UvA ECONOMETRICS

Discussion Paper: 2004/07

# Bartlett Correction in the Stable AR(1) Model with Intercept and Trend

Noud P.A. van Giersbergen

**www.fee.uva.nl/ke/UvA-Econometrics**

Department of Quantitative Economics
Faculty of Economics and Econometrics
Universiteit van Amsterdam
Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA UNIVERSITEIT VAN AMSTERDAM

# Bartlett Correction in the Stable AR(1) Model with Intercept and Trend*

Noud P.A. van Giersbergen
Department of Quantitative Economics,
Universiteit van Amsterdam
Roetersstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: N.P.A.vanGiersbergen@uva.nl

December 10, 2004

**Abstract**

The Bartlett correction is derived for testing hypotheses about the autoregressive parameter $\rho$ in the stable: (i) AR(1) model; (ii) AR(1) model with intercept; (iii) AR(1) model with intercept and linear trend. The correction is found explicitly as a function of $\rho$. In the models with deterministic terms, the correction factor is asymmetric in $\rho$. Furthermore, the Bartlett correction is monotonic increasing in $\rho$ and tends to infinity when $\rho$ approaches the stability boundary of 1. Simulation results indicate that the Bartlett corrections are useful in controlling the size of the LR statistic in small samples.

*Keywords:* Autoregressive models, Bartlett correction, small-sample properties, likelihood ratio statistic
*JEL classification*: C13; C22.

*This is a completely revised and extended version of the paper 'Bartlett correction in the stable AR(1) model with intercept', UvA-Econometrics Discussion Paper 2003/09. Universiteit van Amsterdam.

Helpful comments of participants of the ESEM 2004 meeting (Madrid, Spain) and the UvA-Econometrics seminar (Amsterdam, The Netherlands) are gratefully acknowledged.

## 1 Introduction

Of the three main test principles (likelihood ratio, Lagrange multiplier and Wald), the likelihood ratio (LR) approach is fortunately endowed with some general theory initiated by Bartlett to improve its finite-sample performance. Barlett (1937) has shown that the finite-sample distribution (at least its cumulants to the order $T^{-2}$) of a Barlett-corrected test statistic is closer to the $\chi^2$-distribution than the original test statistic; see also Lawley (1956) and Barndorff-Nielsen and Hall (1988) for the general i.i.d. case. The reader is referred to Cribari-Neto and Cordeiro (1996) for an econometric oriented review about Bartlett corrections.

Recently, Bartlett-type corrections in *un*stable autoregressive models have attracted much attention, see inter alias Bravo (1999), Larsson (1998), Nielsen (1997) and Johansen (2003), 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. However, it appears that a Bartlett-like factor can be calculated and that the use of this factor leads to a reduction of size distortions when testing the unit root hypothesis.

In this paper, we analyse the stable AR(1) model, which was also considered by Taniguchi (1988, 1991) and Omtzigt (2003). Stable AR models are of interest, since they arise as models for first differences, say $\Delta y_t$, when it is known that a time series $y_t$ can be characterised as an autoregressive process containing a unit root. Although in the stable case the inclusion of deterministic components like intercept and trend does not change the asymptotic distributions, these components highly affect the finite-sample behaviour of the LR statistic. Hence, in contrast to the work of Taniguchi, we consider the AR(1) model with intercept (and linear trend). Anticipating the results, we can say that the correction factor in models with deterministic components is considerably different from the factor obtained in the pure AR(1) model. Besides extending the results to a slightly more general model, the method used in this paper is totally different from the rather technical approach used by Taniguchi. Hence, this paper provides an alternative and much simpler proof, which basically involves only summations of geometric series. Although a number of exact inference techniques are available for the AR(1) model, see inter alias Andrews (1993) and Kiviet and Dufour (1997), these methods are based on numerical (Monte Carlo) procedures that do not give sufficient analytical insight into the structure of the finite-sample problem.

The paper is organised as follows. In Section 2, the Bartlett correction is determined in the pure AR(1) model. In Section 3, the results are generalised to the AR(1) model with intercept. The model with intercept and trend is considered in Section 4. Section 5 presents some simulation results to shed some light on the small-sample properties of the Bartlett-corrected tests. The conclusions can be found in the Section 5. Finally, three appendices contain some intermediate results, which are heavily used in the proofs of this paper.

A word on notation. Throughout the paper, the symbol $\overset{1}{=}$ will indicate that we have kept terms of order $T^{-1}$, i.e. if the stochastic expansion is of the form

$$V = V_0 + T^{-1}V_1 + T^{-2}V_2 + ...,$$

then

$$V \overset{1}{=} V_0 + T^{-1}V_1,$$

where $V_i \in O_p(1)$ are random variables. Furthermore, we use $\sum$ to indicate summation over $t = 1, ..., T$.

## 2 AR(1) model without intercept[1]

Consider the stable Gaussian AR(1) model

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

where the starting value $y_0$ is finite and $t = 1, ..., T$. In deriving an expression for the LR statistic for testing the null hypothesis $\rho = \rho_0$ against the two-sided alternative $\rho \neq \rho_0$, model (1) can be treated as just any other (linear) regression model. Therefore, the likelihood ratio (to the power $2/T$) is equal to

$$LR^{2/T} = \frac{SSR_U}{SSR_R}, \tag{2}$$

where $SSR$ denotes the sum of squared residuals and the subindex indicates whether the restricted ($R$) or unrestricted ($U$) residuals are used. Under the null hypothesis, the restricted SSR is equal to

$$SSR_R = \sum (y_t - \rho_0 y_{t-1})^2 = \sum \varepsilon_t^2 = S_{\varepsilon\varepsilon}, \tag{3}$$

where $S_{\varepsilon\varepsilon} \equiv \sum \varepsilon_t^2$. Using $S_{y\varepsilon} \equiv \sum y_{t-1}\varepsilon_t$ and $S_{yy} \equiv \sum y_{t-1}^2$, the unrestricted SSR can be written as

$$SSR_U = \sum (y_t - \hat{\rho} y_{t-1})^2 = S_{\varepsilon\varepsilon} - \frac{S_{y\varepsilon}^2}{S_{yy}}. \tag{4}$$

[1] This section, dealing with the AR(1) model *without* intercept, draws heavily on the unpublished manuscript "Notes on the Bartlett correction of AR(1) processes" by Pieter Omtzigt, although the results in this paper are obtained in a totally different way.

Substituting (3) and (4) in (2), we get

$$LR^{2/T} = 1 - \frac{S_{y\varepsilon}^2}{S_{yy}S_{\varepsilon\varepsilon}} \equiv 1 - Q, \tag{5}$$

where $Q$ is the squared sample correlation coefficient between $\varepsilon_t$ and $y_{t-1}$ assuming zero means, i.e.

$$Q \equiv \frac{S_{y\varepsilon}^2}{S_{yy}S_{\varepsilon\varepsilon}} = \frac{\left(T^{-1/2}S_{y\varepsilon}\right)^2}{T^{-1}S_{yy}S_{\varepsilon\varepsilon}}. \tag{6}$$

The LR test statistic is equal to

$$\begin{aligned} -2\log(LR) &= -T\log(1-Q) \\ &= TQ + \frac{1}{2}TQ^2 + O_p(T^{-2}), \end{aligned} \tag{7}$$

since $Q = O_p(T^{-1})$. Note that $Q$ and therefore the LR statistic is scale invariant, i.e. $Q$ based on $z_t = y_t/\sigma$ is the same as $Q$ based on $y_t$. Hence, we shall continue the analysis assuming $\sigma = 1$.

In order to calculate the Bartlett correction, we have to determine the expectation of the LR test statistic, i.e.

$$\mathbb{E}[-2\log(LR)] \overset{1}{=} \mathbb{E}[TQ] + \mathbb{E}[\frac{1}{2}TQ^2]. \tag{8}$$

The right-hand side of (8) consists of two terms. The last term is easily calculated, since

$$TQ^2 = \frac{1}{T}\frac{S_{y\varepsilon}^4}{S_{yy}^2}\frac{1}{(T^{-1}S_{\varepsilon\varepsilon})^2} \qquad \text{and} \qquad T^{-1}S_{\varepsilon\varepsilon} \overset{0}{=} 1. \tag{9}$$

This allows us to write

$$\mathbb{E}[\frac{1}{2}TQ^2] \overset{1}{=} \frac{1}{2T}\mathbb{E}\left[\frac{S_{y\varepsilon}^4}{S_{yy}^2}\right]. \tag{10}$$

Furthermore, it is known for a stable AR(1) process that

$$\frac{S_{y\varepsilon}}{\sqrt{S_{yy}}} = \frac{T^{-1/2}S_{y\varepsilon}}{\sqrt{T^{-1}S_{yy}}} \overset{d}{\to} \mathcal{N}(0,1), \tag{11}$$

so that

$$\mathbb{E}[\frac{1}{2}TQ^2] \overset{1}{=} \frac{3}{2T}. \tag{12}$$

The first term on the right-hand side of (8) is somewhat more difficult to determine. First, the two factors in the denominator are written in deviation from the expected values, i.e.

$$TQ = \left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(1-(1-\frac{1}{T}S_{\varepsilon\varepsilon})\right)^{-1}\left(\tau^2-(\tau^2-\frac{1}{T}S_{yy})\right)^{-1}, \tag{13}$$

where $\tau^2 = (1-\rho^2)^{-1}$ denotes the (asymptotic) unconditional variance of $y_t$. Next, the use of the series approximation (where $a$ denotes a constant and $x$ is small) for the two terms in the denominator

$$(a-x)^{-1} = a^{-1} + a^{-2}x + a^{-3}x^2 + ..., \tag{14}$$

results in the following approximation

$$
\begin{aligned}
TQ &\stackrel{1}{=} \left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2 \left(1 + (1 - \frac{1}{T}S_{\varepsilon\varepsilon}) + (1 - \frac{1}{T}S_{\varepsilon\varepsilon})^2\right) \\
&\quad \times \left(\tau^{-2} + \tau^{-4}(\tau^2 - \frac{1}{T}S_{yy}) + \tau^{-6}(\tau^2 - \frac{1}{T}S_{yy})^2\right). \qquad (15)
\end{aligned}
$$

Expanding this expression would lead to nine terms. Making use of the fact that

$$(1 - \frac{1}{T}S_{\varepsilon\varepsilon})^2 = O_p(T^{-1}) \qquad \text{and} \qquad (\tau^2 - \frac{1}{T}S_{yy})^2 = O_p(T^{-1}), \qquad (16)$$

three of the nine terms are smaller than $O_p(T^{-1})$. Hence, we get

$$
\begin{aligned}
\mathbb{E}[TQ] &\stackrel{1}{=} \mathbb{E}\left[\tau^{-2}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\right] + && \text{(C1)} \\
&\quad \mathbb{E}\left[\tau^{-2}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(1 - \frac{1}{T}S_{\varepsilon\varepsilon}\right)\right] + && \text{(C2)} \\
&\quad \mathbb{E}\left[\tau^{-4}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(\tau^2 - \frac{1}{T}S_{yy}\right)\right] + && \text{(C3)} \\
&\quad \mathbb{E}\left[\tau^{-4}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(\tau^2 - \frac{1}{T}S_{yy}\right)\left(1 - \frac{1}{T}S_{\varepsilon\varepsilon}\right)\right] + && \text{(C4)} \\
&\quad \mathbb{E}\left[\tau^{-2}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(1 - \frac{1}{T}S_{\varepsilon\varepsilon}\right)^2\right] + && \text{(C5)} \\
&\quad \mathbb{E}\left[\tau^{-6}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\left(\tau^2 - \frac{1}{T}S_{yy}\right)^2\right] && \text{(C6)} \\
&\equiv \sum_{i=1}^{6} C_i && (17)
\end{aligned}
$$

The Bartlett correction in the pure AR(1) model is stated in Theorem 1.

**Theorem 1** *In the stable Gaussian AR(1) model, given in (1), the expectation of the likelihood ratio test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\mathbb{E}[-2\log(LR)] = 1 - \frac{2}{T} + \frac{3}{2T} = 1 - \frac{1}{2}\frac{1}{T}. \qquad (18)$$

**Proof of Theorem 1.** In the proof, the six terms shown in (C1)-(C6) will be referred to as $C_1$ through $C_6$. Using formula (C.3) in Appendix C, i.e. $\mathbb{E}[S_{y\varepsilon}^2] \stackrel{1}{=} \tau^2 T$, the first term $C_1$ can be approximated by

$$C_1 \equiv \mathbb{E}\left[\tau^{-2}\left(\frac{1}{\sqrt{T}}S_{y\varepsilon}\right)^2\right] = \frac{1}{\tau^2 T}\mathbb{E}[S_{y\varepsilon}^2] \stackrel{1}{=} \frac{\tau^2 T}{\tau^2 T} = 1. \qquad (19)$$

The second term can be approximated by

$$C_2 = \frac{1}{\tau^2 T}\mathbb{E}\left[S_{y\varepsilon}^2\right] - \frac{1}{\tau^2 T^2}\mathbb{E}\left[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}\right] \stackrel{1}{=} -\frac{4}{T} \qquad (20)$$

using formulae (C.3) and (C.4) in Appendix C. The third term can be written as

$$C_3 = \frac{1}{\tau^2 T}\mathbb{E}\left[S_{y\varepsilon}^2\right] - \frac{1}{\tau^4 T^2}\mathbb{E}\left[S_{y\varepsilon}^2 S_{yy}\right] \stackrel{1}{=} -\frac{4(1+2\rho^2)\tau^2}{T} \tag{21}$$

using (C.3) and (C.5). The fourth term can be approximated by

$$\begin{aligned} C_4 &= \frac{1}{\tau^2 T}\mathbb{E}[S_{y\varepsilon}^2] - \frac{1}{\tau^4 T^2}\mathbb{E}[S_{y\varepsilon}^2 S_{yy}] - \frac{1}{\tau^2 T^2}\mathbb{E}[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}] + \frac{1}{\tau^4 T^3}\mathbb{E}[S_{y\varepsilon}^2 S_{yy} S_{\varepsilon\varepsilon}] \\ &\stackrel{1}{=} -\frac{4(1+2\rho^2)\tau^2}{T} - 1 - \frac{4}{T} + 1 + \frac{6+4(1+2\rho^2)\tau^2}{T} = \frac{2}{T} \end{aligned} \tag{22}$$

using the results (C.3) to (C.6). The fifth term can be written as

$$\begin{aligned} C_5 &= \frac{1}{\tau^2 T}\mathbb{E}[S_{y\varepsilon}^2] - \frac{2}{\tau^2 T^2}\mathbb{E}[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}] + \frac{1}{\tau^2 T^3}\mathbb{E}[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}^2] \\ &\stackrel{1}{=} 1 - 2\,(1+\frac{4}{T}) + 1 + \frac{10}{T} = \frac{2}{T} \end{aligned} \tag{23}$$

using (C.3), (C.4) and (C.7). The sixth term can be approximated by

$$\begin{aligned} C_6 &= \frac{1}{\tau^2 T}\mathbb{E}[S_{y\varepsilon}^2] - \frac{2}{\tau^4 T^2}\mathbb{E}[S_{y\varepsilon}^2 S_{yy}] + \frac{1}{\tau^6 T^3}\mathbb{E}[S_{y\varepsilon}^2 S_{yy}^2] \\ &\stackrel{1}{=} 1 - 2(1+\frac{4(1+2\rho^2)\tau^2}{T}) + (1+\frac{2(5+13\rho^2)\tau^2}{T}) \\ &= \frac{2(1+5\rho^2)\tau^2}{T} \end{aligned} \tag{24}$$

due to (C.3), (C.5) and (C.8).

Carrying out the summation completes the proof. □

The Bartlett correction agrees with the factor calculated by Taniguchi (1988, 1991), although his results are based on a different technique. Contrary to the first-order bias expansion of the OLS estimator $\hat{\rho}$ that is given by

$$\mathbb{E}[\hat{\rho}] - \rho \stackrel{1}{=} -\frac{2\rho}{T}, \tag{25}$$

see e.g. Marriot and Pope (1954), the Bartlett correction turns out to be independent of the AR(1) parameter $\rho$. 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.

# 3 AR(1) model with intercept

In this section, a constant is added to the statistical model. Consider the AR(1) model with an intercept

$$y_t = \rho y_{t-1} + \mu + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, \sigma^2), \qquad |\rho| < 1, \tag{26}$$

where the starting value $y_0$ is finite. First, we derive an expression for the LR statistic, which is based on the ratio of the two sums of squared residuals. After some lengthy but not very difficult algebra, the following two expressions for the SSRs under the null hypothesis $\rho = \rho_0$ are obtained

$$SSR_R = S_{\varepsilon\varepsilon} - T^{-1} S_\varepsilon^2 \tag{27}$$

and

$$SSR_U = \frac{S_{\varepsilon\varepsilon} S_y^2 - 2 S_\varepsilon S_y S_{y\varepsilon} + S_\varepsilon^2 S_{yy} + S_{y\varepsilon}^2 T - S_{\varepsilon\varepsilon} S_{yy} T}{S_y^2 - S_{yy} T}. \tag{28}$$

Simplifying the ratio of the two SSRs, the likelihood ratio can be written as

$$(LR^\tau)^{2/T} \stackrel{1}{=} 1 - \frac{(S_{y\varepsilon}^c)^2}{S_{yy}^c S_{\varepsilon\varepsilon}^c} \equiv 1 - Q^c, \tag{29}$$

where

$$S_{y\varepsilon}^c \equiv S_{y\varepsilon} - T^{-1} S_\varepsilon S_y, \tag{30a}$$

$$S_{yy}^c \equiv S_{yy} - T^{-1} S_y^2, \tag{30b}$$

$$S_{\varepsilon\varepsilon}^c \equiv S_{\varepsilon\varepsilon} - T^{-1} S_\varepsilon^2. \tag{30c}$$

As before, $Q^c$ in formula (29) can be interpreted as a squared sample correlation coefficient between $\varepsilon_t$ and $y_{t-1}$, but now corrected for non-zero means. The LR statistic is given by

$$-2\log(LR^c) = -T \log(1 - Q^c). \tag{31}$$

Adding a constant to the AR(1) model, makes the LR statistic invariant with respect to location and scale. To verify this, suppose that $z_t = \phi y_t + \psi$, for some constants $\phi$ and $\psi$. Obviously, $S_{\varepsilon\varepsilon}^c$ remains unaffected by the transformation, whereas

$$S_{z\varepsilon}^c = \phi S_{y\varepsilon}^c \qquad \text{and} \qquad S_{zz}^c = \phi^2 S_{yy}^c, \tag{32}$$

so that

$$1 - \frac{(S_{z\varepsilon}^c)^2}{S_{zz}^c S_{\varepsilon\varepsilon}^c} = 1 - \frac{(S_{y\varepsilon}^c)^2}{S_{yy}^c S_{\varepsilon\varepsilon}^c}. \tag{33}$$

Choosing $\phi = \sigma^{-1}$ and $\psi = -\mu/((1-\rho)\sigma)$, shows that the distribution of the LR statistic does not depend on the nuisance parameter $\mu$ and $\sigma$. Hence, the analysis is carried out under the assumption that $\mu = 0$ and $\sigma = 1$.

As before, the Bartlett correction of the LR statistic can be determined by

$$\mathbb{E}[-2\log(LR^c)] \stackrel{1}{=} \mathbb{E}[T Q^c] + \mathbb{E}[\frac{1}{2} T\, (Q^c)^2]. \tag{34}$$

The right-hand side of (34) consists of two expectations. Analogous to the analysis in the previous section, the last expectation can be approximated by

$$\mathbb{E}[\frac{1}{2}T(Q^c)^2] \stackrel{1}{=} \frac{3}{2T}. \tag{35}$$

The first term on the right-hand side of (34) is somewhat more difficult to determine. Since

$$\mathbb{E}\left[\frac{1}{T}S^c_{\varepsilon\varepsilon}\right] \stackrel{0}{=} \mathbb{E}\left[\frac{1}{T}S_{\varepsilon\varepsilon}\right] \qquad \text{and} \qquad \mathbb{E}\left[\frac{1}{T}S^c_{yy}\right] \stackrel{0}{=} \mathbb{E}\left[\frac{1}{T}S_{yy}\right],$$

the expansion of $TQ^c$ is analogous to the expansion of $TQ$, but now in terms of $S^c_{y\varepsilon}$ $S^c_{yy}$ and $S^c_{\varepsilon\varepsilon}$. As before, we are left with six relevant components and we can write

$$\mathbb{E}[TQ^c] = \sum_{i=1}^{6} C^c_i = \sum_{i=1}^{6} \left(C_i + D^c_i\right), \tag{36}$$

where $D^c_i \equiv C^c_i - C_i$ denotes the difference between the coefficients derived in the AR(1) model with intercept and the AR(1) model without intercept. Next, the Bartlett correction factor is given in Theorem 2.

**Theorem 2** *In the stable Gaussian AR(1) model with intercept, given in (26), the expectation of the likelihood ratio test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\mathbb{E}[-2\log(LR^c)] \stackrel{1}{=} 1 - \frac{1}{2T} + \frac{1+3\rho}{(1-\rho)T} = 1 + \frac{1+7\rho}{2(1-\rho)}\frac{1}{T}. \tag{37}$$

**Proof of Theorem 2.** In the proof, each of the components $D^c_i$ and $C^c_i$ will be discussed. Using (C.9) and (C.10) from Appendix C, we get

$$\begin{aligned} D^c_1 &\equiv \frac{1}{\tau^2 T}\mathbb{E}\left[\left(S^c_{y\varepsilon}\right)^2 - S^2_{y\varepsilon}\right] \\ &= -\frac{4\theta^2}{\tau^2 T} + \frac{3\theta^2}{\tau^2 T} = -\frac{\theta^2}{\tau^2 T} \qquad \text{and} \qquad C^c_1 = 1 - \frac{\theta^2}{\tau^2 T}, \end{aligned} \tag{38}$$

where $\theta \equiv 1/(1-\rho)$. The second component $D^c_2$ can be decomposed in 7 terms

$$\begin{aligned} D^c_2 \equiv\ & \underset{(1)}{\frac{\mathbb{E}[S^4_\varepsilon S^2_y]}{\tau^2 T^5}} - \underset{(2)}{\frac{2\mathbb{E}[S^3_\varepsilon S_{y\varepsilon} S_y]}{\tau^2 T^4}} + \underset{(3)}{\frac{\mathbb{E}[S^2_\varepsilon S^2_{y\varepsilon}]}{\tau^2 T^3}} + \\ & \underset{(4)}{\frac{\mathbb{E}[S^2_\varepsilon S^2_y]}{\tau^2 T^3}} - \underset{(5)}{\frac{\mathbb{E}[S^2_\varepsilon S_{\varepsilon\varepsilon} S^2_y]}{\tau^2 T^4}} + \underset{(6)}{\frac{2\mathbb{E}[S^2_\varepsilon S_{\varepsilon\varepsilon} S_{y\varepsilon} S^2_y]}{\tau^2 T^3}} - \underset{(7)}{\frac{2\mathbb{E}[S_\varepsilon S_{y\varepsilon} S_y]}{\tau^2 T^2}}. \end{aligned} \tag{39}$$

The first two terms are 'too small' to contribute to the first-order expansion since

$$\frac{S^4_\varepsilon S^2_y}{T^5} = T^{-2}\left(\frac{S_\varepsilon}{\sqrt{T}}\right)^4\left(\frac{S_y}{\sqrt{T}}\right)^2 = O_p(T^{-2}) \tag{40}$$

and

$$\frac{S_\varepsilon^3 S_{y\varepsilon} S_y}{T^4} = T^{-3/2}\left(\frac{S_\varepsilon}{\sqrt{T}}\right)^3\left(\frac{S_{y\varepsilon}}{\sqrt{T}}\right)\left(\frac{S_y}{\sqrt{T}}\right) = O_p(T^{-3/2}). \tag{41}$$

Furthermore, using the fact that $T^{-1}S_{\varepsilon\varepsilon} \to 1$, we get

$$\frac{S_\varepsilon^2 S_{\varepsilon\varepsilon} S_y^2}{T^4} = \frac{S_{\varepsilon\varepsilon}}{T}\frac{S_\varepsilon^2 S_y^2}{T^3} \stackrel{1}{=} \frac{S_\varepsilon^2 S_y^2}{T^3} \tag{42}$$

and

$$\frac{S_\varepsilon^2 S_{\varepsilon\varepsilon} S_{y\varepsilon} S_y^2}{T^3} = \frac{S_{\varepsilon\varepsilon}}{T}\frac{S_\varepsilon^2 S_{y\varepsilon} S_y^2}{T^2} \stackrel{1}{=} \frac{S_\varepsilon^2 S_{y\varepsilon} S_y^2}{T^2}, \tag{43}$$

from which we deduce that the net contribution of the last four terms is negligible for the first-order expansion. This leaves us with only one contributing component, viz the third term. Making use of formula (C.11) in Appendix C shows that

$$D_2^c \stackrel{1}{=} \frac{\mathbb{E}[S_\varepsilon^2 S_{y\varepsilon}^2]}{\tau^2 T^3} \stackrel{1}{=} \frac{\tau^2 T^2}{\tau^2 T^3} = \frac{1}{T} \qquad \text{and} \qquad C_2^c = -\frac{3}{T}. \tag{44}$$

For components $D_3^c$ to $D_6^c$, we will only give the terms that make a non-zero contribution to the first-order expansion. Just like $D_2^c$, $D_3^c$ consist of 7 terms. Only 3 of these are interesting, namely

$$D_3^c \stackrel{1}{=} \frac{\mathbb{E}[S_{y\varepsilon}^2 S_y^2]}{\tau^4 T^3} + \frac{2\mathbb{E}[S_\varepsilon S_{y\varepsilon} S_y S_{yy}]}{\tau^4 T^3} - \frac{2\mathbb{E}[S_\varepsilon S_{y\varepsilon} S_y]}{\tau^2 T^2}. \tag{45}$$

Applying the three approximations (C.12)-(C.14) in Appendix C gives

$$D_3^c \stackrel{1}{=} \frac{\theta^2\tau^2 T^2}{\tau^4 T^3} + \frac{2(2\theta^2\tau^2 + 2\rho\theta\tau^4)T^2}{\tau^4 T^3} - \frac{2(2\theta^2 T)}{\tau^2 T^2} = \frac{\theta^2 + 4\rho\theta\tau^2}{\tau^2 T} \tag{46}$$

so that

$$C_3^c \stackrel{1}{=} \frac{\theta^2 + 4\rho\theta\tau^2 - 4(1+2\rho^2)\tau^4}{\tau^2 T}. \tag{47}$$

Components $D_4^c$ up to $D_6^c$ consist of sums of 23, 15 and 15 terms respectively, although none of them contribute to the first-order expansion. This leads to

$$D_i^c \stackrel{1}{=} 0 \qquad \text{and} \qquad C_i^c \stackrel{1}{=} C_i \qquad \text{for } i = 4, 5, 6. \tag{48}$$

The total additional term is equal to

$$\sum_{i=1}^{6} D_i^c \stackrel{1}{=} \frac{1 + 4\rho\theta}{T}\Big|_{\theta=(1-\rho)^{-1}} = \frac{1+3\rho}{(1-\rho)T}. \tag{49}$$

Adding (49) to (18) completes the proof. □

Contrary to the case when there is no intercept, i.e. formula (18), we now see that the factor depends on the AR(1) parameter $\rho$. For $\rho > -1/3$, 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 $\rho$ and hence it is asymmetric with respect to the origin; it even has an asymptote for $\rho \uparrow 1$; see Figure 1 for a graph when $T = 20$. This is in contrast to the bias expression of the OLS estimator $\hat{\rho}$, which is given by

$$\mathbb{E}[\hat{\rho}] - \rho \stackrel{1}{=} -\frac{1+3\rho}{T}, \tag{50}$$

see Kendall (1954).

Insert Figure 1 about here.

# 4 AR(1) model with intercept and Trend

In this section, the AR(1) model with an intercept and a linear trend is considered, i.e.

$$y_t = \rho y_{t-1} + \mu + \beta t + \varepsilon_t, \qquad \varepsilon_t \sim \text{NID}(0, \sigma^2), \qquad |\rho| < 1, \tag{51}$$

where the starting value $y_0$ is finite. After some lengthy and tedious but not very difficult algebra, the SSR under the null hypothesis $\rho = \rho_0$ can be written as

$$SSR_R = \frac{S_{\varepsilon\varepsilon} S_\tau^2 - 2S_\varepsilon S_\tau S_{\tau\varepsilon} + S_\varepsilon^2 S_{\tau\tau} + (S_{\tau\varepsilon}^2 - S_{\varepsilon\varepsilon} S_{\tau\tau})T}{S_\tau^2 - S_{\tau\tau} T}, \tag{52}$$

where $S_\tau \equiv \sum t$, $S_{\tau\tau} \equiv \sum t^2$ and $S_{\tau\varepsilon} \equiv \sum t\varepsilon_t$. Although an explicit expression for the unrestricted SSR can be found, the lengthy formula does not reveal any insight and is not shown here. To shorten the formulae considerably, the following two approximations are used

$$S_\tau = \frac{1}{2}T^2 + O(T) \quad \text{and} \quad S_{\tau\tau} = \frac{1}{3}T^3 + O(T^2), \tag{53}$$

without effecting the first-order correction. Using (53) and subsequently simplifying the ratio of the two SSRs, the likelihood ratio can be written as follows

$$(LR^\tau)^{2/T} \stackrel{1}{=} 1 - \frac{(S_{y\varepsilon}^\tau)^2}{S_{yy}^\tau S_{\varepsilon\varepsilon}^\tau} \equiv 1 - Q^\tau, \tag{54}$$

where

$$S_{y\varepsilon}^\tau \equiv S_{y\varepsilon} - \frac{12 S_{\tau\varepsilon} S_{\tau y}}{T^3} + \frac{6(S_\varepsilon S_{\tau y} + S_{\tau\varepsilon} S_y)}{T^2} - \frac{4 S_\varepsilon S_y}{T}, \tag{55a}$$

$$S_{yy}^\tau \equiv S_{yy} - \frac{12 S_{\tau y}^2}{T^3} + \frac{12 S_y S_{\tau y}}{T^2} - \frac{4 S_y^2}{T}, \tag{55b}$$

$$S_{\varepsilon\varepsilon}^\tau \equiv S_{yy} - \frac{12 S_{\tau\varepsilon}^2}{T^3} + \frac{12 S_\varepsilon S_{\tau\varepsilon}}{T^2} - \frac{4 S_\varepsilon^2}{T}. \tag{55c}$$

The structure of the likelihood ratio shown in (54) is very similar to the expressions obtained in the pure AR(1) model, cf (5), and AR(1) model with intercept, cf (29). In line with the previous section, the Bartlett correction will be derived under the assumptions $\mu = \beta = 0$ and $\sigma^2 = 1$. In model (51), the expected value of the LR statistic can be decomposed into

$$\mathbb{E}[-2\log(LR^{\tau})] \overset{1}{=} \mathbb{E}[TQ^{\tau}] + \mathbb{E}[\frac{1}{2}T\,(Q^{\tau})^2]. \tag{56}$$

The right-hand side of (56) consists of two terms of which the latter term can be approximated by

$$\mathbb{E}[\frac{1}{2}T(Q^{\tau})^2] \overset{1}{=} \frac{3}{2T}. \tag{57}$$

The first term on the right-hand side of (56) can be further decomposed into the following six components

$$\mathbb{E}[TQ^{\tau}] = \sum_{i=1}^{6} C_i^{\tau} = \sum_{i=1}^{6} \left(C_i + D_i^{\tau}\right). \tag{58}$$

The coefficients $D_i^{\tau}$ will be calculated explicitly in the proof of Theorem 3. This theorem 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 (51), the expectation of the likelihood ratio test for the null hypothesis $\rho = \rho_0$ has the expansion*

$$\mathbb{E}[-2\log(LR^{\tau})] \overset{1}{=} 1 - \frac{1}{2T} + \frac{4+8\rho}{(1-\rho)T} = 1 + \frac{7+17\rho}{2(1-\rho)}\frac{1}{T}. \tag{59}$$

**Proof of Theorem 3.** We have to determine the appropriate approximations of the 'additional' components $D_i^{\tau}$ $(i = 1, ..., 6)$ of formula (58). Component $D_1^{\tau}$ consists of 13 distinct terms,

$$\begin{aligned} D_1^{\tau} &\equiv \frac{1}{\tau^2 T}\mathbb{E}\left[\left(S_{y\varepsilon}^{\tau}\right)^2 - S_{y\varepsilon}^2\right] \\ &= \frac{144\mathbb{E}[S_{\tau\varepsilon}^2 S_{\tau y}^2]}{\tau^2 T^7} - \frac{144\mathbb{E}[S_{\varepsilon} S_{\tau\varepsilon} S_{\tau y}^2]}{\tau^2 T^6} - \frac{144\mathbb{E}[S_{\tau\varepsilon}^2 S_{\tau y} S_y]}{\tau^2 T^6} + \frac{36\mathbb{E}[S_{\varepsilon}^2 S_{\tau y}^2]}{\tau^2 T^5} + \\ &\quad \frac{168\mathbb{E}[S_{\varepsilon} S_{\tau\varepsilon} S_{\tau y} S_y]}{\tau^2 T^5} + \frac{36\mathbb{E}[S_{\tau\varepsilon}^2 S_y^2]}{\tau^2 T^5} - \frac{48\mathbb{E}[S_{\varepsilon}^2 S_{\tau y} S_y]}{\tau^2 T^4} - \frac{48\mathbb{E}[S_{\varepsilon} S_{\tau e} S_y^2]}{\tau^2 T^4} - \\ &\quad \frac{24\mathbb{E}[S_{\tau\varepsilon} S_{\tau y} S_{y\varepsilon}]}{\tau^2 T^4} + \frac{16\mathbb{E}[S_{\varepsilon}^2 S_y^2]}{\tau^2 T^3} + \frac{12\mathbb{E}[S_{\varepsilon} S_{\tau y} S_{y\varepsilon}]}{\tau^2 T^3} + \frac{12\mathbb{E}[S_{\tau\varepsilon} S_y S_{y\varepsilon}]}{\tau^2 T^3} - \frac{8\mathbb{E}[S_{\varepsilon} S_y S_{y\varepsilon}]}{\tau^2 T^2}, \end{aligned} \tag{60}$$

which are all of order $T^{-1}$. To evaluate these terms, we make use of various relationships that exist between expectations involving sums of $S_{\tau\varepsilon}$, $S_{\tau y}$, etcetera and expectations involving only $S_{\varepsilon}$, $S_y$. For example, we can approximate $\mathbb{E}[S_{\tau\varepsilon}^2 S_{\tau y}^2]$ by $\frac{1}{9}\mathbb{E}[S_{\varepsilon}^2 S_y^2]T^4 + O(T^5)$; see (C.15) in Appendix C for an overview of such relationships. Using (C.15) and simplifying gives

$$D_1^{\tau} \overset{1}{=} \frac{8\mathbb{E}[S_{\varepsilon}^2 S_y^2]}{3\tau^2 T^3} - \frac{4\mathbb{E}[S_{\varepsilon} S_y S_{y\varepsilon}]}{\tau^2 T^2} \overset{1}{=} 0, \tag{61}$$

where the last approximation uses (C.10) and (C.9).

The second component $D_2^\tau$ consists of 52 distinct terms of which 26 terms remain after the use of $T^{-1}S_{\varepsilon\varepsilon} \to 1$. Only 3 terms make a contribute to the first-order correction, which is given by

$$
\begin{aligned}
D_2^\tau &\overset{1}{=} \frac{12\mathbb{E}[S_{\tau\varepsilon}^2 S_{y\varepsilon}^2]}{\tau^2 T^5} - \frac{12\mathbb{E}[S_\varepsilon S_{\tau\varepsilon} S_{y\varepsilon}^2]}{\tau^2 T^4} + \frac{4\mathbb{E}[S_\varepsilon^2 S_{y\varepsilon}^2]}{\tau^2 T^3} \\
&\overset{1}{=} \frac{2\mathbb{E}[S_\varepsilon^2 S_{y\varepsilon}^2]}{\tau^2 T^3} \overset{1}{=} \frac{2}{T},
\end{aligned}
\tag{62}
$$

since $\mathbb{E}[S_{\tau\varepsilon}^2 S_{y\varepsilon}^2] = \frac{1}{3}\mathbb{E}[S_\varepsilon^2 S_{y\varepsilon}^2]T^2 + O(T^3)$ and $\mathbb{E}[S_\varepsilon S_{\tau\varepsilon} S_{y\varepsilon}^2] = \frac{1}{2}\mathbb{E}[S_\varepsilon^2 S_{y\varepsilon}^2]T + O(T^2)$.

Expanding $D_3^\tau$ leads to 52 distinct terms of which 29 are of order $T^{-1}$. After using the relationships shown in (C.15) and (C.16), only 5 different terms remain

$$
\begin{aligned}
D_3^\tau &\overset{1}{=} -\frac{8\mathbb{E}[S_\varepsilon^2 S_y^2 S_{yy}]}{3\tau^4 T^4} + \frac{2\mathbb{E}[S_y^2 S_{y\varepsilon}^2]}{\tau^4 T^3} + \frac{4\mathbb{E}[S_\varepsilon S_y S_{y\varepsilon} S_{yy}]}{\tau^4 T^3} + \frac{8\mathbb{E}[S_\varepsilon^2 S_y^2]}{3\tau^2 T^3} - \frac{4\mathbb{E}[S_\varepsilon S_y S_{y\varepsilon}]}{\tau^2 T^2} \\
&\overset{1}{=} -\frac{8\theta^2}{\tau^2 T} + \frac{2\theta^2}{\tau^2 T} + \frac{4(2\theta^2\tau^2 + 2\rho\theta\tau^4)}{\tau^4 T} + \frac{8\theta^2}{\tau^2 T} - \frac{8\theta^2}{\tau^2 T} \\
&\overset{1}{=} \frac{2(\theta^2 + 4\rho\theta\tau^2)}{\tau^2 T}.
\end{aligned}
\tag{63}
$$

Components $D_4^\tau$ up to $D_6^\tau$ consist of sums of 206, 129 and 129 different terms respectively, although none of them contribute to the first-order expansion. This leads to

$$
D_i^\tau \overset{1}{=} 0 \qquad \text{for } i = 4, 5, 6. \tag{64}
$$

The total additional term is equal to

$$
\sum_{i=1}^{6} D_i^\tau \overset{1}{=} \frac{2(\theta^2 + \tau^2)}{\tau^2 T}\bigg|_{\theta=(1-\rho)^{-1}} = \frac{4 + 8\rho}{(1-\rho)T}. \tag{65}
$$

Adding (65) to (18) 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 $\rho$ and goes to infinity for $\rho \uparrow 1$; see Figure 1 for a graph. 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 server than in the model without trend. Using Theorem 1 of Kiviet and Phillips (1993), the estimation bias of the AR(1) coefficient estimator can be shown to be

$$
\mathbb{E}[\hat{\rho}] - \rho \overset{1}{=} -\frac{2 + 4\rho}{T}. \tag{66}
$$

Note that the absolute value of the estimation bias of $\hat{\rho}$ in the model with trend, which is shown in (66), is uniformly (in $\rho$) larger than the estimation bias in the model without a trend, which is shown in (50).

## 5 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 Matlab.

Observations were generated according to an AR(1) process. All reported results are based on $10^5$ replications and four sample sizes were considered: $T \in \{20, 40, 80, 100\}$. The AR(1) parameter was taken as $\rho \in \{-0.99, 0.98, ..., 0.99\}$. Since the LR statistic is invariant with respect to $\sigma$ (and $\mu$ in the model with intercept), we set $\sigma = 1$ (and $\mu = 0$) without loss of generality. In addition, $\beta = 0$ in the model with a trend. The starting value was set to the expected value of the stationary distribution, i.e. $y_0 = 0$. In this way, the variance of $y_0$ remains constant as $\rho$ varies. The simulations were also carried out for $y_0 \sim \mathcal{N}(0, 1/(1-\rho^2))$, which did not change the results significantly. The nominal significance level was taken to be 5%. Results remain qualitatively the same for the 1% and 10% level.

Figure 2 shows the rejection frequencies for the pure AR(1) model. In this model, the LR test based on the asymptotic $\chi^2$-distribution seems to perform reasonably well. Only when $|\rho|$ is very close to 1, the rejection frequencies rise to around 6.5%. Note that the reject error seems to be an even function of $\rho$. Since the Bartlett correction is constant for a given sample size, the whole curve of rejection frequencies is clearly shifted towards the nominal significance level for $T = 20$.

Insert Figure 2 about here.

Figure 3 shows the rejection frequencies when an intercept is added to the AR(1) model. As expected, the discrepancy between the rejection frequencies and the nominal level is larger than in Figure 2. Furthermore, the reject error is more pronounced for $\rho$ close to $+1$ than for $\rho$ close to $-1$. The Bartlett correction seems to work well for $\rho \in (-0.9, 0.7)$ when $T = 20$. In case $T = 100$, the frequencies are only 1% different from the nominal level for $\rho < 0.92$. Since the Bartlett factor tends to infinity for $\rho \uparrow 1$, the corrected test becomes very conservative for values of $\rho$ close to 1.

The rejection frequencies for the AR(1) model with intercept and trend are shown in Figure 4. When $T = 20$, the ordinary LR test massively rejects the true null hypothesis for large positive values of $\rho$ (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 $T = 100$, the rejection frequencies are within 1 percentage point of the nominal significance level for $\rho < 0.83$.

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 critically on the deterministic terms included in the estimation model.

Insert Figures 3 and 4 about here.

# 6 Conclusion

In this paper, the Bartlett correction is derived for testing hypotheses about the autoregressive parameter $\rho$ in the stable AR(1) model with and without an intercept and linear trend. In case deterministic components are present, it is found that the correction factor is asymmetric in $\rho$. Furthermore, the Bartlett correction is monotonic increasing in $\rho$ and tends to infinity when $\rho$ approaches the stability boundary of $+1$. Hence, the Bartlett factor overcorrects for large positive values of $\rho$.

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, although the range of the parameter space critically depends upon the deterministic components in the estimation model. Hence, these Bartlett corrections are not the ultimate panacea for the finite-sample problems that exists in autoregressive models.

Since the models analysed in this paper are too simple to be useful in describing many economic time series, allowing for more dynamics seems highly desirable.

# References

Andrews, D. W. K. (1993). Exactly Median-unbiased Estimation of First-order Autoregressive/ Unit-root Models. *Econometrica*, **61**, 139–165.

Barlett, M. S. (1937). Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London, Series A*, **160**, 268–282.

Barndorff-Nielsen, O. E. and P. Hall (1988). On the Level Error After Bartlett Adjustment of the Likelihood Statistics. *Biometrika*, **75**, 374–378.

Bravo, F. (1999). A Correction Factor for Unit Root Test Statistics. *Econometric Theory*, **15**, 218–227.

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

Jensen, J. L. and 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. (2003). A Small Sample Correction of the Dickey-Fuller Test. Preprint 2003, Department of Applied Mathematics and Statistics, University of Copenhagen.

Kendall, M. G. (1954). Note on the Bias in the Estimation of Auto Correlation. *Biometrika*, **61**, 403–404.

Kiviet, J. F. and J.-M. Dufour (1997). Exact Tests in Single Equation Autoregressive Distributed Lag Models. *Journal of Econometrics*, **80**, 325–353.

Kiviet, J. F. and G. D. A. Phillips (1993). Alternative Bias Approximations in Regressions with a Lagged Dependent Variable. *Econometric Theory*, **9**, 62–80.

Kiviet, J. F. and G. D. A. Phillips (1998). Higher-order Asymptotic Expansions of the Least-squares Estimation Bias in First-order Dynamic Regression Models. Discussion paper TI 96-167/7, Amsterdam: Tinbergen Institute. Revised in October 1998.

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.

Magnus, J. R. (1978). The Moments of Products of Quadratic Forms in Normal Variables. *Statistica Neerlandica*, **32**, 201–210.

Marriot, F. H. C. and J. A. Pope (1954). Bias in the Estimation of Autocorrelations. *Biometrika*, **61**, 393–403.

Nielsen, B. (1997). Bartlett Correction of the Unit Root Test in Autoregressive Models. *Biometrika*, **84**, 500–504.

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

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*. New York: Springer-Verlag.

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

## Appendix A: Some auxiliary results

Let $A_1, ..., A_4$ be real *symmetric* $T \times T$ matrices. In addition, let the $T \times 1$ random vector $\varepsilon$ be such $\varepsilon \sim \mathcal{N}(0, I_T)$. In a paper by Magnus (1978), the following expressions were derived for the expectations of the form $\prod_{j=1}^{s} \varepsilon' A_j \varepsilon$ when $s \leq 4$:

$$\begin{aligned}
\mathbb{E}[\textstyle\prod_{i=1}^{2} \varepsilon' A_i \varepsilon] &= a_1 a_2 + 2a_{12}, & \text{(A.1)}\\
\mathbb{E}[\textstyle\prod_{i=1}^{3} \varepsilon' A_i \varepsilon] &= a_1 a_2 a_3 + 2(a_1 a_{23} + a_2 a_{13} + a_3 a_{12}) + 8a_{123}, & \text{(A.2)}\\
\mathbb{E}[\textstyle\prod_{i=1}^{4} \varepsilon' A_i \varepsilon] &= a_1 a_2 a_3 a_4 + 8(a_1 a_{234} + a_2 a_{134} + a_3 a_{124} + a_4 a_{123}), &\\
&\quad +4(a_{12}a_{34} + a_{13}a_{24} + a_{14}a_{23}) + 16(a_{1234} + a_{1243} + a_{1324}) &\\
&\quad +2(a_1 a_2 a_{34} + a_1 a_3 a_{24} + a_1 a_4 a_{23} + a_2 a_3 a_{14} + a_2 a_4 a_{13} + a_3 a_4 a_{12}), & \text{(A.3)}
\end{aligned}$$

where $a_i = Tr(A_i)$, $a_{ij} = Tr(A_i A_j)$, $a_{ijl} = Tr(A_i A_j A_l)$, $a_{ijkl} = Tr(A_i A_j A_l A_k)$ and $Tr(\cdot)$ denotes the trace operator.

## Appendix B: Some intermediate results

In calculating the various terms, it will be convenient to write the expectations in vector-form. Hence, define

$$\varepsilon = (\varepsilon_1, ..., \varepsilon_T)'$$

and

$$y_{-1} = (y_0, ..., y_{T-1})' \qquad \text{and} \qquad y = (y_1, ..., y_T)',$$

so that $S_{\varepsilon\varepsilon} = \varepsilon'\varepsilon$, $S_{y\varepsilon} = \varepsilon' y_{-1}$ and $S_{yy} = y'_{-1} y_{-1}$. In matrix notation, the pure AR(1) model can be written as $y = \rho y_{-1} + \varepsilon$. To express $y_{-1}$ in terms of the innovations, define the $T \times T$ matrix and $T \times 1$ vector

$$C = \begin{bmatrix} 0 & 0 & \cdots & & \cdots & 0 \\ 1 & 0 & & & & \vdots \\ \rho & 1 & 0 & & & \\ \rho^2 & \rho & 1 & 0 & & \\ \vdots & & & \ddots & \ddots & \vdots \\ \rho^{T-2} & \cdots & \rho^2 & \rho & 1 & 0 \end{bmatrix} \qquad \text{and} \qquad F = \begin{bmatrix} 1 \\ \rho \\ \rho^2 \\ \rho^3 \\ \vdots \\ \rho^{T-1} \end{bmatrix}.$$

The vector $y_{-1}$ may be decomposed as

$$y_{-1} = C\varepsilon + Fy_0. \tag{B.1}$$

Although the second-order estimation bias is likely to involve the starting value $y_0$, see for instance Kiviet and Phillips (1998) for the ARX(1) model with fixed regressors, the first-order bias does not depend on $y_0$ (provided that $y_0$ is finite). Hence, we shall ignore $y_0$ in the remainder of the analysis and we simply have that $y_{-1} = C\varepsilon$.

Even though the following trace results can be obtained by using the properties of geometric series, we have used the computer algebra system Mathematica© 4.0 for computation; see Wolfram

(1991).

$$
\begin{aligned}
Tr(CC') &= \sum_{j=0}^{T-2}\sum_{i=0}^{j}\rho^{2i} \\
&= \sum_{j=0}^{T-2}\frac{1-\rho^{2(j+1)}}{1-\rho^2} = \frac{(1-\rho^2)T+\rho^{2T}-1}{(1-\rho^2)^2} \\
&\underset{0}{=} \frac{1}{1-\rho^2}T = \tau^2 T, \qquad \text{(B.2)}
\end{aligned}
$$

where $\tau^2 \equiv (1-\rho^2)^{-1}$ and $\underset{0}{=}$ indicates that terms of order $O(T^0)$ or smaller are neglected. Note that the results in Appendix A only apply to symmetric matrices. Hence, let $A = (C + C')/2$. Using (B.2), we find

$$
\begin{aligned}
Tr(AA) &= \frac{1}{4}\left[Tr(CC) + Tr(CC') + Tr(C'C) + Tr(C'C')\right] \\
&= \frac{1}{2}Tr(CC') \underset{0}{=} \frac{1}{2}\tau^2 T, \qquad \text{(B.3)}
\end{aligned}
$$

since $Tr(CC) = Tr(C'C') = 0$ and $Tr(CC') = Tr(C'C)$.

$$
\begin{aligned}
Tr(AAC'C) &= \frac{1}{4}\left[Tr(CCC'C) + Tr(CC'C'C) + Tr(C'CC'C) + Tr(C'C'C'C)\right] \\
&\underset{0}{=} \frac{1}{4}[2\rho^2 + 2(1+\rho^2)]\tau^6 T = \frac{1}{2}(1+2\rho^2)\tau^6 T, \qquad \text{(B.4)}
\end{aligned}
$$

since

$$
\begin{aligned}
Tr(CCC'C) &= Tr(C'C'C'C) \underset{0}{=} \rho^2\tau^6 T, \\
Tr(CC'C'C) &\underset{0}{=} Tr(C'CC'C) \underset{0}{=} (1+\rho^2)\tau^6 T.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
Tr(AC'C) &= \frac{1}{2}Tr(CC'C) + \frac{1}{2}Tr(C'C'C) \\
&\underset{0}{=} \rho\tau^4 T, \qquad \text{(B.5)}
\end{aligned}
$$

since $Tr(CC'C) = Tr(C'C'C) \underset{0}{=} \rho\tau^4 T$.

For the AR(1) model with intercept, the following results are interesting

$$
Tr(A\iota\iota'C) \underset{0}{=} Tr(C'\iota\iota'C) \underset{0}{=} Tr(\iota\iota'AA) \underset{0}{=} \theta^2 T, \qquad \text{(B.6)}
$$

$$
Tr(C'\iota\iota'C\iota\iota') \underset{1}{=} \theta^2 T^2, \qquad \text{(B.7)}
$$

where $\theta^2 \equiv (1-\rho)^{-2}$ and $\underset{1}{=}$ indicates that terms of order $O(T^1)$ or smaller are neglected Finally,

$$
Tr(AAC'\iota\iota'C) \underset{0}{=} Tr(\iota\iota'CAC'A) \underset{0}{=} \theta^4 T. \qquad \text{(B.8)}
$$

## Appendix C: Some expectations

We are interested in calculating expectations such as

$$\mathbb{E}[S_{y\varepsilon}^2] = \mathbb{E}[\varepsilon' C\varepsilon\varepsilon' C\varepsilon]. \tag{C.1}$$

To apply the results of Magnus (1978) as reproduced in Appendix A, define $A = (C + C')/2$. For quadratic forms, it is well known that $\varepsilon' C\varepsilon = \varepsilon' A\varepsilon$. From (A.1) using $A_1 = A_2 = A$, we get

$$\mathbb{E}[\varepsilon' C\varepsilon\varepsilon' C\varepsilon] = \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon] = Tr(A)^2 + 2Tr(AA). \tag{C.2}$$

Note that $Tr(A) = 0$, since all elements on the diagonal are zero. Using result (B.3), we finally get the desired first-order approximation

$$\mathbb{E}[S_{y\varepsilon}^2] = \frac{1}{\tau^2 T} Tr(CC') \underset{0}{=} \tau^2 T, \tag{C.3}$$

where $\underset{0}{=}$ indicates that terms of order $O(T^0)$ or smaller are neglected. Again using (B.3), it follows that

$$\begin{aligned}\mathbb{E}\left[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}\right] &= E[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' I\varepsilon] \\ &= 2Tr(I)Tr(AA) + 8Tr(AA) \underset{0}{=} \tau^2 T^2 + 4\tau^2 T.\end{aligned} \tag{C.4}$$

With the use of (B.2) to (B.4), we find

$$\begin{aligned}\mathbb{E}\left[S_{y\varepsilon}^2 S_{yy}\right] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' C' C\varepsilon] \\ &= 2Tr(C'C)Tr(AA) + 8Tr(AAC'C) \\ &\underset{0}{=} \tau^4 T^2 + 4(1 + 2\rho^2)\tau^6 T.\end{aligned} \tag{C.5}$$

Application of (A.3) and (B.2) to (B.4) shows that

$$\begin{aligned}\mathbb{E}[S_{y\varepsilon}^2 S_{yy} S_{\varepsilon\varepsilon}] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' C' C\varepsilon\varepsilon' I\varepsilon] \\ &= 8\left[Tr(C'C)Tr(AA) + Tr(I)Tr(AAC'C)\right] + 4\,Tr(AA)Tr(C'C) + \\ &\quad 2\,Tr(C'C)Tr(I)Tr(AA) + 3 \cdot 8\,Tr(AAC'C) \\ &\underset{1}{=} \tau^4 T^3 + (6\tau^4 + 4(1 + 2\rho^2)\tau^6)T^2.\end{aligned} \tag{C.6}$$

By applying (A.3) and using (B.3), we get

$$\begin{aligned}\mathbb{E}[S_{y\varepsilon}^2 S_{\varepsilon\varepsilon}^2] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' I\varepsilon\varepsilon' I\varepsilon] \\ &= 2 \cdot 8\,Tr(I)Tr(AA) + 4\,Tr(AA)Tr(I) + 2\,Tr(I)^2 Tr(AA) + 3 \cdot 8\,Tr(AA) \\ &\underset{1}{=} 8\tau^2 T^2 + 2\tau^2 T^2 + \tau^2 T^3 + 12\tau^2 T = \tau^2 T^3 + 10\tau^2 T^2.\end{aligned} \tag{C.7}$$

For the component $C_6$, we need

$$\begin{aligned}\mathbb{E}[S_{y\varepsilon}^2 S_{yy}^2] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' C' C\varepsilon\varepsilon' C' C\varepsilon] \\ &= 2 \cdot 8\,Tr(C'C)Tr(AAC'C) + 4[Tr(AA)tr(C'CC'C) + 2Tr(AC'C)^2] + \\ &\quad 2Tr(C'C)^2 Tr(AA) + 16[2Tr(AAC'CC'C) + Tr(AC'CAC'C)] \\ &= 8(1 + 2\rho^2)\tau^8 T^2 + 4[\frac{1}{2}(1 + \rho^2)\tau^8 T^2 + 2\rho^2\tau^8 T^2] + \tau^6 T^3 + O(T) \\ &\underset{1}{=} \tau^6 T^3 + 2(5 + 13\rho^2)\tau^8 T^2,\end{aligned} \tag{C.8}$$

since $Tr(AAC'CC'C)$ and $Tr(AC'CAC'C)$ are only of order $O(T)$.

Using (B.6) and (B.7), we get

$$\begin{aligned}
\mathbb{E}[S_{y\varepsilon}S_{\varepsilon}S_{y}] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon'\imath\imath' C\varepsilon] = 2Tr(A\imath\imath' C) \underset{0}{=} 2\theta^2 T, && \text{(C.9)}\\
\mathbb{E}[S_{\varepsilon}^2 S_{y}^2] &= \mathbb{E}[\varepsilon'\imath\imath'\varepsilon\varepsilon' C'\imath\imath' C\varepsilon] \\
&= Tr(C'\imath\imath' C)Tr(\imath\imath') + 2Tr(C'\imath\imath' C\imath\imath') \underset{1}{=} 3\theta^2 T^2, && \text{(C.10)}\\
\mathbb{E}[S_{\varepsilon}^2 S_{y\varepsilon}^2] &= E[\varepsilon'\imath\imath'\varepsilon\varepsilon' A\varepsilon\varepsilon' A\varepsilon] \\
&= 2Tr(\imath\imath')Tr(AA) + 8Tr(\imath\imath' AA) \\
&= 2T\tfrac{1}{2}\tau^2 T + 8\theta^2 T \underset{1}{=} \tau^2 T^2. && \text{(C.11)}
\end{aligned}$$

For approximating $D_3^c$, we have to determine the following three expectations

$$\begin{aligned}
\mathbb{E}[S_{y\varepsilon}^2 S_{y}^2] &= \mathbb{E}[\varepsilon' A\varepsilon\varepsilon' A\varepsilon\varepsilon' C'\imath\imath' C\varepsilon] \\
&= 2Tr(C'\imath\imath' C)Tr(AA) + 8Tr(AAC'\imath\imath' C) \\
&\underset{0}{=} 2\theta^2 T\tfrac{1}{2}\tau^2 T + 8\theta^4 T \underset{1}{=} \theta^2\tau^2 T^2, && \text{(C.12)}\\
\mathbb{E}[S_{\varepsilon}S_{y\varepsilon}S_{y}S_{yy}] &= E[\varepsilon'\imath\imath' C\varepsilon\varepsilon' A\varepsilon\varepsilon' C' C\varepsilon] \\
&= 2[Tr(\imath\imath' C)Tr(AC'C) + Tr(C'C)Tr(\imath\imath' CA)] + 8Tr(\imath\imath' CAC'A) \\
&\underset{0}{=} 2[\theta T\tau^4\rho T + \tau^2 T\theta^2 T] + 8\theta^4 T \underset{1}{=} (2\theta^2\tau^2 + 2\rho\theta\tau^4)T^2, && \text{(C.13)}\\
\mathbb{E}[S_{\varepsilon}S_{y}S_{y\varepsilon}] &= E[\varepsilon'\imath\imath' C\varepsilon\varepsilon' A\varepsilon] = 2Tr(\imath\imath' CA) \underset{0}{=} 2\theta^2 T. && \text{(C.14)}
\end{aligned}$$

For approximating $D_1^\tau$ and $D_3^\tau$, we the following relationships are useful

$$\begin{aligned}
\mathbb{E}[S_{\tau\varepsilon}^2 S_{\tau y}^2 S_{\chi}] &\underset{5}{=} \frac{1}{9}\mathbb{E}[S_{\varepsilon}^2 S_{y}^2 S_{\chi}]T^4, && \text{(C.15a)}\\
\mathbb{E}[S_{\varepsilon}S_{\tau\varepsilon}S_{\tau y}^2 S_{\chi}] &\underset{4}{=} \mathbb{E}[S_{\tau\varepsilon}^2 S_{\tau y}S_{y}S_{\chi}] \underset{4}{=} \frac{1}{6}\mathbb{E}[S_{\varepsilon}^2 S_{y}^2 S_{\chi}]T^3, && \text{(C.15b)}\\
\mathbb{E}[S_{\varepsilon}^2 S_{\tau y}^2 S_{\chi}] &\underset{3}{=} \mathbb{E}[S_{\tau\varepsilon}^2 S_{y}^2 S_{\chi}] \underset{3}{=} \mathbb{E}[S_{\varepsilon}S_{\tau\varepsilon}S_{\tau y}S_{y}S_{\chi}] \underset{3}{=} \frac{5}{18}\mathbb{E}[S_{\varepsilon}^2 S_{y}^2 S_{\chi}]T^2, && \text{(C.15c)}\\
\mathbb{E}[S_{\varepsilon}^2 S_{\tau y}S_{y}S_{\chi}] &\underset{2}{=} \mathbb{E}[S_{\varepsilon}S_{\tau\varepsilon}S_{y}^2 S_{\chi}] \underset{2}{=} \frac{1}{2}\mathbb{E}[S_{\varepsilon}^2 S_{y}^2 S_{\chi}]T, && \text{(C.15d)}\\
\mathbb{E}[S_{\tau\varepsilon}S_{\tau y}S_{y\varepsilon}S_{\chi}] &\underset{2}{=} \frac{1}{3}\mathbb{E}[S_{\varepsilon}S_{y}S_{y\varepsilon}S_{\chi}]T^2, && \text{(C.15e)}\\
\mathbb{E}[S_{\varepsilon}S_{\tau y}S_{y\varepsilon}S_{\chi}] &\underset{1}{=} \mathbb{E}[S_{\tau\varepsilon}S_{y}S_{y\varepsilon}S_{\chi}] \underset{1}{=} \frac{1}{2}\mathbb{E}[S_{\varepsilon}S_{y}S_{y\varepsilon}S_{\chi}]T, && \text{(C.15f)}
\end{aligned}$$

where $S_\chi \in \{1, \tau^{-2}T^{-1}S_{yy}\}$. The coefficients $\frac{1}{2}, \frac{1}{3}, \frac{1}{6} = \frac{1}{2}\frac{1}{3}$ and $\frac{1}{9} = \left(\frac{1}{3}\right)^2$ are recognizable as (combinations of) the leading coefficients in the approximations $\sum t \approx \frac{1}{2}T^2$ and $\sum t^2 \approx \frac{1}{3}T^3$. The coefficient $\frac{5}{18}$ in (C.15c), however, is somewhat unexpected. Hence, we shall derive this relationship explicitly.

$$\begin{aligned}
\mathbb{E}[S_{\varepsilon}^2 S_{\tau y}^2] &= \mathbb{E}[\varepsilon'\imath\imath'\varepsilon\varepsilon' C'\pi\pi' C\varepsilon] \\
&= tr(\imath\imath')tr(C'\pi\pi' C) + 2tr(\imath\imath' C'\pi\pi' C) \\
&\underset{3}{=} T\cdot\frac{1}{3}\theta^2 T^3 + 2\cdot\frac{1}{4}\theta^2 T^4 = \frac{5}{6}\theta^2 T^4,
\end{aligned}$$

where $\pi' = (1, 2, ..., T)$. Due to (C.10), we thus have that $\mathbb{E}[S_\varepsilon^2 S_{\tau y}^2] = \frac{5}{3\cdot 6}\mathbb{E}[S_\varepsilon^2 S_y^2]T^2 + O(T^3)$. The following additional relationships are useful for approximating $D_3^\tau$

$$\mathbb{E}[S_{\tau y}^2 S_{y\varepsilon}^2] \underset{3}{=} \frac{1}{3}\mathbb{E}[S_y^2 S_{y\varepsilon}^2]T^2, \tag{C.16a}$$

$$\mathbb{E}[S_{\tau y} S_y S_{y\varepsilon}^2] \underset{2}{=} \frac{1}{2}\mathbb{E}[S_y^2 S_{y\varepsilon}^2]T. \tag{C.16b}$$

Figure 1: The Bartlett correction factor for the AR(1) model with intercept (and trend).

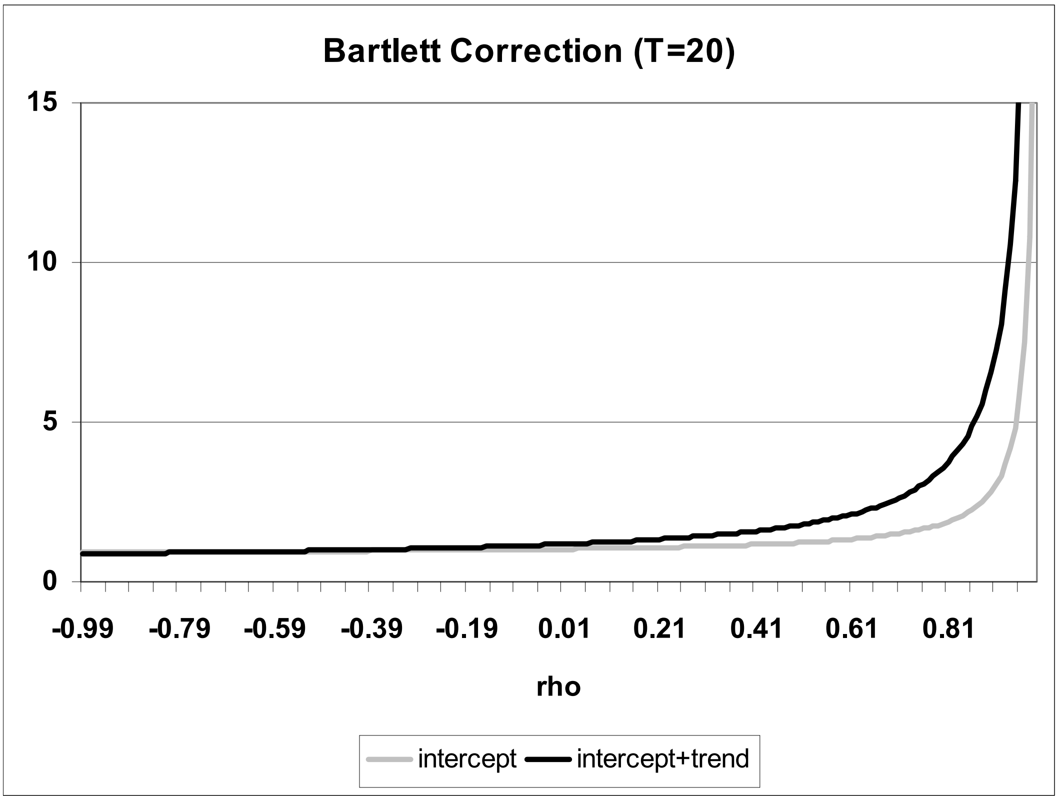

Figure 2: Rejection frequencies in the pure AR(1) model: $y_t = \rho y_{t-1} + \varepsilon_t$. The (gray) LR line is based on the asymptotic $\chi^2$-distribution, while the (black) LR* line is based on the Bartlett corrected critical values.

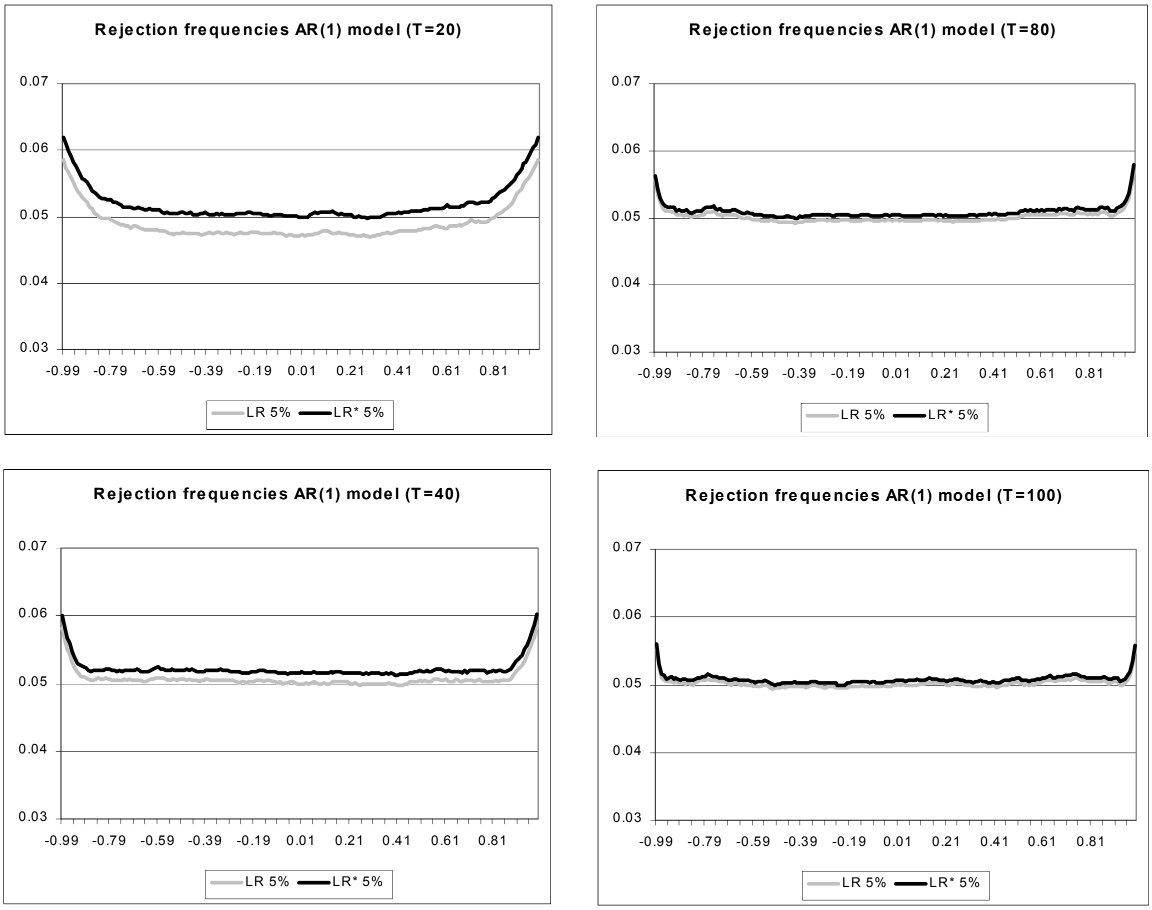

Figure 3: Rejection frequencies in the AR(1) model with intercept: $y_t = \rho y_{t-1} + \mu + \varepsilon_t$. The (gray) LR line is based on the asymptotic $\chi^2$-distribution, while the (black) LR* line is based on the Bartlett corrected critical values.

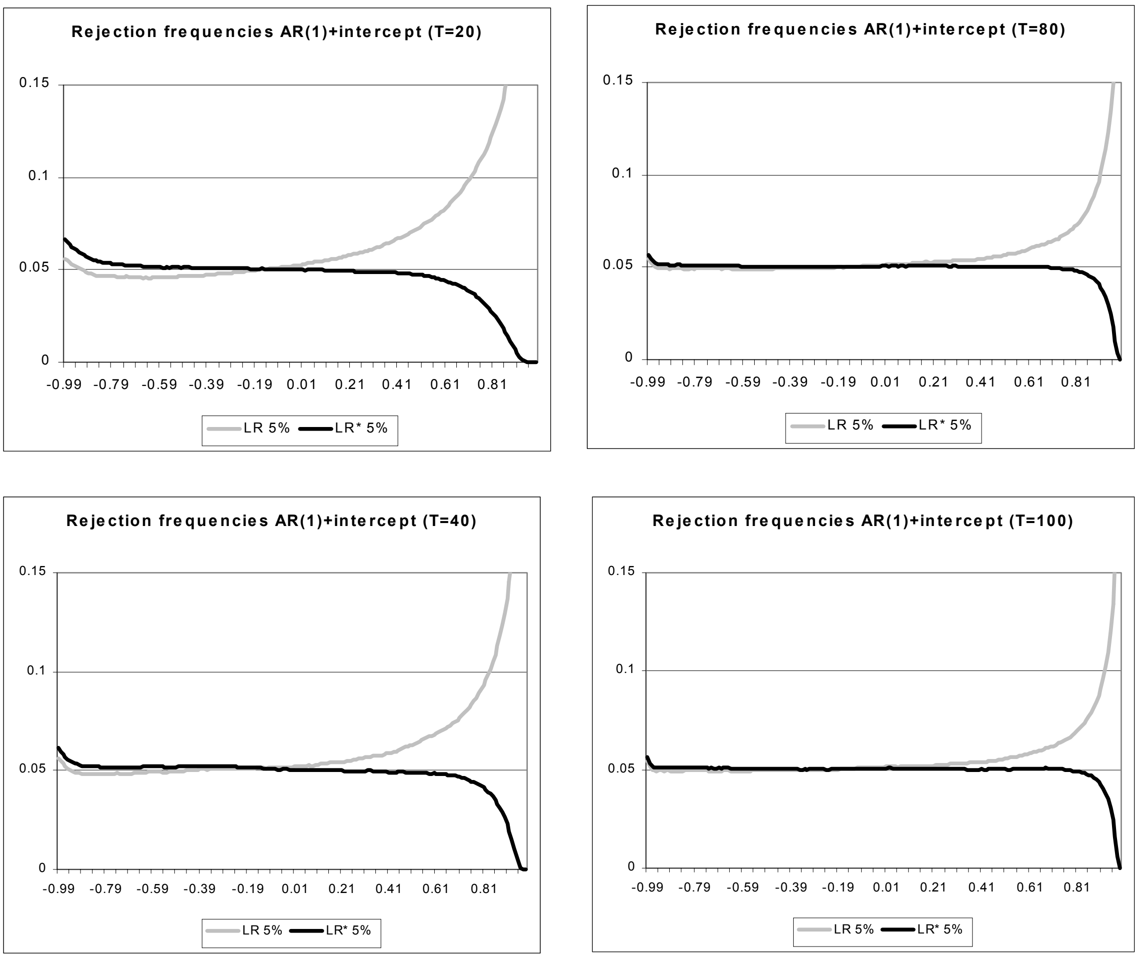

Figure 4: Rejection frequencies in the AR(1) model with intercept and trend: $y_t = \rho y_{t-1} + \mu + \beta t + \varepsilon_t$. The (gray) LR line is based on the asymptotic $\chi^2$-distribution, while the (black) LR* line is based on the Bartlett corrected critical values.