Bartlett Corrections in Cointegration Testing

Tor Jacobson and Rolf Larsson

First version: October 1996

Abstract

When testing for cointegration, the asymptotic inference typically in use can be plagued by size distortion due to an inadequate first order approximation. Hence, for practical purposes the inference can be completely misleading and result in false conclusions regarding the presence of long-run relationships in the data. Which, of course, in many applications is a key issue. We explore the potentials of Bartlett correction of two cointegration test statistics. The idea is to multiply the test statistic by a correcting factor derived from an asymptotic expansion of its expectation. As a consequence, the reference distribution should then provide a closer approximation to the resulting adjusted statistic in comparison with the unadjusted statistic. In a simple bivariate framework we derive a likelihood ratio test, as well as a first order approximation thereof, for testing the null hypothesis of no cointegration. Suitable Bartlett corrections for the two tests are suggested and using Monte Carlo simulation we evaluate the effectiveness of the proposed methods.

Key words: Bartlett correction; Cointegration testing; Monte Carlo simulation
JEL classification: C12; C15

[0] Correspondence to: Rolf Larsson, Department of Statistics, University of Stockholm, S-106 91 Stockholm, Sweden, voice +46 8 16 29 93, e-mail rolf.larsson@stat.su.se.

Tor Jacobson, Department of Economic Statistics, Stockholm School of Economics, Box 6501, S-113 83 Stockholm, Sweden, voice +46 8 736 92 41, e-mail sttj@hhs.se.

We would like to thank Karim Abadir, Herman van Dijk, Sune Karlsson, Helmuth Lütkepohl and Anders Warne for valuable comments and discussion. All remaining errors are the authors responsibility. Financial support from *Stiftelsen Riksbankens Jubileumsfond* and *Jan Wallanders och Tom Hedelius' stiftelse för samhällsvetenskaplig forskning* is gratefully acknowledged.

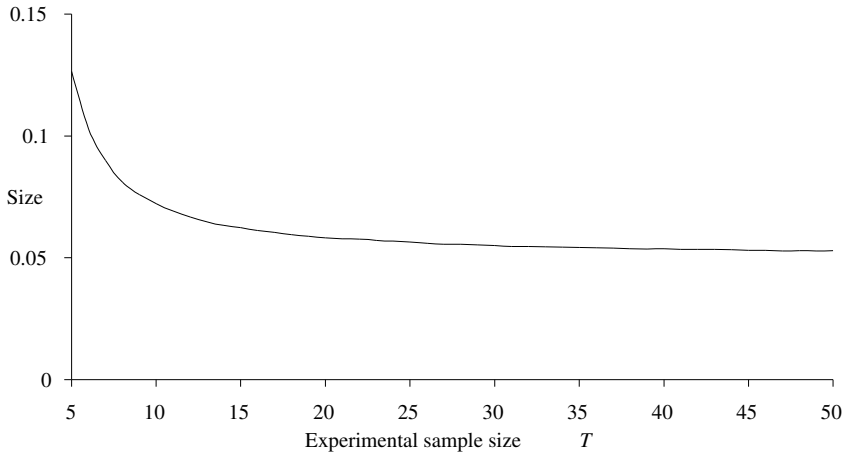
1. Introduction

It is probably fair to say that testing for cointegration, along with the related problem of testing for the presence of unit roots, has been one of the more frequent exercises performed in applied time-series econometrics, or even applied econometrics for that matter, since the introduction of cointegration analysis in the early 1980s. In this paper we are concerned with small sample properties of cointegration tests. Available tests are generally asymptotic in the sense that they rely on the convergence of the distribution for the test statistic to some known limit distribution as the sample size goes to infinity. However, Monte Carlo simulation has demonstrated the inadequacy of such approximations for finite samples, i.e. considerable size distortion has typically been found, see e.g. Banerjee, Dolado, Hendry and Smith (1986), Cheung and Lai (1993) and Jacobson (1995). Figure 1.1 below provides an illustration of the size problem for the likelihood ratio cointegration test subsequently considered in the paper. We see that for a small number of observations the empirical test size is more than twice the desired nominal size. As the number of observations increase the problem is gradually reduced. Hence, for practical purposes asymptotic inference in small samples can be completely misleading and result in false conclusions regarding the presence of long-run relationships in the data. Which, of course, is a key issue in many applications.

One could describe the problem as one of lacking coherence between the test statistic and its reference distribution and there are, in principle, two distinct routes to alleviate the problem; either for given test statistic correct the reference distribution, or, for given reference distribution correct the statistic in use. (A third possibility, somewhat less serious and certainly less appealing, is to postpone the analysis in wait for a sufficient sample size to accrue!!)

Thus, in the former case we would consider a corrected distribution for the test statistic at hand, that is replace the critical values of the limit distribution with such that will generate an actual test size closer to the nominal one. Analytically this amounts to Edgeworth expansions, or related techniques, of the distribution function, see Barndorrf-Nielsen and Cox (1989) or Hall (1992) for overviews. Bootstrap hypothesis testing is a plausible numerical alternative, which in fact can be expressed and interpreted in terms of Edgeworth expansions as shown

Figure 1.1: Monte Carlo estimated empirical sizes for a 5% nominal sized likelihood ratio cointegration test, $T = 5,...,50$.



by Hall. Although the consistency of bootstrapping in the unit root context is still unclear, Harris (1992) has evaluated bootstrapping of Dickey-Fuller unit root tests and Giersbergen (1996) has recently presented promising results for the multivariate maximum likelihood trace test for cointegrating rank. Larsson (1995) uses saddlepoint techniques to approximate small sample corrections of the lower tails of the distributions for some unit root test statistics.

The second route involves correction of the test statistic so as to make the known limit distribution an acceptable approximation. This is the route we will adhere to in the current study and we will, specifically, evaluate what is widely known as Bartlett correction applied to two tests of the null hypothesis of no cointegration against a cointegrating alternative: (i) a likelihood ratio test and (ii) a closely related test which is given by a Taylor expansion of the likelihood ratio test. However, due to the intricate analysis involved we do *not* treat the general case of a p-dimensional system allowing for a cointegrating rank $r, 1 \leq r < p$, as suggested by Johansen (1988,1991). Rather, the paper can be seen as an extension of the univariate analysis in Larsson (1994), where Bartlett corrections for unit root test statistics are obtained, to a restricted bivariate analysis. Thus, we derive and analyse the likelihood ratio test in a simple two dimensional data

generating process (DGP) which has been analyzed by i.a. Banerjee et al. (1986), Kremers, Ericsson, and Dolado (1992) and Campos, Ericsson, and Hendry (1996). In particular, Kremers et al. (1992) discuss the properties of two approaches to cointegration testing; the static residual-based test due to Engle and Granger (1987) and the dynamic error correction model t-test suggested by Banerjee et al. (1986).

We will close this section by a brief presentation of the idea underlying Bartlett correction. The next section covers maximum likelihood estimation and testing; test statistics and asymptotic distributions. In section 3 we present results on Bartlett corrections for the likelihood ratio and the approximated cointegration tests. A Monte Carlo experiment is performed in section 4 to evaluate the effectiveness of correcting the two tests. We will also present numerical results of the methods proposed in section 3, as well as comparisons with simulation based estimates of quantiles and expectations for the cointegration test.

Bartlett's (1937) pioneering paper conveyed the following idea. Consider a test statistic C_T that converges to C_∞, with error terms of orders T^{-1}, say, or smaller. C_∞ has a known distribution which provides the critical values for the asymptotic test. Now, we would like to obtain a transformed test statistic C_T^*, such that C_T^* converges to C_∞, and only with error terms of order T^{-2} or less at play. In other words, we want a correction of C_T which eliminates the influence of error terms of orders T^{-1}. Such a correction could be based on the expectation of C_T, recognizing that $\frac{C_T}{EC_T}$ tends to $\frac{C_\infty}{EC_\infty}$ as $T \to \infty$, and hence $C_T \approx EC_T \frac{C_\infty}{EC_\infty}$. Since EC_T is not known in general, the method hinges on the existence of a series expansion under the null hypothesis

$$EC_T = EC_\infty + \frac{R}{T} + O\left(\frac{1}{T^2}\right),$$

where R is either known or can be consistently estimated. Ignoring the $O(T^{-2})$-term, we get

$$C_T \approx \left(EC_\infty + \frac{R}{T}\right)\frac{C_\infty}{EC_\infty},$$

and the Bartlett corrected test statistic is

$$C_T^* = C_T\left(1 + \frac{R_0}{T}\right)^{-1},$$

where $R_0 = \frac{R}{EC_\infty}$. Under mild regularity restrictions, C_∞ is typically χ^2-distributed for the likelihood ratio test statistic and in such "standard" cases the correction

has been shown to correct not only the first, but also higher moments and fractiles (cf. Jensen (1993) and references therein). However, ours is not a "standard" case, and in fact, Jensen and Wood (1995) have shown that in the AR(1) case Bartlett correction only corrects the first moment. It is very likely, judging from Jensen and Wood, that the optimal correction for higher moments and fractiles in the cointegration case is not Bartlett but some other. Nevertheless, according to the Monte Carlo results in Nielsen (1994) for a multivariate cointegrated DGP, a simulation based Bartlett correction (i.e. EC_T above is estimated by use of simulation) improves the asymptotic approximation considerably. Moreover, the method is relatively easy to implement, i.e. in relation to other approaches. For a slightly different application of a similar Bartlett-type idea, cf. Lucas (1996).

2. Maximum likelihood estimation and testing; test statistics and asymptotic distributions

In this section we will begin by presenting the framework for our analytical and simulation based results, i.e. a first-order bivariate data generating process. Next we consider the log likelihood function and the implied maximum likelihood estimators, and, subsequently, likelihood ratio testing for cointegration.

2.1. The data generating process

The DGP in (2.1) has been frequently analyzed in the literature in various, slightly modified forms. Recently Kremers et al. (1992) have used it for investigation of the power functions for single-equation cointegration tests, and Campos et al. (1996) evaluate cointegration tests in presence of structural shifts. In its error correction representation the process can be written

$$\begin{cases} \Delta y_t = a\Delta z_t + b\left(y_{t-1} - \lambda z_{t-1}\right) + \varepsilon_t, \\ \Delta z_t = u_{t,} \end{cases} \tag{2.1}$$

where $\{\varepsilon_t\}$ and $\{u_t\}, t = 1, 2, \ldots$ are independent sequences of independent random normal variables with mean 0 and variances σ_ε^2 and σ_u^2, respectively, and $y_0 = z_0 = 0$. The parameters a and b are unknown with the restrictions $0 \leq a \leq 1$ and $-1 \leq b \leq 0$, whereas λ is assumed to be known. Both variables are integrated of order one and when the error correction coefficient b is strictly less than zero the two are cointegrated with cointegrating vector $[1 \quad -\lambda]$. As pointed out by Campos et al. (1996), a and λ can be interpreted as short- and long-run elasticities, with a typically smaller than λ.

2.2. The maximum likelihood estimators

Having taken observations up to time T and defining $x_t \overset{\text{def}}{=} y_t - \lambda z_t, \forall t$, the log likelihood function $l\left(a, b, \sigma_\varepsilon^2, \sigma_u^2\right)$ can be written

$$\begin{aligned} l\left(a, b, \sigma_\varepsilon^2, \sigma_u^2\right) &= -\frac{T}{2}\log\left(2\pi\right) - \frac{T}{2}\left(\log\sigma_\varepsilon^2 + \log\sigma_u^2\right) - \\ &\quad -\frac{1}{2\sigma_\varepsilon^2}\sum\left(\Delta y_t - a\Delta z_t - b x_{t-1}\right)^2 - \frac{1}{2\sigma_u^2}\sum\left(\Delta z_t\right)^2, \end{aligned} \tag{2.2}$$

where the summation, here and henceforth, goes from $t = 1$ to T unless otherwise is explicit. Differentiating in (2.2), we get the following relations for the maximum likelihood estimates $\widehat{a}, \widehat{b}, \widehat{\sigma}_{\varepsilon}^2$ and $\widehat{\sigma}_u^2$, where

$$\widehat{a}_0 \overset{\text{def}}{=} \frac{\sum \Delta z_t \Delta y_t}{\sum (\Delta z_t)^2} \tag{2.3}$$

is the maximum likelihood estimator of a if $b = 0$:

$$\widehat{a} = \widehat{a}_0 - \widehat{b} \frac{\sum x_{t-1} \Delta z_t}{\sum (\Delta z_t)^2}, \tag{2.4}$$

$$\widehat{b} = \frac{\sum x_{t-1} (\Delta y_t - \widehat{a} \Delta z_t)}{\sum x_{t-1}^2}, \tag{2.5}$$

$$\widehat{\sigma}_{\varepsilon}^2 = \frac{1}{T} \sum \left(\Delta y_t - \widehat{a} \Delta z_t - \widehat{b} x_{t-1} \right)^2, \tag{2.6}$$

$$\widehat{\sigma}_u^2 = \frac{1}{T} \sum (\Delta z_t)^2 . \tag{2.7}$$

Thus, from (2.4) and (2.5), we find the estimator of the error correction coefficient b in (2.2) as given by:[1]

$$\widehat{b} = \frac{\sum x_{t-1} \Delta y_t \sum (\Delta z_t)^2 - \sum \Delta z_t \Delta y_t \sum x_{t-1} \Delta z_t}{\sum x_{t-1}^2 \sum (\Delta z_t)^2 - (\sum x_{t-1} \Delta z_t)^2}. \tag{2.8}$$

2.3. Likelihood ratio testing

Next we consider the likelihood ratio test for cointegration in (2.2), i.e. a test of the null hypothesis $H_0 : b = 0$ (no cointegration) against the alternative $H_1 : b < 0$ (cointegration). For the sake of simplicity and with no loss of generality, we will from now and onwards assume unit variances for the error processes, i.e. $\sigma_{\varepsilon}^2 = \sigma_u^2 = 1$. Let us first define the following three random walks: $S_{t,1} \overset{\text{def}}{=} \sum_{i=1}^{t} \varepsilon_i$, $S_{t,2} \overset{\text{def}}{=} \sum_{i=1}^{t} u_i$ and $S_{t,3} \overset{\text{def}}{=} \sum_{i=1}^{t} v_i$, where $v_t \overset{\text{def}}{=} (a - \lambda) u_t + \varepsilon_t, t = 1, \ldots, T$, and the corresponding Wiener processes $W_{t,k}, k = 1, 2, 3$, such that $T^{-1/2} S_{t,k}$ converges to $W_{t',k}$ as $T \to \infty, t' = t/T$. The theory of weak convergence (cf. Billingsley, 1968) implies that, under H_0,

$$\frac{1}{T} \sum x_{t-1} \Delta y_t = \frac{1}{T} \sum S_{t-1,3} (a u_t - \varepsilon_t) \xrightarrow{\text{d}} a \int_0^1 W_{t,3} dW_{t,2} + \int_0^1 W_{t,3} dW_{t,1}, \tag{2.9}$$

[1] For a different presentation of these results, see Banerjee and Hendry (1992).

$$\frac{1}{T}\sum(\Delta z_t)^2 = \frac{1}{T}\sum u_t^2 \xrightarrow{\mathrm{d}} 1, \tag{2.10}$$

$$\frac{1}{T}\sum \Delta z_t \Delta y_t = \frac{1}{T}\left(a\sum u_t^2 + \sum \varepsilon_t u_t\right) \xrightarrow{\mathrm{d}} a \tag{2.11}$$

$$\frac{1}{T}\sum x_{t-1}\Delta z_t = \frac{1}{T}\sum S_{t-1,3}u_t \xrightarrow{\mathrm{d}} \int_0^1 W_{t,3}dW_{t,2}, \text{ and finally} \tag{2.12}$$

$$\frac{1}{T^2}\sum x_{t-1}^2 = \frac{1}{T^2}\sum S_{t-1,3}^2 \xrightarrow{\mathrm{d}} \int_0^1 W_{t,3}^2 dt \tag{2.13}$$

as $T \to \infty$, where $\xrightarrow{\mathrm{d}}$ signifies weak convergence. Hence, equations (2.9)-(2.13) in conjunction with (2.8) and some simplification give the following convergence result:[2]

$$T\widehat{b} = \frac{\frac{1}{T}\sum S_{t-1,3}\varepsilon_t}{\frac{1}{T^2}\sum S_{t-1,3}^2} + o_p(1) \xrightarrow{\mathrm{d}} \frac{\int_0^1 W_{t,3}dW_{t,1}}{\int_0^1 W_{t,3}^2 dt} \quad \text{as} \quad T \to \infty \tag{2.14}$$

Now, turning to the likelihood ratio test, Q_T say, it follows that

$$Q_T = \frac{\left(\widehat{\sigma}_{\varepsilon,0}^2\right)^{-\frac{T}{2}} e^{\left\{-\frac{1}{2\widehat{\sigma}_{\varepsilon,0}^2}\sum\left(\Delta y_t - \widehat{a}_0\Delta z_t\right)^2 - \frac{1}{2\widehat{\sigma}_u^2}\sum(\Delta z_t)^2\right\}}}{\left(\widehat{\sigma}_{\varepsilon}^2\right)^{-\frac{T}{2}} e^{\left\{-\frac{1}{2\widehat{\sigma}_{\varepsilon}^2}\sum\left(\Delta y_t - \widehat{a}\Delta z_t - \widehat{b}x_{t-1}\right)^2 - \frac{1}{2\widehat{\sigma}_u^2}\sum(\Delta z_t)^2\right\}}} = \left(\frac{\widehat{\sigma}_{\varepsilon,0}^2}{\widehat{\sigma}_{\varepsilon}^2}\right)^{-\frac{T}{2}},$$

and so

$$-2\log Q_T = -T\log\left(\frac{\widehat{\sigma}_{\varepsilon}^2}{\widehat{\sigma}_{\varepsilon,0}^2}\right), \tag{2.15}$$

where $\widehat{\sigma}_{\varepsilon,0}^2 = T^{-1}\sum(\Delta y_t - \widehat{a}_0\Delta z_t)^2$ is the maximum likelihood estimate of σ_ε^2 under the null hypothesis. It follows from (2.4)-(2.6) that

$$T\widehat{\sigma}_{\varepsilon}^2 = \sum(\Delta y_t - \widehat{a}\Delta z_t)^2 - \frac{\left(\sum x_{t-1}\left(\Delta y_t - \widehat{a}\Delta z_t\right)\right)^2}{\sum x_{t-1}^2}$$

[2] c.f. Billingsley (1968) and the corresponding result for t statistics in Banerjee and Hendry (1992). The notation $X_t = o_p(T^\alpha)$, where $\{X_t\}$ is a sequence of random variables, is taken to mean that $T^{-\alpha}X_t$ tends to zero in probability as $T \to \infty$. Replacing o_p with O_p changes the interpretation to: $T^{-\alpha}X_t$ tends to a non-degenerate random variable as $T \to \infty$.

and furthermore, by (2.4) and (2.5),

$$\sum \left(\Delta y_t - \widehat{a}\Delta z_t\right)^2 = \sum \left(\left(\Delta y_t - \widehat{a}_0 \Delta z_t\right) + \widehat{b}\frac{\sum x_{t-1\Delta z_t}}{\sum \left(\Delta z_t\right)^2}\Delta z_t \right)^2 =$$
$$= \sum \left(\Delta y_t - \widehat{a}_0 \Delta z_t\right)^2 + \widehat{b}^2 \frac{\left(\sum x_{t-1}\Delta z_t\right)^2}{\sum \left(\Delta z_t\right)^2},$$

and so,

$$\frac{\widehat{\sigma}_{\varepsilon}^2}{\widehat{\sigma}_{\varepsilon,0}^2} = 1 - \frac{M_T}{T} \tag{2.16}$$

where

$$M_T = T\left(\frac{\left(\sum x_{t-1}\left(\Delta y_t - \widehat{a}\Delta z_t\right)\right)^2}{\sum \left(\Delta y_t - \widehat{a}_0 \Delta z_t\right)^2 \sum x_{t-1}^2} - \widehat{b}^2 \frac{\left(\sum x_{t-1}\Delta z_t\right)^2}{\sum \left(\Delta y_t - \widehat{a}_0 \Delta z_t\right)^2 \sum \left(\Delta z_t\right)^2} \right), \tag{2.17}$$

a quantity which we will refer to as the approximated test statistic. By Taylor expansion (we subsequently show that $M_T = O_p\left(1\right)$), (2.15) and (2.16) yield

$$-2\log Q_T = M_T + \frac{M_T^2}{2T} + O_p\left(T^{-2}\right) \tag{2.18}$$

The lemma below gives representations of the approximated and the likelihood ratio test statistics.

Lemma 2.1. *Under $H_0 : b = 0$ (no cointegration) in (2.1),*

$$M_T = Z_T + \frac{1}{T}R_T + O_p\left(T^{-2}\right), \tag{2.19}$$

where

$$Z_T = T\frac{\left(\sum S_{t-1,3}\varepsilon_t\right)^2}{\sum \varepsilon_t^2 \sum S_{t-1,3}^2}, \tag{2.20}$$

and

$$R_T = Z_T - 2R_T^{(1)} + R_T^{(2)}, \tag{2.21}$$

with

$$R_T^{(1)} \overset{\text{def}}{=} \frac{\sum S_{t-1,3}\varepsilon_t \sum S_{t-1,3}u_t}{\sum S_{t-1,3}^2} \sum \varepsilon_t u_t \qquad \textit{and} \tag{2.22}$$

$$R_T^{(2)} \overset{\text{def}}{=} \frac{\left(\sum S_{t-1,3}\varepsilon_t\right)^2 \left(\sum S_{t-1,3}u_t\right)^2}{\left(\sum S_{t-1,3}^2\right)^2}. \tag{2.23}$$

Furthermore,

$$-2\log Q_T = Z_T + \frac{1}{T}\left(R_T + \frac{1}{2}Z_T^2\right) + o_p\left(T^{-1}\right) \tag{2.24}$$

The quantities $Z_T, R_T^{(1)}$ *and* $R_T^{(2)}$ *are all* $O_p(1)$. □

Remark 1: It can be shown that, if a is assumed known, the approximated test statistic actually equals Z_T given in (2.20). □

Remark 2: $R_T^{(1)}$ in (2.22) may appear to be of order $\sqrt{T}$. However, it can be shown that $\sum \varepsilon_t u_t$ is asymptotically uncorrelated with $\sum S_{t-1,3}\varepsilon_t, \sum S_{t-1,3}u_t$ and $\sum S_{t-1,3}^2$, implying that $R_T^{(1)}$ is of a smaller order (cf. Lemma 4.2 of Larsson (1994) and Theorem 3.1 below). □

Proof: We have from (2.4) that

$$\sum x_{t-1}\left(\Delta y_t - \widehat{a}\Delta z_t\right) = \sum x_{t-1}\Delta y_t - \widehat{a}_0 \sum x_{t-1}\Delta z_t + \widehat{b}\frac{\left(\sum x_{t-1}\Delta z_t\right)^2}{\sum \left(\Delta z_t\right)^2},$$

and since $\Delta y_t = au_t + \varepsilon_t$, (2.3) and (2.9)-(2.12) yield

$$\frac{1}{T}\left(\sum x_{t-1}\Delta y_t - \widehat{a}_0 \sum x_{t-1}\Delta z_t\right) =$$

$$= \frac{1}{T}\sum S_{t-1,3}\left(au_t + \varepsilon_t\right) - \frac{\frac{1}{T}\left(a\sum u_t^2 + \sum \varepsilon_t u_t\right)}{\frac{1}{T}\sum u_t^2}\frac{1}{T}\sum S_{t-1,3}u_t =$$

$$= \frac{1}{T}\sum S_{t-1,3}\varepsilon_t - \frac{1}{T}\frac{\sum \varepsilon_t u_t}{\sum u_t^2}\sum S_{t-1,3}u_t.$$

(Note that since $\sum \varepsilon_t u_t$ has mean 0 and variance T, it follows from (2.10) that $\sum \varepsilon_t u_t / \sum u_t^2$ is $o_p(1)$.) Hence, from (2.10), (2.12) and (2.14),

$$\left(\frac{1}{T}\sum x_{t-1}\left(\Delta y_t - \widehat{a}\Delta z_t\right)\right)^2 = \tag{2.25}$$

$$= \left(\frac{1}{T}\sum S_{t-1,3}\varepsilon_t - \frac{1}{T}\frac{\sum \varepsilon_t u_t}{\sum u_t^2}\sum S_{t-1,3}u_t + \frac{1}{T}\frac{\frac{1}{T}\sum S_{t-1,3}\varepsilon_t}{\frac{1}{T^2}\sum S_{t-1,3}^2}\frac{\left(\frac{1}{T}\sum S_{t-1,3}u_t\right)^2}{\frac{1}{T}\sum u_t^2}\left(1+o_p(1)\right)\right)^2 =$$

$$= \left(\frac{1}{T}\sum S_{t-1,3}\varepsilon_t\right)^2 + \frac{1}{T}r_T + o_p\left(T^{-1}\right),$$

where

$$r_T \overset{\text{def}}{=} 2\frac{\sum S_{t-1,3}\varepsilon_t}{\sum u_t^2}\left(\frac{1}{T}\sum S_{t-1,3}u_t\right)\left(-\sum \varepsilon_t u_t + \frac{\sum S_{t-1,3}\varepsilon_t \sum S_{t-1,3}u_t}{\sum S_{t-1,3}^2}\right), \tag{2.26}$$

which is $O_p(1)$. Furthermore, by (2.3),

$$\sum (\Delta y_t - \widehat{a}_0 \Delta z_t)^2 = \sum (\Delta y_t)^2 - \frac{\left(\sum \Delta z_t \Delta y_t\right)^2}{\sum (\Delta z_t)^2},$$

and so, since

$$\sum (\Delta y_t)^2 = a^2 \sum u_t^2 + 2a \sum \varepsilon_t u_t + \sum \varepsilon_t^2,$$

(2.10) and (2.11) imply, after simplification

$$\frac{1}{T} \sum (\Delta y_t - \widehat{a}_0 \Delta z_t)^2 = \frac{1}{T} \sum \varepsilon_t^2 - \frac{1}{T} \frac{\left(\sum \varepsilon_t u_t\right)^2}{\sum u_t^2}, \tag{2.27}$$

where it should be observed that

$$\frac{1}{T} \left(\sum \varepsilon_t u_t\right)^2 \xrightarrow{\mathrm{d}} 1 \quad \text{as } T \to \infty, \tag{2.28}$$

so by (2.10), the second term in the r.h.s. of (2.27) is $T^{-1} + O_p\left(T^{-2}\right)$. Thus, plugging in (2.10), (2.13), (2.14), (2.25), (2.27) and (2.28) into (2.17), we get

$$\begin{aligned} M_T &= \left(\frac{1}{T} \sum \varepsilon_t^2 - \frac{1}{T}\right)^{-1} \times \\ &\times \left(\frac{\left(\frac{1}{T} \sum S_{t-1,3} \varepsilon_t\right)^2 + \frac{1}{T} r_T}{\frac{1}{T^2} \sum S_{t-1,3}^2} - \frac{1}{T} \frac{\left(\frac{1}{T} \sum S_{t-1,3} \varepsilon_t\right)^2}{\left(\frac{1}{T^2} \sum S_{t-1,3}^2\right)^2} \frac{\left(\frac{1}{T} \sum S_{t-1,3} u_t\right)^2}{\frac{1}{T} \sum u_t^2} \right) + o_p\left(T^{-1}\right), \end{aligned}$$

and so by Taylor expansion,

$$M_T = Z_T + \frac{1}{T} R_T + o_p\left(T^{-1}\right),$$

with Z_T as in (2.20), which clearly is $O_p(1)$, and

$$\begin{aligned} R_T &= Z_T + \frac{r_T}{\left(\frac{1}{T} \sum \varepsilon_t^2\right)\left(\frac{1}{T^2} \sum S_{t-1,3}^2\right)} - \frac{\left(\frac{1}{T} \sum S_{t-1,3} \varepsilon_t\right)^2}{\left(\frac{1}{T} \sum \varepsilon_t^2\right)\left(\frac{1}{T^2} \sum S_{t-1,3}^2\right)^2} \frac{\left(\frac{1}{T} \sum S_{t-1,3} u_t\right)^2}{\frac{1}{T} \sum u_t^2} = \\ &= Z_T - 2R_T^{(1)} + R_T^{(2)}, \end{aligned}$$

which is (2.21), with $Z_T, R_T^{(1)}$ and $R_T^{(2)}$ as in (2.22) and (2.23), because of (2.10) and the fact that $T^{-1} \sum \varepsilon_t^2 \xrightarrow{\mathrm{d}} 1$ as $T \to \infty$, and obviously all these terms are $O_p(1)$. Finally, (2.24) follows directly from (2.18) and (2.19), which completes the proof of the lemma. □

3. Bartlett corrections

In order to find the Bartlett corrections for the approximated and the likelihood ratio test statistics it is necessary to calculate the asymptotic expectation of Z_T to the second order, i.e. including a T^{-1} term, as well as first order asymptotic expectations of $Z_T, Z_T^2, R_T^{(1)}$ and $R_T^{(2)}$. Analytical representations of the expectations of interest are listed in the following theorem.

Theorem 3.1. *Using the Laplace transform*

$$\varphi(s,u,v,w,r) \stackrel{\text{def}}{=} E\left(e^{\left\{-s\sum S_{t-1,3}^2 - u\sum \varepsilon_t^2 - v\sum S_{t-1,3}\varepsilon_t - w\sum S_{t-1,3}u_t - r\sum \varepsilon_t u_t\right\}}\right),$$

we may represent the expectations as:

$$E(Z_T) = T\int_0^\infty\int_0^\infty \frac{\partial^2}{\partial v^2}\varphi(s,u,v,w,0)\,ds\,du \tag{3.1}$$

$$E\left(Z_T^2\right) = T^2\int_0^\infty\int_0^\infty su\frac{\partial^4}{\partial v^4}\varphi(s,u,v,w,0)\,ds\,du \tag{3.2}$$

$$E\left(R_T^{(1)}\right) = -\int_0^\infty \frac{\partial^3}{\partial v\partial u\partial r}\varphi(s,0,v,w,r)\,ds \tag{3.3}$$

$$E\left(R_T^{(2)}\right) = \int_0^\infty s\frac{\partial^4}{\partial v^2\partial w^2}\varphi(s,0,v,w,0)\,ds \tag{3.4}$$

where the derivatives are taken at $\{v=0, w=0, r=0\}$. *Furthermore, the derivatives are given by*

$$\frac{\partial^2}{\partial v^2}\varphi(s,u,v,w,r) = \frac{1}{\sqrt{\det P_0}}\left(\frac{1}{4}\operatorname{tr}^2\left(P_0^{-1}H_1\right) + \frac{1}{2}\operatorname{tr}\left(\left(P_0^{-1}H_1\right)^2\right)\right), \tag{3.5}$$

$$\begin{aligned}\frac{\partial^3}{\partial v\partial u\partial r}\varphi(s,0,v,w,r) = \frac{-1}{8\sqrt{\det P_0}}\Big\{&\operatorname{tr}\left(P_0^{-1}H_1\right)\operatorname{tr}\left(P_0^{-1}H_2\right)\operatorname{tr}\left(P_0^{-1}H_3\right) + \\ &+2\left(\operatorname{tr}\left(P_0^{-1}H_1\right)\operatorname{tr}\left(P_0^{-1}H_2P_0^{-1}H_3\right) + \operatorname{tr}\left(P_0^{-1}H_2\right)\operatorname{tr}\left(P_0^{-1}H_1P_0^{-1}H_3\right) + \right. \\ &\left. + \operatorname{tr}\left(P_0^{-1}H_3\right)\operatorname{tr}\left(P_0^{-1}H_1P_0^{-1}H_2\right)\right) + 8\operatorname{tr}\left(P_0^{-1}H_1P_0^{-1}H_2P_0^{-1}H_3\right)\Big\}\end{aligned} \tag{3.6}$$

$$\frac{\partial^4}{\partial v^4}\varphi(s,0,v,w,0) = \frac{1}{16\sqrt{\det P_0}}\left\{\mathrm{tr}^{\,4}\left(P_0^{-1}H_1\right)+32\,\mathrm{tr}\left(P_0^{-1}H_1\right)\mathrm{tr}\left(P_0^{-1}H_1\right)^3+\right.$$
$$\left.+12\,\mathrm{tr}^{\,2}\left(\left(P_0^{-1}H_1\right)^2\right)+12\,\mathrm{tr}^{\,2}\left(P_0^{-1}H_1\right)\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^2\right)+48\,\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^4\right)\right\}, \quad (3.7)$$

and

$$\frac{\partial^4}{\partial v^2\partial w^2}\varphi(s,0,v,w,0) = \frac{1}{16\sqrt{\det P_0}}\left\{\mathrm{tr}^{\,2}\left(P_0^{-1}H_1\right)\mathrm{tr}^{\,2}\left(P_0^{-1}H_2\right)+\right. \quad (3.8)$$
$$+16\,\mathrm{tr}\left(P_0^{-1}H_1\right)\mathrm{tr}\left(P_0^{-1}H_1\left(P_0^{-1}H_2\right)^2\right)+16\,\mathrm{tr}\left(P_0^{-1}H_2\right)\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^2P_0^{-1}H_2\right)+$$
$$+4\,\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^2\right)\mathrm{tr}\left(\left(P_0^{-1}H_2\right)^2\right)+8\,\mathrm{tr}^{\,2}\left(P_0^{-1}H_1P_0^{-1}H_2\right)+$$
$$+16\,\mathrm{tr}\left(\left(P_0^{-1}H_1P_0^{-1}H_2\right)^2\right)+2\,\mathrm{tr}^{\,2}\left(P_0^{-1}H_1\right)\mathrm{tr}\left(\left(P_0^{-1}H_2\right)^2\right)+$$
$$+2\,\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^2\right)\mathrm{tr}^{\,2}\left(P_0^{-1}H_2\right)+8\,\mathrm{tr}\left(P_0^{-1}H_1\right)\mathrm{tr}\left(P_0^{-1}H_2\right)\mathrm{tr}\left(P_0^{-1}H_1P_0^{-1}H_2\right)+$$
$$\left.+32\,\mathrm{tr}\left(\left(P_0^{-1}H_1\right)^2\left(P_0^{-1}H_2\right)^2\right)\right\},$$

where the $2T\times 2T$ *block matrix*

$$P_0 \stackrel{\mathrm{def}}{=} \begin{pmatrix} \overline{P}_{10} & 2s\,(a-\lambda)\,I_- \\ 2s\,(a-\lambda)\,I_- & P_{20} \end{pmatrix},$$

where the $T\times T$ *matrices* $\overline{P}_{10}=(1+2u)\,P_{10}$,

$$P_{i0} \stackrel{\mathrm{def}}{=} \begin{pmatrix} \alpha_i & -1 & 0 & \dots & & & 0 \\ -1 & \alpha_i & -1 & & & & \vdots \\ 0 & -1 & \alpha_i & & & & \\ \vdots & & & \ddots & & & \\ & & & & \alpha_i & -1 & 0 \\ & & & & -1 & \alpha_i & -1 \\ 0 & & & \dots & 0 & -1 & 1 \end{pmatrix}, \qquad i=1,2,$$

where $\alpha_1=2\,(1+s/\,(1+2u))\,,\alpha_2=2\left(1+(a-\lambda)^2\,s\right)$,

$$I_- \stackrel{\mathrm{def}}{=} \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \\ & & & 1 & 0 \\ & \dots & & 0 & 0 \end{pmatrix},$$

where the $2T \times 2T$ *block matrices*

$$H_1 \stackrel{\text{def}}{=} \begin{pmatrix} h_1 & (a-\lambda)\, h_2' \\ (a-\lambda)\, h_2 & 0 \end{pmatrix}, H_2 \stackrel{\text{def}}{=} \begin{pmatrix} 0 & h_2 \\ h_2' & (a-\lambda)\, h_1 \end{pmatrix}, H_3 \stackrel{\text{def}}{=} \begin{pmatrix} 0 & h_3 \\ h_3 & 0 \end{pmatrix},$$

where the $T \times T$ *matrices are*

$$h_1 \stackrel{\text{def}}{=} \begin{pmatrix} -2 & 1 & 0 & \dots & & & 0 \\ 1 & -2 & 1 & & & & \vdots \\ 0 & 1 & -2 & & & & \\ \vdots & & & \ddots & & & \\ & & & & -2 & 1 & 0 \\ & & & & 1 & -2 & 1 \\ 0 & & & \dots & 0 & 1 & 0 \end{pmatrix}, h_2 \stackrel{\text{def}}{=} \begin{pmatrix} -1 & 1 & 0 & \dots & & & 0 \\ 0 & -1 & 1 & & & & \vdots \\ 0 & 0 & -1 & & & & \\ \vdots & & & \ddots & & & \\ & & & & -1 & 1 & 0 \\ & & & & 0 & -1 & 1 \\ 0 & & & \dots & 0 & 0 & 0 \end{pmatrix},$$

$$h_3 \stackrel{\text{def}}{=} \begin{pmatrix} 2 & -1 & 0 & \dots & & & 0 \\ -1 & 2 & -1 & & & & \vdots \\ 0 & -1 & 2 & & & & \\ \vdots & & & \ddots & & & \\ & & & & 2 & -1 & 0 \\ & & & & -1 & 2 & -1 \\ 0 & & & \dots & 0 & -1 & 1 \end{pmatrix}.$$

□

Proof of Theorem 3.1: The proof uses the technique of Larsson (1994, 1997) and i.a. Evans and Savin (1981), Mikulski and Monsour (1994), Le Breton and Pham (1989) and Abadir (1993). Indeed, via (2.20), (3.1) and (3.2) follow from (3.5) and (3.7), respectively, by applying the identity

$$(xy)^{-n} = \int_0^\infty \int_0^\infty s^{n-1} u^{n-1} e^{-sx-uy} ds du, \qquad n = 1, 2,$$

setting $x = \sum S_{t-1,3}^2$, $y = \sum \varepsilon_t^2$, multiplying by $e^{-v \sum S_{t-1,3}^2 e_t - w \sum S_{t-1,3} u_t - r \sum e_t u_t}$, taking expectations and differentiating w.r.t. v. In a similar fashion, (3.3) and (3.4) follow from (2.22), (2.23), (3.6) and (3.8). Moreover

$$\varphi(s,u,v,w,r) = \int \dots \int (2\pi)^{-T} e^{-A} dx_{1,1} \dots dx_{T,1} dx_{1,2} \dots dx_{T,2}, \qquad (3.9)$$

where

$$\begin{aligned} A \stackrel{\text{def}}{=} & -s\sum\left(x_{t-1,1}+(a-\lambda)\,x_{t-1,2}\right)^2 - u\sum\left(x_{t,1}-x_{t-1,1}\right)^2 - \\ & -v\sum\left(x_{t-1,1}+(a-\lambda)\,x_{t-1,2}\right)\left(x_{t,1}-x_{t-1,1}\right) - \\ & -w\sum\left(x_{t-1,1}+(a-\lambda)\,x_{t-1,2}\right)\left(x_{t,2}-x_{t-1,2}\right) - \\ & -r\sum\left(x_{t,1}-x_{t-1,1}\right)\left(x_{t,2}-x_{t-1,2}\right) - \\ & -\frac{1}{2}\sum\left(x_{t,1}-x_{t-1,1}\right)^2 - \frac{1}{2}\sum\left(x_{t,2}-x_{t-1,2}\right)^2 \\ = & -\frac{1}{2}\underline{x}'P\underline{x}, \end{aligned}$$

where $\underline{x}' \stackrel{\text{def}}{=} (x_{1,1},\ldots,x_{T,1},x_{1,2},\ldots,x_{T,2})$ and P is a $2T\times 2T$ block matrix

$$P \stackrel{\text{def}}{=} \begin{pmatrix} P_1 & P_3 \\ P_3' & P_2 \end{pmatrix},$$

where P_1, P_2 and P_3 are $T\times T$ matrices

$$P_i \stackrel{\text{def}}{=} \begin{pmatrix} p_{i1} & p_{i2} & 0 & \ldots & & & 0 \\ p_{i3} & p_{i1} & p_{i2} & & & & \vdots \\ 0 & p_{i3} & p_{i1} & & & & \\ \vdots & & & \ddots & & & \\ & & & & p_{i1} & p_{i2} & 0 \\ & & & & p_{i3} & p_{i1} & p_{i2} \\ 0 & & & \ldots & 0 & p_{i3} & p_{i4} \end{pmatrix}, \qquad i=1,2,3,$$

where $p_{11}=2\,(1+s+2u-v)$, $p_{12}=p_{13}=-\,(1+2u-v)$, $p_{14}=1+2u$, $p_{21}=2\left(1+s\,(a-\lambda)^2-w\,(a-\lambda)\right)$, $p_{22}=p_{23}=-\,(1-w\,(a-\lambda))$, $p_{24}=1$, $p_{31}=2s\,(a-\lambda)-v\,(a-\lambda)-w+2r$, $p_{32}=w-r$, $p_{33}=v\,(a-\lambda)-r$ and $p_{34}=r$.

Now, from (3.9),

$$\varphi\,(s,u,v,w,r)=\frac{1}{\sqrt{\det P}}, \tag{3.10}$$

and it turns out to be convenient to represent the P matrix as

$$P=P_0+H, \tag{3.11}$$

where

$$H=vH_1+wH_2+rH_3, \tag{3.12}$$

with P_0, H_1, H_2 and H_3 as in the theorem.

Furthermore, (3.10) and (3.11) imply that

$$\varphi(s,u,v,w,r) = \frac{1}{\sqrt{\det P_0}} \int \cdots \int e^{\left\{-\frac{1}{2}\underline{x}'H\underline{x}\right\}} \sqrt{\det P_0}\,(2\pi)^{-T} \cdot$$
$$\cdot e^{\left\{-\frac{1}{2}\underline{x}'H\underline{x}\right\}} dx_{1,1}\ldots dx_{T,1}dx_{1,2}\ldots dx_{T,2} = \frac{1}{\sqrt{\det P_0}} E\left(e^{-\frac{1}{2}\underline{X}'H\underline{X}}\right),$$

taking expectation with respect to the $2T$ dimensional random vector $\underline{X}$, which is normally distributed with covariance matrix P_0^{-1}. A Taylor expansion of the exponential yields

$$\varphi(s,u,v,w,r) = \frac{1}{\sqrt{\det P_0}}\left(1 - \frac{1}{2}m_1 + \frac{1}{2^2 2!}m_2 - \frac{1}{2^3 3!}m_3 + \frac{1}{2^4 4!}m_4 - \ldots\right), \tag{3.13}$$

where $m_i = E\left((\underline{X}'H\underline{X})^i\right),\ i = 1,2,\ldots$ are moments of a Wishart distribution, as calculated by Magnus (1978). In particular

$$m_2 = \mathrm{tr}^{\,2}\left(P_0^{-1}H\right) + 2\,\mathrm{tr}\left(\left(P_0^{-1}H\right)^2\right), \tag{3.14}$$

$$m_3 = \mathrm{tr}^{\,3}\left(P_0^{-1}H\right) + 6\,\mathrm{tr}\left(P_0^{-1}H\right)\mathrm{tr}\left(\left(P_0^{-1}H\right)^2\right) + 8\,\mathrm{tr}\left(\left(P_0^{-1}H\right)^3\right) \tag{3.15}$$

and

$$\begin{aligned} m_4 \;=\; & \mathrm{tr}^{\,4}\left(P_0^{-1}H\right) + 32\,\mathrm{tr}\left(P_0^{-1}H\right)\mathrm{tr}\left(\left(P_0^{-1}H\right)^3\right) + 12\,\mathrm{tr}^{\,2}\left(\left(P_0^{-1}H\right)^2\right) + \\ & + 12\,\mathrm{tr}^{\,2}\left(P_0^{-1}H\right)\mathrm{tr}\left(\left(P_0^{-1}H\right)^2\right) + 48\,\mathrm{tr}\left(\left(P_0^{-1}H\right)^4\right) \end{aligned} \tag{3.16}$$

Since v, w and r enter these expressions only through $H = vH_1 + wH_2 + rH_3$, (3.14)-(3.16) are polynomials in v, w and r, and by inserting these polynomials into (3.13) it is possible to identify the φ derivatives of Theorem 3.2. This is how (3.5)-(3.8) are obtained, and the proof is completed. □

The results of Theorem 3.1 can be used for numerical computations of the expectations in (3.1)-(3.4) for a fixed T, and via Lemma 2.1 and Theorem 3.1, we can obtain approximate expectations of the approximated and the likelihood ratio test statistics. As an example, consider the calculation of (3.1):

1. Compute (3.5) for a (sufficiently large) grid of s and u values.
2. The mean taken over this grid gives an approximation of the integral. This

approximation can be improved by refining the grid and by using numerical extrapolation techniques, see e.g. Dahlqvist et al. (1974).
Handling (3.3) and (3.4) is simpler because we will only need a grid of s values. In the next section we will present numerical results and a comparison with Monte Carlo simulation based results.

In order to reduce the dimensionality of the matrices involved in the calculation of (3.5) (and hence, in the double integral of (3.1)), the following corollary is useful.

Corollary 3.2.

$$\frac{\partial^2}{\partial v^2}\varphi(s,u,v,w,0) = \frac{1}{\sqrt{\det \overline{P}_{10} \det A_2}} \cdot \tag{3.17}$$

$$\cdot \left(\frac{1}{4}\left(\operatorname{tr} C_{11} + 2(a-\lambda)\left(\operatorname{tr} C_{12} + \operatorname{tr} C_{21}^*\right)\right)^2 + \frac{1}{2}\operatorname{tr}\left(C_{11}^2\right) + (a-\lambda)\operatorname{tr}\left(C_{11}C_{12}\right) + \right.$$

$$\left. + \frac{1}{2}(a-\lambda)^2 \operatorname{tr}\left(C_{12}^2\right) + (a-\lambda)\operatorname{tr}\left(C_{11}^* C_{12}\right) + (a-\lambda)^2 \operatorname{tr}\left(C_{11}^* C_{22}\right) + \frac{1}{2}(a-\lambda)^2 \operatorname{tr}\left(C_{21}^{*\,2}\right) \right)$$

where the derivative is taken at $\{v=0, w=0\}$. *Here,* $C_{11} \overset{\text{def}}{=} B_{11}h_1$, $C_{12} \overset{\text{def}}{=} B_{12}h_2$, $C_{21} \overset{\text{def}}{=} B_{21}h_1$, $C_{22} \overset{\text{def}}{=} B_{22}h_2$, $C_{11}^* \overset{\text{def}}{=} B_{11}h_2'$, $C_{21}^* \overset{\text{def}}{=} B_{21}h_2'$, $P_0^{-1} \overset{\text{def}}{=} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ *and the* B_{ij} $(i,j=1,2)$ *matrices are* $T \times T$, *and given by* $B_{11} = A_1^{-1}$, $B_{12} = -2s(a-\lambda)\overline{P}_{10}^{-1} I_- A_2^{-1}$, $B_{21} = -2s(a-\lambda) P_{20}^{-1} I_- A_1^{-1}$, $B_{22} = A_2^{-1}$, *with* $A_1 \overset{\text{def}}{=} \overline{P}_{10} - 4s^2(a-\lambda)^2 I_- P_{20}^{-1} I_-$ *and* $A_2 \overset{\text{def}}{=} P_{20} - 4s^2(a-\lambda)^2 I_- \overline{P}_{10}^{-1} I_-$. □

Proof: The result follows directly from (3.5) and standard linear algebra. □

Remark 1: By inspection of (3.17), it is seen that the expectations in (3.1) and (3.4) are symmetric functions of $(a-\lambda)$, (This turns out to be true also for the expectation in (3.3).). This is an intuitive consequence of the model (2.1), since a change in the sign of $(a-\lambda)$ corresponds to a change of sign for the $\{u_t\}$ sequence, and clearly that does not affect the probabilistic properties of the model. Also, note that for $a-\lambda=0$, the results of Theorem 3.1 and Corollary 3.2 specialize to the results in Larsson (1994, 1997). □

Remark 2: The methods in Larsson (1994, 1997) could possibly be developed further in order to obtain more tractable integral expressions in (3.1)-(3.4). An issue for future work on the subject. □

Remark 3: As pointed out by Nielsen (1995), the quantities $\sum \varepsilon_t^2$ and Z_T are independent if $a - \lambda = 0$. Hence, in this special case, we have for $n \geq 1$,

$$E\left(\left(\frac{\left(\sum S_{t-1,3}\varepsilon_t^2\right)}{\sum S_{t-1,3}^2}\right)^n\right) = E\left(\left(\frac{\sum \varepsilon_t^2}{T}\right)^n\right) E\left(Z_T^n\right) = \frac{2^n \Gamma\left(\frac{T}{2}+n\right)}{T^n \Gamma\left(\frac{T}{2}\right)} E\left(Z_T^n\right),$$

simplifying (3.1) and (3.2) to

$$E\left(Z_T\right) = \int_0^\infty \frac{\partial^2}{\partial v^2} \varphi\left(s,u,v,w,0\right) ds$$

and

$$E\left(Z_T^2\right) = \frac{T}{T+2} \int_0^\infty s \frac{\partial^4}{\partial v^4} \varphi\left(s,u,v,w,0\right) ds,$$

respectively. □

4. Numerical evaluation

The purpose of this section is to evaluate the performance of the approach suggested in the previous section. We will, in particular, examine the size distortion problems for the likelihood ratio cointegration test in (2.15) and the approximated test in (2.17). Furthermore, we will calculate the expectation of Z_T, Z_T^2 $R_T^{(1)}$, and $R_T^{(2)}$, for sample sizes $T = 5, \ldots, 50$ and $a = 1.0$ and $a = 0.5$ using the numerical method. A comparison of the numerical results with simulated counterparts will give guidance on the appropriateness of the suggested method. Finally, simulations of the Bartlett corrected tests will show to what extent the size problem can be alleviated in this particular context. It is useful to disentangle this issue in two. First, what is the effect of correcting an oversized test by its expectation, i.e. does the Bartlett idea hold? We evaluate this by using a simulation estimated expectation as a basis for the Bartlett correction. Second, how does the Bartlett correction based on the proposed numerical method perform? In effect, we will, using simulation, estimate empirical sizes for the uncorrected tests, denoted by $-2\log Q_T$ and M_T, and compare with empirical sizes for simulation based corrections, $-2\log Q_T^{BCsim}$ and M_T^{BCsim}, as well as numerically corrected tests, $-2\log Q_T^{BC}$ and M_T^{BC}.

All calculations have been carried out on a Digital Venturis 5120 running at 120 MHz. The code is in double precision FORTRAN and the experiments involve $R = 1,000,000$ replications throughout, except for the asymptotic distributions which involve $T = 10,000$ observations and $R = 500,000$ replications. The pseudo $\mathrm{N}(0,1)$ variates are generated using the ISML routine DRNNOF. In order to reduce the Monte Carlo random variability across experiments, a sample of size $T = 50$ is generated in each replication and test statistics for all experimental sample sizes are calculated recursively on that sample.

Monte Carlo simulation estimates of the empirical sizes for the likelihood ratio test and the approximated test require asymptotic critical values, and hence, we must first simulate the asymptotic distributions. Tables A1a and A1b, show that the asymptotic distribution is not invariant to a. Moreover, as predicted by Lemma 2.1, the asymptotic distributions of $-2\log Q_T, Z_T$, and M_T are indeed

equivalent. The numerical method gives an asymptotic EZ_T of 1.142 when $a = 1$, and is hence in agreement with the simulation result in Table A1a.

Tables A2a and A2b confirm the accuracy of the numerical representations (3.1)-(3.4) in Theorem 3.1. When comparing with the simulated moments, the standard errors of the latter imply confidence intervals that in general will cover the numerical expectations. It is only for the cases of EZ_T^2 and larger sample sizes that a slight divergence occurs.

Now, to the heart of the matter. Tables A3a and A3b, A4a and A4b, and Figures A1-A8 provide partly overlapping evidence on the size behaviour of the unadjusted and adjusted tests. Apparently the parameter a makes little difference and we can hence concentrate on the results for $a = 0.5$ in Figures A1-A4. The two unadjusted tests are both distorted in size, but in opposite directions. Whereas the likelihood ratio test is oversized for small T, and severely so for a 1% nominal size, the approximated test is undersized, in particular for a 1% nominal size where the null remains unrejected throughout 1,000,000 samples for the very small T. The effectiveness of Bartlett correction is also different for the two tests. In the case of the likelihood ratio test we find that the numerical correction is at best of no value, whereas the simulated correction is almost perfect. In the case of the approximated test the roles are reversed, the simulated correction only marginally reduces the distortion problem while the numerical correction has an effect for the 1% nominal size.

Summarizing the numerical evidence we find that the idea of Bartlett correction, i.e. adjusting by using the expectation of the test statistic, functions very well for the likelihood ratio test but not for the approximated test. Unfortunately, the numerical approach suggested in section 3 proves to be ineffective, whereas basing the correction on a simulated expectation yields a remarkably accurate likelihood ratio test, even for as few observations as 5. The failure of the numerical method can be attributed to the relative impacts of the second and third terms in (2.18), that is, the first order correction given by the second term does not overcome the effects from the $O\left(T^{-2}\right)$ error term.

References

[1] Abadir, K.M. (1993). "OLS bias in nonstationary autoregression", *Econometric Theory*, Vol 9, pp. 81-93.

[2] Bartlett, M.S. (1937). "Properties of sufficiency and statistical tests", *Proceedings of the Royal Statistical Society of London*, Series A, Vol. 160, pp. 268-282.

[3] Banerjee, A., Dolado, J.J., Hendry, D.F. and Smith, G.W. (1986). "Exploring Equilibrium Relationships in Econometrics through Static Models: some Monte Carlo Evidence, *Oxford Bulletin of Economics and Statistics*, Vol 48, pp. 253-277.

[4] Barndorff-Nielsen, O.E. and Cox, D.R. (1989). "*Asymptotic Techniques for Use in Statistics*", London, Chapman and Hall.

[5] Campos, J., Ericsson, N.R. and Hendry, D.F. (1996). "Cointegration tests in presence of structural breaks", *Journal of Econometrics*, Vol. 70, pp. 187-220.

[6] Cheung, Y-W. and Lai, K.S. (1993). "Finite Sample Sizes of Johansens Likelihood Ratio Tests for Cointegration", *Oxford Bulletin of Economics and Statistics*, Vol 55, pp. 313-328.

[7] Dahlqvist, G., Björck, Å and Anderson, N. (1974). "*Numerical methods*", Prentice-Hall.

[8] Engle, R.F. and Granger, C.W.J. (1987). "Co-integration and Error Correction: Representation, Estimation and Testing", *Econometrica,* Vol. 55, pp. 251-276.

[9] Evans, G.B.A. and Savin, N.E. (1981). "Testing for unit roots: 1", *Econometrica*, Vol 49, pp. 753-779.

[10] Giersbergen, N.P.A. van (1996). "Bootstrapping the trace statistics in VAR models: Monte Carlo results and applications", *Oxford Bulletin of Economics and Statistics*, Vol 58, pp. 391-408.

[11] Hall, P. (1992). "*The Bootstrap and Edgeworth Expansion*", New York, Springer-Verlag.

[12] Harris, R.I.D. (1992). "Small Sample Testing For Unit Roots", *Oxford Bulletin of Economics and Statistics*, Vol 54, pp. 615-625.

[13] Jensen, J.L. (1993). "A historical sketch and some new results on the improved log likelihood ratio test statistic", *Scandinavian Journal of Statistics*, Vol 20, pp. 1-15.

[14] Jensen, J.L. and Wood, A.T.A. (1995). "*On the non-existence of Bartlett corrections for unit root tests*", Unpublished manuscript.

[15] Jacobson, T. (1995). "Simulating Small-Sample Properties of the Maximum Likelihood Cointegration Method: Estimation and Testing", *Finnish Economic Papers*, Vol 8, pp. 96-107.

[16] Johansen, S. (1988). "Statistical Analysis of Cointegrating Vectors", *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231-254.

[17] Johansen, S. (1991). "Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, Vol. 59, pp. 1551-1580.

[18] Kremers, J.J.M., Ericsson, N.R. and Dolado, J.J. (1992). "Testing Integration and Cointegration: an Overview, *Oxford Bulletin of Economics and Statistics*, Vol 54, pp. 325-348.

[19] Larsson, R. (1994). "*Bartlett Corrections for Unit Root Test Statistics*", Institute of Mathematical Statistics, University of Copenhagen, Preprint no. 2, 1994.

[20] Larsson, R. (1995). "*Small Sample Corrections for Unit Root Test Statistics*", Department of Mathematics, Uppsala University, Report no. 1995:13.

[21] Larsson, R. (1997). "On the asymptotic expectations of some unit root tests in a first order autoregressive process in the presence of a trend". To appear in the *Annals of the Institute of Statistical Mathematics*.

[22] Le Breton, A. and Pham, D.T. (1989). "On the bias of the least squares estimator for the first order autoregressive process", *Annals of the Institute of Statistical Mathematics*, Vol 3, pp. 555-563.

[23] Lucas, A. (1996). "Outlier Robust Unit Root Analysis", Tinbergen Institute, Erasmus University, Research Series no: 111.

[24] Magnus, J.R. (1978). "The moments of products of quadratic forms in normal variables", *Statistica Nederlandica*, Vol 32, pp. 201-210.

[25] Mikulski, P.W. and Monsour, M.J. (1994). "Moments of the limiting distribution for the boundary case in first order autoregressive processes", *American Journal of Mathematical and Management Sciences*, Vol 14, pp. 327-347.

[26] Nielsen, B. (1994). "*Bartlett correction in the cointegration model*", Unpublished manuscript, Institute of Mathematical Statistics, University of Copenhagen,

[27] Nielsen, B. (1995). "*Bartlett correction of the unit root test in autoregressive models*", Unpublished manuscript, Institute of Mathematical Statistics, University of Copenhagen,

Appendix: Tables and Figures

Table A1a: Expectation and associated standard error, variance, and percentiles of the asymptotic distributions of $-2\log Q_T$, Z_T, and M_T. $T = 10{,}000, R = 500,000, a = 1.0$.

	E	*s.e.*	V	50th	80th	90th	95th	99th
$-2\log Q_T$	1.1395	.0015	1.1765	.604	1.887	2.977	4.109	6.822
Z_T	1.1392	.0015	1.1743	.603	1.888	2.980	4.108	6.825
M_T	1.1393	.0015	1.1752	.604	1.887	2.977	4.108	6.820

Table A1b: Expectation and associated standard error, variance, and percentiles of the asymptotic distributions of $-2\log Q_T$, Z_T, and M_T. $T = 10{,}000, R = 500,000, a = 0.5$.

	E	*s.e.*	V	50th	80th	90th	95th	99th
$-2\log Q_T$	1.1125	.0016	1.2108	.550	1.856	2.961	4.121	6.887
Z_T	1.1122	.0016	1.2097	.551	1.855	2.961	4.117	6.888
M_T	1.1124	.0016	1.2095	.550	1.856	2.960	4.120	6.884

Table A2a: Small sample expectation for Z_T and Z_T^2 evaluated by way of Monte Carlo simulation (sim) and the numerical method (num).

		EZ_T			EZ_T^2		
T	a	num	sim	s.e.	num	sim	s.e.
10	1	0.9991	0.9994	.0016	2.2909	2.2893	.0070
	.5	0.9991	1.0010	.0017	2.3873	2.3899	.0075
20	1	1.0609	1.0582	.0018	2.7862	2.7581	.0097
	.5	1.0469	1.0460	.0019	2.8251	2.8099	.0101
30	1	1.0837	1.0819	.0019	3.0039	2.9700	.0110
	.5	1.0642	1.0661	.0019	3.0189	2.9981	.0114
40	1	1.0944	1.0962	.0019	3.1287	3.1009	.0120
	.5	1.0715	1.0769	.0020	3.1316	3.0978	.0121
50	1	1.0996	1.1039	.0020	3.2116	3.1825	.0125
	.5	1.0740	1.0825	.0020	3.2071	3.1640	.0126

Table A2b: Small sample expectation for $R_T^{(1)}$ and $R_T^{(2)}$ evaluated by way of Monte Carlo simulation (sim) and the numerical method (num).

		$ER_T^{(1)}$			$ER_T^{(2)}$		
T	a	num	sim	s.e.	num	sim	s.e.
10	1	0.9996	0.9994	.0057	0.9996	0.9951	.0038
	.5	0.9971	1.0027	.0057	0.9599	0.9632	.0038
20	1	1.0622	1.0583	.0075	1.0623	1.0559	.0040
	.5	1.0423	1.0323	.0074	1.0124	1.0077	.0039
30	1	1.0866	1.0743	.0089	1.0869	1.0792	.0041
	.5	1.0586	1.0509	.0087	1.0349	1.0333	.0040
40	1	1.0995	1.0998	.0101	1.0998	1.0968	.0041
	.5	1.0676	1.0692	.0099	1.0474	1.0485	.0040
50	1	1.1075	1.1024	.0112	1.1079	1.1038	.0041
	.5	1.0730	1.0657	.0109	1.0554	1.0487	.0040

Table A3a: Percentiles associated with the simulation based Bartlett corrected and the numerically Bartlett corrected likelihood ratio test, in comparison with the simulated asymptotic and small sample percentiles.

		95:th percentile				99:th percentile			
T	a	asymp	simul	sim.BC	num.BC	asymp	simul.	sim.BC	num.BC
10	1	4.109	4.816	4.768	4.016	6.822	8.186	7.916	6.667
	.5	4.121	4.891	4.901	4.128	6.887	8.352	8.191	6.899
20	1	4.109	4.376	4.341	4.077	6.822	7.372	7.209	6.768
	.5	4.121	4.436	4.426	4.135	6.887	7.457	7.396	6.911
30	1	4.109	4.270	4.244	4.088	6.822	7.211	7.046	6.788
	.5	4.121	4.305	4.303	4.127	6.887	7.284	7.191	6.897
40	1	4.109	4.270	4.205	4.087	6.822	7.152	6.981	6.786
	.5	4.121	4.234	4.250	4.113	6.887	7.142	7.103	6.874
50	1	4.109	4.235	4.190	4.081	6.822	7.069	6.956	6.775
	.5	4.121	4.208	4.223	4.097	6.887	7.095	7.058	6.847

Table A3b: Percentiles associated with the simulation based Bartlett corrected and the numerically Bartlett corrected approximated test, in comparison with the simulated asymptotic and small sample percentiles.

		95:th percentile				99:th percentile			
T	a	asymp	simul	sim.BC	num.BC	asymp	simul.	sim.BC	num.BC
10	1	4.108	3.822	4.041	3.603	6.820	5.589	6.709	5.981
	.5	4.120	3.868	4.138	3.686	6.884	5.662	6.915	6.158
20	1	4.108	3.930	4.014	3.825	6.820	6.166	6.665	6.351
	.5	4.120	3.978	4.084	3.873	6.884	6.224	6.824	6.472
30	1	4.108	3.980	4.030	3.908	6.820	6.410	6.690	6.487
	.5	4.120	4.010	4.081	3.940	6.884	6.467	6.819	6.583
40	1	4.108	4.050	4.045	3.946	6.820	6.549	6.715	6.551
	.5	4.120	4.018	4.086	3.968	6.884	6.541	6.827	6.630
50	1	4.108	4.061	4.061	3.965	6.820	6.590	6.742	6.582
	.5	4.120	4.035	4.092	3.977	6.884	6.615	6.838	6.646

Table A4a: Estimated sizes for the uncorrected likelihood ratio test, $-2\log Q_T$, the numerically Bartlett corrected likelihood ratio test, $-2\log Q_T^{BC}$, and the simulation wise Bartlett corrected likelihood ratio test, $-2\log Q_T^{BCsim}$. Nominal test sizes are 5% and 1%.

		$-2\log Q_T$		$-2\log Q_T^{BCsim}$		$-2\log Q_T^{BC}$	
T	a	5%	1%	5%	1%	5%	1%
10	1	.0723	.0191	.0518	.0114	.0757	.0206
	.5	.0734	.0194	.0501	.0107	.0732	.0193
20	1	.0582	.0133	.0511	.0107	.0593	.0137
	.5	.0590	.0136	.0501	.0104	.0586	.0134
30	1	.0551	.0122	.0510	.0107	.0557	.0124
	.5	.0553	.0123	.0499	.0105	.0551	.0122
40	1	.0537	.0177	.0508	.0107	.0544	.0119
	.5	.0536	.0155	.0497	.0102	.0538	.0116
50	1	.0530	.0155	.0505	.0106	.0539	.0118
	.5	.0531	.0112	.0499	.0102	.0538	.0115

Table A4b: Estimated sizes for the uncorrected approximated test, M_T, the numerically Bartlett corrected approximated test, M_T^{BC}, and the simulation wise Bartlett corrected approximated test, M_T^{BCsim}. Nominal test sizes are 5% and 1%.

		M_T		M_T^{BCsim}		M_T^{BC}	
T	a	5%	1%	5%	1%	5%	1%
10	1	.0400	.0022	.0423	.0026	.0602	.0065
	.5	.0411	.0024	.0404	.0023	.0581	.0058
20	1	.0444	.0059	.0473	.0067	.0540	.0086
	.5	.0451	.0061	.0462	.0063	.0535	.0084
30	1	.0463	.0074	.0487	.0081	.0527	.0093
	.5	.0465	.0075	.0477	.0078	.0522	.0092
40	1	.0471	.0081	.0491	.0087	.0523	.0097
	.5	.0471	.0080	.0481	.0083	.0519	.0094
50	1	.0478	.0087	.0492	.0091	.0523	.0101
	.5	.0478	.0085	.0486	.0087	.0524	.0099

Figure A1: Monte Carlo estimated empirical sizes for $-2\log Q_T$, $-2\log Q_T^{BCsim}$ and $-2\log Q_T^{BC}$. The nominal size is 5%, $a = 0.5$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

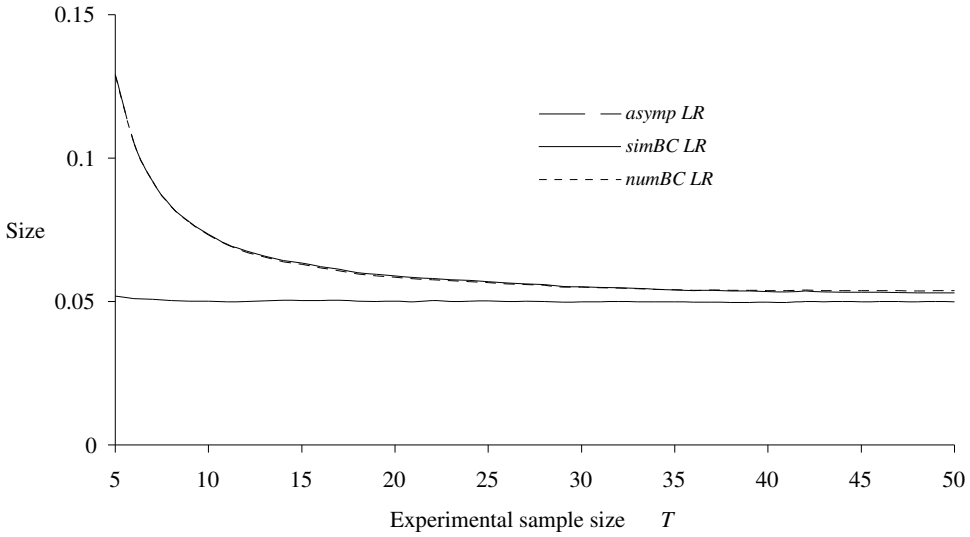


Figure A2: Monte Carlo estimated empirical sizes for $-2\log Q_T$, $-2\log Q_T^{BCsim}$ and $-2\log Q_T^{BC}$. The nominal size is 1%, $a = 0.5$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

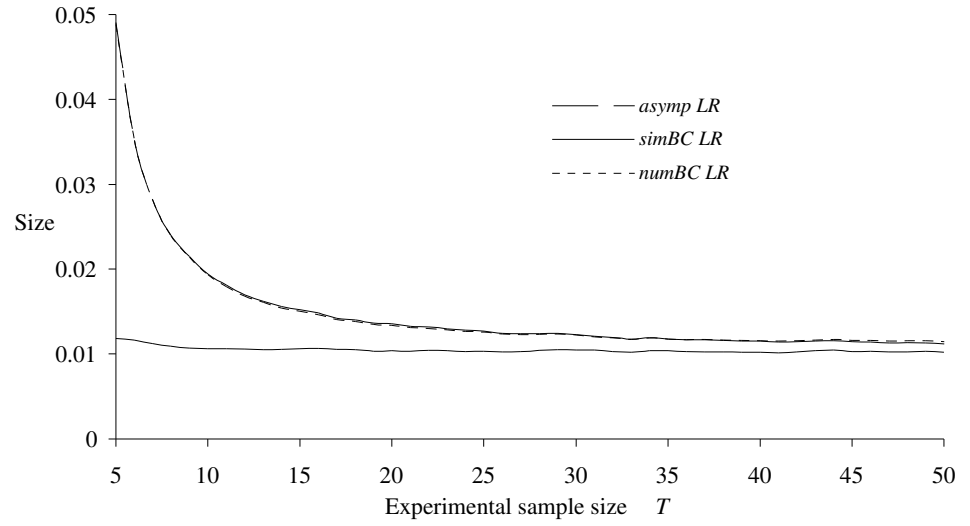


Figure A3: Monte Carlo estimated empirical sizes for M_T, M_T^{BCsim} and M_T^{BC}. The nominal size is 5%, $a = 0.5$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

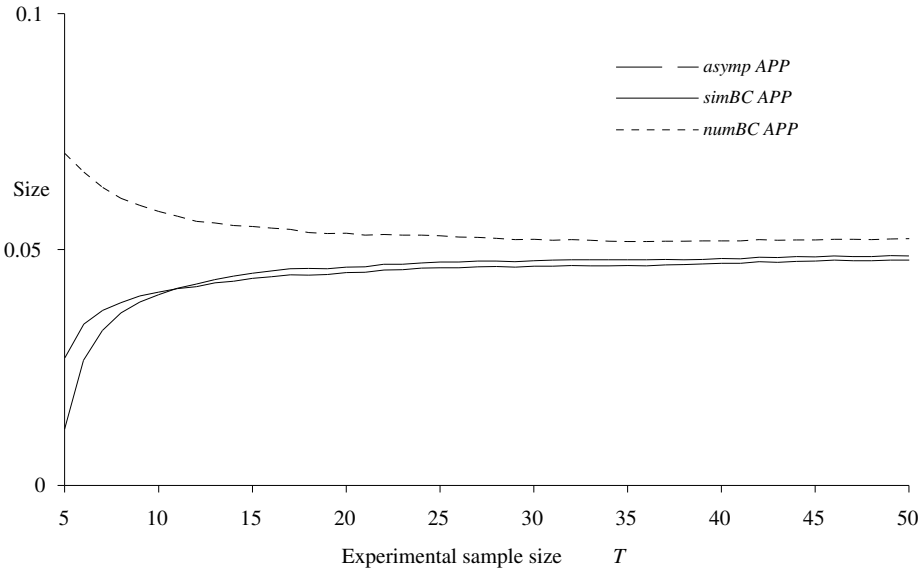


Figure A4: Monte Carlo estimated empirical sizes for M_T, M_T^{BCsim} and M_T^{BC}. The nominal size is 1%, $a = 0.5$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

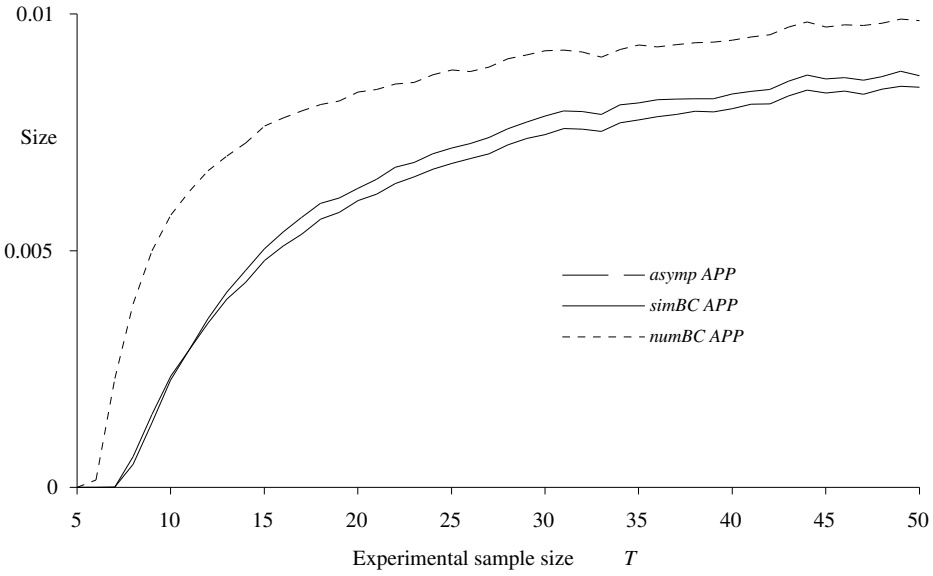


Figure A5: Monte Carlo estimated empirical sizes for $-2\log Q_T$, $-2\log Q_T^{BCsim}$ and $-2\log Q_T^{BC}$. The nominal size is 5%, $a = 1.0$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

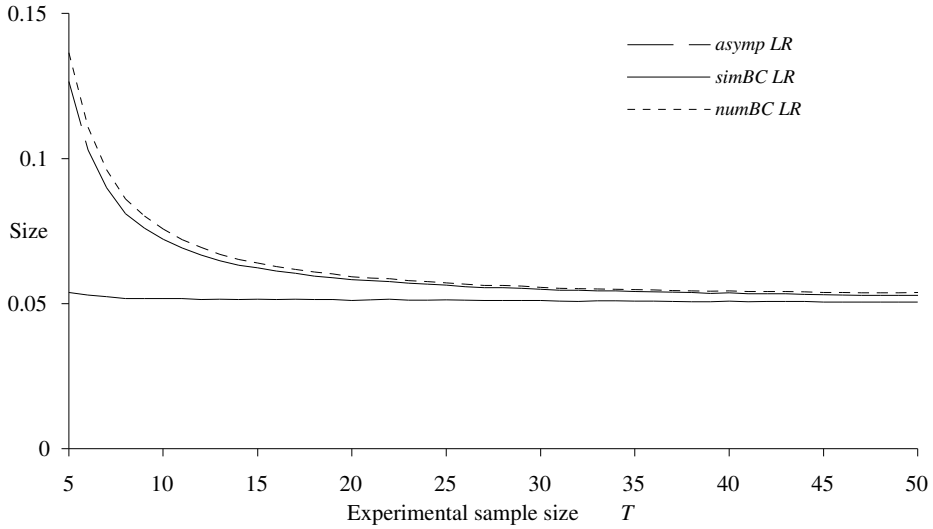


Figure A6: Monte Carlo estimated empirical sizes for $-2\log Q_T$, $-2\log Q_T^{BCsim}$ and $-2\log Q_T^{BC}$. The nominal size is 1%, $a = 1.0$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

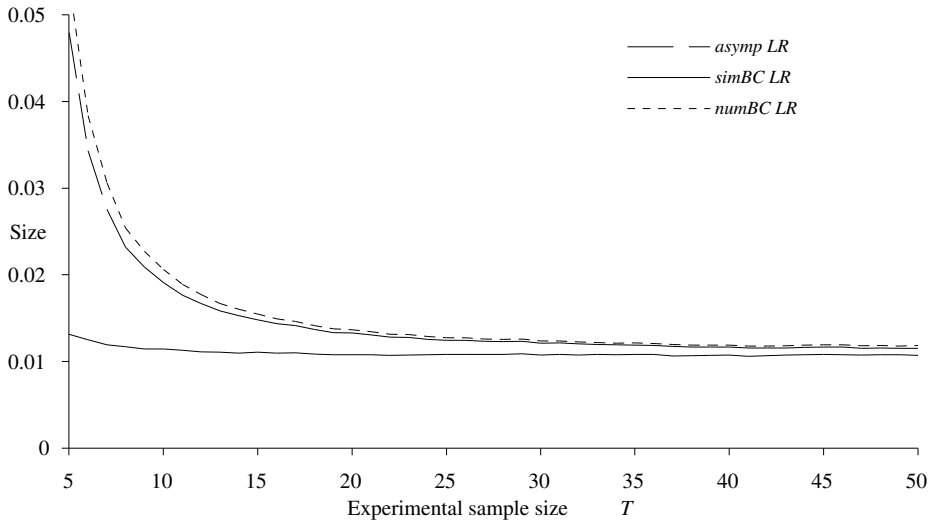


Figure A7: Monte Carlo estimated empirical sizes for M_T, M_T^{BCsim} and M_T^{BC}. The nominal size is 5%, $a = 1.0$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.

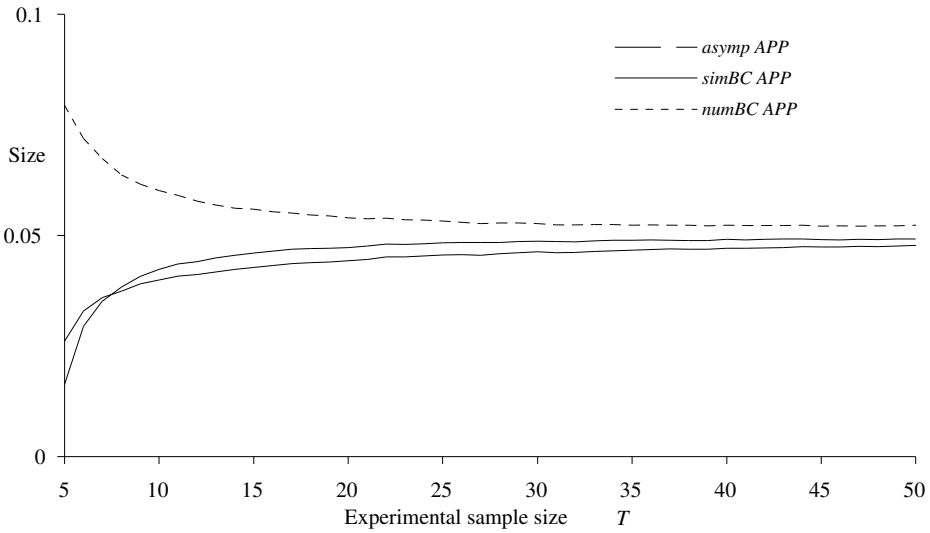
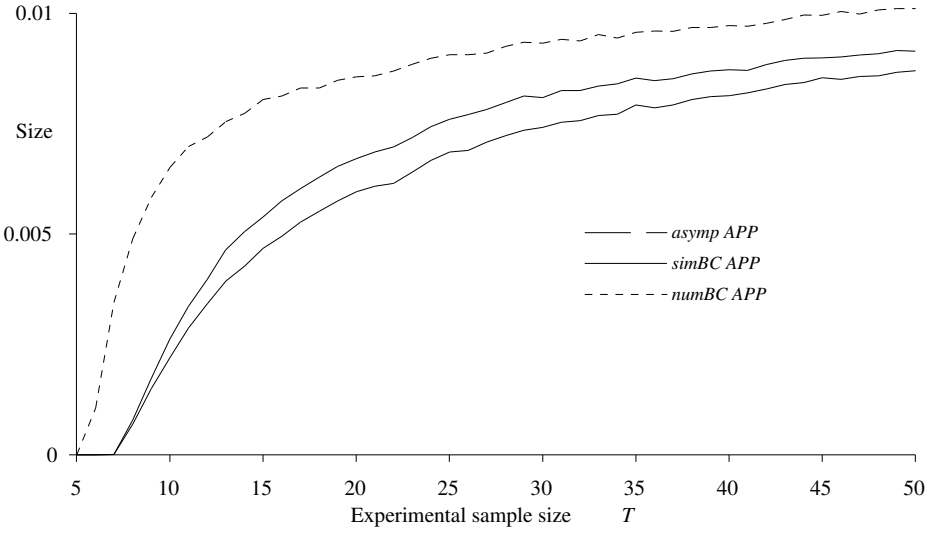



Figure A8: Monte Carlo estimated empirical sizes for M_T, M_T^{BCsim} and M_T^{BC}. The nominal size is 1%, $a = 1.0$, $T = 5, \ldots, 50$, and the number of replications is $1,000,000$.