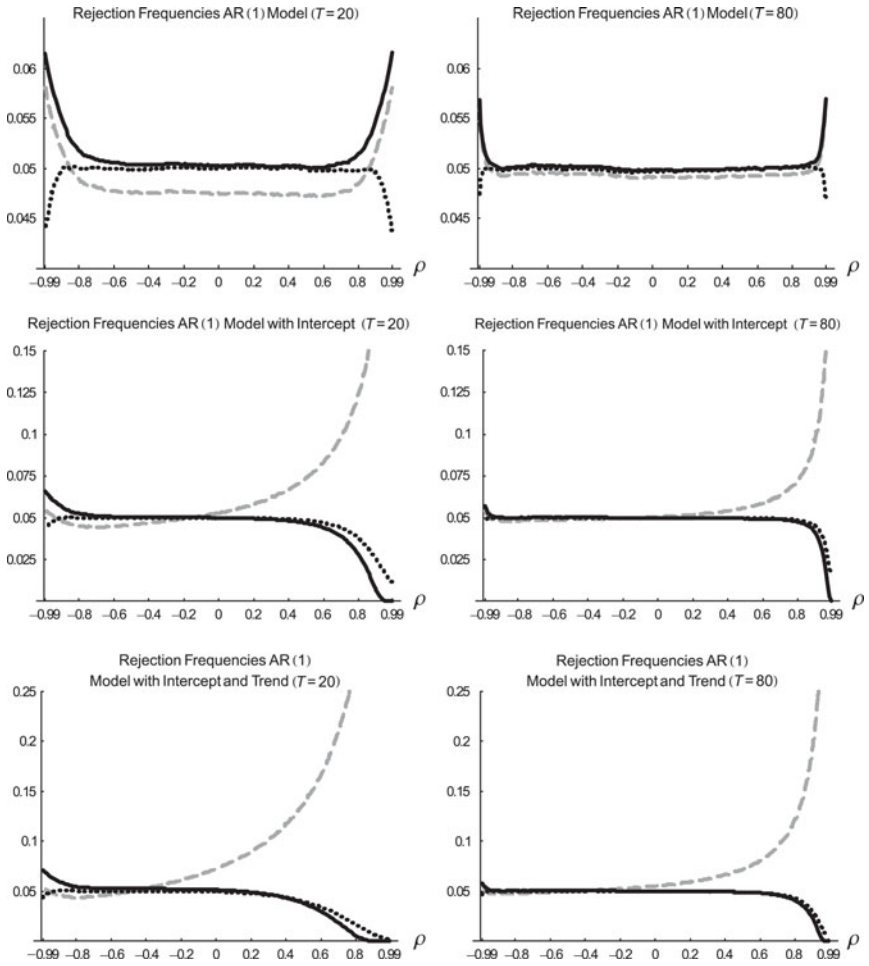


FIGURE 2. Rejection frequencies in the AR(1) model with possible intercept and trend. The dashed gray line is based on the LR statistic and the asymptotic χ^2-distribution, and the black line is based on the Bartlett-corrected LR statistic using the adjustments shown in Theorems 1–3. The dotted line uses the simulated value of E[LR] (top), E[LR^c] (middle), or E[LR^τ] (bottom) as correction factor.

rejection frequencies and the nominal level is larger than in the model without intercept. Furthermore, the rejection error rate is more pronounced for ρ close to $+1$ than for ρ close to -1. The Bartlett correction seems to work well for $\rho \in (-0.9, 0.7)$ when $T = 20$. For $T = 80$, the frequencies are only 1% different from the nominal level if $\rho \leq 0.9$. Because the Bartlett factor tends to infinity as $\rho \uparrow 1$, the corrected test becomes very conservative for values of ρ close to 1. The use of the simulated values of E[LR^c] leads to a slight improvement in comparison to the analytical Bartlett factor.

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

TABLE 1. Rejection frequencies (in percentages) in the AR(1) model with possible intercept and trend for various starting values. The column labeled χ^2 is based on the asymptotic χ^2-distribution. The columns labeled Th. 1–Th. 3 are based on the Bartlett corrections shown in Theorems 1–3, whereas the columns labeled (14a)–(14c) are based on the Bartlett corrections shown in formulas (14a)–(14c) using $\hat{\gamma} = (y_0 - \hat{\mu})^2/(s^2 T)$ where $\hat{\mu}$ and s^2 are based on the OLS estimators of μ and σ^2 (and $\hat{\mu} = 0$ in the pure AR(1) model)

	AR(1)			AR(1) + intercept			AR(1) + intercept + trend		
ρ	χ^2	Th. 1	(14a)	χ^2	Th. 2	(14b)	χ^2	Th. 3	(14c)
$T = 20$				$y_0 = 3/\sqrt{1-\rho^2}$					
0	5.1	5.4	5.0	5.8	5.5	5.1	7.4	5.2	5.2
0.6	5.2	5.5	5.0	7.9	4.2	5.0	14.4	2.5	4.4
0.8	5.4	5.7	5.0	10.2	2.3	4.7	24.1	0.5	2.1
0.9	5.8	6.1	5.2	12.9	0.5	3.4	37.1	0.0	0.5
0.95	6.3	6.6	5.5	15.9	0.0	1.6	48.2	0.0	0.1
0.99	6.3	6.6	5.4	24.7	0.0	0.1	59.0	0.0	0.0
$T = 20$				$y_0 = 6/\sqrt{1-\rho^2}$					
0	5.5	5.8	5.0	6.2	5.9	5.1	7.4	5.3	5.3
0.6	5.6	5.9	5.0	7.3	3.8	5.2	11.4	1.9	5.5
0.8	5.7	6.0	5.0	8.3	1.8	5.3	17.5	0.3	4.3
0.9	6.0	6.3	5.2	9.1	0.4	5.1	30.0	0.0	1.4
0.95	6.1	6.5	5.3	10.0	0.0	3.4	45.1	0.0	0.2
0.99	6.0	6.4	5.1	16.9	0.0	0.3	58.9	0.0	0.0
$T = 80$				$y_0 = 3/\sqrt{1-\rho^2}$					
0	4.9	5.0	5.0	5.1	5.0	5.0	5.5	5.0	5.0
0.6	5.0	5.0	5.0	5.8	4.9	5.0	7.7	4.6	4.9
0.8	5.0	5.1	5.0	6.8	4.5	4.9	11.0	3.6	4.5
0.9	5.0	5.1	5.0	8.3	3.5	4.7	16.7	1.8	3.3
0.95	5.1	5.1	5.0	10.3	1.8	4.1	26.0	0.3	1.2
0.99	5.7	5.8	5.5	17.0	0.0	0.7	51.4	0.0	0.0
$T = 80$				$y_0 = 6/\sqrt{1-\rho^2}$					
0	5.0	5.1	5.0	5.2	5.1	5.0	5.5	5.0	5.0
0.6	5.0	5.1	5.0	5.7	4.8	5.0	7.1	4.2	5.0
0.8	5.1	5.1	5.0	6.3	4.1	5.0	9.1	2.8	5.0
0.9	5.1	5.2	5.0	6.9	2.8	5.0	12.3	1.2	4.7
0.95	5.1	5.2	5.0	7.6	1.2	5.0	18.7	0.2	3.0
0.99	5.4	5.5	5.2	9.6	0.0	2.1	49.5	0.0	0.0

The lower two graphs of Figure 2 show the results for the AR(1) model with intercept and trend. When $T = 20$, the ordinary LR test massively rejects the true null hypothesis for large positive values of ρ (its empirical size is higher than 20% for $\rho > 0.7$). The Bartlett-corrected LR test, however, becomes very conservative for these parameter values. For $T = 80$, the rejection frequencies are within 1

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

percentage point of the nominal significance level for $\rho \leq 0.8$. Again, adjusting using simulated values of $\mathrm{E}[LR^r]$ leads to more accurate inference for values of ρ close to 1.

Next, we investigate the effect of the starting value on the simulation results for $T \in \{20, 80\}$. We consider starting values that are 3 and 6 times the unconditional standard deviation away from the mean, i.e., $y_0 = \delta/\sqrt{1-\rho^2}$ for $\delta \in \{3, 6\}$. Table 1 shows the rejection frequencies (in percentages). Overall, the starting value has little impact on the estimated sizes of tests using the χ^2-distribution and the Bartlett corrections shown in Theorems 1–3. However, the approximations based on the formulas (14a)–(14c) are able to correct the undersized Bartlett corrections for a large range of ρ values. For instance, when $\rho = 0.9$ and $T = 20$ in the AR(1) model with intercept, the estimated size increases from 0.5% to 3.4% for $\delta = 3$ and from 0.4% to 5.1% for $\delta = 6$. When $T = 80$, the estimated size is very close to the nominal significance level for $\rho = 0.9$ even in the AR(1) model with intercept and trend.

Last, we comment on the (size-adjusted) power properties of the Bartlett-corrected tests. In this paragraph, ρ represents the true value, and ρ_0 is the value under test. Let $LR_{\alpha,\rho,T}$ denote the critical value of the LR test statistic with size α. If $B_T(\rho) = 1 + k(\rho)/T$ denotes the Bartlett correction, then the critical value of the Bartlett-corrected test statistic with size α is equal to $LR_{\alpha,\rho,T}/B_T(\rho)$. Because $B(\rho_0)$ is used as the correction factor when testing $H_0 : \rho = \rho_0$, the power of the Bartlett-corrected test statistic $LR^* = LR/B_T(\rho_0)$ depends upon the ratio $B_T(\rho_0)/B_T(\rho)$. In the pure AR(1) model, LR^* inherits the power properties of the underlying test statistic LR as a result of the fact that $B_T(\rho_0)/B_T(\rho) = 1$ because the Bartlett correction does not depend upon ρ. In the model with deterministic terms, however, the Bartlett correction is increasing in ρ, which leads to $B_T(\rho_0)/B_T(\rho) < 1$ for $\rho_0 < \rho$ and $B(\rho_0)/B(\rho) > 1$ for $\rho_0 > \rho$. So, if deterministic terms are included, the power of the Bartlett-corrected test statistic LR^* will be higher than the power of LR for $\rho_0 < \rho$, whereas the power of LR^* will be lower than the power of LR for $\rho_0 > \rho$. Moreover, because of the explosive nature of $B_T(\rho_0)$ as ρ_0 tends to 1, the power might not increase as ρ_0 moves away from ρ toward 1.

Overall, we conclude that the Bartlett correction works well within a reasonable range of the parameter space. However, the ability to control the size depends on the deterministic terms included in the estimation model and the treatment of the starting value.

7. CONCLUSION

In this paper, the Bartlett correction is derived for testing hypotheses about the autoregressive parameter ρ in the stable AR(1) model with and without an intercept and linear trend. Although we have focused on hypothesis testing, accurate confidence intervals for ρ can be constructed by inverting a test statistic with good finite-sample properties. In the model without deterministic terms, the Bartlett

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

correction does not depend on ρ, and its magnitude is small. In case deterministic components are present, it is found that the correction factor is asymmetric in ρ. Furthermore, the Bartlett correction is monotonically increasing in ρ and tends to infinity when ρ approaches the stability boundary of $+1$. Hence, the Bartlett factor overcorrects for large positive values of ρ. At least theoretically, the (first-order) Bartlett correction does not depend on the treatment of the starting value, although it can be important in finite samples.

The simulation results indicate that the Bartlett corrections are useful for controlling the size of the LR statistic in the models considered. The empirical size is close to the nominal significance level (only 1 percentage point deviation) for a large part of the parameter space, even though the range of the parameter space critically depends upon the deterministic components in the estimation model. Although useful, these Bartlett corrections are not the ultimate panacea for the finite-sample problems that exist in autoregressive models.

REFERENCES

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

Bartlett, M.S. (1937) Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London, Series A* 160, 268–282.

Bravo, F. (1999) A correction factor for unit root test statistics. *Econometric Theory* 15, 218–227.

Chen, W.W. & R.S. Deo (2006) A Smooth Transition to the Unit Root Distribution via the Chi-Square Distribution with Interval Estimation for Nearly Integrated Autoregressive Processes. Working paper (MPRA 1215) presented at the ESEM 2007 (Budapest).

Cribari-Neto, F. & G. Cordeiro (1996) On Bartlett and Bartlett-type corrections. *Econometric Reviews* 15, 339–367.

Evans, G.B.A. & N.E. Savin (1981) Testing for unit roots: 1. *Econometrica* 49, 753–779.

Hayakawa, T. (1977) The likelihood ratio criterion and the asymptotic expansion of its distribution (see *Annals of the Institute of Statistical Mathematics*, 39, 681 for correction). *Annals of the Institute of Statistical Mathematics* 29, 359–378.

Jensen, J.L. & A.T.A. Wood (1997) On the non-existence of a Bartlett correction for unit root tests. *Statistics and Probability Letters* 35, 181–187.

Johansen, S. (2004) Asymptotic and finite sample properties of the Dickey-Fuller test. In A. Welfe (ed.), *New Directions in Macromodelling*, vol. 269, pp. 49–68. Elsevier Science.

Kiviet, J.F. & J.-M. Dufour (1997) Exact tests in single equation autoregressive distributed lag models. *Journal of Econometrics* 80, 325–353.

Lagos, B.M. & P.A. Morettin (2004) Improvement of the likelihood ratio test statistic in ARMA models. *Journal of Time Series Analysis* 25, 83–101.

Larsson, R. (1998) Bartlett corrections for unit root test statistics. *Journal of Time Series Analysis* 19, 425–438.

Lawley, D.N. (1956) A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika* 43, 295–303.

Nielsen, B. (1997) Bartlett correction of the unit root test in autoregressive models. *Biometrika* 84, 500–504.

Nielsen, B. (1999) The likelihood-ratio test for rank in bivariate canonical correlation analysis. *Biometrika* 86, 279–288.

Omtzigt, P.H. (2003) Notes on the Bartlett Correction of AR(1) processes. Manuscript, European University Institute, Florence.

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press

Phillips, P.C.B. (1977) Approximations to some finite sample distributions associated with a first order stochastic difference equation. *Econometrica* 45, 463–486.

Sul, D., P.C.B. Phillips & C.-Y. Choi (2005) Prewhitening bias in HAC estimation. *Oxford Bulletin of Economics and Statistics* 67, 517–546.

Taniguchi, M. (1988) Asymptotic expansions of the distribution of some test statistics for Gaussian ARMA processes. *Journal of Multivariate Analysis* 27, 494–511.

Taniguchi, M. (1991) *Higher Order Asymptotic Theory for Time Series Analysis*. Springer-Verlag.

van Giersbergen, N.P.A. (2004) Bartlett Correction in the Stable AR(1) Model with Intercept and Trend. UvA-Econometrics, Working paper 2004/07.

Wolfram, S. (1991) *Mathematica: A System for Doing Mathematics by Computer*. Addison-Wesley.

https://doi.org/10.1017/S0266466609090690 Published online by Cambridge University Press