Estimation of seasonal autoregressive time series

Edward Henry Lee
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
http://lib.dr.iastate.edu/rtd/7442

This Dissertation is brought to you for free and open access by Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University Microfilms International
300 N. ZEEB RD., ANN ARBOR, MI 48106
Lee, Edward Henry

ESTIMATION OF SEASONAL AUTOREGRESSIVE TIME SERIES

Iowa State University

Ph.D. 1981

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages
2. Colored illustrations, paper or print
3. Photographs with dark background
4. Illustrations are poor copy
5. Pages with black marks, not original copy
6. Print shows through as there is text on both sides of page
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements
9. Tightly bound copy with print lost in spine
10. Computer printout pages with indistinct print
11. Page(s) lacking when material received, and not available from school or author.
12. Page(s) seem to be missing in numbering only as text follows.
13. Two pages numbered . Text follows.
14. Curling and wrinkled pages
15. Other

University
Microfilms
International
Estimation of seasonal autoregressive
time series
by
Edward Henry Lee

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1981
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF LITERATURE</td>
<td>5</td>
</tr>
<tr>
<td>III</td>
<td>ESTIMATORS FOR THE PARAMETER $\rho$ IN THE FIRST-ORDER SEASONAL AUTOREGRESSIVE MODEL</td>
<td>41</td>
</tr>
<tr>
<td>IV</td>
<td>ESTIMATORS FOR THE PARAMETERS OF HIGHER ORDER SEASONAL AUTOREGRESSIVE MODELS</td>
<td>66</td>
</tr>
<tr>
<td>V</td>
<td>A MONTE CARLO STUDY</td>
<td>81</td>
</tr>
<tr>
<td>VI</td>
<td>SUMMARY</td>
<td>132</td>
</tr>
<tr>
<td>VII</td>
<td>EXAMPLES</td>
<td>139</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>147</td>
</tr>
<tr>
<td>A</td>
<td>THE EXACT MOMENTS OF THE LEAST SQUARES ESTIMATOR FOR THE FIRST-ORDER SEASONAL AUTOREGRESSIVE PROCESS</td>
<td>154</td>
</tr>
<tr>
<td>B</td>
<td>THE COVARIANCES OF ESTIMATED AUTOCOVARIANCES FOR STATIONARY AUTOREGRESSIVE PROCESSES</td>
<td>164</td>
</tr>
<tr>
<td>C</td>
<td>EMPIRICAL MEANS, VARIANCES AND MEAN SQUARE ERRORS OF ALTERNATIVE ESTIMATORS OF $\rho$ AND ONE AND THREE PERIOD PREDICTORS</td>
<td>171</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
<td>178</td>
</tr>
</tbody>
</table>
CHAPTER I. INTRODUCTION

Economic data displaying seasonal variations is commonly analyzed using the time series approach as developed by Box and Jenkins (1976). The seasonal behavior of the observed series is commonly due to a varying mean, and a seasonal means model is used to account for the seasonal fluctuations. A second method of incorporating seasonality into a structural model is to consider the economic series to be generated by a multiplicative time series model of the type suggested by Box and Jenkins (1976). The multiplicative seasonal autoregressive processes with seasonal means seem to be reasonable approximations for many practical applications.

Let \( Y_t \) satisfy the stochastic difference equation

\[
(1-\phi_1B-\phi_2B^2 - \ldots - \phi_pB^p)(1-\alpha_1B^k - \ldots - \alpha_rB^{kr})Y_t = \sum_{i=0}^{k-1} \delta_{it}\theta_i + e_t, \quad (1.1)
\]

where

\[
\delta_{it} = 1 \quad \text{if} \quad i = (t-1) \mod k
\]

\[
= 0 \quad \text{otherwise},
\]

\( \{e_t\} \) is a sequence of independent normal \((0, \sigma^2)\) random variables and
B is the usual backshift operator defined so that \( B^j Y_t = Y_{t-j} \). Then \( \{Y_t\} \) is a multiplicative seasonal autoregressive process of order \((p, 0)x(r, 0)\) with period \( k \) and seasonal intercepts \( \theta_0, \theta_1, \ldots, \theta_{k-1} \). The value of \( r \) is generally less than 3 and the value of \( k \) is most commonly equal to 12 which corresponds to monthly observations. It is assumed that the roots of the polynomial equations in \( m \), \( m^{kr} - \alpha_1 m^{k(r-1)} - \ldots - \alpha_r = 0 \) and \( m^p - \phi_1 m^{p-1} - \ldots - \phi_p = 0 \), are less than unity in modulus. The parameter space is restricted by the requirements that the roots of the two polynomial equations in \( m \) are constrained to lie within the unit circle.

Given a realization \( \{Y_t; t = 1, 2, \ldots, nk\} \) of \( nk \) observations, the least squares procedure is commonly used to estimate the parameters of the seasonal autoregressive process. The method of maximum likelihood is appealing under the assumption of normality, but is difficult and expensive to compute in all but the simplest case of a first-order seasonal autoregressive process with known means. The least squares estimators are consistent and asymptotically normal, but are biased in finite samples. In econometric work, small sample sizes ranging from 5 to 20 years are frequently encountered such that the least squares bias of the autoregressive coefficients is appreciable in magnitude.

Consider the first-order seasonal autoregressive process \( \{Y_t\} \) with period \( k \) which satisfies the stochastic difference equation

\[
Y_t = \sum_{i=0}^{k-1} \delta_i \theta_i + \rho Y_{t-k} + \epsilon_t, \tag{1.2}
\]
where $\delta_{it}$ and $\{e_t\}$ are given in (1.1). This model is a special case of (1.1) with $p = 0$ and $r = 1$. Approximate expressions for the mean and variance of the least squares estimator of $\rho$ are obtained and some modified least squares estimators which correct for the least squares bias are proposed. In the case of the first-order seasonal autoregressive process, the least squares estimator for $\rho$ is a ratio of two quadratic forms and some exact results for the moments of the least squares estimator are possible. Other methods including the maximum likelihood procedure, are also considered.

For higher order seasonal autoregressive processes with seasonal means, the least squares procedure is customarily used to estimate the autoregressive coefficients. The method of obtaining approximate expressions for the first-order seasonal autoregressive coefficient is extended to include higher order seasonal autoregressive processes. Approximate expressions for the biases of the least squares estimators for the parameters of a stationary normal second-order autoregressive process with seasonal means correct to terms of order $n^{-1}$ are derived. Some modifications to the least squares estimators which adjust for the biases in the case of the second-order seasonal autoregressive process are proposed. Although it is seldom necessary to consider seasonal autoregressive processes of order greater than two, the methods of modifying the least squares estimators can be extended to include higher order processes.

The various estimators for the autoregressive coefficients are asymptotically equivalent, but are expected to behave differently in
finite samples. The adequacy of approximating the small sample properties of the estimators by asymptotic properties requires investigation.

A Monte Carlo study to evaluate the small sample behavior of the various estimators for the parameter of the normal first-order seasonal autoregressive process with seasonal means was carried out. The Monte Carlo study also investigated the adequacy of approximating the null distributions of the regression "t-statistics" by the Student's t distribution and compares the alternative predictors with the least squares predictor.

Examples of seasonal autoregressive processes for which the above results are applicable, are presented.
CHAPTER II. REVIEW OF LITERATURE

The majority of economic data display periodic fluctuations which recur every year with similar timing and intensity. Such behavior is commonly called seasonality. Considerable literature on analyzing economic time series is concerned with methods of dealing with seasonality. In recent years, autoregressive moving average processes have been proposed for analyzing economic data. See Box and Jenkins (1976), Box, Hillmer and Tiao (1976), Fuller (1976), Jenkins and Watts (1968), and Parzen and Pagano (1977). With the advent of the computer, the autoregressive moving average schemes are widely accepted as a reliable method for estimating and predicting the behavior of a real process.

Yule (1927), Walker (1931) and Slutsky (1937) first formulated the concept of autoregressive moving average schemes. The major contribution came in 1938 when Wold (1954) obtained a general representation for time series. Since then, a considerable body of literature in the area of time series dealing with parameter estimation and the order determination of time series models has appeared. More recently, Jenkins and Watts (1968) and Box and Jenkins (1976) extended the autoregressive moving averages to include seasonal time series.

Most of the results in time series deal with stationary processes. A stochastic process is called strictly stationary if the distribution of \((Y_{t_1}, Y_{t_2}, ..., Y_{t_m})\) is the same as the distribution of \((Y_{t_1+h}, Y_{t_2+h}, ..., Y_{t_m+h})\) for every set \(\{t_1, t_2, ..., t_m\}\) and for every \(h\) such that
$t^i, t^i+h \in T, \ i = 1, 2, \ldots, m,$ where $T$ is the set of time points at which measurements are taken. In many situations the form of the distribution is unknown and the lower order moments are used to characterize the process. A stochastic process is defined to be weakly stationary if its first and second moments exist and

\begin{align*}
&i) \ E(Y_t) = \mu \text{ for all } t \text{ in } T, \text{ and} \\
&ii) \ E[(Y_t - \mu)(Y_{t+h} - \mu)] = \gamma(h)
\end{align*}

for all $t, t + h$ in $T$. The autocorrelation function is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}. \quad (2.2)$$

Autocorrelations are measures of the relationship between successive values of a variable ordered in time and are of considerable interest in time series analysis. Much of the early work in time series was concerned with estimating the autocorrelations and deriving tests of independence between successive values of a variable.

Anderson (1942), Dixon (1944), Durbin and Watson (1951), Geisser (1956), Hart and von Neumann (1942), Koopmans (1942), Leipnik (1947), Moran (1948), von Neumann (1942), Rubin (1945), Shenton and Johnson (1965) and White (1957, 1961) investigated the problem of testing for zero autocorrelation.
Anderson (1942) considered the distribution of the first circular serial correlation

\[ r_1 = \frac{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}{n} \]  

in samples from an independent normal series. He derived the exact distribution of \( r_1 \) and higher order circular serial correlations. The first two moments of \( r_1 \) for the first-order stationary autoregressive process with \( Y_0 = Y_n \) are

\[ E(r_1) = -\frac{1}{n-1}, \]  

and

\[ \text{var}(r_1) = \frac{n(n-3)}{(n+1)(n-1)^2}. \]

Dixon (1944) obtained an approximate form for Anderson's distribution by smoothing the characteristic function. The first four moments of the approximation agreed exactly with the noncircular moments. In fact, the distribution of \( r_1^2 \) is approximately distributed as the squared ordinary correlation coefficient in samples of \( n + 2 \) from an uncorrelated normal population. Koopmans (1942) reached the same result by
a different method. See also Rubin (1945), Madow (1945), Leipnik (1947), Quenouille (1948), Jenkins (1954) and Kendall (1957).

Durbin and Watson (1951) have shown that certain modified noncircular definitions of the serial correlation coefficients have Anderson's distribution in the uncorrelated case.

Daniels (1956) derived sampling distributions of serial correlations using a saddlepoint method. In the case of a circular definition serial correlation, he obtains the Madow-Leipnik distribution. He was able to derive an approximate distribution for the noncircular statistic

\[
r = \frac{2}{n} \sum_{t=2}^{n} Y_t Y_{t-1} \frac{\Sigma_{t=2}^{n} Y^2_t}{\Sigma_{t=2}^{n} Y^2_{t-1} + \Sigma_{t=2}^{n} Y^2_t}
\]

for the first-order autoregressive process with known mean. Daniels finds the approximate distribution of \( r \) given in (2.6) as

\[
h(r) = \frac{n}{2\pi} \frac{\Gamma\left(\frac{n}{2}\right) - 1}{\Gamma\left(\frac{n}{2} - 2\right)} \frac{(1-p^2)^{\frac{1}{2}}(1-r^2)^{\frac{n}{2}} - 1}{(1-pr)(1-2pr+p^2)^{\frac{n}{2}} - 1} \left(1+O(n^{-1})\right).
\]

where \( p \) is the parameter of the first-order autoregressive process.

When the mean is unknown, he considers the noncircular statistic
where \( \overline{Y}_0 = [2(n-1)]^{-1} \left( \sum_{t=2}^{n} Y_{t-1} + \sum_{t=2}^{n} Y_t \right) \). The approximate distribution of \( r \) is

\[
h(r) = \frac{\Gamma \left( \frac{N+3}{2} \right)}{2\pi T \left[ \frac{N}{2} \right] [N(1-p) - (1+p)]} \frac{N^2 - 1}{(1-r)(1-r^2)^{N/2}} \left( 1 + O(n^{-2}) \right),
\]

where \( N = n - 1 + (1-p^2)^{-1} p^2 \). Daniels also considered the general autoregressive process circularly defined. See also Phillips (1977, 1978).

Reeves (1972) gives a method of obtaining values of the distribution function of the statistic

\[
\rho = \frac{\sum_{t=2}^{n} Y_t Y_{t-1}}{\sum_{t=2}^{n} Y_t^2},
\]

where \( Y_t \) is a first-order autoregressive process of zero mean. In general, the distribution function involves approximating the distribution
of a linear combination of chi-squared variates. An exact result is obtained in a special case.

Bias in Estimation of the Autocorrelation Function

The estimation of autocorrelations plays a key role in analyzing time series. Various definitions of the serial correlation coefficients have been proposed and extensively investigated. The exact distribution of the estimated autocorrelations remains unknown despite the efforts of many investigators. Anderson (1942) obtained the exact distribution of a circularly defined serial correlation coefficient for a series of random normal deviates. He showed that the bias of the serial correlation given in (2.3) is \((n-1)^{-1}\). Moran (1948) showed that the bias in all serial correlations of a random series using deviations from the mean is \((n-1)^{-1}\) for both circular and noncircular definitions. It was realized that estimates of autocorrelations derived from certain types of stationary series systematically underestimate or overestimate the true values. Methods of deriving the extent of the biases have generally relied on approximations.

The autocorrelation function of a stationary process is commonly estimated by the statistic

\[
 r(h) = \frac{n^{-1} \sum_{t=1}^{n-h} (Y_t - \mu)(Y_{t+h} - \mu)}{n^{-1} \sum_{t=1}^{n} (Y_t - \mu)^2} \quad (2.9)
\]
when the mean is known. Many modifications have been suggested that consist of changes in the end terms, $Y_1, \ldots, Y_n, Y_{n-1}$; other modifications involve multiplication by the factor $n(n-h)^{-1}$ for $0 \leq h \leq n-1$. However, the estimator in (2.9) guarantees the positive definiteness of the estimated autocorrelation function. See Jenkins and Watts (1968).

If the mean is unknown, the serial correlation may be defined by any of the following statistics:

\begin{equation}
    r_1(h) = \frac{(n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t Y_{t+h} \right] - (n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_{t+h} \right]^2 \left[ \sum_{t=1}^{n-h} Y_t^2 \right]^{1/2}}{(n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t^2 \right] - (n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t \right]^2},
\end{equation}

\begin{equation}
    r_2(h) = \frac{(n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t Y_{t+h} \right] - (n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_{t+h} \right]^2 \left[ \sum_{t=1}^{n-h} Y_t^2 \right]^{1/2}}{(n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t^2 \right] - (n-h)^{-1} \left[ \sum_{t=1}^{n-h} Y_t \right]^2},
\end{equation}

or

\begin{equation}
    r_3(h) = \frac{(n-h)^{-1} \sum_{t=1}^{n-h} Y_t Y_{t+h} - n^{-1} \sum_{t=1}^{n} Y_t^2 \left[ \sum_{t=1}^{n} Y_t \right]^2}{n^{-1} \sum_{t=1}^{n} Y_t^2 - n^{-1} \left[ \sum_{t=1}^{n} Y_t \right]^2},
\end{equation}
Various modifications in the end terms as well as modifications in the estimators of the mean terms have been considered. Numerous definitions of the serial correlation coefficients have been adopted by various authors.

Orcutt (1948) first described the two sources of bias in serial correlation coefficients. If the true mean of the series were known, unbiased estimators of the numerator and denominator of $\rho(h)$ are available; however, the expectation of their ratio is not the ratio of their expectation in general. This is the first source of bias which is inherent in ratio estimators. The second source of bias results from estimating the mean when the true mean is unknown.

Hurwicz (1950) considered the first-order autoregressive process with mean level known and initial condition $Y_0 = 0$. He showed that the least squares estimator of $\rho$ is biased in finite samples, and evaluated the bias exactly for samples of size 3. For larger sample sizes, he demonstrated the existence of a bias that tends to zero as the autoregressive parameter goes to zero.

Sastry (1951) derived the biases of the serial correlation coefficient by assuming that

$$E\left(\frac{A}{B}\right) = \frac{E(A)}{E(B)}$$

for any $A$ and $B$ such that $E(B) \neq 0$. He considered the two definitions of the serial correlation coefficient in (2.10) and (2.12) and evaluated the second source of bias of each estimator. Sastry concluded
that $r_1(h)$ was a more acceptable estimator than $r_3(h)$, but indicated the bias of $r_1(h)$ can be large for moderate sample sizes.

Marriott and Pope (1954) considered the three definitions of the serial correlation coefficient given in (2.10), (2.11) and (2.12). Of the three autocorrelation estimators, only $r_2(h)$ seemed reasonable as it employs a sensible correction for the mean in the numerator and denominator and is not cumbersome to calculate. Using $r_2(h)$ as an estimator, they obtained approximate expressions for the bias of the serial correlation coefficient to terms of order $n^{-1}$ for a first order autoregressive process with normal random variables.

Using the notation of Marriott and Pope, denote the numerator and denominator of $r_2(k)$ by

$$N_k = \frac{1}{n^{-k}} \sum_{t=1}^{n-k} Y_t Y_{t+k} - \frac{1}{(n-k)^2} \left( \sum_{t=1}^{n-k} Y_t \right) \left( \sum_{t=1}^{n-k} Y_{t+k} \right)$$

and

$$D = \frac{1}{n} \sum_{t=1}^{n} Y_t^2 - \frac{1}{n^2} \left( \sum_{t=1}^{n} Y_t \right)^2,$$

respectively. They expand $r_2(k)$ in a binomial series to the second order of approximation to obtain
For a first-order autoregressive process with zero mean and normal random error, the authors consider the terms

$$E(r_2(k)) = \frac{E(N_k)}{E(D)} \cdot \frac{\text{Cov}(N_k, D)}{[E(D)]^2} + \frac{E(N_k) \text{Var}(D)}{[E(D)]^3}. \quad (2.13)$$

To terms of order $n^{-1}$, the authors evaluate the means of $N_k$ and $D$ as

$$E(N_k) = \frac{\rho^k}{1-\rho^2} - \frac{1}{n(1-\rho^2)} + o(n^{-2}) \quad (2.14)$$

and

$$E(D) = \frac{1}{1-\rho^2} - \frac{1}{n(1-\rho^2)} + o(n^{-2}). \quad (2.15)$$
The second moments of \( N'_k \) and \( D' \) are derived to terms of order \( n^{-1} \) as

\[
\text{var}(D') = \frac{2(1+p^2)}{n(1-p^2)^3} + O(n^{-2}) .
\] (2.16)

and

\[
\text{cov}(N'_k, D') = \frac{2\rho^k [(k+1)-(k-1)p^2]}{n(1-p^2)^3} + O(n^{-2}) .
\] (2.17)

The second moments of \( N'_k \) and \( D \) are equal to the second moments of \( N'_k \) and \( D' \) to order \( n^{-1} \).

Substituting these results in (2.13), led Marriott and Pope to the approximation

\[
E\{r^2(k)\} = \rho k - \frac{1}{n} \left( \frac{(1+p)(1-p^k)}{1-p} + 2k\rho^k \right) + O(n^{-2}) .
\] (2.18)

When the true mean is known to be zero, the authors obtain the approximate mean of the corresponding value for \( r^*(k) \),

\[
E\{r^*(k)\} = \rho k - 2n^{-1} k\rho^k + O(n^{-2}) .
\] (2.19)

If \( k = 1 \),
\[ E\{r_2(1)\} = \rho - n^{-1}(1 + 3\rho) + O(n^{-2}) \quad (2.20) \]

and

\[ E\{r^*(1)\} = \rho - 2n^{-1} \rho + O(n^{-2}) . \quad (2.21) \]

An asymptotic expression for \( \text{var}(r_2(k)) \) to order \( n^{-1} \) has been derived by Bartlett (1946). The variance of \( r_2(k) \) is

\[
\text{var}(r_2(k)) = \frac{1}{n} \left[ \frac{(1+\rho^2)(1-\rho^{2k})}{(1-\rho^2)} - 2k\rho \right] + O(n^{-2}) . \quad (2.22)
\]

This expression also holds for \( \text{var}(r^*(k)) \).

Kendall (1954), in a related note, obtained more general results in evaluating the intermediate results in deriving the approximate bias of serial correlation coefficients. He evaluated the approximate mean of \( r_1(k) \) given in (2.10) to order \( n^{-1} \) as

\[
E(r_1(k)) = \rho^k - \frac{1}{n-k} \left[ \frac{1+\rho}{(1-\rho)} (1-\rho^k) + 2k\rho \right] . \quad (2.23)
\]

This is seen to be the result obtained by Marriott and Pope for \( r_2(k) \).

Kendall stated that such expressions are probably satisfactory for values of \( \rho \) near zero, but are of very doubtful validity for \( \rho \) near
to unity. The distribution of the serial correlation coefficient is highly skewed such that the use of expectations as a criteria of bias is open to question. Considering terms to order \( n^{-2} \) or \( n^{-3} \) does not necessarily give better results.

White (1961) considers the first serial correlation given in (2.8) for a first-order autoregressive process of zero mean. Denoting the numerator and denominator of \( \hat{\rho} \) by \( U = \sum_{t=2}^{n} Y_t Y_{t-1} \), and \( V = \sum_{t=2}^{n} Y_{t-1}^2 \), respectively, he obtains the first two moments of \( \hat{\rho} \) by considering the joint moment generating function of \( U \) and \( V \). The integrals involved in evaluating \( E(\rho^k) \) are computed by expanding the integrand in a Maclaurin series and then integrating termwise. When \( Y_0 = 0 \), White obtains, to terms of order \( \rho^4 n^{-3} \), the first two moments of \( \hat{\rho} \) as

\[
E(\hat{\rho}) = \left( 1 - \frac{2}{n} + \frac{4}{n^2} - \frac{2}{n^3} \right) \rho + \frac{12}{n} \rho^3 + \frac{18}{n^3} \rho^5 + O(n^{-4}) , \tag{2.24}
\]

and

\[
\text{var}(\hat{\rho}) = \left( \frac{1}{n} - \frac{1}{n^2} + \frac{5}{n^3} \right) - \left( \frac{1}{n} + \frac{14}{n^2} - \frac{73}{n^3} \right) \rho^2 + \frac{3}{n^3} \rho^4 + O(n^{-4}) . \tag{2.25}
\]

When \( Y_0 \) is a normal random variable with mean zero and variance \( \sigma^2 (1-\rho^2)^{-1} \), the first two moments of \( \hat{\rho} \) are
\[ E(\hat{\rho}) = \left( 1 - \frac{2}{n} + \frac{4}{n^2} - \frac{2}{n^3} \right) \rho + \frac{2}{n^2} \rho^3 + \frac{2}{n^2} \rho^5 + O(n^{-4}) , \quad (2.26) \]

and

\[ \text{var}(\hat{\rho}) = \left( \frac{1}{n} - \frac{1}{n^2} + \frac{5}{n^3} \right) - \left( \frac{1}{n} - \frac{13}{n^2} + \frac{69}{n^3} \right) \rho^2 - \frac{20}{n^3} \rho^4 + O(n^{-4}) . \quad (2.27) \]

To terms of order \( n^{-1} \) the first two moments of \( \hat{\rho} \) for the two models with differing initial conditions agree. The results derived by Marriott and Pope (1954) and Kendall (1954) are in agreement as well.

Estimation of the Parameters of an Autoregressive Process

The sampling theory approach to the estimation problem of an autoregressive process has generally been analogous to the treatment of the univariate regression model. While there are a variety of estimators recommended for these models on the basis of their asymptotic characteristics, the small sample properties of these techniques have proved difficult to derive analytically. In the past, there have been a number of Monte Carlo studies examining their respective small sample performance patterns. See Marriott and Pope (1954), Copas (1966), Thornber (1967), Orcutt and Winokur (1969), Salem (1971), Min (1975) and Bora-Senta and Kounias (1980). In economic data, small samples are frequently encountered in practice. Generally, 5 to 20 years of monthly or
quarterly data are commonly available for time series analysis. The systematic bias arising from well-known estimation procedures can be substantial when the series length is moderately small.

The autoregressive time series of order $p$ is defined by the stochastic difference equation

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + e_t, \quad (2.28)$$

t = p+1, p+2, \ldots, \text{ where the } e_t \text{ are uncorrelated } (0, \sigma^2) \text{ random variables and } Y_1, Y_2, Y_3, \ldots, Y_p \text{ are initial conditions. It is assumed that } a_p \neq 0 \text{ and the roots of the characteristic equation}

$$m^p - a_1 m^{p-1} - a_2 m^{p-2} - \ldots - a_p = 0 \quad (2.29)$$

are less than one in absolute value. The parameters of the model and the variance of $e_t$ are to be estimated from an observed sequence $Y_1, Y_2, \ldots, Y_n$.

The first-order autoregressive process with unknown level has received considerable attention in the literature. The process can be represented by

$$(Y_t - \mu) = \rho (Y_{t-1} - \mu) + e_t, \quad (2.30)$$
\( t = 2, 3, 4, \ldots \), where \( Y_1 \) is the initial condition and \( p \) is less than one in absolute value. When \( Y_1 = \mu \), the first-order autoregressive process will be denoted as model \( A \); when \( Y_1 \) is a normal \(( \mu, \sigma^2(1-p^2)^{-1})\) random variable, the process will be denoted as model \( B \).

Multiplying equation (2.28) by \((Y_{t-h} - \mu)\) for \( h \geq 0 \) and taking expectations of both sides, one obtains a system of equations relating the autocovariance function to the coefficients of the model. The equations corresponding to \( h = 1, 2, \ldots, p \) are

\[
\begin{align*}
\gamma(1) &= a_1 \gamma(0) + \alpha_2 \gamma(1) + \ldots + \alpha_p \gamma(p-1) \\
\gamma(2) &= a_1 \gamma(1) + \alpha_2 \gamma(0) + \ldots + \alpha_p \gamma(p-2) \\
\vdots \\
\gamma(p) &= a_1 \gamma(p-1) + \alpha_2 \gamma(p-2) + \ldots + \alpha_p \gamma(0)
\end{align*}
\]

which is a system of \( p \) simultaneous equations known as the Yule-Walker equations. See Yule (1926), Walker (1931). By dividing (2.31) by \( \gamma(0) \), the autocorrelation function is similarly related to the coefficients of the model.

Yule (1927), Wold (1954) and Kendall (1947) suggested the use of \( r_3(1)(1-n^{-1}) \) as an estimate of \( p \) in model (2.30). For a second-order autoregressive process, the Yule-Walker equations are of the form

\[
\begin{align*}
\rho(1) &= a_1 + \alpha_2 \rho(1) \\
\rho(2) &= \alpha_2 \rho(1) + \alpha_2
\end{align*}
\]
The authors propose substituting \( r_3(1)(1-n^{-1}) \) and \( r_3(2)(1-2n^{-1}) \) for \( \rho(1) \) and \( \rho(2) \), respectively and solving the equations for \( \alpha_1 \) and \( \alpha_2 \). This method gives the Yule-Walker estimates of \( \alpha_1 \) and \( \alpha_2 \) as

\[
\alpha_1 = \frac{r(1) - r(1)r(2)}{1 - [r(1)]^2}
\]

and

\[
\alpha_2 = \frac{r(2) - [r(1)]^2}{1 - [r(1)]^2}.
\]

Levinson (1947) and Durbin (1960) give a recursive procedure for obtaining the Yule-Walker estimates of a \( p \)-th order autoregression.

The studies on the serial correlation coefficients are related to the estimation of parameters of autoregressive processes as seen by the Yule-Walker estimates. In particular, the parameter of the first-order autoregressive model is estimated by the first serial correlation coefficient. From the previous section, the estimators of \( \rho \) are biased for the autoregressive parameter.

The estimation of the parameters of an autoregressive process is generally treated as an estimation problem in a univariate regression model. The ordinary least squares procedure provides the best linear unbiased estimators in the classical linear regression model. Mann and Wald (1943), in a classical paper, have shown that the least squares
estimators are consistent and asymptotically normally distributed in the autoregressive case. However, the assumptions of the Gauss Markov theorem are not met in the autoregressive case since lagged values of the dependent variable are not distributed independently of the error term for all lags. For finite samples, the least squares estimators are generally biased.

For the first-order autoregressive process, the least squares estimator of $\rho$ is

$$\hat{\rho} = \frac{n}{\sum_{t=2}^{n} \left( \sum_{t=1}^{n} Y_t Y_{t-1} \right)}$$

(2.34)

when the mean is known and equal to zero. When the mean is unknown the least squares estimator is

$$\hat{\rho} = \frac{n}{\sum_{t=2}^{n} \left( \sum_{t=2}^{n} (Y_t - \bar{Y}_0)(Y_{t-1} - \bar{Y}_1) \right)}$$

(2.35)

where $\bar{Y}_0 = (n-1)^{-1} \sum_{t=2}^{n} Y_t$ and $\bar{Y}_1 = (n-1)^{-1} \sum_{t=2}^{n} Y_{t-1}$. It is well-known that the conditional maximum likelihood estimators, conditional on $Y_1, Y_2, \ldots, Y_p$, of the autoregressive parameters lead to the least squares estimators.
Koopmans (1942) considered the maximum likelihood estimator of the parameter $\rho$ in model $B$ with zero mean. The maximum likelihood estimator of $\rho$ is obtained as the solutions to the equation

$$g(\rho) = a_3 \rho^3 + a_2 \rho^2 + a_1 \rho + a_0 = 0, \quad (2.36)$$

where

$$a_3 = \left(\frac{n-1}{n}\right)^n \sum_{t=2}^{n-1} y_t^2,$$

$$a_2 = -\left(\frac{n-2}{n}\right)^n \sum_{t=2}^{n} y_{t-1} y_t,$$

$$a_1 = -\frac{n-1}{n} \sum_{t=2}^{n} y_t^2 - \frac{1}{n} \sum_{t=1}^{n} y_t^2,$$

$$a_0 = \sum_{t=2}^{n} y_{t-1} y_t.$$

For a stationary first-order autoregressive process, the maximum likelihood estimator of $\rho$ is defined as the unique root of (2.36) between $-1$ and $1$, where

$$\hat{\rho}_{MLE} = \frac{2}{3} a_3^{-1} (a_2 - 3a_1 a_3)^{1/6} \cos \theta - \frac{1}{3} a_2 a_3^{-1} \quad (2.37)$$

and
\[ \theta = \frac{1}{3} \arccos \left[ -\frac{1}{4} (2a_2 - 9a_1 a_2 a_3 + 27a_0 a_2^2) (a_2^2 - 3a_1 a_3)^{-1} \left( \frac{3}{2} \right) \right] + \frac{4\pi}{3} . \]

For a discussion concerning the roots, see White (1961), Anderson (1971) and Hasza (1980).

For higher order autoregressive processes, the maximum likelihood procedure has no closed form solutions. Box and Jenkins (1976) proposed a method that gives the approximate maximum likelihood estimators in the case of normally distributed errors. The estimates \( \hat{a}_0, \hat{a}_1, \ldots, \hat{a}_p \) minimize the sum of squares

\[ S(a) = \sum_{t=-k}^{n} [e_t]^2 , \]

where \( [e_t] = y_t - a_0 - a_1 y_{t-1} - \ldots - a_p y_{t-p} \), \( t = p+1, p+2, \ldots, n \) and \( [e_p], [e_{p-1}], \ldots, [e_{-k}] \) are found from

\[ [e_t] = y_t - a_0 - a_1 y_{t-1} - \ldots - a_p y_{t-p} \), \( t = p, p-1, \ldots, -k \),

where

\[ y_t = a_0 + a_1 y_{t+1} + a_2 y_{t+2} + \ldots + a_p y_{t+p} \), \( t \leq 0 \).

Recursive algorithms such as Marquardt's (1963) algorithm are used to perform the iterations.
Burg (1967, 1968) suggested a method of estimating the autoregressive parameters based on the Levinson (1947)-Durbin (1960) procedure used in computing the Yule-Walker estimates. Increasing orders of autoregressions are fit in a stepwise fashion. Denote the estimate of the k-th coefficient when fitting an autoregression of order p as $a_k(p)$, and the estimate of $\sigma^2$ as $S_p$. The recursion begins with

$$ S_0 = \frac{1}{n} \sum_{t=1}^{n} y_t^2 $$

$$ a_1(1) = \frac{2 \sum_{t=2}^{n} y_t y_{t-1}}{n \sum_{t=2}^{n} y_t^2 + \sum_{t=2}^{n} y_t^2} $$

$$ S_1 = \frac{1}{n} \sum_{t=1}^{n} y_t^2 (1 - a_1^2(1)) $$

when the mean is known and equal to zero. At the p-th stage, define the residuals from a p-th order autoregression by

$$ e_{t}(p) = y_t - \sum_{k=1}^{p} a_k(p) y_{t-k}, \quad t = p+1, \ldots, n $$

$$ = y_t - \sum_{k=1}^{p-1} (a_k(p-1) - a_p(p) a_{p-k}(p-1)) y_{t-k} - a_p(p) y_{t-p}. $$
Similarly the backward residuals are

\[ h_t(p) = y_t - \sum_{k=1}^{p-1} (a_k(p-1) - a_p(p)a_{p-k}(p-1))y_{t+k} - a_p(p)y_{t+p}, \]

\[ t = 1, 2, \ldots, n-p. \]

The coefficient \( a_p(p) \) minimizes the sum of squares

\[ \sum_{t=p+1}^{n} [e_t(p)]^2 + \sum_{t=1}^{n-p} [h_t(p)]^2 \]

giving

\[ a_p(p) = \frac{2 \sum_{t=p+1}^{n} e_t(p-1)h_{t-p}(p-1)}{\sum_{t=1}^{n-p} [h_t(p-1)]^2 + \sum_{t=p+1}^{n} [e_t(p-1)]^2}. \]

(2.38)

The other coefficients are updated by

\[ a_k(p) = a_k(p-1) - a_p(p)a_{p-k}(p-1), \quad k = 1, 2, \ldots, p-1 \]

and

\[ s_p = s_{p-1}(1 - a^2_p(p)). \]
When the mean is unknown, $Y_t - \bar{Y}$ is substituted in place of $Y_t$ and the recursion proceeds as before. See Burg (1975), Ulrych and Bishop (1975), Jones (1978), and Robinson and Silvia (1980).

The various estimators considered are asymptotically equivalent but behave differently in small samples. It is well-known that the estimation methods are biased in finite samples although the exact distributions of the estimators are not known. In the case of the first-order autoregressive process, a number of methods have been proposed for reducing the bias in the estimate of $\rho$.

Quenouille (1949) suggested a method of removing the bias in the least squares estimators of autoregressive parameters. Assuming the bias is proportional to $n^{-1}$, the method consists of dividing the series into halves and estimating the autoregressive parameters using the whole series and each half separately. An estimator of $\rho$ unbiased to order $n^{-1}$ is obtained as

$$\hat{\rho} = 2\hat{\rho} - \frac{1}{2}(\hat{\rho}' + \hat{\rho}'') ,$$

where $\hat{\rho}'$ and $\hat{\rho}''$ are the least squares estimator of $\rho$ for the first and second halves, respectively.

For the first-order autoregressive process with unknown mean, a nearly unbiased estimator of $\rho$, based on the work of Marriott and Pope is
\[ \hat{\rho}_{mp} = \frac{n-1}{n-4} \hat{\rho} + \frac{1}{n-4}, \] (2.40)

where \( \hat{\rho} \) is the ordinary least squares estimator. When the true value of \( \rho \) is zero, \( \hat{\rho}_{mp} \) is an unbiased estimator of \( \rho \).

Salem (1971) suggested two methods of reducing the bias in the least squares estimator of \( \rho \) for the first-order autoregressive process with unknown mean. The first estimator which corrects for the bias in estimating the mean is given as

\[ \hat{\rho}_1 = \frac{\hat{\rho} + \frac{1}{n-1} \left( \frac{1+\hat{\rho}}{1-\hat{\rho}} \right)}{1 + \frac{1}{n-1} \left( \frac{1+\hat{\rho}}{1-\hat{\rho}} \right)}. \] (2.41)

The second estimator is based on a ratio estimator suggested by Beale which is nearly unbiased. The form of the estimator is

\[ \hat{\rho}_2 = \frac{\hat{\rho} + \frac{1}{n-1} \left[ \frac{1+\hat{\rho} + \frac{4\hat{\rho}^2}{1-\rho}}{1-\rho} \right]}{1 + \frac{1}{n-1} \left[ \frac{1+\hat{\rho} + 2(1+\hat{\rho}^2)}{1-\rho} \right]}, \] (2.42)

Bora-Senta and Kounias (1980) recently proposed a method for parameter estimation of an autoregressive model with unknown mean. The authors propose an iterative procedure using modified estimators of the autocorrelations.
\[ \hat{\rho}_h = \frac{\hat{\Lambda}}{\hat{\gamma}_0} + \frac{n}{n-h} r_h \left( 1 - \frac{\hat{\Lambda}}{n\hat{\gamma}_0} \right), \]  

(2.43)

where

\[ c_h = n^{-1} \sum_{t=1}^{n-h} (Y_t - \bar{Y})(Y_{t+h} - \bar{Y}) \]

\[ r_h = \frac{c_h}{c_0} \]

\[ \frac{\hat{\Lambda}}{\hat{\gamma}_0} = \frac{1 - \hat{\alpha}_1 \hat{\alpha}_1 - \hat{\alpha}_2 \hat{\alpha}_2 - \ldots - \hat{\alpha}_p \hat{\alpha}_p}{(1 - \hat{\alpha}_1 - \hat{\alpha}_2 - \ldots - \hat{\alpha}_p)^2}. \]  

(2.44)

The iteration proceeds as follows:

i) As a first approximation, compute \( \hat{\rho}_{h,1} = r_h, h = 1, 2, \ldots, p \).

ii) Using \( \hat{\rho}_{h,1}, h = 1 \) to \( p \), compute the Yule-Walker type estimates \( \hat{\alpha}_{1,1}, \hat{\alpha}_{2,1}, \ldots, \hat{\alpha}_{p,1} \).

iii) Calculate \( \frac{\hat{\Lambda}}{\hat{\gamma}_0} \) from (2.44) using the estimates \( \hat{\alpha}_{h,1}, \hat{\rho}_{h,1}, h = 1 \) to \( p \).

iv) Obtain second approximations \( \hat{\rho}_{h,2}, h = 1, 2, \ldots, p \), using \( \frac{\hat{\Lambda}}{\hat{\gamma}_0} \) in (2.43).

v) Check the conditions for stationarity. If violated, take the previous estimates.
vi) If not violated, continue until the sum of squares

\[ J_1 = \sum_{h=1}^{p} (\hat{\alpha}_{h,1+1} - \hat{\alpha}_{h,1})^2 \]

is less than a quantity \( \alpha \).

For a first-order autoregressive process, the authors consider the estimator

\[ \bar{\beta} = \beta_1(1 + \frac{2}{n}) + \frac{1}{n} . \quad (2.45) \]

Marriott and Pope (1954) studied the bias of the first-order serial correlation in a limited Monte Carlo study. Empirical results indicated that the approximate bias underestimated the true biases in small samples. The approximate variances are not satisfactory in the first order autoregressive process and tend to underestimate the sampling variance.

Copas (1966) compares the performance of a Bayesian estimator of the first-order autoregressive parameter with zero mean to the least squares, maximum likelihood and the bias corrected estimator associated with Marriott and Pope. Barnard et al. (1962) proposed a Bayesian estimator corresponding to a uniform prior defined as

\[ \bar{\beta}_1 = \frac{\int_{\rho} L(\rho) d\rho}{\int_{\rho} L(\rho) d\rho} . \quad (2.46) \]
where $L(p)$ is the likelihood function of the series. The "mean likelihood" estimator $\bar{p}_1$, gave the smallest average mean square error when averaged over the values of $p$ considered. For values of $-0.30 \leq p \leq 0.6$, $\bar{p}_1$ gave the least mean square error, the least squares estimator being slightly better for $p > 0.6$. For values of $n = 10$ and $20$, the Marriott and Pope adjusted estimator had a large mean square error because the reduction in bias did not compensate for the increase in variance.

Thornber (1967) compares the small sample properties of the least squares, weighted least squares, maximum likelihood and Bayesian estimators of $p$ from model B with $\mu = 0$. The weighted least squares estimator

$$\hat{p}_2 = \frac{\sum_{t=2}^{n} Y_t Y_{t-1}}{\sum_{t=3}^{n} Y_t^2}$$

is derived by minimizing the sum of squares

$$V = (1-p^2)Y_1^2 + \sum_{t=2}^{n} (Y_t - pY_{t-1})^2.$$

Based on an expected risk measure, the Bayes estimator is optimal. However, the least squares estimator is nearly as good and easier to compute.
Orcutt and Winokur (1969) evaluated several estimates for the first order autoregressive process with unknown mean. The least squares estimator of $\rho$ and two bias corrected estimators, based on the work of Marriott and Pope (1954) and Quenouille (1949), are compared using Monte Carlo techniques. Estimates of the expected value and variance of $\hat{\rho}$ for each of 48 combinations of $(\rho, n)$ are based on a sample of 1,000 series. Both modified least squares estimators were essentially unbiased for all combinations of $(\rho, n)$ considered. For values of $\rho$ near zero, the least squares estimator had the lowest mean square error; for larger values of $\rho$, the modified least squares estimator, based on Marriott and Pope, had a smaller mean square error. The modified least squares estimator, based on Quenouille, had uniformly larger mean square error than the least squares or Marriott and Pope's estimators.

Predictions using the three estimates of $\rho$ were also considered by Orcutt and Winokur. All three predictors were nearly unbiased and performed equally well for large samples. The least squares predictor had the smallest variance for all values of $n$ and $\rho$ considered. The authors conclude that standard least squares prediction appears nearly optimal in small samples.

Salem (1971) compared the least squares estimator of $\rho$ and three modified least squares estimators given in (2.40), (2.41) and (2.42). Five values of $n$ ranging from 6 to 36 and 9 values of $\rho$ between -0.99 and 0.99 were used. The sample mean, variance and mean square error of each estimator were calculated for each $(\rho, n)$ combination by summing over 1,000 samples. The modified estimators, based on Marriott
and Pope and Beale are nearly unbiased for all values of $n$ and $-1.0 < \rho < 0.9$. For values of $-0.9 < \rho < 0.0$, the least squares estimator had the smallest mean square error. Little differences in mean square error among the estimators were noted for values of $-0.99 < \rho < -0.60$. For values of $0.0 < \rho < 0.90$, the estimator $\hat{\rho}_1$ had the smallest mean square error, the Marriott and Pope estimator $\hat{\rho}_{mp}$ doing slightly better for $0.90 < \rho < 0.99$.

Sawa (1978) evaluated the exact moments of the least squares estimator for the stationary first-order autoregressive process with a normal error process. Based on the moment generating function of the numerator and denominator of $\hat{\rho}$, the moments of the least squares estimator are obtained by numerically integrating the partial derivatives of the moment generating function. Kendall's (1954) approximation to the mean and Bartlett's (1946) approximation to the variance are shown to be satisfactory unless the parameter value is close to the boundary of the stationary region.

De Gooijer (1980) derived the first four exact moments of the sample autocorrelations for a stationary autoregressive moving average process using Sawa's (1978) approach. The author concludes that exact formulae for the moments of the sample autocorrelations are not analytically tractable. The approximate moments of the serial correlations are adequate for relatively small sample sizes.

Ansley and Newbold (1980) analyzed numerically the small sample properties of the exact maximum likelihood, exact least squares and conditional least squares estimators for stationary autoregressive moving
average time series with zero mean. For a series length of 50, the 3 estimators were very similar in terms of prediction mean square error for autoregressive processes. The authors preferred the maximum likelihood estimators in general, based on its superior performance for autoregressive moving average processes. Some results for seasonal models are presented and the maximum likelihood estimator is still preferred.

Min (1975) evaluated the small sample properties of various estimates for model A with zero mean. Empirical means and variances of estimates for various values of \( \rho \) and \( n = 100 \) are obtained by averaging over a sample of 100 series. An estimator suggested by Quenouille (1956) had the smallest bias for most values of \( \rho \) considered. For \( n = 100 \), little differences between the means, variances and mean square errors of the various estimators were observed.

Bora-Senta and Kounias (1980) evaluated the small sample performance of their proposed modified method of moments procedure and four other methods which include the least squares and Yule-Walker estimators. Three different parameter values for each autoregressive model of orders 1, 2 and 3 with series length of 20 and 200 were used. The least squares estimates had the smallest variances but systematically underestimated the true values. The proposed modified method of moments estimates had the smallest biases with only a slight increase in variance for \( n = 20 \). For large samples, the time domain estimates were nearly identical in both means and variances.
Predictions for Autoregressive Processes

Time series models have been used extensively by practitioners in varying disciplines to provide adequate forecasts based on historic data. The main objective of many time series analysis is to predict future values in the short term to perhaps aid in managerial decision-making. There has been a recent interest in deriving expressions for the mean square of the prediction error for autoregressive processes. The sampling theory approach to obtaining forecasts of autoregressive processes is analogous to least squares predictors of univariate regression models.

For a stationary p-th order autoregressive process with known parameters, the one period ahead predictor with minimum mean square error is

\[ \hat{Y}_{n+1} = \alpha_0 + \sum_{j=1}^{p} \alpha_j y_{n-j}. \]

The h-step ahead predictor with minimum mean square error is obtained recursively from

\[ \hat{Y}_{n+h} = \alpha_0 + \sum_{j=1}^{p} \alpha_j \hat{Y}_{n+h-j}, \]  

where \( \hat{Y}_{n+j} = y_{n+j} \) for \( j \leq 0 \). It is well-known that the error of predicting \( h \) steps ahead is given by
\[ Y_{n+h} - \hat{Y}_{n+h} = \sum_{j=0}^{h-1} w_j e_{n+j}, \quad h = 1, 2, 3, \ldots, \]

where \( w_0 = 1 \) and \( w_1, w_2, w_3, \ldots \), are coefficients determined in the infinite moving average representation of \( Y_t \),

\[ Y_t = \sum_{j=0}^{\infty} w_j e_{t-j}. \]

The mean squared error of the \( h \)-step ahead predictor is

\[ \sigma_h^2 = \sigma^2 \sum_{j=0}^{h-1} w_j^2, \quad (2.49) \]

where \( \sigma^2 = \text{E}(e^2_t) \).

The coefficients of the autoregressive process are seldom known and must be estimated from an observed portion of the series. The predictor of \( Y_{n+h} \) is obtained by substituting the estimates \( \hat{\alpha}_i \) for \( \alpha_i \), \( i = 0 \) to \( p \) in (2.48). Manageable expressions for the mean square error of multi-step predictors from an estimated autoregressive model have been derived by a number of authors. See Davisson (1965), Box and Jenkins (1976), Bloomfield (1972), Bhansali (1974), Schmidt (1974), Yamamoto (1976), Phillips (1979) and Fuller and Hasza (1980). Most authors made the assumption that the observations used for prediction are independent from those used in estimating the autoregressive parameters. The
exceptions are Davisson, Bloomfield, Phillips and Fuller and Hasza. Only Fuller and Hasza permit a nonzero mean. Fuller and Hasza (1978) have shown that the least squares predictor is unbiased for the first-order autoregressive process with symmetrically distributed error process.

Davisson (1965) derived an expression for the asymptotic mean square error of a one period ahead predictor for a p-th order autoregressive process with zero mean. He showed that

\[
E\{(Y_{n+1} - \hat{Y}_{n+1})^2\} = \sigma^2(1 + n^{-1}p) + O(n^{-2}),
\]

where \( n \) is the length of the series used to construct estimates of the autoregressive parameters. Bloomfield (1972) obtained the same result using a different technique.

Box and Jenkins (1976) pointed out the difficulty in obtaining expressions for the asymptotic mean square error of h-step ahead predictors. For the simplest first-order autoregressive process with mean zero, the authors derived the mean square error as

\[
E\{(Y_{n+h} - \hat{Y}_{n+h})^2\} = \sigma_h^2 + \frac{\sigma^2 \rho^2 (h-1)}{n} + O(n^{-2}),
\]

where \( \sigma_h^2 \) is the mean square error of predicting \( h \) steps ahead using known coefficients given in (2.49).
Bhansali (1974) obtained the asymptotic mean square error of h-step ahead predictors for a stationary p-th order autoregressive process as

$$E\{ (\hat{Y}_{n+h} - \hat{Y}_{n+h}^* )^2 \} = \sigma_h^2 (1 + n^{-1} p) + O(n^{-2}) . \quad (2.52)$$

The result agrees with Bloomfield's result when h = 1 but differs notably from the result of Box and Jenkins when h > 1. Schmidt (1974) derived the asymptotic mean square error of multi-step predictors for a more general simultaneous autoregressive model with exogenous variables.

Yamamoto (1976) derives an alternative expression for the asymptotic mean square error of h-step ahead predictors which is compatible to the results of Bloomfield and Box and Jenkins. Letting $Y_t = (Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p+1})$, the linear least squares predictor of $Y_{n+h}$ is $\hat{Y}_{n+h}^* = a'(h)_{n+h}$ when the autoregressive parameters are known, and is $\hat{Y}_{n+1}^* = a'(h)_{n+1}$ when the coefficients $a' = (a_1, a_2, \ldots, a_p)$ are estimated. Let

$$[\frac{3a'(h)}{3a}] = M_h .$$

Then the prediction error $(\hat{Y}_{n+h}^* - Y_{n+h})$ is given by

$$\hat{Y}_{n+h}^* - Y_{n+h} = - \sum_{j=0}^{h-1} w_j e_{n+h-j} + (a-a)'M_{n+h}Y_n + O(n^{-1}) . \quad (2.53)$$

The asymptotic mean square error of prediction, to order $n^{-1}$, is obtained as

$$E\{ (\hat{Y}_{n+h}^* - Y_{n+h} )^2 \} = \sigma_h^2 + \frac{\sigma^2}{n} \text{tr} [M_h^{-1} M_h' \Gamma] , \quad (2.54)$$
where $\Gamma = \mathbb{E}(Y_t Y_t)$, when $\hat{a}$ is independent of $\{e_t\}$.

Phillips (1979) considered the distribution of the forecast error for the stationary first-order autoregressive process with zero mean conditional on the last observation $Y_n$. Assuming that the least squares estimator of $\rho$ and $Y_n$ are independent, Phillips derived an approximation of the conditional distribution of $\hat{Y}_{n+h}$ using an Edgeworth-type expansion.

Fuller and Hasza (1980) studied the mean square error of the $h$-step ahead predictor conditional on the last $p$ observations from a stationary $p$-th order autoregressive process with normal errors. The usual least squares procedure provides a consistent estimator of the conditional mean square error of the one-step ahead least squares predictor. The authors obtain as an expression for the conditional mean square error of $\hat{Y}_{n+h}$ given $Y_n$

$$
E\{ (Y_{n+h} - \hat{Y}_{n+h} )(Y_{n+h} - \hat{Y}_{n+h} )' | Y_n \}
$$

$$
= \sigma^2 \sum_{j=0}^{h-1} \sum_{k=0}^{h-1-j} A_j^i M_{jk}^a + n^{-1} \sum_{j=0}^{h-1} \sum_{k=0}^{h-1-j} \sum_{l=0}^{h-1-k} Y_{n+h-j-l} Y_{n+h-k-l} A_j^i M_{jk}^a
$$

$$
+ R_n, \quad (2.55)
$$

where
\( A = \begin{bmatrix}
\alpha_1 & \alpha_2 & \cdots & \alpha_{p-1} & \alpha_p & \alpha_0 \\
1 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & \cdots & \cdots & \cdots & \cdots \\
0 & 0 & \cdots & 1 & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 1 \\
\end{bmatrix} \),

\[ \hat{Y}_{n+j} = A^j Y_n \quad \text{for} \quad j = 0, 1, 2, \ldots, h, \quad \hat{Y}_{n+h} = \hat{A}^h Y_n \]

and \( M \) is a matrix with 1 in the upper left hand corner and zeroes elsewhere. The results of Phillips, Yamamoto and Davisson are obtained as special cases of (2.53).
CHAPTER III. ESTIMATORS FOR THE PARAMETER $p$
IN THE FIRST-ORDER SEASONAL AUTOREGRESSIVE MODEL

The estimation of the parameters of autoregressive processes has many analogies to the sampling theory approach to univariate regression models. Various methods based on the least squares procedure are commonly used to estimate the autoregressive parameters. With the additional assumption of normality, maximum likelihood estimators have been constructed by various authors. Whittle (1951) has shown the maximum likelihood estimators to be consistent and asymptotically normal. See also Mann and Wald (1943). In the absence of normality, the least squares estimators have been proven to have the same asymptotic distribution as the maximum likelihood estimators. See Walker (1964) and Whittle (1962). Theoretical studies of the maximum likelihood estimators and the least squares estimators have shown the asymptotic properties of the various estimators to be equivalent. For samples of the size generally encountered in econometric work, it is expected that the estimators will differ appreciably from one another. In this chapter, alternative estimators for the parameter of the first-order seasonal autoregressive process are derived.

The maximum likelihood and least squares methods of estimation are biased. The seriousness of the bias depends on whether the mean of the process is known or is estimated from the observed series. The estimators of the parameters of seasonal time series models with unknown seasonal intercepts will be shown to have substantial biases.
Consider the first-order seasonal autoregressive process \{Y_t\} of period \(k\) which satisfies the stochastic difference equation

\[
Y_t = \sum_{i=0}^{k-1} \delta_{it} \theta_i + \rho Y_{t-k} + e_t, \quad t = k+1, k+2, \ldots, \tag{3.1}
\]

where

\[
\delta_{it} = 1 \text{ if } i = (t-1) \mod k
\]

\[
= 0 \quad \text{otherwise},
\]

and \(Y_1, Y_2, \ldots, Y_k\) are initial conditions and \(\{e_t\}\) is a sequence of independent normal \((0, \sigma^2)\) random variables. The parameter \(\rho\) is assumed to be strictly less than one in absolute value. The parameters \(\theta_i, i = 0, 1, \ldots, k-1\), are the seasonal intercepts associated with the various periods. The first-order stationary autoregressive process with nonzero mean is a special case of (3.1) corresponding to \(k = 1\).

A more useful representation of the process \(\{Y_t\}\) is obtained by making use of the double subscripts notation \(Y_{ij}\), where \(t = (i+1) + k(j-1)\). The variable \(Y_{ij}\) denotes the sampled value for the \(i\)-th period of the \(j\)-th cycle and satisfies the stochastic difference equation

\[
Y_{ij} = \theta_i + \rho Y_{i,j-1} + e_{ij}, \tag{3.2}
\]
\( j = 2,3,4,..., \) and \( i = 0,1,2,..., k-1 \). The observations for period \( i \) constitute a realization from a first-order autoregressive process with a nonzero intercept \( \theta_i \) and parameter \( \rho \).

Given a realization of nk observations from the model (3.2), the least squares estimator of \( \rho \) is given by

\[
\hat{\rho}_{\text{OLS}} = \frac{\left[ (n-1)k \right]^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - \overline{Y}_{i10})(Y_{i,j-1} - \overline{Y}_{i1})}{\left[ (n-1)k \right]^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j-1} - \overline{Y}_{i1})^2}, \quad (3.3)
\]

where

\[
\overline{Y}_{i10} = (n-1)^{-1} \sum_{j=2}^{n} Y_{ij}
\]

and

\[
\overline{Y}_{i1} = (n-1)^{-1} \sum_{j=2}^{n} Y_{i,j-1}, \quad i = 0,1,2,..., k-1.
\]

By conditioning on the initial \( k \) observations, the maximum likelihood estimator of \( \rho \) is the same as the least squares estimator of \( \rho \) given in (3.3). The estimator \( \hat{\rho}_{\text{OLS}} \) can lie outside the stationary region \((-1,1)\) even though the observed series is stationary.
Despite the efforts of numerous investigators, the exact distribution of $\hat{\rho}_{\text{OLS}}$ remains unknown. Methods of deriving expressions for the first two moments of $\hat{\rho}_{\text{OLS}}$ have relied on series expansions and large sample theory. Sawa (1978) gives a method for computing the exact moments of the least squares estimator of a stationary first-order autoregressive parameter but no closed form expressions are given. However, the mean of $\hat{\rho}_{\text{OLS}}$ can be evaluated exactly if the series is random, i.e., $\rho = 0$. The result is given in the following lemma and is based on the work of Moran (1948).

**Lemma 3.1.** Let $Y_{ij}$ be a sequence of independent $(\mu_i, \sigma^2)$ random variables. Then the expectation of $\hat{\rho}_{\text{OLS}}$ given in (3.3) is

\[
E(\hat{\rho}_{\text{OLS}}) = -\frac{1}{(n-1)^{-1}}. \tag{3.4}
\]

**Proof.** The distribution of $\hat{\rho}_{\text{OLS}}$ is independent of the parameters $\theta_i$, $i = 0$ to $k-1$, since the variables $Y_{ij} - \bar{Y}_{in}$ are independently distributed of $\theta_i$, $i = 0$ to $k-1$. Without loss of generality the seasonal intercepts $\theta_i$ are assumed to be zero. The expected value of $\hat{\rho}_{\text{OLS}}$ is evaluated as

\[
E(\hat{\rho}_{\text{OLS}}) = E\left\{ \frac{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - \bar{Y}_{10})(Y_{i,j-1}-\bar{Y}_{11})}{k-1 \cdot \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j}-\bar{Y}_{11})^2} \right\}
\]
The mean of \( \hat{\rho}_{\text{OLS}} \) for a random series can also be derived exactly using Sawa's technique which is given in Appendix A. The least squares estimator has a downward bias in the case of independent observations with unknown means. In the case of a random series with known means, the least squares estimator of \( \rho \) is given as

\[
\hat{\rho}_{\text{OLS}} = \frac{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j} - \mu_{i})(Y_{i,j-1} - \mu_{i})}{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j} - \mu_{i})^2}.
\] (3.5)

The method of obtaining the mean of \( \hat{\rho}_{\text{OLS}} \) is not useful in deriving the mean of \( \rho^* \), but Sawa's method can be used to obtain the moments of \( \rho^* \) exactly. The first two moments of \( \rho^* \) are given in the following lemma.

**Lemma 3.2.** Let \( Y_{ij} \) be a sequence of independent normal \((\mu_i, \sigma^2)\) random variables. Then the first two moments of \( \rho^* \) are given as

\[
E(\rho^*) = 0,
\]

and

\[
E(\rho^*_2) = \frac{[k(n-1) + 2]^{-1} + 4(n-1)^{-1}[k^2(n-1)^2 - 4]^{-1}}. \] (3.6)
Proof. Let the matrix \( L_n \) be the \( n \times n \) matrix with elements equal to one along the first upper diagonal and zeroes elsewhere. Let the matrix \( B_n \) be the \( n \times n \) matrix with the first \( n - 1 \) diagonal elements equal to one and zeroes elsewhere. Let \( Z_i = (Y_{i1} - \mu_1, Y_{i2} - \mu_2, \ldots, Y_{in} - \mu_n) \), \( i = 0, 1, 2, \ldots, k-1 \), then the least squares estimator given in (3.5) has the representation

\[
\hat{\beta}_{OLS} = \frac{\sum_{i=0}^{k-1} Z_i' (L_n + L_n') Z_i}{\sum_{i=0}^{k-1} Z_i' B_n Z_i},
\]

Each \( Z_i \), \( i = 0 \) to \( k-1 \), is an \( n \times 1 \) random vector which is distributed as \( N(0, \sigma^2) \). Since \( \hat{\beta}_{OLS} \) is distributed independently of \( \sigma^2 \), we can assume \( \sigma^2 = 1 \) without loss of generality. The \( n \times n \) diagonal matrix \( A \) containing the eigenvalues of \( B_n \) is equal to \( B_n \) and the matrix \( C \) is equal to \( 2^{-1} (L_n + L_n') \). Using Theorem A.2, the first moment of \( \hat{\beta}_{OLS} \) is equal to zero since \( C_{jj} = 0 \), \( j = 1, 2, \ldots, n \). The second moment of \( \hat{\beta}_{OLS} \) is computed using equation (A.13) as

\[
E(\hat{\beta}_{OLS}^2) = k \sum_{i=1}^{n} \sum_{j=1}^{n} 2C_{ij} \int_{0}^{\infty} \frac{x dx}{(1+2\lambda_1 x)(1+2\lambda_j x)(1+2x)^{k(n-1)/2}}
\]

\[
= k \left\{ \left( \int_{0}^{\infty} \frac{x dx}{(1+2x)^{[k(n-1)+1]}/2} + \int_{0}^{\infty} \frac{x dx}{(1+2x)^{[k(n-1)+2]}/2} \right) \right\}
\]
\[
\begin{align*}
&= k \left\{ (n-2)[k(n-1)+2]^{-1}[k(n-1)]^{-1} \\
&\quad + [k(n-1)]^{-1}[k(n-1)-2]^{-1} \right\} \\
&= [k(n-1) + 2]^{-1} + 4(n-1)^{-1}[k^2(n-1)^2 - 4]^{-1}.
\end{align*}
\]

In the case of the first-order seasonal autoregressive process with known means, the least squares estimator of \( \rho \) is unbiased when the true value of \( \rho \) is zero; when the means are unknown the least squares estimator of \( \rho \) displays a negative bias which is attributable to the estimation of the means. Exact expressions for the first two moments of \( \hat{\rho}_{OLS} \) are considerably more difficult to derive using the moment generating function technique of Sawa. Useful approximations to the first two moments of \( \hat{\rho}_{OLS} \) are derived based on the work of Marriott and Pope (1954). Some results given by Bartlett (1946) and Parzen (1957) which will be useful in obtaining the moments of \( \hat{\rho}_{OLS} \) are presented.

**Theorem 3.1.** Let \( \{Y_t\} \) be a stationary time series with absolutely summable covariance function. Then

\[
\lim_{n \to \infty} n \text{var}(\overline{Y}_n) = \sum_{h=-\infty}^{\infty} \gamma(h),
\]

where \( \overline{Y}_n \) denotes the sample mean, 
\[
\overline{Y}_n = n^{-1} \sum_{t=1}^{n} Y_t.
\]
Proof. See Fuller (1976, p. 232) or Anderson (1971, p. 459). \[\square\]

For a stationary time series with known mean, an unbiased estimator of the autocovariance at lag \( h \) is

\[
C(h) = (n-h)^{-1} \sum_{t=1}^{n-h} (Y_t - \mu)(Y_{t+h} - \mu) .
\]

(3.8)

In general, the mean \( \mu \) is unknown and an estimator of \( \gamma(h) \) is

\[
\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (Y_t - \overline{Y})(Y_{t+h} - \overline{Y})
\]

(3.9)

which has a bias of order \( n^{-1} \).

Theorem 3.2. Let \( \{Y_t\} \) be a stationary time series defined by

\[
Y_t = \sum_{j=0}^{\infty} \alpha_j e_{t-j} ,
\]

where the sequence \( \{\alpha_j\} \) is absolutely summable and the \( e_t \) are independent \( (0,\sigma^2) \) random variables with \( E(e_t^h) = \eta_0^h \). Given fixed \( h > q > 0 \),

\[
E(\hat{\gamma}(h) - \gamma(h)) = -|h| n^{-1} \gamma(h) - (n+|h|) n^{-1} \text{var}(\overline{Y}) + o(n^{-2}) ,
\]

(3.10)

and
\[ \lim_{n \to \infty} n^2 (n-h)^{-1} \text{cov}(\hat{\gamma}(h), \hat{\gamma}(q)) = (n-3)\gamma(h)\gamma(q) \]

\[ + \sum_{j=-\infty}^{\infty} \gamma(j)\gamma(j-h+q) + \gamma(j+q)\gamma(j-h) \]. \hspace{1cm} (3.11)

**Proof.** See Fuller (1976, p. 239) or Anderson (1971, p. 448). \[ \]

Other estimators of the autocovariance function are handled similarly. The second moments of the sample autocovariance function \( C(h) \) when the mean is known is equal, to order \( n^{-1} \), to the second moments of \( \hat{\gamma}(h) \) given in (3.11). However, the estimation of the mean introduces a bias in the estimators \( \hat{\gamma}(h) \) which can be substantial for small samples. This is the second source of bias in \( \hat{\rho}_{\text{OLS}} \) first described by Orcutt (1948). Using Theorems 3.1 and 3.2 and the results of Appendix B, close approximations to the first two moments of \( \hat{\rho}_{\text{OLS}} \) are obtained. The results are presented in the following theorem.

**Theorem 3.3.** Let \( \{Y_{ij}\} \) be a stationary time series satisfying the stochastic difference equation

\[ Y_{ij} = \theta_i + \rho Y_{i,j-1} + e_{ij} , \]

where the \( e_{ij} \) are independent normal \( (0,\sigma^2) \) random variables and \( |\rho| < 1 \). Then

\[ E(\hat{\rho}_{\text{OLS}}) = \rho - (n-1)^{-1}[1+k^{-1}(k+2)\rho] + O(n^{-2}) \], \hspace{1cm} (3.12)
\[
\text{var}(\hat{\rho}_{OLS}) = [k(n-1)]^{-1}(1-\rho^2) + O(n^{-2}).
\] (3.13)

**Proof.** Let the numerator and denominator of \( \hat{\rho}_{OLS} \) given in (3.3) be \( N = \sum_{i=0}^{k-1} \gamma_{i}^{(1)} \) and \( D = \sum_{i=0}^{k-1} \gamma_{i}^{(0)} \), where

\[
\gamma_{i}^{(1)} = (n-1)^{-1} \sum_{j=2}^{n} (Y_{i,j} - \bar{Y}_{11})(Y_{i,j-1} - \bar{Y}_{11})
\]

and

\[
\gamma_{i}^{(0)} = (n-1)^{-1} \sum_{j=2}^{n} (Y_{i,j} - \bar{Y}_{11})^2, \quad i = 0, 1, 2, \ldots, k-1.
\]

Let \( \gamma(h) \) be the autocovariance function for a first-order stationary autoregressive process with parameter \( \rho \). For a first-order seasonal autoregressive process the random vectors \( \{Y_{11}, Y_{12}, \ldots, Y_{in}\} \), \( i = 0, 1, 2, \ldots, k-1 \), are mutually independent and are realizations from first-order autoregressive processes with common parameter \( \rho \) and differing intercepts \( \theta_{i} \), \( i = 0, 1, 2, \ldots, k-1 \). Therefore, the random vectors \( \{\gamma_{i}^{(0)}, \gamma_{i}^{(1)}\} \), \( i = 0 \) to \( k-1 \), are mutually independent and identically distributed. By Theorem 3.2, \( \gamma_{i}(h) - E(\gamma_{i}(h)) = O_p(n^{-1/2}) \).
h = 0,1, and \( i = 0,1, \ldots, k-1 \). Therefore, \( \text{N-E(N)} = O_p(n^{-\frac{1}{2}}) \) and \\
\( \text{D-E(D)} = O_p(n^{-\frac{1}{2}}) \).

Let \( f(x_1, x_2) = x_1 x_2^{-1} \), where \((x_1, x_2)\) belong to any closed sphere \\
\( S \) about \((\text{E(N)}, \text{E(D)})\) which is bounded away from the line \((x,0): x \in \mathbb{R})\). Then \( f(x_1, x_2) \) is a continuous function with bounded deriva­
atives over the closed sphere \( S \). Let \( X_n = (\text{N,D}) \). By Taylor's 

theorem, \( \hat{\rho}_{\text{OLS}} = f(X_n) \) can be expanded about \((\text{E(N)}, \text{E(D)})\) to give

\[
\hat{\rho}_{\text{OLS}} = \frac{\text{E(N)}}{\text{E(D)}} + \frac{\text{N-E(N)}}{\text{E(D)}} - \frac{\text{E(N)}[\text{D-E(D)}]}{[\text{E(D)}]^2} - \frac{[\text{N-E(N)}][\text{D-E(D)}]}{[\text{E(D)}]^2} \\
+ \frac{\text{E(N)}[\text{D-E(D)}]^2}{[\text{E(D)}]^3} + O_p(n^{-3/2}).
\]

(3.14)

The integrability of the least squares estimator has been established by 
Fuller and Hasza (1981) so that the first two moments of \( \hat{\rho}_{\text{OLS}} \) exist 
and are

\[
\text{E}[\hat{\rho}_{\text{OLS}}] = \frac{\text{E(N)}}{\text{E(D)}} - \frac{\text{cov}(\text{N,D})}{[\text{E(D)}]^2} + \frac{\text{E(N)}\var(D)}{[\text{E(D)}]^3} + O(n^{-2}),
\]

(3.15)

and

\[
\text{var}[\hat{\rho}_{\text{OLS}}] = \frac{\text{var}^2(N)}{[\text{E(D)}]^2} + \frac{[\text{E(N)}]^2}{[\text{E(D)}]^4} \text{var}(D) - \frac{2\text{E(N)}\text{cov}(\text{N,D})}{[\text{E(D)}]^6} \\
+ \frac{2}{[\text{E(D)}]^8} + O(n^{-2}).
\]

(3.16)
Expressions for the first two moments of $N$ and $D$ are obtained using Theorems 3.1 and 3.2 which are correct to terms of order $n^{-1}$.

Using Theorem 3.1, the variance of $\tilde{Y}_{11}$ is

$$\text{var}(\tilde{Y}_{11}) = (n-1)^{-1} \sum_{j=-\infty}^{\infty} \gamma(j) + o(n^{-2})$$

$$= (n-1)^{-1}(1-p)^{-2} \sigma^{2} + o(n^{-2}) \quad (3.17)$$

From equation (3.10) of Theorem 3.2 and (3.17), the means of $\tilde{\gamma}_{1}(0)$ and $\tilde{\gamma}_{1}(1)$ are

$$E\{\tilde{\gamma}_{1}(h)\} = \gamma(h) - (n-1)^{-1}(1-p)^{-2} \sigma^{2} + o(n^{-2})$$

$h = 0,1$ and $i = 0,1,2,\ldots, k-1$. Under normality, $E\{e_{i}^{h}\} = 3\sigma^{h}$ and the second moments of $\tilde{\gamma}_{1}(0)$ and $\tilde{\gamma}_{1}(1)$ are computed using (3.11) of Theorem 3.2 with $\eta = 3$. A general procedure for obtaining expressions for the covariances of sample autocovariances is given in Appendix B.

From equations (B.9), (B.10), and (B.11) of Appendix B with $\alpha_{1} = \rho$ and $\alpha_{2} = 0$, the second moments of $[\tilde{\gamma}_{1}(0), \tilde{\gamma}_{1}(1)]$ for normal $Y_{1j}$ are

$$\text{cov}\{\tilde{\gamma}_{1}(0), \tilde{\gamma}_{1}(1)\} = (n-1)^{-1}(1-p^{2})^{-1} 4\rho \gamma^{2}(0) + o(n^{-2})$$

$$\text{var}\{\tilde{\gamma}_{1}(0)\} = (n-1)^{-1}(1-p^{2})^{-1} 2(1+p^{2}) \gamma^{2}(0) + o(n^{-2})$$
and

\[ \text{var}(\gamma_i(1)) = (n-1)^{-1}(1-p^2)^{-1}(1+q^2-p^4)\gamma^2(0) + O(n^{-2}) . \]

From these results and the independence of \([\gamma_i(0), \gamma_i(1)]\), \(i = 0, 1, \ldots, k-1\), the first two moments of \((N,D)\) are

\[
\begin{align*}
E(N) & = E\{k^{-1} \sum_{i=0}^{k-1} \gamma_i(1)\} \\
& = \gamma(1) - (n-1)^{-1}(1-p)^{-1}(1+p)\gamma(0) + O(n^{-2}) , \\
E(D) & = E\{k^{-1} \sum_{i=0}^{k-1} \gamma_i(0)\} \\
& = \gamma(0) - (n-1)^{-1}(1-p)^{-1}(1+p)\gamma(0) + O(n^{-2}) , \\
\text{var}(N) & = \text{var}\{k^{-1} \sum_{i=0}^{k-1} \gamma_i(1)\} \\
& = [k(n-1)]^{-1}(1-p^2)^{-1}(1+q^2-p^4)\gamma^2(0) + O(n^{-2}) , \\
\text{var}(D) & = \text{var}\{k^{-1} \sum_{i=0}^{k-1} \gamma_i(0)\} \\
& = [k(n-1)]^{-1}(1-p^2)^{-1}2(1+p^2)\gamma^2(0) + O(n^{-2}) ,
\end{align*}
\]
\[ \text{cov}(N,D) = \text{cov}\{k^{-1} \sum_{i=0}^{k-1} \gamma_i(1) , k^{-1} \sum_{i=0}^{k-1} \gamma_i(0)\} \]

\[ = [k(n-1)]^{-1}(1-p^2)^{-1} 4p \gamma^2(0) + O(n^{-2}) . \quad (3.22) \]

From (3.18) and (3.19) the first term in the approximate mean of \( \hat{\rho}_{OLS} \) given in (3.15) is

\[ E(N)[E(D)]^{-1} = \rho - (n-1)^{-1}(1+p) + O(n^{-2}) . \quad (3.23) \]

The other two terms of (3.13) are evaluated using equations (3.18) to (3.22) and neglecting terms of order \( n^{-2} \) as

\[ -\text{cov}(N,D) \quad (E(D))^2 \]

\[ \frac{-\text{cov}(N,D)}{[E(D)]^2} = -[k(n-1)]^{-1}(1-p^2)^{-1} 4p + O(n^{-2}) , \quad (3.24) \]

and

\[ \frac{E(N)\text{var}(D)}{[E(D)]^3} = [k(n-1)]^{-1}(1-p^2)^{-1} 2p(1+p^2) + O(n^{-2}) . \quad (3.25) \]

From (3.23), (3.24) and (3.25) the approximate mean of \( \hat{\rho}_{OLS} \) correct to order \( n^{-1} \) is

\[ E[\hat{\rho}_{OLS}] = \rho - (n-1)^{-1}[1 + (k+2)k^{-1}p] + O(n^{-2}) . \]
Similarly, the variance of $\hat{\rho}_{OLS}$ using (3.14) and (3.16) to (3.20) is, after some algebraic simplification,

$$\text{var}(\hat{\rho}_{OLS}) = [k(n-1)]^{-1}(1-\rho^2) + O(n^{-2}).$$

From Lemmas 3.1 and 3.2, the means of $\hat{\rho}_{OLS}$ and $\rho_{OLS}^*$ when $\rho = 0$ differed by terms of order $n^{-1}$. The approximate mean and variance of $\rho_{OLS}^*$ can be obtained in an analogous manner.

**Corollary 3.3.** Let the assumptions of Theorem 3.3 be satisfied with $\theta_i = 0$, $i = 0, 1, \ldots, k-1$. Then

$$E(\rho_{OLS}^*) = \rho - [k(n-1)]^{-1}2\rho + O(n^{-2}) \quad (3.26)$$

and

$$\text{var}(\rho_{OLS}^*) = [k(n-1)]^{-1}(1-\rho^2) + O(n^{-2}). \quad (3.27)$$

**Proof.** Let the numerator and denominator of $\rho_{OLS}^*$ given in (3.5) be $N^* = k^{-1} \sum_{i=0}^{k-1} Y_i(1)$ and $D^* = k^{-1} \sum_{i=0}^{k-1} \hat{Y}_i(0)$, where $\hat{Y}_i(h) = \sum_{j=1}^{n-h} Y_{ij}Y_{i,h+j}$, $h = 0, 1$ and $i = 0, 1, 2, \ldots, k-1$. By the arguments used in the proof of Theorem 3.3, $\rho_{OLS}^*$ can be expanded in a Taylor series about $[E(N^*), E(D^*)]$ and the first two moments of $\rho_{OLS}^*$ are analogous to (3.15) and (3.16). Now $E(N^*) = \gamma(1)$ and $E(D^*) = \gamma(0)$ and the second moments of $(N^*, D^*)$ are the same as the second moments.
of \((N,D)\) to terms of order \(n^{-1}\). The results follow from the proof of Theorem 3.3.

When \(p = 0\), the approximate means of \(\hat{\rho}_{\text{OLS}}^*\) and \(\rho_{\text{OLS}}^*\) are in agreement to terms of order \(n^{-1}\) with the exact results given in Lemmas 3.1 and 3.2. Also, the approximate variance of \(\rho_{\text{OLS}}^*\) differs from the exact variance of \(\rho_{\text{OLS}}^*\) given in (3.6) by terms of order \((nk)^{-2}\). From these observations, the approximate moments of \(\hat{\rho}_{\text{OLS}}^*\) and \(\rho_{\text{OLS}}^*\) are expected to be satisfactory for values of \(p\) near zero.

Copas (1966) and Orcutt and Winokur (1969) considered inverting equation (3.12) to obtain a nearly unbiased estimator of \(\rho\) in the first-order autoregressive process \((k=1)\) with unknown mean. Applying the same approach to the first-order seasonal autoregressive process, the adjusted least squares estimator of \(\rho\) which is unbiased to order \(n^{-2}\) is

\[
\hat{\rho}_{\text{MP}} = \left[1 - (n-1)^{-1} k^{-1}(k+2)\right]^{-1}\left[\hat{\rho}_{\text{OLS}} + (n-1)^{-1}\right],
\]

\[
\hat{\rho}_{\text{OLS}} \in [-1 + 2(n-1)^{-1} k^{-1}, 1 - 2(n-1)^{-1} k^{-1}(k+1)]
\]

\[
= 1 \text{ with } \hat{\rho}_{\text{OLS}} \geq 1 - 2(n-1)^{-1} k^{-1}(k+1)
\]

\[
= -1 \text{ with } \hat{\rho}_{\text{OLS}} \leq -1 + 2(n-1)^{-1} k^{-1}.
\] (3.28)

The estimator is denoted by \(\hat{\rho}_{\text{MP}}^*\) because it is based on a rather direct application of the approximate mean of \(\hat{\rho}_{\text{OLS}}\) originally derived by Marriott and Pope (1954). By Lemma 3.1, the unconstrained version of \(\hat{\rho}_{\text{MP}}^*\)
is exactly unbiased when the true value of \( \rho \) is equal to zero. Since \( \hat{\rho}_{MP} \) is a multiple of \( \hat{\rho}_{OLS} \), where the multiplier is greater than one, the variance of \( \hat{\rho}_{MP} \) is larger than the variance of \( \hat{\rho}_{OLS} \).

When the means are known, the bias of \( \hat{\rho}_{OLS}^* \) is of the order \( (nk)^{-1} \). When the means are unknown and estimated, the bias of \( \hat{\rho}_{OLS}^* \) is considerably larger than the bias of \( \hat{\rho}_{OLS} \) and the difference in biases can be attributed to the estimation of the means. From the derivation of the approximate mean of \( \hat{\rho}_{OLS} \), the portion of the bias of \( \hat{\rho}_{OLS}^* \) due to the sample means can be accounted for by obtaining nearly unbiased estimators for the numerator and denominator of \( \rho \). From equations (3.18) and (3.19), the biases of \( N \) and \( D \) are given as

\[
E(N-\gamma(1)) = -(n-1)^{-1}(1-\rho)^{-1}(1+\rho)\gamma(0) + O(n^{-2}) ,
\]

and

\[
E(D-\gamma(0)) = -(n-1)^{-1}(1-\rho)^{-1}(1+\rho)\gamma(0) + O(n^{-2}) .
\]

By substituting \( \hat{\rho}_{OLS} \) and \( D \) for \( \rho \) and \( \gamma(0) \) in the expressions for the biases, nearly unbiased estimators of \( \gamma(1) \) and \( \gamma(0) \) are obtained as

\[
\hat{\gamma}(1) = N + (n-1)^{-1}(1-\hat{\rho}_{OLS})^{-1}(1+\hat{\rho}_{OLS})D , \quad |\hat{\rho}_{OLS}| < 1 \quad (3.29)
\]

\[
= N , \quad |\hat{\rho}_{OLS}| \geq 1
\]
and

$$\tilde{\gamma}(0) = D + (n-1)^{-1}(1-\hat{\rho}_{OLS})^{-1}(1+\hat{\rho}_{OLS})D, \quad |\hat{\rho}_{OLS}| < 1 \quad (3.30)$$

$$= D \quad , \quad |\hat{\rho}_{OLS}| \geq 1 .$$

The modified least squares estimator of $\rho$ is obtained as the ratio of $\tilde{\gamma}(1)$ and $\tilde{\gamma}(0)$. For computational purposes, the estimator has the form

$$\hat{\rho}_1 = \frac{\hat{\rho}_{OLS}(1-\hat{\rho}_{OLS}) + (n-1)^{-1}(1+\hat{\rho}_{OLS})}{(1-\hat{\rho}_{OLS}) + (n-1)^{-1}(1+\hat{\rho}_{OLS})}, \quad \hat{\rho}_{OLS} \in (-1, 1)$$

$$= 1 \quad , \quad \hat{\rho}_{OLS} \geq 1$$

$$= -1 \quad , \quad \hat{\rho}_{OLS} \leq -1 . \quad (3.31)$$

Let $V = (n-1)^{-1}(1-\hat{\rho}_{OLS})^{-1}(1+\hat{\rho}_{OLS})D$. Then

$$V = (n-1)^{-1}(1-\hat{\rho})^{-1}(1+\hat{\rho})\gamma(0) + \hat{\rho}(n^{-3/2}) . \quad (3.32)$$

From (3.18), (3.19) and (3.32), $E(\tilde{\gamma}(0)) = \gamma(0) + (n^{-2})$ and $E(\tilde{\gamma}(1)) = \gamma(1) + O(n^{-2})$. Also, the second moments of $[\tilde{\gamma}(1), \gamma(0)]$ are the same as the second moments of $(N,D)$ to terms of order $n^{-1}$.

In particular, the first and second moments of the random vector $[\tilde{\gamma}(1), \gamma(0)]$ are the same as the first and second moments of $(N^*,D^*)$ to terms of order $n^{-1}$. By considering expansions of the form (3.15)
and (3.16), it is noted that the first two moments of $\hat{\rho}_1$ are equal to the first two moments of $\rho_{OLS}$, to order $n^{-1}$. Taking the difference between $\hat{\rho}_1$ and $\hat{\rho}_{OLS}$ gives

$$\hat{\rho}_1 - \hat{\rho}_{OLS} = \frac{(n-1)^{-1}(1+\rho_{OLS})(1-\rho_{OLS})}{(1-\rho_{OLS}) + (n-1)^{-1}(1+\rho_{OLS})}$$

$$= o_p(n^{-1}).$$

Hence, the estimators $\hat{\rho}_1$ and $\hat{\rho}_{OLS}$ are asymptotically equivalent and the limiting distribution of $n^{1/2}(\hat{\rho}_{OLS} - \rho)$ is equal to the limiting distribution of $n^{1/2}(\hat{\rho}_1 - \rho)$.

Sastry (1951), Box and Jenkins (1976, p. 200) and Bora-Senta and Kounias (1980) considered using $E(N)[E(D)]^{-1}$ as an approximation to $E(\hat{\rho}_{OLS})$ in the case of a first-order autoregressive process, but this approximation is not very appropriate for values of $\rho$ near $\pm 1$. This approximation only accounts for the bias attributed to the estimation of the mean and ignores the first source of bias first described by Orcutt (1948). In the case of the first-order seasonal autoregressive process of period $k$, it is a reasonable approximation since the last two terms of (3.15) are of the order $(nk)^{-1}$ instead of $n^{-1}$ in the case of the first-order autoregressive process. The value of $k$ is often as large as $n$ in many economic time series with monthly observations.
Another modified least squares estimator of $\rho$ which is suggested by the derivation of the bias of $\hat{\rho}_{OLS}$ is

$$\hat{\rho}_2 = \hat{\rho}_{OLS}[1+(n-2)^{-1}] + (n-2)^{-1}, \quad \hat{\rho}_{OLS} \in (-1, 1-2(n-1)^{-1})$$

$$= 1 \quad \hat{\rho}_{OLS} \geq 1-2(n-1)^{-1}$$

$$= -1 \quad \hat{\rho}_{OLS} \leq -1 . \quad (3.33)$$

This estimator is obtained by approximating the mean of $\hat{\rho}_{OLS}$ by $E(N)[E(D)]^{-1}$. Bora-Senta and Kounias (1980) considered a similar estimator of $\rho$ in the case of the first-order autoregressive process; however, as noted earlier the approximation ignores terms of order $n^{-1}$ and is poor for values of $\rho$ near unity. Since $\hat{\rho}_2$ is a linear function of $\hat{\rho}_{OLS}$ with multiplier greater than one, the variance of $\hat{\rho}_2$ is larger than the variance of $\hat{\rho}_{OLS}$. However, the difference is $O(n^{-2})$. The approximate mean of $\hat{\rho}_2$ is equal to the approximate mean of $\rho^*$ given in (3.26).

The various estimators of $\rho$ considered so far are asymptotically equivalent and have the same variance to order $n^{-1}$. None of these four estimators is guaranteed to satisfy the stationary restriction that $\hat{\rho} \in (-1, 1)$. Burg (1967, 1968) proposed a method of estimating autoregressive parameters which guarantees that the estimators satisfy the stationary restrictions. The form of the estimator for the first-order seasonal autoregressive process is
\[
\hat{\rho}_{\text{SYM}} = \frac{[k(n-1)]^{-1} \sum_{i=0}^{n-1} \sum_{j=2}^{k-1} (Y_{i,j} - \overline{Y}_i)(Y_{i,j-1} - \overline{Y}_i)}{[2k(n-1)]^{-1} \left( \sum_{i=0}^{n-1} \sum_{j=2}^{k-1} (Y_{i,j} - \overline{Y}_i)^2 + \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j-1} - \overline{Y}_i)^2 \right)},
\]

(3.34)

where \( \overline{Y}_i \) denotes the \( i \)-th sample mean, \( n^{-1} \sum_{j=1}^{n} Y_{i,j} \). The technique of Marriott and Pope (1954) and Kendall (1954) can be used to show the first two moments of \( \hat{\rho}_{\text{SYM}} \) agree with the first two moments of \( \hat{\rho}_{\text{OLS}} \), to order \( n^{-1} \).

The conditional maximum likelihood estimator of \( \rho \), conditional on the initial \( k \) observations, of a normal first-order seasonal autoregressive process is the same as the least squares estimator of \( \rho \). For a zero mean normal stationary first-order autoregressive process, a closed-form expression for the maximum likelihood estimator of \( \rho \) exists and is given in (2.37). In the case of a normal stationary first-order seasonal autoregressive process with zero means, the maximum likelihood estimator of \( \rho \) is

\[
\rho_{\text{MLE}} = \frac{2}{3} a_3^{-1} (a_2^2 - 3a_1 a_3)^{1/2} \cos \theta - \frac{1}{3} a_2 a_3^{-1} ,
\]

(3.35)
\[ e = \frac{1}{3} \arccos\left[-\frac{1}{2}(2a_2^3 - 9a_1a_2a_3 + 27a_0a_3^2)(a_2^2 - 3a_1^2)\right]^{3/2} + \frac{4\pi}{3}, \]

\[ a_0 = \sum_{i=0}^{k-1} \sum_{j=2}^{n} Y_{ij} Y_{i},j-1, \]

\[ a_1 = -\sum_{i=0}^{k-1} \sum_{j=2}^{n} Y_{ij}^2 - n^{-1} \sum_{i=0}^{k-1} \sum_{j=1}^{n} Y_{ij}^2, \]

\[ a_2 = -(n-2)n^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} Y_{ij} Y_{i},j-1, \]

and

\[ a_3 = (n-1)n^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} Y_{ij}^2. \]

By construction, the maximum likelihood estimator satisfies the stationary restriction that \( \rho_{\text{MLE}} \in (-1, 1) \). When the means are nonzero unknown parameters to be estimated, no closed-form expression for the maximum likelihood estimator of \( \rho \) is possible. Recursive algorithms such as the Newton-Raphson and scoring iterative procedures are used to obtain the maximum likelihood estimator. See Anderson (1971). An approximate maximum likelihood estimator of \( \rho \) is obtained by substituting \( Y_{ij} - \overline{Y}_i \).
for $Y_{ij}$ in (3.35) and is denoted by $\hat{\rho}_{\text{MLE}}$. This estimator can be shown to satisfy the stationary restriction. The proof follows Hasza (1980).

**Lemma 3.3.** Let $Y_{ij}$, $i = 0,1,\ldots,k-1$, and $j = 1,2,\ldots,n$, be $nk \geq 3$ observations on a normal stationary first-order seasonal autoregressive process with unknown means. Then with probability one, the estimator $\hat{\rho}_{\text{MLE}}$ lies in $(-1,1)$.

**Proof.** Let $a_0, a_1, a_2$ and $a_3$ be defined as in (3.35) with $Y_{ij} - \bar{Y}_i$ replacing $Y_{ij}$. Let $g(p)$ be given by

$$g(p) = a_3p^3 + a_2p^2 + a_1p + a_0 .$$

Then with probability one

$$\lim_{\rho \to \infty} g(p) = +\infty,$$

$$\lim_{\rho \to -\infty} g(p) = -\infty,$$

$$g(+1) = -n^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - Y_{i,j-1})^2 < 0 ,$$

$$g(-1) = n^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} + Y_{i,j-1})^2 > 0 .$$
It follows that \( g(\rho) = 0 \) has three real roots; one in each of the intervals \((-\infty, -1), (-1, 1)\) and \((1, \infty)\). The roots of \( g(\rho) \) are expressed as

\[
r_j = 2^{a_j^{1/2}} \cos \frac{1}{3} \arccos \left[ \frac{b_j}{2} \left( \frac{3}{a_j} \right)^{3/2} \right] + \frac{2(j-1)\pi}{3} - \frac{p}{3},
\]

\( j = 1, 2, 3, \)

where \( a = 3^{-1}(p^2 - 3q) \), \( b = -27^{-1}(2p^3 - 9pq + 27r) \), \( p = a_0a_3^{-1} \),
\( q = a_1a_3^{-1} \) and \( r = a_2a_3^{-1} \). For \( x \in [-1, 1] \), \( \arccos(x) \in [0, \pi] \) and
\[
\cos \left( \frac{2\pi}{3} + \frac{1}{3} \arccos(x) \right) \leq \cos \left( \frac{4\pi}{3} + \frac{1}{3} \arccos(x) \right) \leq \cos \left( \frac{\pi}{3} \arccos(x) \right).
\]

Therefore, \( r_2 \leq r_3 \leq r_1 \). The result follows because \( \hat{\rho}_{\text{MLE}} \) is equal to \( r_3 \).

It is observed from the proof of Lemma 3.3 that other estimators of \( \mu_i \), \( i = 0, 1, \ldots, k-1 \), can be used in constructing an approximate maximum likelihood estimator of \( \rho \) which satisfies the stationary restriction that \( \rho \in (-1, 1) \).
CHAPTER IV. ESTIMATORS FOR THE PARAMETERS OF HIGHER ORDER SEASONAL AUTOREGRESSIVE MODELS

The methods of least squares and maximum likelihood estimation are commonly used to construct estimators for the parameter of a stationary normal first-order seasonal autoregressive process with zero means. The least squares procedure easily extends to include higher order autoregressive processes with unknown means. However, even in the simplest case of a first-order autoregressive process with unknown mean, no closed analytical form for the maximum likelihood estimator exists. The complexity of the equations obtained by setting the derivatives of the likelihood function equal to zero increases with the order of the autoregressive process. The small sample properties of the least squares estimators have received considerable attention in the case of the first-order autoregressive process. However, the small sample properties of the least squares estimators for the parameters of higher order autoregressive processes have received little attention. Salem (1971) extended the method of Marriott and Pope (1954) to obtain expressions for the approximate biases of the least squares estimators. Bora-Senta and Kounias (1980) considered an iterative method of moments procedure as an alternative to the least squares estimation procedure for higher order autoregressive processes. In this chapter the approximate moments of the least squares estimators for the parameters of a stationary normal second-order seasonal autoregressive process with unknown means are derived. Modifications to the least squares estimators for the autoregressive coefficients which correct for the least squares biases are proposed.
Consider the stationary normal p-th order autoregressive process \( \{Y_t\} \) which satisfies the stochastic difference equation

\[
Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + e_t,
\]

where \( \{e_t\} \) is a sequence of independent normal \((0, \sigma^2)\) random variables. It is assumed that the roots of the characteristic equation,

\[
m^p - \alpha_1 m^{p-1} - \ldots - \alpha_p = 0,
\]

are less than unity in modulus. The Yule-Walker equations associated with the p-th order autoregressive process are given in (2.31) and are denoted by \( H_\alpha = N \), where \( \alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_p) \), \( N' = [\gamma(1), \gamma(2), \ldots, \gamma(p)] \) and

\[
H = \begin{pmatrix}
\gamma(0) & \gamma(1) & \ldots & \gamma(p-1) \\
\gamma(1) & \gamma(0) & & \\
& & \ddots & \\
\gamma(p-1) & \gamma(p-2) & & \gamma(0)
\end{pmatrix},
\]

The matrix \( H \) is a nonsingular \( p \times p \) matrix and \( N \) is a \( p \times 1 \) vector whose elements are the autocovariances of \( Y_t \).

By Theorem 3.2, consistent estimators for \( H \) and \( N \) are available and are denoted by \( \hat{H} \) and \( \hat{N} \), respectively. Let \( A \) be the \( p \times p \) matrix given by \( A = \hat{H} - H \) and let \( d \) be the \( p \times 1 \) vector given by \( d = \hat{N} - N \). Then \( E(A) = o(n^{-1}) \), \( E(d) = o(n^{-1}) \) and \( E(A^2) = o(n^{-1}) \), \( E(dd') = o(n^{-1}) \), and \( E(Ad) = o(n^{-1}) \). A common solution to the estimation problem would take \( \hat{\alpha} = H^{-1} \hat{N} \) as the estimator for \( \alpha \).
The method of obtaining an approximate expression for the moments of the least squares estimator for the first-order seasonal autoregressive coefficient is generalized to include higher order autoregressive processes.

Let \( f(x) = [H(x)]^{-1} N(x) \), where \( x' = (x_0, x_1, \ldots, x_p) \), \( N'(x) = (x_1, x_2, \ldots, x_p) \) and

\[
H(x) = \begin{pmatrix}
  x_0 & x_1 & \cdots & x_{p-1} \\
  x_1 & x_0 & x_{p-2} & \vdots \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{p-1} & x_{p-2} & \cdots & x_0
\end{pmatrix}
\]

so that \( x \) belongs to any closed sphere \( \mathcal{S} \) about \( r' = [\gamma(0), \gamma(1), \ldots, \gamma(p)] \) which is bounded away from the region \( \{x \in \mathbb{R}^{p+1} : H'(x) \text{ is not positive definite}\} \). Then \( f(x) \) is a \( p \times 1 \) vector whose elements are continuous functions with bounded derivatives over the closed sphere \( \mathcal{S} \). Let \( x'_n = (\hat{h}_{11}, \hat{h}_{12}, \ldots, \hat{h}_{1p}, \hat{n}_p) \), where \( \hat{h}_{ij} \) is the \((i,j)\)-th element of \( \hat{H} \) and \( \hat{n}_p \) is the \( p \)-th element of \( \hat{N} \). By Taylor's theorem, \( \hat{a} = \hat{H}^{-1} \hat{N} \) can be expanded about \( \Gamma \) to give

\[
\hat{a} = (H + A)^{-1}(N + d)
\]

\[
= H^{-1}(I - AH^{-1} + AH^{-1}AH^{-1} + \mathcal{O}_p(n^{-3/2}))N + d
\]

\[
= a - H^{-1} a + H^{-1} d - H^{-1} d + \mathcal{O}_p(n^{-3/2}) \quad (4.3)
\]
The integrability of the least squares estimators for normal autoregressive parameters has been established by Fuller and Hasza (1981) so that the first two moments of $\hat{\alpha}$ exist. From (4.3) the mean of $\hat{\alpha}$ is

$$E(\hat{\alpha}) = \alpha + E(-H^{-1}A\alpha + H^{-1}AH^{-1}A\alpha + H^{-1}d - H^{-1}AH^{-1}d) + O(n^{-2}).$$

(4.4)

Bartlett (1966) obtained the covariance matrix of $\hat{\alpha}$ to be

$$E((\hat{\alpha} - \alpha)(\hat{\alpha} - \alpha)') = n^{-1}H^{-1} \sigma^2 + O(n^{-2}).$$

(4.5)

Consider the stationary normal second-order seasonal autoregressive process with seasonal means and period $k$ satisfying the stochastic difference equation

$$Y_{ij} = \theta_i + \alpha_1 Y_{i,j-1} + \alpha_2 Y_{i,j-2} + e_{ij},$$

(4.6)

where $\{e_{ij}\}$ is a sequence of independent normal $(0, \sigma^2)$ random variables. It is assumed that the roots of the characteristic equation, $\lambda^2k - \lambda_1^{(m)} - \lambda_2 = 0$, are less than unity in modulus. Given a realization of $nk$ observations from model (4.6), the least squares estimator for $\hat{\alpha}' = (\hat{\alpha}_1, \hat{\alpha}_2)$ is $\hat{\alpha}_{OLS} = H^{-1}\hat{\eta}$, where
\[ \hat{N} = [k(n-2)]^{-1} \]

\[
\begin{align*}
&\left[ k-1 \sum_{i=0}^{n} \sum_{j=3}^{n} (Y_{i,j-1} - \bar{Y}_{i1})^2 \\
&\quad + k-1 \sum_{i=0}^{n} \sum_{j=3}^{n} (Y_{i,j-1} - \bar{Y}_{i1})(Y_{i,j-2} - \bar{Y}_{i2}) \right] \\
&\quad + k-1 \sum_{i=0}^{n} \sum_{j=3}^{n} (Y_{i,j-1} - \bar{Y}_{i1})(Y_{i,j-2} - \bar{Y}_{i2})^2 \\
&\quad \text{and } Y_{i,j-h} = (n-2)^{-1} \sum_{j=3}^{n} Y_{i,j-h}, \ h = 0,1,2 . \ \text{Let} \\
&\hat{\gamma}(h) = [k(n-2)]^{-1} \sum_{i=0}^{n} \sum_{j=3}^{n} (Y_{i,j} - \bar{Y})(Y_{i,j-h} - \bar{Y}_{ih}), \ h = 0,1,2 . \\
&\text{Then } \hat{N}' = [\hat{\gamma}(1), \hat{\gamma}(2)] \text{ and} 
\end{align*}
\]
The approximate means of the least squares estimators for the parameters of the normal second-order seasonal autoregressive process with period $k$ are given in the following theorem.

**Theorem 4.1.** Let $\{Y_{ij}\}$ be a stationary time series satisfying the stochastic difference equation

$$Y_{ij} = \theta_1 + \alpha_1 Y_{i-1,j-1} + \alpha_2 Y_{i-2,j-2} + e_{ij},$$

where the $e_{ij}$ are independent normal $(0, \sigma^2)$ random variables and the roots of the polynomial equation, $m^{2k} - \alpha_1^m k - \alpha_2 = 0$, are less than unity in modulus. Then

$$E\{\hat{\alpha}_{1,OLS}\} = \alpha_1 - (n-2)^{-1}(1+\alpha_2) + [k(n-2)]^{-1}[1-\rho^2(1)]^{-3}\rho(1)$$

$$\cdot (1+\alpha_2)^{-1}[-1+7\alpha_2^2+37\alpha_2^4+4\alpha_2^5]\rho^2(1)$$

$$-3(1-\alpha_2^2)\rho^4(1) + (1-\alpha_2^2)\rho^6(1) + O(n^{-2}), \quad (4.10)$$

$$E\{\hat{\alpha}_{2,OLS}\} = \alpha_2 - (n-2)^{-1}(1+\alpha_2) + [k(n-2)]^{-1}[1-\rho^2(1)]^{-3}$$

$$\cdot [-1+3\alpha_2] + [-12(1+\alpha_2)^{-1} + 15 - 3\alpha_2 - 2\alpha_2^2]\rho^2(1)$$
\[-[3 + 9\alpha_2 + 2\alpha_2^2]\rho^h(1) + [1 + 3\alpha_2]\rho^6(1)\]

\[+ O(n^{-2}), \quad (4.11)\]

where \(\rho(1) = \alpha_1 (1-\alpha_2)^{-1}\).

**Proof.** Let \(\gamma(h)\) be the autocovariance function for a stationary normal second-order autoregressive process. Using Theorem 3.1, \(\text{var}(\bar{Y}_{10})\)

\[= (n-2)^{-1} \sum_{h=-\infty}^{\infty} \gamma(h) + O(n^{-2}) . \quad (4.12)\]

For a stationary second-order autoregressive process,

\[\sum_{h=-\infty}^{\infty} \gamma(h) = \gamma(0)(1+\alpha_2)(1-\alpha_2+\alpha_1)((1-\alpha_2)(1-\alpha_2-\alpha_1))^{-1}. \quad (4.12)\]

Using (3.10) of Theorem 3.2, the means of \(\hat{H}\) and \(\hat{N}\) defined in (4.7) and (4.8) are

\[E(\hat{H}) = \begin{bmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{bmatrix} - \text{var}(\bar{Y}_{10}) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + O(n^{-2}), \quad (4.13)\]

and
\[
E(\hat{N}) = \begin{bmatrix}
\gamma(1) \\
\gamma(2)
\end{bmatrix}
- \text{var}\{Y_{10}\}
\begin{bmatrix}
1 \\
1
\end{bmatrix}
+ O(n^{-2}) .
\] (4.14)

Also, \( E\{A\} = E\{\hat{H} - H\} \) and \( E\{d\} = E\{\hat{N} - N\} \).

The second moments of \([\hat{\gamma}(0), \hat{\gamma}(1), \hat{\gamma}(2)]\) are derived using (3.11) of Theorem 3.2 and the computations are given in Appendix B. Let

\[
A(0) = \sum_{h=0}^{\infty} n^2(h) \quad \text{and} \quad A(1) = \sum_{h=0}^{\infty} n(1) n(h+1) .
\]

Then

\[
\text{var}\{\hat{\gamma}(0)\} = 2[k(n-2)]^{-1}[2A(0) - \gamma^2(0)] + O(n^{-2}) ,
\] (4.15)

\[
\text{var}\{\hat{\gamma}(1)\} = [k(n-2)]^{-1}[2(1+\alpha_2)A(0) + 2\alpha_1 A(1) + \gamma^2(1) - \gamma^2(0)] + O(n^{-2}) ,
\] (4.16)

\[
\text{cov}\{\hat{\gamma}(0), \hat{\gamma}(1)\} = 4[k(n-2)]^{-1}A(1) + O(n^{-2}) ,
\] (4.17)

\[
\text{cov}\{\hat{\gamma}(1), \hat{\gamma}(2)\} = 2[k(n-2)]^{-1}[\alpha_1 \alpha_2 A(0) + (1+\alpha_2+\alpha_1^2)A(1) + \alpha_1 \gamma^2(1) + \alpha_2 \gamma(0) \gamma(1)] + O(n^{-2}) ,
\] (4.18)

\[
\text{cov}\{\hat{\gamma}(0), \hat{\gamma}(2)\} = 2[k(n-2)]^{-1}[2\alpha_2 A(0) + 2\alpha_1 A(1) + \gamma^2(1)] + O(n^{-2}) .
\] (4.19)

Let \( G = E(AH^{-1}A) \) and \( D = E(AH^{-1}d) \). Let the elements of \( G \) be \( G_{ij} \) and the elements of \( H^{-1} \) be \( h_{ij} \), where
From (4.15) to (4.20) it follows that

\[ G_{11} = h^{11}[\text{var}(\hat{\gamma}(0)) + \text{var}(\hat{\gamma}(1))] + 2h^{12}\text{cov}(\hat{\gamma}(0), \hat{\gamma}(1)) \]

\[ = [\gamma^2(0) - \gamma^2(1)]^{-1}[k(n-2)]^{-1}\gamma(0)[(6+2\alpha_2)A(0) \]

\[ + [2\alpha_1 - 8\rho(1)]A(1) + \gamma^2(1) - 3\gamma^2(0)] + O(n^{-2}) \]  

and

\[ G_{12} = h^{12}[\text{var}(\hat{\gamma}(0)) + \text{var}(\hat{\gamma}(1))] + 2h^{11}\text{cov}(\hat{\gamma}(0), \hat{\gamma}(1)) \]

\[ = [\gamma^2(0) - \gamma^2(1)]^{-1}[k(n-2)]^{-1}\gamma(0)[-\rho(1) (6+2\alpha_2)A(0) \]

\[ + [8 - 2\alpha_1 \rho(1)]A(1) - \rho(1)[\gamma^2(1) - 3\gamma^2(0)]] + O(n^{-2}) \]  

By symmetry, it is observed that \( G_{11} = G_{22} \) and \( G_{21} = G_{12} \). Let the elements of \( D \) be \( D_1 \). Then

\[ D_1 = h^{11}[\text{cov}(\hat{\gamma}(0), \hat{\gamma}(1)) + \text{cov}(\hat{\gamma}(1), \hat{\gamma}(2))] \]

\[ + h^{12}[\text{var}(\hat{\gamma}(1)) + \text{cov}(\hat{\gamma}(0), \hat{\gamma}(2))] \]
\[
\begin{align*}
&= [\gamma^2(0) - \gamma^2(1)]^{-1} [k(n-2)]^{-1} \gamma(0) [2a_1 a_2 - \rho(1) (2+6a_2)] A(0) \\
&\quad + [6+2a_1^2 + 2a_2 - 6a_1 \rho(1)] A(1) + \rho(1) \gamma^2(0) \\
&\quad + 2a_2 \gamma(0) \gamma(1) + [2a_1 - 3\rho(1)] \gamma^2(1) + \delta(n^{-2}), \quad (4.23)
\end{align*}
\]

and
\[
\begin{align*}
D_2 &= h^{11} \text{var} [\gamma(1)] + \text{cov} [\gamma(0), \gamma(2)] \\
&\quad + h^{12} [\text{cov} [\gamma(0), \gamma(1)] + \text{cov} [\gamma(1), \gamma(2)]] \\
&= [\gamma^2(0) - \gamma^2(1)]^{-1} [k(n-2)]^{-1} \gamma(0) [2+6a_2 - \rho(1) 2a_1 a_2] A(0) \\
&\quad + [6a_1 - \rho(1) (6+2a_1^2+2a_2)] A(1) - \gamma^2(0) \\
&\quad - 2a_2 \rho(1) \gamma(0) \gamma(1) + (3-\rho(1)a_1) \gamma^2(1) + \delta(n^{-2}). \quad (4.24)
\end{align*}
\]

Let \( W = H^{-1} (G_\alpha - D) \) with elements \( W_1 \). Then
\[
\begin{align*}
W_1 &= [\gamma^2(0) - \gamma^2(1)]^{-1} \gamma(0) [G_{11} [a_1 - \rho(1)a_2] \\
&\quad + G_{12} [a_2 - \rho(1)a_1] - D_1 + \rho(1) D_2] \\
&= [\gamma^2(0) - \gamma^2(1)]^{-2} [k(n-2)]^{-1} \gamma^2(0) [6a_1 + 4(1-a_2) \rho(1) \\
&\quad + 6a_1 \rho^2(1)] A(0) - [6(1-a_2) + 4a_1 (1+a_2) \rho(1) + 6(1-a_2) \rho^2(1)] A(1) \\
&\quad + [-3a_1 + (-2+4a_2) \rho(1) - 4a_1 \rho^2(1) + (6-4a_2) \rho^3(1) - a_1 \rho^4(1)] \gamma^2(0)]
\end{align*}
\]
\[ W_2 = [\gamma^2(0) - \gamma^2(1)]^{-1} \gamma(0) \{G_{11}[\alpha_2 - \alpha_1 \rho(1) + G_{12}[\alpha_1 - \alpha_2 \rho(1)] - D_2 + \rho(1)D_1 \}
\]
\[ = [\gamma^2(0) - \gamma^2(1)]^{-2}[k(n-2)]^{-1} \gamma^2(0) \{-2(1-\alpha_2)[1+\alpha_2+6\rho^2(1)
+ (1+\alpha_2)\rho^2(1)]A(0) + \rho(1)(1-\alpha_2)[14+2\alpha_2+2(1+\alpha_2)\rho^2(1)]A(1)
+ [1-3\alpha_2+4(1-\alpha_2)\rho^2(1) - (1+\alpha_2)\rho^4(1)]\gamma^2(0) \} + O(n^{-2})
\]
\[ = [1-\rho^2(1)]^{-3}[k(n-2)]^{-1}[-(1+3\alpha_2)+[-12(1+\alpha_2)]^{-1}
+ 15-3\alpha_2-2\alpha_2^2] \rho^2(1) - [3+9\alpha_2+2\alpha_2^2] \rho^4(1)
+ [1+3\alpha_2] \rho^6(1) \} + O(n^{-2}) \].

From (4.13) and (4.14) it follows that
Let \( U = H^{-1} E(d - Aa) \) and let the elements of \( U \) be \( U_1 \). Then

\[
U_1 = -(n-2)^{-1}(1+a_2) + (n^{-2}) \quad \text{and} \quad U_2 = U_1.
\]

The results of the theorem follow since \( E(\hat{a}_{1,OLS}) = a_1 + U_1 + W_1 \) and \( E(\hat{a}_{2,OLS}) = a_2 + U_2 + W_2 \).

The complexity of the approximate means of the least squares estimators for the parameters of the second-order seasonal autoregressive process prohibits the use of a convenient method of correcting for the biases. However, by ignoring terms of order \((nk)^{-1}\) in (4.10) and (4.11), modified least squares estimators which correct for the bias attributed to the estimated means can be obtained. From (4.10) and (4.11), the means of \( \hat{a}_{1,OLS} \) and \( \hat{a}_{2,OLS} \) are

\[
E(\hat{a}_{1,OLS}) = a_1 - (n-2)^{-1}(1+a_2) + O(n^{-1}k^{-1}), \tag{4.28}
\]

and

\[
E(\hat{a}_{2,OLS}) = a_2 - (n-2)^{-1}(1+a_2) + O(n^{-1}k^{-1}). \tag{4.29}
\]
The modified least squares estimators for the second-order seasonal autoregressive coefficients are

\[ \hat{\alpha}_2 = \hat{\alpha}_{2, OLS} \left[ 1 + (n-3)^{-1} \right] + (n-3)^{-1}, \quad \hat{\alpha}_{2, OLS} \leq -1 \]

\[ \hat{\alpha}_2 = 1 \quad \hat{\alpha}_{2, OLS} \geq 1 - 2(n-2)^{-1} \quad (4.30) \]

and

\[ \hat{\alpha}_1 = \hat{\alpha}_{1, OLS} \left[ 1 + (n-2)^{-1} \right], \quad (4.31) \]

where \( \hat{\alpha}_1 \) is truncated to lie within or on the boundaries of the stationary region. The variance of \( \hat{\alpha}_1, \hat{\alpha}_2 \) is equal to the variance of \( \hat{\alpha}_{1, OLS}, \hat{\alpha}_{2, OLS} \) to terms of order \( n^{-1} \). These estimators are expected to be similar to the modified method of moments estimators proposed by Bora-Senta and Kounias (1980) since both methods adjust for the mean bias.

Another method of adjusting for the mean bias is to correct for the biases of \( \hat{\alpha}_1, \hat{\alpha}_2 \) given in (4.13) and (4.14), respectively. If the least squares estimators for \( \hat{\alpha} \) lie within the stationary region, an estimator for \( \text{var}(\hat{\gamma}_10) \) is obtained by substituting \( \hat{\alpha}_{OLS} \) and \( \hat{\gamma}(0) \) for \( \alpha \) and \( \gamma(0) \) in (4.12). If the least squares estimators for \( \hat{\alpha} \) lie outside the stationary region, then the estimators, \( \hat{\alpha}_{OLS} \), are merely truncated to lie on the boundaries of the stationary region and no further modification is considered. However, assuming the least squares
estimators obey the stationary conditions, let \( \hat{V} \) denote the estimator for \( \text{var}\{\hat{Y}_{10}\} \). Then nearly unbiased estimators for \( H \) and \( N \) are

\[
\hat{H} = \hat{H} + \hat{V} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},
\]

and

\[
\hat{N} = \hat{N} + \hat{V} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.
\]

The procedure gives \( \hat{\alpha} = \hat{H}^{-1} \hat{N} \) as the modified estimators for \( \alpha \), although it is necessary to truncate the estimators, \( \hat{\alpha} \), to lie within or on the boundaries of the stationary region. To terms of order \( n^{-1} \), the variance of \( \hat{\alpha} \) agrees with the variance of \( \hat{\alpha}_{OLS} \) given in (4.5). This procedure is the extension of the procedure which gives \( \hat{\rho}_1 \) in the case of the first-order seasonal autoregressive process.

Salem (1971) proposed another modification to the least squares estimators which gives nearly unbiased estimators to terms of order \( n^{-1} \). Assuming \( \hat{\alpha}_{OLS} \) obey the stationary conditions, estimators for \( G = E(\hat{A} \hat{H}^{-1} A) \) and \( D = E(\hat{A} \hat{H}^{-1} d) \) are obtained by substituting \( \hat{\alpha}_{OLS} \) and \( \hat{\gamma}(0) \) for \( \alpha \) and \( \gamma(0) \) in the approximate expressions for \( G \) and \( D \), respectively. Let \( \hat{G} \) and \( \hat{D} \) denote the estimators of \( G \) and \( D \).
Then an unbiased estimator for \( a \) to order \( n^{-1} \) is given by

\[
\hat{a}^* = [\hat{H} + \hat{G}]^{-1}[\hat{N} + \hat{U}],
\]

where \( \hat{H} \) and \( \hat{N} \) are given in (4.32) and (4.33). The difficulty in obtaining approximate expressions for \( G \) and \( D \) for higher order stationary autoregressive processes make this procedure infeasible for practical purposes, and it is believed that this procedure is unstable for values of \( a_{OLS} \) near the boundaries of the stationarity region. It is noted that the expressions for \( G \) and \( D \) given by Salem (1971) are not correct and so are not in agreement with the expressions given in (4.21) to (4.24) with \( k = 1 \).

In practical situations, it is seldom necessary to postulate a seasonal autoregressive process of order greater than two. However, the modified least squares estimators which correct for the mean bias extends to higher order autoregressive processes. The iterative procedure given by Bora-Senta and Kounias (1980) is seen as a third alternative to obtaining modified estimators for stationary autoregressive coefficients which correct for the mean bias.
CHAPTER V. A MONTE CARLO STUDY

Expressions for the biases of the least squares estimators of autoregressive parameters have been derived using a series expansion and asymptotic consideration. In the autoregressive case, the ordinary least squares estimators systematically overestimate or underestimate the unknown values. The magnitude of the biases can be substantial for the moderately small samples generally encountered in economics. It is important to investigate the accuracy of the approximate expressions for the bias in finite samples. The behavior in finite samples is investigated empirically by generating time series of known structure. Modified least squares estimators with corrections for the bias are compared with the least squares estimator to determine the practical value of adjusting for the bias.

To generate the normal random variates, a sequence of pseudo-random uniform \((0, 1)\) deviates was generated using Marsaglia and Bray's (1968) algorithm and transformed to standard normal \((0, 1)\) deviates by inverting the normal cumulative distribution function. The simulation was performed using the latest version of the IMSL package and all computations were performed using double precision arithmetic. Standard normal error processes were used throughout the study.

The first-order seasonal autoregressive model with nonzero mean has the form

\[ y_{ij} = \theta_i + \rho y_{i,j-1} + e_{ij}, \quad (5.1) \]
where \( e_{ij} \) is a sequence of normal \((0, 1)\) variates. The subscript \((i, j)\) for \(i = 0, 1, \ldots, k-1\) and \(j = 2, 3, \ldots, n\) denotes the \(i\)-th period in the \(j\)-th cycle. The seasonal intercepts \(\theta_0, \theta_1, \ldots, \theta_{k-1}\) were set equal to zero in the simulation with no loss of generality. For stationary processes, the initial observations were generated by

\[
y_{i1} = (1-\rho^2)^{-\frac{1}{2}} e_{i1}, \quad i = 0, 1, \ldots, k-1.
\]  

(5.2)

For nonstationary processes, \(y_{i1}\) was set equal to \(e_{i1}\) for \(i = 0\) to \(k-1\).

In the analysis of economic time series, observations are commonly reported at monthly intervals. For such data, a 12-month cycle is often observed and the seasonal autoregressive model in (5.1) with \(k = 12\) is fitted. The seasonal period \(k\) is set equal to 12 in this study, but similar results are expected for values of \(k\) greater than one. To reflect the sample sizes often encountered in economic applications, 3 series lengths of 5, 10, and 20 years of monthly observations are considered. For each sample size, 15 values of \(\rho\) ranging from \(-1.0\) to \(1.0\) are used.

For each \((\rho, n)\) combination, various point estimates were computed using the same set of observations. This was repeated 1,000 times with a different set of generated observations each time. Sample biases, variances and mean square errors for each estimator were obtained by averaging over the 1,000 replications. The numerical results are reported in the following tables and graphs.
For a first-order seasonal autoregressive process, a great number of estimators could have been included in the study. The least squares estimator of \( \rho \) defined in (3.3) is consistent and asymptotically normal but is seriously biased in sample sizes typically dealt with in economics. Two bias correction procedures based on the work of Marriott and Pope (1954) and Salem (1971) are considered as competitors for the least squares estimator. These modified estimators are denoted by \( \hat{\rho}_{\text{MP}} \) and \( \hat{\rho}_1 \) and are defined in equations (3.27) and (3.30), respectively. Also, the symmetric estimator \( \hat{\rho}_{\text{SYM}} \) given in (3.33) is compared with the least squares and the modified least squares estimators of \( \rho \). The results of a Monte Carlo study of the properties of these four estimators of \( \rho \) are presented in this section.

Five other estimators of \( \rho \) were included in the Monte Carlo study and the results for these estimators are presented in Appendix C. The second group of estimators consists of the maximum likelihood estimator \( \hat{\rho}_{\text{MLE}}^\ast \) given in (3.35), the approximate maximum likelihood estimator \( \hat{\rho}_{\text{MLE}}^\ast \) considered in Lemma 3.3, the least squares estimator \( \hat{\rho}_{\text{OLS}} \) given in (3.5), the modified least squares estimator \( \hat{\rho}_2 \) defined in (3.33) and the modified symmetric estimator \( \hat{\rho}_{\text{SYM,MP}} \) given by

\[
\hat{\rho}_{\text{SYM,MP}} = [1-(n-1)^{-1} k^{-1} (k+2)]^{-1} [\hat{\rho}_{\text{SYM}} + (n-1)^{-1} ]
\]

\[
\hat{\rho}_{\text{SYM}} \in [-1+2k^{-1}(n-1)^{-1}, 1-2k^{-1}(k+1)(n-1)^{-1}]
\]

\[
= -1 \quad \text{for} \quad \hat{\rho}_{\text{SYM}} \leq -1+2k^{-1}(n-1)^{-1}
\]

\[
= 1 \quad \text{for} \quad \hat{\rho}_{\text{SYM}} \geq 1-2k^{-1}(k+1)(n-1)^{-1}
\]

(5.3)
One of the more important issues faced in estimation of autoregressive parameters concerns the assumption of stationarity. In many economic applications the assumption of stationarity is often appropriate on a priori grounds. With the exception of the symmetric estimator, the least squares estimator and its two modifications are not constrained to lie in the interval \((-1, 1)\). Various authors have considered this problem and in most cases truncated the estimators to lie within the stationary region. Ansley and Newbold (1980) consider arbitrarily placing the estimate on the boundary of the region for which the moduli of the roots of the characteristic equation were all less than or equal to 0.999. However, this procedure results in a discontinuity at the boundary points. In this study, a more reasonable method of truncation is used. The constrained estimators are defined by

\[
\hat{\rho} = \begin{cases} 
-1 & \text{if } \hat{\rho} \leq -1, \\
1 & \text{if } \hat{\rho} \geq 1, \\
\hat{\rho} & \text{otherwise}.
\end{cases}
\]

Tables 1 to 3 contain the Monte Carlo biases of the 4 estimators of \(\rho\) for sample sizes of 5, 10 and 20 years, respectively. For \(\rho > -0.90\) the unadjusted estimators, \(\hat{\rho}_{\text{OLS}}\) and \(\hat{\rho}_{\text{SYM}}\), have considerably larger absolute biases than either \(\hat{\rho}_{\text{MP}}\) or \(\hat{\rho}_{1}\). The absolute bias of \(\hat{\rho}_{1}\) generally lies between the absolute biases of \(\hat{\rho}_{\text{MP}}\) and \(\hat{\rho}_{\text{OLS}}\). For
\(-1.0 \leq \rho \leq -0.90\), slight differences between the biases of the three least squares based estimators are found. When \(\rho = -1.0\), the symmetric estimator displays a much larger bias than the other three estimators. In cases where the absolute bias of \(\hat{\rho}_{\text{MP}}\) is not the smallest, the differences are small. The bias for the unadjusted estimators exceeds 50% of the true value when \(\rho > -0.50\) and \(n = 5\); when \(\rho > -0.50\), the bias exceeds 30% of the true value of \(\rho\) for \(n = 10\) and 14% of the true value for \(n = 20\) years.

The biases of \(\hat{\rho}_{\text{OLS}}\) and \(\hat{\rho}_{\text{SYM}}\) have very similar trends with the magnitude of the bias of \(\hat{\rho}_{\text{SYM}}\) being larger than the bias of \(\hat{\rho}_{\text{OLS}}\) for \(0.0 < \rho \leq 1.0\). For \(-0.90 < \rho \leq 1.0\), the unadjusted estimators systematically underestimate the true value. For all values of \(n\) considered, the biases of \(\hat{\rho}_{\text{OLS}}\) and \(\hat{\rho}_{\text{SYM}}\) are decreasing functions of \(\rho\). The large sample theory gives \(-{(n-1)}^{-1}(1 + 7\rho/6)\) as the bias of \(\hat{\rho}_{\text{OLS}}\) and \(\hat{\rho}_{\text{SYM}}\). The Monte Carlo results are slightly below the theoretical results for values of \(0.1 < \rho \leq 1.0\). See Figures 1 and 2. The theoretical expression for the bias does not consider the constraint placed on the estimators by restricting them to \([-1, 1]\). Therefore, the differences between the theoretical and empirical results for values of \(\rho\) near one are partially explained by the truncation to \([-1, 1]\).

For \(n = 5\), the estimated standard errors of the biases of the various estimators range from 0.002 to 0.007. The estimated standard errors of the biases of the various estimators were about 0.003 for \(n = 10\) and about 0.001 for \(n = 20\). For \(-0.1 \leq \rho \leq 0.1\), the biases of \(\hat{\rho}_{\text{MP}}\) are within two standard errors of zero, the theoretical bias of
Table 1. Empirical biases of various estimators of $\rho$ for $n = 5$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>.023</td>
<td>.016</td>
<td>.029</td>
<td>.193</td>
</tr>
<tr>
<td>-0.99</td>
<td>.003</td>
<td>-.009</td>
<td>.007</td>
<td>.074</td>
</tr>
<tr>
<td>-0.90</td>
<td>-.014</td>
<td>-.028</td>
<td>.008</td>
<td>.047</td>
</tr>
<tr>
<td>-0.70</td>
<td>-.063</td>
<td>-.028</td>
<td>-.006</td>
<td>-.014</td>
</tr>
<tr>
<td>-0.50</td>
<td>-.109</td>
<td>-.026</td>
<td>-.018</td>
<td>-.083</td>
</tr>
<tr>
<td>-0.30</td>
<td>-.157</td>
<td>-.005</td>
<td>-.034</td>
<td>-.143</td>
</tr>
<tr>
<td>-0.10</td>
<td>-.226</td>
<td>-.008</td>
<td>-.078</td>
<td>-.219</td>
</tr>
<tr>
<td>0.00</td>
<td>-.248</td>
<td>.007</td>
<td>-.087</td>
<td>-.248</td>
</tr>
<tr>
<td>0.10</td>
<td>-.271</td>
<td>.012</td>
<td>-.097</td>
<td>-.277</td>
</tr>
<tr>
<td>0.30</td>
<td>-.351</td>
<td>-.028</td>
<td>-.160</td>
<td>-.370</td>
</tr>
<tr>
<td>0.50</td>
<td>-.426</td>
<td>-.032</td>
<td>-.221</td>
<td>-.464</td>
</tr>
<tr>
<td>0.70</td>
<td>-.504</td>
<td>-.067</td>
<td>-.291</td>
<td>-.567</td>
</tr>
<tr>
<td>0.90</td>
<td>-.578</td>
<td>-.123</td>
<td>-.363</td>
<td>-.676</td>
</tr>
<tr>
<td>0.99</td>
<td>-.611</td>
<td>-.147</td>
<td>-.398</td>
<td>-.726</td>
</tr>
<tr>
<td>1.00</td>
<td>-.621</td>
<td>-.162</td>
<td>-.408</td>
<td>-.736</td>
</tr>
</tbody>
</table>
Table 2. Empirical biases of various estimators of $\rho$ for $n = 10$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>.015</td>
<td>.010</td>
<td>.017</td>
<td>.096</td>
</tr>
<tr>
<td>-0.99</td>
<td>.002</td>
<td>-.007</td>
<td>.002</td>
<td>.003</td>
</tr>
<tr>
<td>-0.90</td>
<td>.002</td>
<td>-.009</td>
<td>.009</td>
<td>.004</td>
</tr>
<tr>
<td>-0.70</td>
<td>-.021</td>
<td>-.001</td>
<td>.009</td>
<td>-.014</td>
</tr>
<tr>
<td>-0.50</td>
<td>-.051</td>
<td>-.005</td>
<td>-.003</td>
<td>-.045</td>
</tr>
<tr>
<td>-0.30</td>
<td>-.065</td>
<td>.008</td>
<td>.002</td>
<td>-.063</td>
</tr>
<tr>
<td>-0.10</td>
<td>-.101</td>
<td>-.003</td>
<td>-.019</td>
<td>-.100</td>
</tr>
<tr>
<td>0.00</td>
<td>-.117</td>
<td>-.006</td>
<td>-.027</td>
<td>-.116</td>
</tr>
<tr>
<td>0.10</td>
<td>-.124</td>
<td>.000</td>
<td>-.026</td>
<td>-.125</td>
</tr>
<tr>
<td>0.30</td>
<td>-.161</td>
<td>-.013</td>
<td>-.051</td>
<td>-.165</td>
</tr>
<tr>
<td>0.50</td>
<td>-.188</td>
<td>-.014</td>
<td>-.069</td>
<td>-.195</td>
</tr>
<tr>
<td>0.70</td>
<td>-.229</td>
<td>-.031</td>
<td>-.106</td>
<td>-.243</td>
</tr>
<tr>
<td>0.90</td>
<td>-.282</td>
<td>-.064</td>
<td>-.162</td>
<td>-.303</td>
</tr>
<tr>
<td>0.99</td>
<td>-.310</td>
<td>-.088</td>
<td>-.195</td>
<td>-.340</td>
</tr>
<tr>
<td>1.00</td>
<td>-.313</td>
<td>-.091</td>
<td>-.199</td>
<td>-.345</td>
</tr>
</tbody>
</table>
Table 3. Empirical biases of various estimators of $\rho$ for $n = 20$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{OL}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>.010</td>
<td>.008</td>
<td>.011</td>
<td>.052</td>
</tr>
<tr>
<td>-0.99</td>
<td>.002</td>
<td>-.004</td>
<td>.003</td>
<td>.002</td>
</tr>
<tr>
<td>-0.90</td>
<td>.000</td>
<td>-.002</td>
<td>.006</td>
<td>.002</td>
</tr>
<tr>
<td>-0.70</td>
<td>-.010</td>
<td>-.001</td>
<td>.004</td>
<td>-.009</td>
</tr>
<tr>
<td>-0.50</td>
<td>-.022</td>
<td>.000</td>
<td>.003</td>
<td>-.021</td>
</tr>
<tr>
<td>-0.30</td>
<td>-.032</td>
<td>.002</td>
<td>.002</td>
<td>-.031</td>
</tr>
<tr>
<td>-0.10</td>
<td>-.050</td>
<td>-.004</td>
<td>-.007</td>
<td>-.049</td>
</tr>
<tr>
<td>0.00</td>
<td>-.053</td>
<td>-.001</td>
<td>-.006</td>
<td>-.053</td>
</tr>
<tr>
<td>0.10</td>
<td>-.058</td>
<td>.001</td>
<td>-.006</td>
<td>-.058</td>
</tr>
<tr>
<td>0.30</td>
<td>-.073</td>
<td>-.002</td>
<td>-.014</td>
<td>-.074</td>
</tr>
<tr>
<td>0.50</td>
<td>-.086</td>
<td>-.003</td>
<td>-.021</td>
<td>-.088</td>
</tr>
<tr>
<td>0.70</td>
<td>-.104</td>
<td>-.012</td>
<td>-.034</td>
<td>-.106</td>
</tr>
<tr>
<td>0.90</td>
<td>-.132</td>
<td>-.028</td>
<td>-.066</td>
<td>-.136</td>
</tr>
<tr>
<td>0.99</td>
<td>-.156</td>
<td>-.048</td>
<td>-.096</td>
<td>-.164</td>
</tr>
<tr>
<td>1.00</td>
<td>-.161</td>
<td>-.052</td>
<td>-.102</td>
<td>-.167</td>
</tr>
</tbody>
</table>
Table 4. Empirical variances of various estimators of \( \rho \) for \( n = 5 \)

<table>
<thead>
<tr>
<th>True value of ( \rho )</th>
<th>( \hat{\rho}_{OLS} )</th>
<th>( \hat{\rho}_{MP} )</th>
<th>( \hat{\rho}_1 )</th>
<th>( \hat{\rho}_{SYM} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>.0251</td>
<td>.0238</td>
<td>.0388</td>
<td>.0236</td>
</tr>
<tr>
<td>-0.99</td>
<td>.0026</td>
<td>.0002</td>
<td>.0041</td>
<td>.0003</td>
</tr>
<tr>
<td>-0.90</td>
<td>.0391</td>
<td>.0548</td>
<td>.0604</td>
<td>.0113</td>
</tr>
<tr>
<td>-0.70</td>
<td>.0932</td>
<td>.2020</td>
<td>.1412</td>
<td>.0523</td>
</tr>
<tr>
<td>-0.50</td>
<td>.1356</td>
<td>.2660</td>
<td>.2010</td>
<td>.0942</td>
</tr>
<tr>
<td>-0.30</td>
<td>.1689</td>
<td>.3483</td>
<td>.2424</td>
<td>.1371</td>
</tr>
<tr>
<td>-0.10</td>
<td>.1842</td>
<td>.3671</td>
<td>.2570</td>
<td>.1429</td>
</tr>
<tr>
<td>0.00</td>
<td>.2042</td>
<td>.3950</td>
<td>.2786</td>
<td>.1666</td>
</tr>
<tr>
<td>0.10</td>
<td>.2200</td>
<td>.4385</td>
<td>.2932</td>
<td>.1766</td>
</tr>
<tr>
<td>0.30</td>
<td>.2343</td>
<td>.4647</td>
<td>.2989</td>
<td>.1733</td>
</tr>
<tr>
<td>0.50</td>
<td>.2560</td>
<td>.5035</td>
<td>.3066</td>
<td>.1822</td>
</tr>
<tr>
<td>0.70</td>
<td>.2573</td>
<td>.4982</td>
<td>.2845</td>
<td>.1635</td>
</tr>
<tr>
<td>0.90</td>
<td>.2884</td>
<td>.4328</td>
<td>.2608</td>
<td>.1590</td>
</tr>
<tr>
<td>0.99</td>
<td>.2903</td>
<td>.2998</td>
<td>.2747</td>
<td>.1587</td>
</tr>
<tr>
<td>1.00</td>
<td>.2984</td>
<td>.3391</td>
<td>.2795</td>
<td>.1581</td>
</tr>
</tbody>
</table>
Figure 1. Empirical biases of various estimators of $\rho$ for $n = 5$
Figure 2. Empirical biases of various estimators of $\rho$ for $n = 20$
\[ \hat{\rho}_{OLS} \] differs by more than two standard errors from the approximate bias, \(-\frac{1}{(n-1)^2}(1+7\rho/6)\).

The bias of the modified estimator \( \hat{\rho}_1 \) is a decreasing function of \( \rho \) with the exception of \( \rho = -0.99 \). For \(-0.10 \leq \rho \leq 1.0\), the estimator \( \hat{\rho}_1 \) systematically underestimates the true value of \( \rho \). The large sample theory gives \(-\frac{1}{12(n-1)} 2\rho\) as the bias of \( \hat{\rho}_1 \), which is an odd function of \( \rho \) about 0. The Monte Carlo biases of \( \hat{\rho}_1 \) are not in close agreement with the approximate biases for sample sizes as large as 20 years. Salem (1971) obtained similar results for \( \hat{\rho}_1 \) in the first-order autoregressive process.

Figures 1 and 2 present the Monte Carlo biases of the various estimators of \( \rho \) for \( 0.0 \leq \rho \leq 1.0 \) and sample sizes of 5 and 20 years, respectively. These two figures give a reasonable representation of the shape of the bias functions of the different estimators of \( \rho \) for series lengths ranging from 5 to 20 years.

The reduction in bias of the modified estimators comes at the expense of variance. This result has been obtained by a number of authors. See Orcutt and Winokur (1969) and Dent and Min (1978). Tables 4 to 6 present the sample variances of the 4 estimators of \( \rho \) for the series lengths of 5, 10 and 20 years, respectively. For \(-0.90 \leq \rho \leq 0.90\), the two unadjusted estimators of \( \rho \) have uniformly smaller variances than the two modified estimators. For values of \( |\rho| \) near one, the adjusted estimators, \( \hat{\rho}_{MP} \) and \( \hat{\rho}_1 \), have smaller variances than the least squares estimator due to the truncation used. In general, the symmetric estimator gave the most stable estimator for all values of \( n \) and \( \rho \). For
n = 20, the least squares estimator differs little from the symmetric estimator with the exception of values of $\rho$ close to one.

Based on the work of Bartlett (1946), the theoretical asymptotic variance of $\hat{\rho}_{OLS}$ is $\left[12(n-1)^{-1}(1-\rho^2)\right]$ for $-1 < \rho < 1$. This is also the large sample variance of $\hat{\rho}_{SYM}$. For negative and small positive $\rho$, the sample variances of the least squares and symmetric estimators appear to underestimate the asymptotic variances, while the reverse is true for values of $0.1 < \rho < 1.0$. The Monte Carlo variances of $\hat{\rho}_{OLS}$ are in fair agreement with the asymptotic variances for $-0.99 < \rho < 0.10$ for all sample sizes considered. The approximation is poor for small-sample sizes and values of $\rho$ near 1.0. The empirical variances are declining much faster than $n^{-1}$ in the range $\rho > 0$. Orcutt and Winokur (1969), Salem (1971), and Dent and Min (1978) obtained similar results for the first-order autoregressive process.

From the derivations of the modified least squares estimators, the variance of $\hat{\rho}_{MP}$ is $\left[1-7[6(n-1)]^{-1}\right]^{-2}$ times the variance of $\hat{\rho}_{OLS}$, while the variance of $\hat{\rho}_{1}$ is approximately $n^2(n-1)^{-2}$ times the least squares variance. The Monte Carlo variance of $\hat{\rho}_{MP}$ is roughly proportional to the variance of $\hat{\rho}_{OLS}$ for values of $-0.90 < \rho < 0.90$ for all sample sizes. The ratio of the Monte Carlo variances of $\hat{\rho}_{OLS}$ to $\hat{\rho}_{1}$ is poorly approximated by $n^2(n-1)^{-2}$ for samples of 5 and 10; for $n = 20$ the approximation of the variance ratio of $\hat{\rho}_{OLS}$ to $\hat{\rho}_{1}$ is satisfactory for $-0.90 < \rho < 0.90$. 
Table 5. Empirical variances of various estimators of $\rho$ for $n = 10$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{\text{OLS}}$</th>
<th>$\hat{\rho}_{\text{MP}}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>0.0066</td>
<td>0.0057</td>
<td>0.0082</td>
<td>0.0083</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.0015</td>
<td>0.0006</td>
<td>0.0020</td>
<td>0.0003</td>
</tr>
<tr>
<td>-0.90</td>
<td>0.0202</td>
<td>0.0245</td>
<td>0.0250</td>
<td>0.0130</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.0423</td>
<td>0.0558</td>
<td>0.0518</td>
<td>0.0373</td>
</tr>
<tr>
<td>-0.50</td>
<td>0.0629</td>
<td>0.0830</td>
<td>0.0766</td>
<td>0.0576</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.0740</td>
<td>0.0983</td>
<td>0.0891</td>
<td>0.0706</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.0841</td>
<td>0.1111</td>
<td>0.1001</td>
<td>0.0811</td>
</tr>
<tr>
<td>0.00</td>
<td>0.0883</td>
<td>0.1169</td>
<td>0.1043</td>
<td>0.0853</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0938</td>
<td>0.1238</td>
<td>0.1096</td>
<td>0.0905</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0948</td>
<td>0.1267</td>
<td>0.1080</td>
<td>0.0905</td>
</tr>
<tr>
<td>0.50</td>
<td>0.0898</td>
<td>0.1206</td>
<td>0.0980</td>
<td>0.0855</td>
</tr>
<tr>
<td>0.70</td>
<td>0.0830</td>
<td>0.1142</td>
<td>0.0842</td>
<td>0.0734</td>
</tr>
<tr>
<td>0.90</td>
<td>0.0794</td>
<td>0.1107</td>
<td>0.0711</td>
<td>0.0623</td>
</tr>
<tr>
<td>0.99</td>
<td>0.0785</td>
<td>0.0788</td>
<td>0.0649</td>
<td>0.0591</td>
</tr>
<tr>
<td>1.00</td>
<td>0.0813</td>
<td>0.0708</td>
<td>0.0682</td>
<td>0.0569</td>
</tr>
</tbody>
</table>
Table 6. Empirical variances of various estimators of $\rho$ for $n = 20$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{\text{OLS}}$</th>
<th>$\hat{\rho}_{\text{MP}}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>.0027</td>
<td>.0023</td>
<td>.0030</td>
<td>.0032</td>
</tr>
<tr>
<td>-0.99</td>
<td>.0009</td>
<td>.0007</td>
<td>.0010</td>
<td>.0003</td>
</tr>
<tr>
<td>-0.90</td>
<td>.0085</td>
<td>.0104</td>
<td>.0095</td>
<td>.0070</td>
</tr>
<tr>
<td>-0.70</td>
<td>.0220</td>
<td>.0250</td>
<td>.0243</td>
<td>.0204</td>
</tr>
<tr>
<td>-0.50</td>
<td>.0285</td>
<td>.0324</td>
<td>.0315</td>
<td>.0281</td>
</tr>
<tr>
<td>-0.30</td>
<td>.0374</td>
<td>.0425</td>
<td>.0412</td>
<td>.0370</td>
</tr>
<tr>
<td>-0.10</td>
<td>.0403</td>
<td>.0457</td>
<td>.0442</td>
<td>.0398</td>
</tr>
<tr>
<td>0.00</td>
<td>.0443</td>
<td>.0503</td>
<td>.0484</td>
<td>.0439</td>
</tr>
<tr>
<td>0.10</td>
<td>.0451</td>
<td>.0512</td>
<td>.0491</td>
<td>.0447</td>
</tr>
<tr>
<td>0.30</td>
<td>.0457</td>
<td>.0519</td>
<td>.0493</td>
<td>.0454</td>
</tr>
<tr>
<td>0.50</td>
<td>.0362</td>
<td>.0411</td>
<td>.0383</td>
<td>.0362</td>
</tr>
<tr>
<td>0.70</td>
<td>.0311</td>
<td>.0410</td>
<td>.0314</td>
<td>.0295</td>
</tr>
<tr>
<td>0.90</td>
<td>.0248</td>
<td>.0304</td>
<td>.0221</td>
<td>.0210</td>
</tr>
<tr>
<td>0.99</td>
<td>.0216</td>
<td>.0215</td>
<td>.0170</td>
<td>.0165</td>
</tr>
<tr>
<td>1.00</td>
<td>.0213</td>
<td>.0194</td>
<td>.0164</td>
<td>.0158</td>
</tr>
</tbody>
</table>
The estimated standard errors for the variances of Tables 4 to 6 were computed using the formula

\[ s.e. (\text{var}) = \left[ \left( \frac{1}{(1000)} \right)^{-1} \left( \frac{1}{(999)} \right)^{-1} \sum_{i=1}^{1000} (\text{var}_i - \text{var})^2 \right]^{1/2} \]  

(5.4)

where \( \text{var}_i \) is the estimated variance of \( \hat{\rho} \) for the \( i \)-th sample and \( \text{var} \) is the sample mean, \( \left( \frac{1}{1000} \right) \sum_{i=1}^{1000} \text{var}_i \), reported in the tables.

The estimated standard errors for \( n = 5 \) were about 0.0011 and 0.0020 for the variances of \( \hat{\rho}_{OLS} \) and \( \hat{\rho}_{MP} \), respectively. The estimated standard errors for the variances of \( \hat{\rho}_{OLS} \) and \( \hat{\rho}_{MP} \) were about 0.00042 and 0.00055 for \( n = 10 \) and about 0.00018 and 0.00021 for \( n = 20 \).

For values of \( \rho \) near zero, the empirical variances of \( \hat{\rho}_{OLS} \) were within two standard errors of the theoretical asymptotic variances.

The mean square error is generally used to compare various estimation procedures. It has been shown that the least biased estimators also have the largest variances in autoregressive processes. The mean square errors of the 4 estimators of \( \rho \) for \( n = 5, 10 \) and 20 years are presented in Tables 7 to 9. In general, the mean square error of the 4 estimators of \( \rho \) increases as \( \rho \) increases. For \( -0.3 < \rho < 1.0 \), the modified least squares estimators dominate the unadjusted estimators for all values of \( n \) considered. For large negative values of \( \rho \), the mean square errors of the four estimators are similar.

For \( n = 5 \), the adjusted estimator \( \hat{\rho}_{MP} \) has the minimum mean square error for 8 of the 15 values of \( \rho \). In estimating \( \rho \), the modified
estimator $\hat{\rho}_I$ exhibits minimum mean-square error in four cases and the symmetric estimator dominates in the other three cases. Even in those seven instances where $\hat{\rho}_{MP}$ is not the minimum, the Monte Carlo mean square error for $\hat{\rho}_{MP}$ is close to the optimum level. A particularly strong mean square error performance of the estimator $\hat{\rho}_{MP}$ is observed for large positive values of $\rho$. In the case of the first-order seasonal autoregressive process, the small sample biases tend to outweigh the small sample variances for the unadjusted estimators. This is particularly evident for values of $\rho$ near 1.0.

For values of $n = 10$ and 20, the relative performance of the various estimators is similar to that for $n = 5$. For negative values of $\rho$ there is little difference in estimator performance. For positive values of $\rho$, the mean square errors of the unadjusted estimators increase at a faster rate than those of either of the modified estimators. The strongest performance in estimating $\rho$ is once again given by $\hat{\rho}_{MP}$.

Figures 3 and 4 represent graphically the Monte Carlo mean square error performances of the 4 estimators of $\rho$ for positive values of the parameter and series lengths of 5 and 20 years, respectively. Figure 5 presents the mean square errors of $\hat{\rho}_{MP}$ and $\hat{\rho}_I$ relative to the mean square error of $\hat{\rho}_{ULS}$ for $n = 10$. These figures emphasize the dominance of $\hat{\rho}_{MP}$ in estimating the parameter $\rho$ and provide a fairly typical representation of the mean square errors for sample sizes between 5 and 20 years.

From the point of view of statistical decision theory, Chernoff and Moses (1959, pp. 119-165) conclude that the average risk is the best
Table 7. Empirical mean square errors of various estimators of $\rho$
for $n = 5$

<table>
<thead>
<tr>
<th>True values of $\rho$</th>
<th>$\hat{\rho}_{\text{OLS}}$</th>
<th>$\hat{\rho}_{\text{MP}}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>0.0304</td>
<td>0.0263</td>
<td>0.0470</td>
<td>0.3941</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.0028</td>
<td>0.0011</td>
<td>0.0046</td>
<td>0.0552</td>
</tr>
<tr>
<td>-0.90</td>
<td>0.0409</td>
<td>0.0627</td>
<td>0.0610</td>
<td>0.0333</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.1326</td>
<td>0.2024</td>
<td>0.1415</td>
<td>0.0543</td>
</tr>
<tr>
<td>-0.50</td>
<td>0.2544</td>
<td>0.2670</td>
<td>0.2041</td>
<td>0.1628</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.4166</td>
<td>0.3486</td>
<td>0.2542</td>
<td>0.3413</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.6962</td>
<td>0.3677</td>
<td>0.3182</td>
<td>0.6234</td>
</tr>
<tr>
<td>0.00</td>
<td>0.8210</td>
<td>0.3955</td>
<td>0.3540</td>
<td>0.7801</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9543</td>
<td>0.4398</td>
<td>0.3874</td>
<td>0.9409</td>
</tr>
<tr>
<td>0.30</td>
<td>1.4639</td>
<td>0.4728</td>
<td>0.5543</td>
<td>1.5405</td>
</tr>
<tr>
<td>0.50</td>
<td>2.0686</td>
<td>0.5136</td>
<td>0.7969</td>
<td>2.3392</td>
</tr>
<tr>
<td>0.70</td>
<td>2.7995</td>
<td>0.5431</td>
<td>1.1325</td>
<td>3.3823</td>
</tr>
<tr>
<td>0.90</td>
<td>3.6242</td>
<td>0.5841</td>
<td>1.5973</td>
<td>4.7276</td>
</tr>
<tr>
<td>0.99</td>
<td>4.0289</td>
<td>0.5159</td>
<td>1.8582</td>
<td>5.4215</td>
</tr>
<tr>
<td>1.00</td>
<td>4.1591</td>
<td>0.6008</td>
<td>1.9456</td>
<td>5.5775</td>
</tr>
</tbody>
</table>
Table 8. Empirical mean square errors of various estimators of $\rho$ for $n = 10$

<table>
<thead>
<tr>
<th>True values of $\rho$</th>
<th>Mean square error multiplied by ten</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\rho}_{OLS}$</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.0089</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.0016</td>
</tr>
<tr>
<td>-0.90</td>
<td>0.0203</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.0469</td>
</tr>
<tr>
<td>-0.50</td>
<td>0.0888</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.1165</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.1861</td>
</tr>
<tr>
<td>0.00</td>
<td>0.2239</td>
</tr>
<tr>
<td>0.10</td>
<td>0.2479</td>
</tr>
<tr>
<td>0.30</td>
<td>0.3536</td>
</tr>
<tr>
<td>0.50</td>
<td>0.4443</td>
</tr>
<tr>
<td>0.70</td>
<td>0.6087</td>
</tr>
<tr>
<td>0.90</td>
<td>0.8724</td>
</tr>
<tr>
<td>0.99</td>
<td>1.0377</td>
</tr>
<tr>
<td>1.00</td>
<td>1.0627</td>
</tr>
</tbody>
</table>
Table 9. Empirical mean square errors of various estimators of $\rho$
for $n = 20$

<table>
<thead>
<tr>
<th>True values of $\rho$</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>0.0038</td>
<td>0.0029</td>
<td>0.0041</td>
<td>0.0306</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0011</td>
<td>0.0003</td>
</tr>
<tr>
<td>-0.90</td>
<td>0.0086</td>
<td>0.0104</td>
<td>0.0098</td>
<td>0.0070</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.0231</td>
<td>0.0250</td>
<td>0.0245</td>
<td>0.0213</td>
</tr>
<tr>
<td>-0.50</td>
<td>0.0335</td>
<td>0.0324</td>
<td>0.0316</td>
<td>0.0325</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.0478</td>
<td>0.0425</td>
<td>0.0413</td>
<td>0.0466</td>
</tr>
<tr>
<td>-0.10</td>
<td>0.0652</td>
<td>0.0459</td>
<td>0.0446</td>
<td>0.0642</td>
</tr>
<tr>
<td>0.00</td>
<td>0.0727</td>
<td>0.0503</td>
<td>0.0488</td>
<td>0.0723</td>
</tr>
<tr>
<td>0.10</td>
<td>0.0783</td>
<td>0.0512</td>
<td>0.0494</td>
<td>0.0783</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0994</td>
<td>0.0519</td>
<td>0.0512</td>
<td>0.1002</td>
</tr>
<tr>
<td>0.50</td>
<td>0.1109</td>
<td>0.0412</td>
<td>0.0425</td>
<td>0.1137</td>
</tr>
<tr>
<td>0.70</td>
<td>0.1383</td>
<td>0.0423</td>
<td>0.0430</td>
<td>0.1418</td>
</tr>
<tr>
<td>0.90</td>
<td>0.1984</td>
<td>0.0384</td>
<td>0.0659</td>
<td>0.2065</td>
</tr>
<tr>
<td>0.99</td>
<td>0.2652</td>
<td>0.0445</td>
<td>0.1098</td>
<td>0.2841</td>
</tr>
<tr>
<td>1.00</td>
<td>0.2801</td>
<td>0.0463</td>
<td>0.1201</td>
<td>0.2960</td>
</tr>
</tbody>
</table>
Figure 3. Empirical mean square errors of various estimators of $\rho$
for $n = 5$
Figure 4. Empirical mean square errors of various estimators of $\rho$ for $n = 20$
Figure 5. Ratio of empirical mean square errors of various estimators of \( \hat{\rho} \) and the empirical mean square error of \( \hat{\rho}_{OLS} \) for \( n=10 \).
available criterion for evaluating the relative performances of the various estimators. Using the squared error loss function, Thornber (1967) compares various estimators of $\rho$ in the first-order autoregressive process using the average risk measure. For a uniform weight function, the average risk measure is the mean of the mean square error averaged over the values of $\rho$ considered. The average risks of the various estimators are presented in Table 10.

Table 10. Average risks of various estimators of $\rho$ multiplied by 10 averaged over the values of $\rho$ considered

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>$\hat{\rho}_1$</th>
<th>$\hat{\rho}_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 5$</td>
<td>1.4329</td>
<td>0.3561</td>
<td>0.6438</td>
<td>1.7583</td>
</tr>
<tr>
<td>$n = 10$</td>
<td>0.3547</td>
<td>0.0969</td>
<td>0.1536</td>
<td>0.3949</td>
</tr>
<tr>
<td>$n = 20$</td>
<td>0.0951</td>
<td>0.0351</td>
<td>0.0458</td>
<td>0.0945</td>
</tr>
</tbody>
</table>

Table 11 presents the relative average risks of the various estimators of $\rho$ expressed as a percentage of the average risk of $\hat{\rho}_{OLS}$. For $n = 5$ and 10, the average risk of $\hat{\rho}_{MP}$ is one-fourth the average risk of $\hat{\rho}_{OLS}$ and is slightly greater than one-third the average risk of $\hat{\rho}_{OLS}$ for $n = 20$. The average risk of the modified least squares estimator $\hat{\rho}_1$ is slightly less than one-half the average risk of $\hat{\rho}_{OLS}$ for the values of $n$ considered. For $n = 5$ and 10, the symmetric estimator $\hat{\rho}_{SYM}$ has a larger average risk than $\hat{\rho}_{OLS}$ and is nearly equal in performance to $\hat{\rho}_{OLS}$ for $n = 20$. The estimator $\hat{\rho}_{MP}$
is judged superior to either of the three estimators based on its average risk. A considerable reduction in average risk is realized by the estimator \( \hat{\rho}_{\text{MP}} \) in comparison with the average risk of \( \hat{\rho}_{\text{OLS}} \).

Table 11. Average risks of various estimators of \( \rho \) as a percentage of the average risk of \( \hat{\rho}_{\text{OLS}} \) averaged over the values of \( \rho \) considered

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \hat{\rho}_{\text{MP}} )</th>
<th>( \hat{\rho}_1 )</th>
<th>( \hat{\rho}_{\text{SYM}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>24.9</td>
<td>44.9</td>
<td>122.7</td>
</tr>
<tr>
<td>10</td>
<td>27.3</td>
<td>43.3</td>
<td>111.3</td>
</tr>
<tr>
<td>20</td>
<td>36.9</td>
<td>48.2</td>
<td>99.4</td>
</tr>
</tbody>
</table>

In time series analysis, a model is often judged by its ability to predict future values. One and three period ahead forecasts were generated by the equations

\[
\hat{Y}_{1,n+1} = \hat{\theta}_1 + \hat{\rho} Y_{1,n} \tag{5.5}
\]

and

\[
\hat{Y}_{1,n+3} = \hat{\theta}_1 (1+\hat{\rho}^2) + \hat{\rho}^3 Y_{1,n} \tag{5.6}
\]

The error in predicting one period ahead given \( Y_n, Y_{n-1}, \ldots, Y_1 \) is
Fuller and Hasza (1978) have shown that the \( s \) period ahead forecasts are unbiased for normal errors. The mean square error of prediction for one period ahead forecast is

\[
E[(X_{i,n+1} - \hat{Y}_{i,n+1})^2] = E[e_{i,n+1}^2] + E[(\theta_i - \hat{\theta}_i)^2] + (\rho - \hat{\rho})Y_{in}^2
\]

and only the distribution of 

\((\theta_i - \hat{\theta}_i) + (\rho - \hat{\rho})Y_{in}\)

requires study. In general, the error in predicting \( s \) periods ahead given \( Y_n, Y_{n-1}, \ldots, Y_1 \) is

\[
Y_{i,n+s} - \hat{Y}_{i,n+s} = \sum_{j=0}^{s-1} \rho^j e_{i,n+s-j} + \theta_i \sum_{j=0}^{s-1} \rho^j \hat{\theta}_i \sum_{j=0}^{s-1} \rho^j Y_{i,n}
\]

and only the distribution of the last two terms needs simulation.
The one period ahead prediction errors, $Y_{i,n+1} - \hat{Y}_{i,n+1}$, are identically distributed for $i = 0,1,2,...,11$. The sampling means and variances of $(\theta_i - \hat{\theta_i}) + (\rho - \hat{\rho})Y_{i,n}$ were calculated by averaging over the 12 months and the 1,000 replications. Similar calculations were done for the three period ahead predictors. The one step ahead predictors constructed from the four estimators of $\rho$ are denoted by $\hat{Y}_{n+1,OLS}$, $\hat{Y}_{n+1,MP}$, $\hat{Y}_{n+1,RHOL}$ and $\hat{Y}_{n+1,SYM}$, respectively.

Various methods of estimating the seasonal intercepts $\theta_i$, $i = 0$ to $11$, needed in the forecasting equations (5.5) and (5.6) were considered. Although the estimators of $\theta_i$ considered are asymptotically equivalent, it is expected that their small sample properties are somewhat different. The methods used in estimating $\theta_i$ are

\begin{align*}
\hat{\theta}_{i,OLS} &= \bar{Y}_{i0} - \hat{\rho}_{OLS}\bar{Y}_{i1}, \\
\hat{\theta}_{i,MP} &= (1-\hat{\rho}_{MP})\left[\frac{Y_{i1} + Y_{i1} + (1-\hat{\rho}_{MP})\sum_{j=2}^{n-1} Y_{ij}}{2 + (1-\hat{\rho}_{MP})(n-2)}\right], \\
\hat{\theta}_{i,RHOL} &= (1-\hat{\rho}_{1})\left[\frac{Y_{i1} + Y_{i1} + (1-\hat{\rho}_{1})\sum_{j=2}^{n-1} Y_{ij}}{2 + (1-\hat{\rho}_{1})(n-2)}\right],
\end{align*}

and

\begin{align*}
\hat{\theta}_{i,SYM} &= (1-\hat{\rho}_{1})\left[\frac{Y_{i1} + Y_{i1} + (1-\hat{\rho}_{1})\sum_{j=2}^{n-1} Y_{ij}}{2 + (1-\hat{\rho}_{1})(n-2)}\right].
\end{align*}
\[ \hat{\theta}_{i,\text{SYM}} = (1 - \hat{\rho}_{\text{SYM}}) \bar{Y}_i. \]  

where

\[ \bar{Y}_{i0} = (n-1)^{-1} \sum_{j=2}^{n} Y_{ij}, \]

\[ \bar{Y}_{i1} = (n-1)^{-1} \sum_{j=2}^{n} Y_{i,j-1}, \]

and \( \bar{Y}_i \) denotes the sample mean, \( n^{-1} \sum_{j=1}^{n} Y_{ij} \). The expressions given in (5.11) and (5.12) are approximate generalized least squares estimators of \( \theta_i \). It is assumed that \( |\rho| \leq 1 \) and that \( \theta_i = 0 \) if \( |\rho| = 1 \).

The estimators of \( \theta_i \) actually considered are defined by

\[ \tilde{\theta}_i = \hat{\theta}_i, \quad \text{if} \quad |\hat{\rho}| < 1, \]

\[ = 0, \quad \text{otherwise}. \]

These constrained estimators of \( \theta_i \) are used in the construction of the one and three period predictors.

Table 13 compares the Monte Carlo variances of the one period predictors for a series length of five years. The various predictors are unbiased and the mean square error of prediction is equal to the variance. The predictors based on the modified estimators \( \hat{\rho}_{\text{MP}} \) and \( \hat{\rho}_\perp \) very
nearly dominate the predictors associated with the estimators \( \hat{\rho}_{\text{OLS}} \) and \( \hat{\rho}_{\text{SYM}} \); in cases where the predictors \( \hat{Y}_{n+1,\text{MP}} \) and \( \hat{Y}_{n+1,\text{RHO1}} \) are inferior, the mean square errors are very near the minimum value. In particular, the predictor \( \hat{Y}_{n+1,\text{MP}} \) has the minimum mean square error in 8 of the 13 cases considered. For positive values of \( \rho \), the one step ahead predictor \( \hat{Y}_{n+1,\text{MP}} \) does considerably better than either \( \hat{Y}_{n+1,\text{OLS}} \) or \( \hat{Y}_{n+1,\text{SYM}} \). The predictor \( \hat{Y}_{n+1,\text{SYM}} \) is to be avoided for small samples based of its inferior performance for values of \( |\rho| \) near 1.0.

The large sample theory gives \( 1 + \frac{13[12(n-1)]^{-1}}{\text{mean square error of the one step predictor for } |\rho| < 1} \) as the mean square error of the one step predictor for \( |\rho| < 1 \). For values of \( |\rho| < 0.90 \) and \( n = 5 \), the Monte Carlo mean square error of \( \hat{Y}_{n+1,\text{OLS}} \) is larger than the theoretical variance. The constraints placed on the estimates produce more accurate forecasts for values of \( |\rho| \) near 1.0. For small negative values of \( \rho \), the Monte Carlo variances of the various predictors are in fair agreement with the approximate variance; however, the approximation worsens for \( 0.0 \leq \rho \leq 0.70 \).

Tables 14 and 15 contain the Monte Carlo mean square errors of one period predictions of the four estimators for values of \( |\rho| \) near 1.0 and sample sizes \( n = 10 \) and 20. The predictor \( \hat{Y}_{n+1,\text{SYM}} \) is dominated by the other one period predictors for the values of \( \rho \) considered. The one step ahead predictor \( \hat{Y}_{n+1,\text{MP}} \) has the minimum mean square error in six of the seven cases. For values of \( |\rho| \) near 1.0, the one step forecast based on \( \hat{\rho}_{\text{MP}} \) has variances very close to the minimum level of 1.0 achieved with known \( \rho \). The Monte Carlo mean square error of
\( \hat{Y}_{n+1,\text{OLS}} \) appears to be slightly larger than the approximate theoretical variance for values of \(-0.90 \leq \rho \leq 0.90\).

Table 16 compares the mean square error of three period predictions of the four estimators for a sample of five years. The Monte Carlo variance of \( \hat{Y}_{n+3,\text{MP}} \) is optimal in 5 of the 13 cases while \( \hat{Y}_{n+3,\text{RHO}} \) and \( \hat{Y}_{n+3,\text{SYM}} \) have minimum mean square error in three and four instances, respectively. In those instances where \( \hat{Y}_{n+3,\text{MP}} \) is not minimum, the variance of \( \hat{Y}_{n+3,\text{MP}} \) is near the minimum value. The least squares predictor is almost uniformly dominated by the modified least squares predictors, \( \hat{Y}_{n+3,\text{MP}} \) and \( \hat{Y}_{n+3,\text{RHO}} \). For values of \(|\rho|\) near 1.0, the mean square error performance of \( \hat{Y}_{n+3,\text{MP}} \) dominates the other three predictors and is very near the minimum level attained in the case of known parameters. The predictor \( \hat{Y}_{n+3,\text{SYM}} \) is to be avoided because of its poor performance near the boundaries of the stationary region.

One reason for the poor performances of \( \hat{Y}_{n+1,\text{SYM}} \) and \( \hat{Y}_{n+3,\text{SYM}} \) is the estimator of \( \theta_1 \) used in constructing the predictor. For \( \rho = -0.99 \), the variance of \( Y_{ij} \) is 50.25 and the variance of \( \bar{Y}_i \) is 10.05 for \( n = 5 \). The variance of \( (1-\hat{\rho}_{\text{SYM}})\bar{Y}_i \) is about double the variance of \( \bar{Y}_i \) when \( \rho = -0.99 \). For \( n = 5 \) and \( \rho = -0.99 \), the large variances of \( \hat{Y}_{n+1,\text{SYM}} \) and \( \hat{Y}_{n+3,\text{SYM}} \) are mainly due to the use of the estimator \( \hat{\theta}_{i,\text{SYM}} \). The one and three period predictors \( \hat{Y}_{n+1,\text{SYM}}^* \) and \( \hat{Y}_{n+3,\text{SYM}}^* \), defined in (5.5) and (5.6) using \( \hat{\rho}_{\text{SYM}} \) and \( \hat{\theta}_{i,\text{SYM}} \) defined by
\[
\theta_{i,\text{SYM}} = (1 - \hat{\rho}_{\text{SYM}}) \frac{\sum_{j=2}^{n-1} Y_{ij}}{2 + (1 - \hat{\rho}_{\text{SYM}})(n-2)}
\]

(5.16)

were also studied. The Monte Carlo variances of the predictors are presented in Table C.4 of Appendix C. The wild values for the variances of \(\hat{Y}_{n+1,\text{SYM}}\) and \(\hat{Y}_{n+3,\text{SYM}}\) when \(\rho = -0.99\) and \(n = 5\) are not as evident in the variances of \(\hat{Y}_{n+1,\text{SYM}}^*\) and \(\hat{Y}_{n+3,\text{SYM}}^*\); however, the variances of \(\hat{Y}_{n+1,\text{SYM}}^*\) and \(\hat{Y}_{n+3,\text{SYM}}^*\) when \(\rho = -0.99\) and \(n = 5\) are considerably larger than the variances of \(\hat{Y}_{n+1,\text{OLS}}\) and \(\hat{Y}_{n+3,\text{OLS}}\). The predictors using \(\hat{\rho}_{\text{SYM}}\) are particularly poor for values of \(|\rho|\) near one.

Table 12 presents the relative average risks of the various predictors of \(Y_{i,n+1}\) and \(Y_{i,n+3}\) expressed as a percentage of the average risks of \(\hat{Y}_{n+1,\text{OLS}}\) and \(\hat{Y}_{n+3,\text{OLS}}\), respectively for \(n = 5\). A slight preference for the predictors of \(\hat{\rho}_{\text{MP}}\) over the predictors of \(\hat{\rho}_{\text{OLS}}\) is indicated by the average risk measure.

Table 12. Average risks of various predictors of \(Y_{i,n+1}\) and \(Y_{i,n+3}\) as a percentage of the average risks of \(\hat{Y}_{n+1,\text{OLS}}\) and \(\hat{Y}_{n+3,\text{OLS}}\) averaged over the values of \(\rho\) considered.

<table>
<thead>
<tr>
<th>(s = 1)</th>
<th>(\hat{Y}_{n+s,\text{MP}})</th>
<th>(\hat{Y}_{n+s,\text{RH01}})</th>
<th>(\hat{Y}_{n+s,\text{SYM}})</th>
<th>(\hat{Y}_{n+s,\text{SYM}}^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(e = 3)</td>
<td>91.5</td>
<td>95.9</td>
<td>192.6</td>
<td>116.2</td>
</tr>
<tr>
<td>96.0</td>
<td>100.1</td>
<td>169.1</td>
<td>122.9</td>
<td></td>
</tr>
</tbody>
</table>
Table 13. Empirical mean square errors of various predictors of $Y_{i,n+1}$ for $n = 5$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+1,OLS}$</th>
<th>$\hat{Y}_{n+1,MP}$</th>
<th>$\hat{Y}_{n+1,RH01}$</th>
<th>$\hat{Y}_{n+1,SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>1.083</td>
<td>1.049</td>
<td>1.086</td>
<td>1.972</td>
</tr>
<tr>
<td>-0.99</td>
<td>1.166</td>
<td>1.021</td>
<td>1.182</td>
<td>11.345</td>
</tr>
<tr>
<td>-0.90</td>
<td>1.251</td>
<td>1.221</td>
<td>1.256</td>
<td>1.904</td>
</tr>
<tr>
<td>-0.70</td>
<td>1.293</td>
<td>1.296</td>
<td>1.276</td>
<td>1.320</td>
</tr>
<tr>
<td>-0.50</td>
<td>1.300</td>
<td>1.294</td>
<td>1.268</td>
<td>1.253</td>
</tr>
<tr>
<td>-0.30</td>
<td>1.309</td>
<td>1.269</td>
<td>1.262</td>
<td>1.248</td>
</tr>
<tr>
<td>0.00</td>
<td>1.322</td>
<td>1.244</td>
<td>1.236</td>
<td>1.254</td>
</tr>
<tr>
<td>0.30</td>
<td>1.360</td>
<td>1.217</td>
<td>1.234</td>
<td>1.323</td>
</tr>
<tr>
<td>0.50</td>
<td>1.363</td>
<td>1.197</td>
<td>1.218</td>
<td>1.380</td>
</tr>
<tr>
<td>0.70</td>
<td>1.347</td>
<td>1.152</td>
<td>1.208</td>
<td>1.468</td>
</tr>
<tr>
<td>0.90</td>
<td>1.275</td>
<td>1.089</td>
<td>1.184</td>
<td>1.569</td>
</tr>
<tr>
<td>0.99</td>
<td>1.205</td>
<td>1.041</td>
<td>1.166</td>
<td>1.607</td>
</tr>
<tr>
<td>1.00</td>
<td>2.209</td>
<td>1.042</td>
<td>1.167</td>
<td>1.612</td>
</tr>
</tbody>
</table>

The large sample theory gives $(1+p^2+p^4) + (n-1)^{-1}[9p^6/12 + (1+p+p^2)^2]$ as the mean square error of three period prediction for $|\rho| < 1$. For $-0.90 < \rho < 0.50$, the Monte Carlo variance of $\hat{Y}_{n+3,OLS}$ is in fair agreement with the approximation. For values of $|\rho|$ near 1.0, the truncation used in the estimation gives more accurate predictors than the corresponding predictors based on
unconstrained estimates. For negative values of $\rho$, the approximate variance underestimates the Monte Carlo variance of $\hat{Y}_{n+3,OLS}$ and overestimates the Monte Carlo variance of $\hat{Y}_{n+3,OLS}$ for positive values of $\rho$.

Table 14. Empirical mean square errors of various predictors of $Y_{1,n+1}$ for $n = 10$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+1,OLS}$</th>
<th>$\hat{Y}_{n+1,MP}$</th>
<th>$\hat{Y}_{n+1,RH01}$</th>
<th>$\hat{Y}_{n+1,SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>1.054</td>
<td>1.034</td>
<td>1.055</td>
<td>1.448</td>
</tr>
<tr>
<td>-0.99</td>
<td>1.094</td>
<td>1.034</td>
<td>1.094</td>
<td>1.199</td>
</tr>
<tr>
<td>-0.90</td>
<td>1.123</td>
<td>1.126</td>
<td>1.123</td>
<td>1.166</td>
</tr>
<tr>
<td>0.70</td>
<td>1.161</td>
<td>1.084</td>
<td>1.091</td>
<td>1.189</td>
</tr>
<tr>
<td>0.90</td>
<td>1.150</td>
<td>1.058</td>
<td>1.089</td>
<td>1.258</td>
</tr>
<tr>
<td>0.99</td>
<td>1.114</td>
<td>1.029</td>
<td>1.089</td>
<td>1.316</td>
</tr>
<tr>
<td>1.00</td>
<td>1.107</td>
<td>1.024</td>
<td>1.088</td>
<td>1.321</td>
</tr>
</tbody>
</table>

Tables 17 and 18 contain the Monte Carlo mean square errors of the various three-step ahead predictors for samples of size 10 and 20 years for values of $\rho$ near 1.0. The three period predictor $\hat{Y}_{n+3,MP}$ has the minimum mean square error in five of the seven cases and is very near the optimal level in the other two cases. It is expected that the mean square error performance of the three period predictors of the four estimates are similar for values of $\rho$ near zero. Also, the Monte Carlo
variances of $\hat{Y}_{n+3,OLS}$ appear to be in close agreement with the asymptotic variance for values of $-0.90 \leq \rho \leq 0.90$.

The estimated standard errors for the mean square errors of Tables 13 to 18 were computed using the formula

$$s.e.(\text{MSE}) = [1000]^{-1}[999]^{-1} \sum_{\lambda=1}^{1000} (\text{MSE}_{\lambda} - \text{MSE})^2, \quad (5.17)$$

where $\text{MSE}$ is the sample mean, $(1000)^{-1} \sum_{\lambda=1}^{1000} \text{MSE}_{\lambda}$, and

$\text{MSE}_{\lambda} = (12)^{-1} \sum_{i=0}^{11} \text{MSE}_{i,\lambda}$, where $\text{MSE}_{i,\lambda}$ denotes the estimated mean square error of prediction for the $i$-th period due to the estimated parameters for the $\lambda$-th sample. The mean square error of the one period predictor for the $i$-th period due to the estimated parameters for the $\lambda$-th sample is

$$\text{MSE}_{1,\lambda} = [\hat{\theta}_{1,\lambda} - \hat{\theta}] + (\rho - \hat{\rho}) (Y_{in})^2,$$

where $(Y_{in})_{\lambda}$ is the observation for the $i$-th period for the $n$-th cycle for the $\lambda$-th sample. The values reported in the tables are not $\text{MSE}$, but $\text{MSE} + \sum_{m=0}^{s-1} \rho^m$ for the $s$-period predictor. The estimated standard errors of the estimated mean square errors of $\hat{Y}_{n+1,OLS}$ and $\hat{Y}_{n+3,OLS}$ for $n = 5$ were about 0.0054 and 0.0069, respectively. The estimated standard errors for $n = 5$ were about 0.0042 and 0.0053 for the mean.
square errors of $\hat{Y}_{n+1,\text{OLS}}$ and $\hat{Y}_{n+1,\text{MP}}$. The estimated standard errors for the mean square errors of $\hat{Y}_{n+1,\text{OLS}}$ and $\hat{Y}_{n+1,\text{MP}}$ were about 0.0025 and 0.0035 for $n = 10$ and about 0.0011 and 0.0025 for $n = 20$. The estimated standard errors for the mean square errors of $\hat{Y}_{n+1,\text{OLS}}$ and $\hat{Y}_{n+1,\text{MP}}$ were slightly less than the estimated standard errors for the mean square errors of $\hat{Y}_{n+1,\text{OLS}}$ and $\hat{Y}_{n+1,\text{MP}}$. The empirical mean square errors of $\hat{Y}_{n+1,\text{OLS}}$ and $\hat{Y}_{n+1,\text{MP}}$ were slightly more than two standard errors from the theoretical mean square errors for most values of $n$ and $\rho$.

Table 15. Empirical mean square errors of various predictors of $Y_{i,n+1}$ for $n = 20$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+1,\text{OLS}}$</th>
<th>$\hat{Y}_{n+1,\text{MP}}$</th>
<th>$\hat{Y}_{n+1,\text{RH01}}$</th>
<th>$\hat{Y}_{n+1,\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>1.030</td>
<td>1.023</td>
<td>1.030</td>
<td>1.235</td>
</tr>
<tr>
<td>-0.99</td>
<td>1.049</td>
<td>1.034</td>
<td>1.049</td>
<td>1.097</td>
</tr>
<tr>
<td>-0.90</td>
<td>1.058</td>
<td>1.059</td>
<td>1.058</td>
<td>1.075</td>
</tr>
<tr>
<td>0.70</td>
<td>1.072</td>
<td>1.049</td>
<td>1.049</td>
<td>1.077</td>
</tr>
<tr>
<td>0.90</td>
<td>1.078</td>
<td>1.037</td>
<td>1.043</td>
<td>1.110</td>
</tr>
<tr>
<td>0.99</td>
<td>1.062</td>
<td>1.019</td>
<td>1.046</td>
<td>1.155</td>
</tr>
<tr>
<td>1.00</td>
<td>1.058</td>
<td>1.015</td>
<td>1.047</td>
<td>1.159</td>
</tr>
</tbody>
</table>
Table 16. Empirical mean square errors of various predictors of $Y_{i,n+3}$ for $n = 5$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+3,OLS}$</th>
<th>$\hat{Y}_{n+3,MP}$</th>
<th>$\hat{Y}_{n+3,RHO1}$</th>
<th>$\hat{Y}_{n+3,SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>3.115</td>
<td>3.075</td>
<td>3.134</td>
<td>5.191</td>
</tr>
<tr>
<td>-0.99</td>
<td>3.171</td>
<td>3.001</td>
<td>3.242</td>
<td>18.085</td>
</tr>
<tr>
<td>-0.90</td>
<td>2.802</td>
<td>2.857</td>
<td>2.817</td>
<td>3.370</td>
</tr>
<tr>
<td>-0.70</td>
<td>2.017</td>
<td>2.045</td>
<td>1.975</td>
<td>1.930</td>
</tr>
<tr>
<td>-0.50</td>
<td>1.523</td>
<td>1.491</td>
<td>1.463</td>
<td>1.460</td>
</tr>
<tr>
<td>-0.30</td>
<td>1.279</td>
<td>1.251</td>
<td>1.249</td>
<td>1.246</td>
</tr>
<tr>
<td>0.00</td>
<td>1.234</td>
<td>1.206</td>
<td>1.198</td>
<td>1.197</td>
</tr>
<tr>
<td>0.30</td>
<td>1.511</td>
<td>1.444</td>
<td>1.440</td>
<td>1.442</td>
</tr>
<tr>
<td>0.50</td>
<td>1.886</td>
<td>1.813</td>
<td>1.768</td>
<td>1.794</td>
</tr>
<tr>
<td>0.70</td>
<td>2.445</td>
<td>2.312</td>
<td>2.310</td>
<td>2.414</td>
</tr>
<tr>
<td>0.90</td>
<td>3.129</td>
<td>2.910</td>
<td>3.085</td>
<td>3.424</td>
</tr>
<tr>
<td>0.99</td>
<td>3.361</td>
<td>3.098</td>
<td>3.504</td>
<td>4.040</td>
</tr>
<tr>
<td>1.00</td>
<td>3.399</td>
<td>3.157</td>
<td>3.547</td>
<td>4.101</td>
</tr>
</tbody>
</table>

The prediction of an aggregate sum of future values is often of interest in economics. The linear combination of future values given by $Y_{i,n+1} + Y_{i,n+2} + Y_{i,n+3}$, $i = 0,1,2,\ldots,11$, is commonly forecast by the predictor $\hat{Y}_{i,n+1} + \hat{Y}_{i,n+2} + \hat{Y}_{i,n+3}$, $i = 0$ to 11, where $\hat{Y}_{i,n+s}$ is the $s$ period predictor given in (5.5) and (5.6) for $s = 1$.
and 3, respectively. From (5.9) the error in predicting the aggregate sum of future values is

\[ \sum_{s=1}^{3} \left( Y_{i,n+s} - \hat{Y}_{i,n+s} \right) = \sum_{s=1}^{3} \sum_{j=0}^{s-1} \rho^j \varepsilon_{i,n+s-j} + \sum_{s=1}^{3} \left[ \sum_{j=1}^{s-1} \theta_i \varepsilon_{i,n+s-j} \right] + (\rho^s - \hat{\rho}^s) Y_{in} \]  

(5.14)

and only the last sum requires simulation. Table 19 presents the empirical variances of the predictors of the aggregate sum using \( \hat{\rho}_{OLS} \) and \( \hat{\rho}_{MP} \). Table 20 compares the empirical variances of the second term in (5.14) for the two predictors.

With the exception of \( \rho = -0.90 \), the predictors of \( \hat{\rho}_{MP} \) have smaller mean square errors than the predictors of \( \hat{\rho}_{OLS} \). Because \( \hat{Y}_{i,n+s} \) is an unbiased estimator for \( Y_{i,n+s} \), the predictors of the aggregate sum \( Y_{i,n+1} + Y_{i,n+2} + Y_{i,n+3} \) are unbiased and the mean square errors of the predictors are also the variances. From Table 20 the portion of the variances of the predictors due to estimation of the parameters is appreciably larger for the predictors using \( \hat{\rho}_{OLS} \) than for the predictors using \( \hat{\rho}_{MP} \). The average risks of the estimation error using \( \hat{\rho}_{MP} \) as a percentage of the average risks of the estimation error using \( \hat{\rho}_{OLS} \) when predicting \( Y_{i,n+1} + Y_{i,n+2} + Y_{i,n+3} \) are 62.9, 57.9 and 58.0 percent for \( n = 5, 10 \) and 20 years, respectively.
Table 17. Empirical mean square errors of various predictors of $Y_{i,n+3}$ for $n = 10$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+3,\text{OLS}}$</th>
<th>$\hat{Y}_{n+3,\text{MP}}$</th>
<th>$\hat{Y}_{n+3,\text{RH01}}$</th>
<th>$\hat{Y}_{n+3,\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>3.086</td>
<td>3.056</td>
<td>3.093</td>
<td>4.224</td>
</tr>
<tr>
<td>-0.99</td>
<td>3.077</td>
<td>3.013</td>
<td>3.084</td>
<td>3.145</td>
</tr>
<tr>
<td>-0.90</td>
<td>2.629</td>
<td>2.652</td>
<td>2.630</td>
<td>2.632</td>
</tr>
<tr>
<td>0.70</td>
<td>2.223</td>
<td>2.090</td>
<td>2.085</td>
<td>2.273</td>
</tr>
<tr>
<td>0.90</td>
<td>3.105</td>
<td>2.822</td>
<td>2.928</td>
<td>3.405</td>
</tr>
<tr>
<td>0.99</td>
<td>3.429</td>
<td>3.122</td>
<td>3.431</td>
<td>4.259</td>
</tr>
<tr>
<td>1.00</td>
<td>3.452</td>
<td>3.149</td>
<td>3.489</td>
<td>4.374</td>
</tr>
</tbody>
</table>

Table 18. Empirical mean square errors of various predictors of $Y_{i,n+3}$ for $n = 20$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+3,\text{OLS}}$</th>
<th>$\hat{Y}_{n+3,\text{MP}}$</th>
<th>$\hat{Y}_{n+3,\text{RH01}}$</th>
<th>$\hat{Y}_{n+3,\text{SYM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>3.056</td>
<td>3.040</td>
<td>3.059</td>
<td>3.731</td>
</tr>
<tr>
<td>-0.99</td>
<td>3.013</td>
<td>3.004</td>
<td>3.015</td>
<td>3.046</td>
</tr>
<tr>
<td>-0.90</td>
<td>2.536</td>
<td>2.545</td>
<td>2.537</td>
<td>2.544</td>
</tr>
<tr>
<td>0.70</td>
<td>2.005</td>
<td>1.946</td>
<td>1.942</td>
<td>2.000</td>
</tr>
<tr>
<td>0.90</td>
<td>2.901</td>
<td>2.640</td>
<td>2.736</td>
<td>3.033</td>
</tr>
<tr>
<td>0.99</td>
<td>3.322</td>
<td>3.079</td>
<td>3.264</td>
<td>3.898</td>
</tr>
<tr>
<td>1.00</td>
<td>3.356</td>
<td>3.113</td>
<td>3.331</td>
<td>3.996</td>
</tr>
</tbody>
</table>
The estimated standard errors for the variances of Tables 19 and 20 were computed using (5.17). The estimated standard errors for the variances of the predictors using \( \hat{\rho}_{OLS} \) and \( \hat{\rho}_{MP} \) were about 0.043 and 0.035 for \( n = 5 \), about 0.030 and 0.020 for \( n = 10 \) and about 0.018 and 0.011 for \( n = 20 \).

Analogous to the univariate regression model, the distribution of the regression "t-statistic" for \( \rho \) is commonly compared with Student's t with \( n(k-1) - (k+1) \) degrees of freedom when testing for \( \rho = \rho_0 \), \( |\rho_0| < 1 \). When testing for \( \rho = \rho_0 \), \( |\rho_0| = 1 \), the regression "t-statistic" has a distribution which is considerably different from Student's t under the null hypothesis. Tables of the empirical percentiles of the "t-statistic" when testing for a unit root are given by Dickey and Fuller (1979). The regression "t-statistic" for an estimator of \( \rho \) is defined by

\[
t = (\hat{\rho} - \rho)(s^2)_{\hat{\rho}_{OLS}}^{-1/2},
\]

where

\[
s^2_{\hat{\rho}_{OLS}} = \frac{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - \hat{\rho}_{i,OLS}Y_{i,j-1})^2}{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j-1} - \overline{Y}_{i}^2)}.
\]

The "t-statistics" for the four estimators of \( \rho \) are denoted by \( t_{OLS} \), \( t_{MP} \), \( t_{RHO1} \) and \( t_{SYM} \).

Tables 21 and 22 present the means and variances of the various "t-statistics" obtained from a sample of 1,000 for each \( (\rho, n) \) combination. The empirical mean of \( t_{OLS} \) is a decreasing function of \( \rho \) with
fairly large negative values for \( \rho \) near one. The mean of \( t_{\text{SYM}} \) has a similar trend with the exception of \( \rho = -0.99 \) and \( n = 5 \). The means of the "t-statistics" for the two modified least squares estimators, \( \hat{\rho}_{MP} \) and \( \hat{\rho}_1 \), have smaller absolute values than either \( t_{\text{OLS}} \) or \( t_{\text{SYM}} \). The empirical means of the \( t_{\text{OLS}} \) and \( t_{\text{SYM}} \) are declining by \( n^{-\frac{1}{2}} \) for \(-0.99 \leq \rho \leq 0.70\) and the empirical means of \( t_{\text{MP}} \) and \( t_{\text{RHO1}} \) are declining by \( n^{-1} \) for \(-0.7 \leq \rho \leq 0.7\).

The estimated standard errors for \( n = 5 \) range from 0.012 to 0.040 for the means of the various "t-statistics." For \( n = 10 \) and 20, the estimated standard errors for the various means range from 0.012 to 0.036. For \(-0.30 \leq \rho < 0.30\), the mean of the "t-statistic" for \( \hat{\rho}_{MP} \) is within two standard errors of zero. The means of \( t_{\text{OLS}}, t_{\text{RHO1}} \) and \( t_{\text{SYM}} \) are significantly different from zero, the theoretical mean of Student's t.

For \( n = 5 \) and \( \rho \geq 0.30 \), the means of \( t_{\text{OLS}} \) and \( t_{\text{SYM}} \) are less than the 2.5-th percentile of Student's t with 35 degrees of freedom. For \( \rho > 0.70 \), the means of \( t_{\text{OLS}} \) and \( t_{\text{SYM}} \) are less than the 0.5-th percentile of Student's t with \( k(n-2)-1 \) degrees of freedom. For \( n = 5 \) and \( \rho = -0.99 \), the mean of \( t_{\text{SYM}} \) exceeds the 99.5-th percentile of Student's t with 35 degrees of freedom. When \( \rho = -1 \), the mean of \( t_{\text{SYM}} \) exceeds the 2.5-th percentile of the "t-statistic" given by Dickey and Fuller (1979). The "t-statistics" are clearly not symmetrically distributed about zero.

The standard errors of the variances of the "t-statistics" were estimated using the standard moment formula given in (5.4). The estimated
Table 19. Empirical variances of various predictors of $Y_{i,n+1} + Y_{i,n+2} + Y_{i,n+3}$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n 5 10 20</td>
<td>n 5 10 20</td>
</tr>
<tr>
<td>-1.00</td>
<td>2.31 2.23 2.13</td>
<td>2.19 2.16 2.09</td>
</tr>
<tr>
<td>-0.99</td>
<td>2.66 2.35 2.18</td>
<td>2.08 2.12 2.12</td>
</tr>
<tr>
<td>-0.90</td>
<td>2.85 2.31 2.06</td>
<td>2.71 2.32 2.06</td>
</tr>
<tr>
<td>-0.70</td>
<td>2.87 2.21 1.95</td>
<td>2.83 2.20 1.95</td>
</tr>
<tr>
<td>-0.30</td>
<td>3.69 2.81 2.45</td>
<td>3.53 2.78 2.44</td>
</tr>
<tr>
<td>0.00</td>
<td>5.17 3.99 3.49</td>
<td>4.88 3.91 3.47</td>
</tr>
<tr>
<td>0.30</td>
<td>7.82 6.15 5.37</td>
<td>7.20 5.93 5.31</td>
</tr>
<tr>
<td>0.70</td>
<td>13.15 11.49 10.15</td>
<td>11.84 10.51 9.77</td>
</tr>
<tr>
<td>0.90</td>
<td>15.80 15.19 14.03</td>
<td>14.00 13.54 13.02</td>
</tr>
<tr>
<td>0.99</td>
<td>16.78 16.28 15.53</td>
<td>14.80 14.63 14.36</td>
</tr>
<tr>
<td>1.00</td>
<td>16.80 16.37 15.59</td>
<td>14.85 14.69 14.47</td>
</tr>
</tbody>
</table>

Table 20. Empirical variances of estimation error in various predictors of $Y_{i,n+1} + Y_{i,n+2} + Y_{i,n+3}$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n 5 10 20</td>
<td>n 5 10 20</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.31 0.23 0.13</td>
<td>0.19 0.16 0.09</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.68 0.37 0.20</td>
<td>0.10 0.14 0.14</td>
</tr>
<tr>
<td>-0.90</td>
<td>1.01 0.48 0.22</td>
<td>0.87 0.48 0.23</td>
</tr>
<tr>
<td>-0.70</td>
<td>1.16 0.50 0.24</td>
<td>1.11 0.49 0.23</td>
</tr>
<tr>
<td>-0.30</td>
<td>1.57 0.70 0.33</td>
<td>1.41 0.67 0.33</td>
</tr>
<tr>
<td>0.00</td>
<td>2.17 0.99 0.49</td>
<td>1.88 0.91 0.47</td>
</tr>
<tr>
<td>0.30</td>
<td>3.20 1.53 0.75</td>
<td>2.58 1.30 0.68</td>
</tr>
<tr>
<td>0.70</td>
<td>4.47 2.81 1.47</td>
<td>3.15 1.83 1.09</td>
</tr>
<tr>
<td>0.90</td>
<td>3.85 3.24 2.07</td>
<td>2.05 1.58 1.07</td>
</tr>
<tr>
<td>0.99</td>
<td>3.00 2.50 1.75</td>
<td>1.02 0.85 0.58</td>
</tr>
<tr>
<td>1.00</td>
<td>2.80 2.37 1.59</td>
<td>0.85 0.69 0.47</td>
</tr>
</tbody>
</table>
standard errors of the variances of $t_{OLS}$ and $t_{MP}$ were about 0.045 and 0.080, respectively, for $n=5$ and about 0.041 and 0.046 for $n=20$. The theoretical variance of a random variable from a Student's t distribution with $m$ degrees of freedom is $m(m-2)^{-1}$ for $m > 2$. For $n = 5$ and $-0.90 \leq \rho \leq 0.70$, the empirical variance of $t_{OLS}$ is about four standard errors from the theoretical variance of 1.061. As $n$ increases, the empirical variances of $t_{OLS}$ are approaching the theoretical variances. The empirical variances of $t_{MP}$ are significantly different from the theoretical variances for $n = 5$ and are in fair agreement with the theoretical variances for $n = 10$ and 20. The empirical variances of $t_{RHO1}$ are in fair agreement with the theoretical variances for $-0.90 \leq \rho \leq 0.99$ and are within four standard errors of the theoretical variances. The empirical variances of $t_{SYM}$ are significantly different from the theoretical variances for $n = 5$ and are in fair agreement with the theoretical variances over a much smaller interval of $\rho$ than the other three "t-statistics" for $n = 10$ and 20.

The "t-statistics" for the various estimators of $\rho$ are compared with the percentiles of a Student's t when testing for $\rho = \rho_0$, $|\rho_0| < 1$, and with the empirical percentiles of the "t-statistic" obtained by Dickey and Fuller (1979) when testing for $\rho = \rho_0$, $|\rho_0| = 1$. The empirical frequencies for the two tailed test are presented in Table 23 and the empirical frequencies for the one tailed test with alternative hypothesis $\rho > \rho_0$ are presented in Table 24.

For $-0.99 \leq \rho \leq -0.70$, the empirical frequencies of $t_{OLS}$ exceeding the critical value for the two tailed test are in reasonable agreement with the expected frequencies of Student's t. For $-0.30 \leq \rho \leq 0.99$,
Table 21. Empirical mean of the "t-statistic" of various estimators of $\rho$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$t_{OLS}$</th>
<th></th>
<th></th>
<th></th>
<th>$t_{MP}$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>-1.00</td>
<td>0.19</td>
<td>0.34</td>
<td>0.41</td>
<td></td>
<td>0.13</td>
<td>0.23</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>-0.99</td>
<td>0.18</td>
<td>0.16</td>
<td>0.13</td>
<td>-0.40</td>
<td>-0.50</td>
<td>-0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.90</td>
<td>-0.27</td>
<td>-0.19</td>
<td>-0.07</td>
<td>-0.50</td>
<td>-0.36</td>
<td>-0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.70</td>
<td>-0.64</td>
<td>-0.42</td>
<td>-0.30</td>
<td>-0.28</td>
<td>-0.14</td>
<td>-0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.30</td>
<td>-1.18</td>
<td>-0.78</td>
<td>-0.54</td>
<td>-0.09</td>
<td>-0.03</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>-1.54</td>
<td>-1.12</td>
<td>-0.77</td>
<td>0.01</td>
<td>-0.03</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.30</td>
<td>-2.09</td>
<td>-1.59</td>
<td>-1.08</td>
<td>-0.09</td>
<td>-0.12</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>-3.07</td>
<td>-2.56</td>
<td>-1.86</td>
<td>-0.43</td>
<td>-0.36</td>
<td>-0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.90</td>
<td>-3.70</td>
<td>-3.52</td>
<td>-3.00</td>
<td>-0.76</td>
<td>-0.76</td>
<td>-0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.99</td>
<td>-4.05</td>
<td>-4.13</td>
<td>-4.11</td>
<td>-1.01</td>
<td>-1.15</td>
<td>-1.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00</td>
<td>-4.10</td>
<td>-4.20</td>
<td>-4.29</td>
<td>-1.04</td>
<td>-1.18</td>
<td>-1.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$t_{\text{RH01}}$</td>
<td></td>
<td>$t_{\text{SYM}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>-------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.23</td>
<td>0.38</td>
<td>0.43</td>
<td>2.38</td>
<td>2.36</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.33</td>
<td>0.26</td>
<td>0.19</td>
<td>4.13</td>
<td>0.45</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td>0.11</td>
<td>0.95</td>
<td>0.09</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.11</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.09</td>
<td>-0.27</td>
<td>-0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.36</td>
<td>-0.09</td>
<td>-0.01</td>
<td>-1.04</td>
<td>-0.73</td>
<td>-0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.54</td>
<td>-0.23</td>
<td>-0.07</td>
<td>-1.57</td>
<td>-1.12</td>
<td>-0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.94</td>
<td>-0.50</td>
<td>-0.18</td>
<td>-2.32</td>
<td>-1.65</td>
<td>-1.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.76</td>
<td>-1.18</td>
<td>-0.59</td>
<td>-3.75</td>
<td>-2.77</td>
<td>-1.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.32</td>
<td>-2.01</td>
<td>-1.49</td>
<td>-4.72</td>
<td>-3.91</td>
<td>-3.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.64</td>
<td>-2.58</td>
<td>-2.50</td>
<td>-5.25</td>
<td>-4.70</td>
<td>-4.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.68</td>
<td>-2.64</td>
<td>-2.67</td>
<td>-5.31</td>
<td>-4.78</td>
<td>-4.58</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 22. Empirical variances of the "t-statistic" of various estimators of $\rho$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$t_{OLS}$</th>
<th></th>
<th></th>
<th>$t_{MP}$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n 5</td>
<td>10</td>
<td>20</td>
<td></td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>-1.00</td>
<td>0.15</td>
<td>0.28</td>
<td>0.33</td>
<td>0.14</td>
<td>0.23</td>
<td>0.24</td>
</tr>
<tr>
<td>-0.99</td>
<td>0.47</td>
<td>0.64</td>
<td>0.82</td>
<td>0.06</td>
<td>0.26</td>
<td>0.61</td>
</tr>
<tr>
<td>-0.90</td>
<td>0.85</td>
<td>1.02</td>
<td>0.98</td>
<td>1.31</td>
<td>1.35</td>
<td>1.13</td>
</tr>
<tr>
<td>-0.70</td>
<td>0.89</td>
<td>0.92</td>
<td>0.93</td>
<td>1.62</td>
<td>1.16</td>
<td>1.03</td>
</tr>
<tr>
<td>-0.30</td>
<td>0.86</td>
<td>0.89</td>
<td>0.95</td>
<td>1.52</td>
<td>1.12</td>
<td>1.05</td>
</tr>
<tr>
<td>0.00</td>
<td>0.84</td>
<td>0.94</td>
<td>0.97</td>
<td>1.51</td>
<td>1.22</td>
<td>1.10</td>
</tr>
<tr>
<td>0.30</td>
<td>0.85</td>
<td>0.85</td>
<td>0.96</td>
<td>1.57</td>
<td>1.16</td>
<td>1.13</td>
</tr>
<tr>
<td>0.70</td>
<td>1.16</td>
<td>0.87</td>
<td>0.94</td>
<td>1.88</td>
<td>1.34</td>
<td>1.26</td>
</tr>
<tr>
<td>0.90</td>
<td>1.29</td>
<td>0.88</td>
<td>0.85</td>
<td>1.59</td>
<td>1.38</td>
<td>1.31</td>
</tr>
<tr>
<td>0.99</td>
<td>1.40</td>
<td>0.98</td>
<td>0.86</td>
<td>1.44</td>
<td>1.21</td>
<td>1.16</td>
</tr>
<tr>
<td>1.00</td>
<td>1.36</td>
<td>0.95</td>
<td>0.89</td>
<td>1.37</td>
<td>1.12</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>$t_{RH01}$</td>
<td></td>
<td></td>
<td>$t_{SYM}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>5</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>0.23</td>
<td>0.35</td>
<td>0.36</td>
<td>0.07</td>
<td>0.15</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>0.74</td>
<td>0.78</td>
<td>0.90</td>
<td>0.92</td>
<td>0.05</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>1.26</td>
<td>1.20</td>
<td>1.05</td>
<td>0.12</td>
<td>0.57</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>1.23</td>
<td>1.06</td>
<td>0.99</td>
<td>0.51</td>
<td>0.80</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>1.13</td>
<td>1.03</td>
<td>1.02</td>
<td>0.71</td>
<td>0.86</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>1.06</td>
<td>1.09</td>
<td>1.06</td>
<td>0.77</td>
<td>0.92</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>1.02</td>
<td>0.99</td>
<td>1.07</td>
<td>0.68</td>
<td>0.82</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>1.18</td>
<td>0.97</td>
<td>1.07</td>
<td>0.58</td>
<td>0.67</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>1.16</td>
<td>0.86</td>
<td>0.90</td>
<td>0.43</td>
<td>0.44</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>1.20</td>
<td>0.86</td>
<td>0.76</td>
<td>0.39</td>
<td>0.32</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>1.16</td>
<td>0.82</td>
<td>0.76</td>
<td>0.38</td>
<td>0.28</td>
<td>0.29</td>
<td></td>
</tr>
</tbody>
</table>
$p < 0.99$, the number of rejections of the null hypothesis using the test statistic $t_{OLS}$ for the two tailed test are larger than the theoretical frequencies of Student's $t$. For values of $\rho$ near one, the number of computed "$t$-statistics" for $t_{OLS}$ exceeding the percentiles of Student's $t$ for the two tailed test is considerably larger than the expected frequencies of Student's $t$. For $n = 0.5$ and $0.30 < \rho < 0.99$, the least squares estimator $\hat{\rho}_{OLS}$ differs from $\rho$ by more than two standard errors in over half of the samples. For $n = 5$ and $\rho = 1.0$, the critical values of the "$t$-statistic" were extrapolated from the tables given by Dickey and Fuller (1979). The empirical frequencies of $t_{OLS}$ exceeding these critical values are in fair agreement with the expected frequencies. For $\rho = 1.0$ and $n = 10$ and 20, the number of times $t_{OLS}$ exceeds the percentiles of the "$t$-statistic" are in close agreement with the expected frequencies. With the exception of $\rho = -0.99$, the agreement between the theoretical and observed frequencies improves as the sample size increases.

With the exception of $\rho = -0.99$ and 1.0, the "$t$-statistic" for $\hat{\rho}_{SYM}$ behaves similarly to the "$t$-statistic" for $\hat{\rho}_{OLS}$. For $n = 5$ and $\rho = -0.99$, the two tailed $t$-test for $\hat{\rho}_{SYM}$ using the 97.5-th percentile of Student's $t$ with 35 degrees of freedom rejects the null hypothesis in 950 of the 1,000 samples. For $\rho = 1.0$ and $n = 10$ and 20, the empirical frequencies of $t_{SYM}$ exceeding the critical values of the "$t$-statistic" are less than the expected frequencies.

The empirical frequencies of $t_{MP}$ and $t_{RHOL}$ for the two tailed test are closer to the theoretical frequencies than either of the
Table 23. Number of "t-statistics" in 1,000 trials exceeding the critical value for the two tailed test

<table>
<thead>
<tr>
<th>True value of ρ</th>
<th>t\textsuperscript{OLS}</th>
<th>t\textsuperscript{MP}</th>
<th>t\textsuperscript{RH01}</th>
<th>t\textsuperscript{SYM}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>-0.99</td>
<td>5</td>
<td>18</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>-0.90</td>
<td>5</td>
<td>19</td>
<td>0</td>
<td>65</td>
</tr>
<tr>
<td>-0.70</td>
<td>5</td>
<td>80</td>
<td>17</td>
<td>128</td>
</tr>
<tr>
<td>-0.30</td>
<td>5</td>
<td>176</td>
<td>46</td>
<td>84</td>
</tr>
<tr>
<td>0.00</td>
<td>5</td>
<td>284</td>
<td>101</td>
<td>107</td>
</tr>
<tr>
<td>0.30</td>
<td>5</td>
<td>517</td>
<td>234</td>
<td>98</td>
</tr>
<tr>
<td>0.70</td>
<td>5</td>
<td>826</td>
<td>628</td>
<td>134</td>
</tr>
<tr>
<td>0.90</td>
<td>5</td>
<td>938</td>
<td>801</td>
<td>157</td>
</tr>
<tr>
<td>0.99</td>
<td>5</td>
<td>964</td>
<td>879</td>
<td>191</td>
</tr>
<tr>
<td>1.00</td>
<td>5</td>
<td>96</td>
<td>36</td>
<td>833</td>
</tr>
<tr>
<td>-0.99</td>
<td>10</td>
<td>23</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>-0.90</td>
<td>10</td>
<td>59</td>
<td>6</td>
<td>106</td>
</tr>
<tr>
<td>-0.70</td>
<td>10</td>
<td>58</td>
<td>12</td>
<td>66</td>
</tr>
<tr>
<td>-0.30</td>
<td>10</td>
<td>110</td>
<td>28</td>
<td>60</td>
</tr>
<tr>
<td>0.00</td>
<td>10</td>
<td>196</td>
<td>62</td>
<td>72</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
<td>339</td>
<td>118</td>
<td>72</td>
</tr>
<tr>
<td>0.70</td>
<td>10</td>
<td>715</td>
<td>458</td>
<td>99</td>
</tr>
<tr>
<td>0.90</td>
<td>10</td>
<td>945</td>
<td>826</td>
<td>151</td>
</tr>
<tr>
<td>0.99</td>
<td>10</td>
<td>983</td>
<td>930</td>
<td>228</td>
</tr>
<tr>
<td>1.00</td>
<td>10</td>
<td>48</td>
<td>15</td>
<td>844</td>
</tr>
<tr>
<td>-0.99</td>
<td>20</td>
<td>34</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>-0.90</td>
<td>20</td>
<td>47</td>
<td>9</td>
<td>64</td>
</tr>
<tr>
<td>-0.70</td>
<td>20</td>
<td>51</td>
<td>18</td>
<td>48</td>
</tr>
<tr>
<td>-0.30</td>
<td>20</td>
<td>92</td>
<td>19</td>
<td>52</td>
</tr>
<tr>
<td>0.00</td>
<td>20</td>
<td>114</td>
<td>32</td>
<td>62</td>
</tr>
</tbody>
</table>
Table 23 (continued)

<table>
<thead>
<tr>
<th>True value of ρ</th>
<th>n</th>
<th>$t_{OLS}$</th>
<th>$t_{MP}$</th>
<th>$t_{RH01}$</th>
<th>$t_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
<td>183</td>
<td>61</td>
<td>62</td>
<td>18</td>
</tr>
<tr>
<td>0.70</td>
<td>20</td>
<td>462</td>
<td>231</td>
<td>83</td>
<td>27</td>
</tr>
<tr>
<td>0.90</td>
<td>20</td>
<td>870</td>
<td>663</td>
<td>118</td>
<td>40</td>
</tr>
<tr>
<td>0.99</td>
<td>20</td>
<td>983</td>
<td>947</td>
<td>241</td>
<td>109</td>
</tr>
<tr>
<td>1.00</td>
<td>20</td>
<td>49</td>
<td>20</td>
<td>839</td>
<td>788</td>
</tr>
</tbody>
</table>

Table 24. Number of computed "t-statistics" in 1,000 trials exceeding the value for the one tailed test with alternative hypothesis $p < p_0$

<table>
<thead>
<tr>
<th>True value of ρ</th>
<th>n</th>
<th>$t_{OLS}$</th>
<th>$t_{MP}$</th>
<th>$t_{RH01}$</th>
<th>$t_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.05</td>
<td>0.025</td>
<td>0.05</td>
<td>0.025</td>
</tr>
<tr>
<td>-1.00</td>
<td>5</td>
<td>9</td>
<td>4</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>-0.99</td>
<td>5</td>
<td>35</td>
<td>18</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>-0.90</td>
<td>5</td>
<td>24</td>
<td>10</td>
<td>46</td>
<td>30</td>
</tr>
<tr>
<td>-0.70</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>54</td>
<td>42</td>
</tr>
<tr>
<td>-0.30</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>34</td>
</tr>
<tr>
<td>0.00</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>85</td>
<td>55</td>
</tr>
<tr>
<td>0.30</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>75</td>
<td>38</td>
</tr>
<tr>
<td>0.70</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>17</td>
</tr>
<tr>
<td>0.90</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.99</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.00</td>
<td>5</td>
<td>93</td>
<td>56</td>
<td>877</td>
<td>832</td>
</tr>
</tbody>
</table>
Table 24 (continued)

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$t_{OLS}$</th>
<th>$t_{MP}$</th>
<th>$t_{RHOL}$</th>
<th>$t_{SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.025</td>
<td>0.05</td>
<td>0.025</td>
</tr>
<tr>
<td>-1.00</td>
<td>10</td>
<td>41</td>
<td>19</td>
<td>32</td>
</tr>
<tr>
<td>-0.99</td>
<td>10</td>
<td>48</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>-0.90</td>
<td>10</td>
<td>31</td>
<td>17</td>
<td>40</td>
</tr>
<tr>
<td>-0.70</td>
<td>10</td>
<td>15</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td>-0.30</td>
<td>10</td>
<td>7</td>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td>0.00</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>59</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>0.70</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>0.90</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0.99</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.00</td>
<td>10</td>
<td>55</td>
<td>22</td>
<td>896</td>
</tr>
<tr>
<td>-1.00</td>
<td>20</td>
<td>69</td>
<td>28</td>
<td>43</td>
</tr>
<tr>
<td>-0.99</td>
<td>20</td>
<td>65</td>
<td>34</td>
<td>18</td>
</tr>
<tr>
<td>-0.90</td>
<td>20</td>
<td>38</td>
<td>18</td>
<td>41</td>
</tr>
<tr>
<td>-0.70</td>
<td>20</td>
<td>17</td>
<td>7</td>
<td>33</td>
</tr>
<tr>
<td>-0.30</td>
<td>20</td>
<td>9</td>
<td>3</td>
<td>51</td>
</tr>
<tr>
<td>0.00</td>
<td>20</td>
<td>11</td>
<td>2</td>
<td>58</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
<td>5</td>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>0.70</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>54</td>
</tr>
<tr>
<td>0.90</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>0.99</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.00</td>
<td>20</td>
<td>60</td>
<td>30</td>
<td>894</td>
</tr>
</tbody>
</table>
empirical frequencies of $t_{OLS}$ or $t_{SYM}$ for $-0.30 \leq \rho \leq 0.99$. The number of rejections of the null hypothesis for the two tailed test using $t_{MP}$ is considerably less than the number of rejections using $t_{OLS}$, $t_{SYM}$ or $t_{RHO1}$ for $0.7 \leq \rho \leq 0.99$, but greater than the expected number of rejections in 1,000 trials. For $\rho = -0.99$, the number of rejections using $t_{MP}$ is less than the expected frequency and the number of rejections using $t_{RHO1}$ agrees reasonably well with the expected frequency. For $\rho = 1.0$, the theoretical percentiles of the "t-statistic" are not appropriate for testing $H_0: \rho = 1.0$ against $H_A: \rho \neq 1.0$ using $t_{MP}$ or $t_{RHO1}$ as the test statistic. For $-0.99 \leq \rho \leq 0.90$, the agreements between the theoretical and empirical frequencies of $t_{MP}$ and $t_{RHO1}$ improve as the sample size increases.

For $n = 5$, the empirical frequencies of $t_{OLS}$ exceeding the 95-th and 97.5-th percentiles of Student's $t$ with 35 degrees of freedom agree reasonably well with the expected frequencies for $-0.99 \leq \rho \leq -0.70$. The number of rejections of the null hypothesis for the one tailed test using $t_{OLS}$ decreases as $\rho$ increases and the null hypothesis is not rejected in any of the 1,000 trials for $-0.30 \leq \rho \leq 0.99$. This is expected because the means of $t_{OLS}$ are less than the 2.5-th percentile of Student's $t$ for $0.30 \leq \rho \leq 0.99$. The agreement between the theoretical and empirical frequencies of $t_{OLS}$ for the one tailed test improves as the sample size increases, but the agreement is still poor for $n = 20$ and $-0.30 \leq \rho \leq 0.99$. With the exception of $\rho = -0.99$ and $1.0$, the empirical frequencies of $t_{SYM}$ are similar to the empirical frequencies of $t_{OLS}$. For $n = 5$ and $\rho = -0.99$, the
critical values for the one tailed test are exceeded by $t_{SYM}$ in all 1,000 samples but are rarely exceeded by $t_{SYM}$ for $n = 10$ and 20 and $\rho = -0.99$. For $\rho = 1.0$, the theoretical percentiles of the "t-statistic" are not appropriate for testing $H_0: \rho = 1$ against $H_A: \rho > 1$ using the test statistic $t_{SYM}$.

The agreements between the theoretical frequencies and the empirical frequencies of $t_{MP}$ and $t_{RHO}$ for the one tail test are reasonable for $-0.99 \leq \rho \leq 0.70$. For $\rho$ near one, the test statistics never exceed the critical values for the one tail test because the estimators of $\rho$ are truncated to lie in the interval $[-1, 1]$. For $\rho = 1.0$, the truncation does not affect the number of rejections because the critical region of the "t-statistic" includes zero.
CHAPTER VI. SUMMARY

The estimation of the parameters of the stationary normal seasonal autoregressive process with seasonal means was investigated. Particular attention was given to the stationary first-order seasonal autoregressive process defined by

$$Y_{ij} = \theta_i + \rho Y_{i,j-1} + e_{ij},$$

where \{e_{ij}\} is a sequence of independent normal \((0, \sigma^2)\) random variables. The variable \(Y_{ij}\) denotes the sampled value for the \(i\)-th period of the \(j\)-th year. The parameter \(\rho\) is assumed to be strictly less than one in absolute value and the parameters \(\theta_i, i = 0\) to \(k - 1\), are the seasonal intercepts.

Given a realization of \(nk\) observations \(\{Y_{ij}: i = 0,1,\ldots, k-1; j = 1,2,\ldots, n\}\), the least squares estimator of \(\rho\) is

$$\hat{\rho}_{OLS} = \frac{\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - \bar{Y}_{1\ell})(Y_{i,j-1} - \bar{Y}_{1\ell})}{\left(\sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{ij} - \bar{Y}_{1\ell})^2\right)^{1/2}},$$

where \(\bar{Y}_{1\ell} = (n-1)^{-1} \sum_{j=2}^{n} Y_{i,j-1}\), \(\ell = 0,1\), and \(i = 0,1,\ldots, k-1\).

Despite the efforts of numerous investigators, the exact distribution of \(\hat{\rho}_{OLS}\) remains unknown. A general procedure for numerically computing
the exact moments of \( \hat{\rho}_{OLS} \) is given in Appendix A, but no exact formulae for the moments of \( \hat{\rho}_{OLS} \) are possible except in the simplest case of \( \rho = 0 \).

For samples of the size frequently encountered in econometrics, the least squares estimator is seriously biased. In Chapter III, approximations to the mean and variance of \( \hat{\rho}_{OLS} \) were derived using a series expansion and large sample theory. The first two moments of \( \hat{\rho}_{OLS} \) are

\[
E(\hat{\rho}_{OLS}) = \rho - (n-1)^{-1}[1+(k+2)k^{-1}\rho] + O(n^{-2}),
\]

and

\[
\text{var}(\hat{\rho}_{OLS}) = [k(n-1)]^{-1}(1-\rho^2) + O(n^{-2}).
\]

A number of alternative estimators of \( \rho \) which are asymptotically equivalent to \( \hat{\rho}_{OLS} \) were considered and the small sample properties of the various estimators were examined theoretically and in a large Monte Carlo study. The least squares estimator was compared with the following three estimators:

1. The modified least squares estimator constructed using the bias expression

\[
\hat{\rho}_{MP} = \left[1-(n-1)^{-1}k^{-1}(k+2)\right]^{-1}[\hat{\rho}_{OLS} + (n-1)^{-1}],
\]

\[
\hat{\rho}_{OLS} \in (-1+2[(n-1)k]^{-1}, 1-2[(n-1)k]^{-1}(k+1)).
\]
2. The modified least squares estimator which corrects for the mean bias

\[
\hat{\rho}_1 = \frac{\hat{\rho}_{OLS} + (n-1)^{-1}(1-\hat{\rho}_{OLS})^{-1}(1+\hat{\rho}_{OLS})}{1 + (n-1)^{-1}(1-\hat{\rho}_{OLS})^{-1}(1+\hat{\rho}_{OLS})}, \quad \hat{\rho}_{OLS} \in (-1, 1)
\]

\[
= 1, \quad \hat{\rho}_{OLS} \geq 1
\]

\[
= -1, \quad \hat{\rho}_{OLS} \leq -1.
\]

3. The symmetric estimator based on the work of Burg (1975)

\[
\hat{\rho}_{SYM} = \frac{[k(n-1)]^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j}-\bar{Y}_1)(Y_{i,j-1}-\bar{Y}_1)}{[2k(n-1)]^{-1} \sum_{i=0}^{k-1} \sum_{j=2}^{n} (Y_{i,j}-\bar{Y}_1)^2 + \sum_{j=2}^{n} (Y_{1,j-1}-\bar{Y}_1)^2}.
\]

The bias in \(\hat{\rho}_{SYM}\) is \(O(n^{-2})\) while the bias in the other two estimators is \(O(n^{-1})\). The bias of \(\hat{\rho}_1\) is less than that of \(\hat{\rho}_{OLS}\) for positive \(\rho\) because the order \(n^{-1}\) portion of the bias introduced by estimating the seasonal means is removed. The results of the simulation study for a second group of estimators are presented in Appendix C.

The least squares procedure applied to the stationary normal first-order seasonal autoregressive process with seasonal means performed poorly in the empirical study in samples of sizes commonly encountered in
econometric work. The approximation to the least squares bias derived in Chapter III is adequate unless the parameter value is close to one. It is recommended that the modified least squares estimator $\hat{\rho}_{MP}$ be used as the method of estimation for stationary normal first-order seasonal autoregressive processes.

The approximation to the variance of the least squares estimator based on the work of Bartlett (1946) is satisfactory unless the parameter value is near one. For positive values of $\rho$, the approximation to the variance improves rapidly as the sample size increases.

Orcutt and Winokur (1969) stated that the least squares predictors seem to be nearly optimal in small samples and seem to have smaller prediction variances than predictors using a corrected least squares estimator in the case of a first-order autoregressive process. The findings of this study indicate that for monthly data the predictors using the proposed estimator, $\hat{\rho}_{MP}$, have smaller prediction variances than the least squares predictors of a first-order seasonal autoregressive process. In particular, the prediction variance due to the estimated parameters is reduced by as much as 75 percent using $\hat{\rho}_{MP}$ rather than $\hat{\rho}_{OLS}$.

For drawing inferences about the autoregressive parameter, the regression "t-statistic" is generally compared with the percentiles of a Student's $t$. However, the results of the Monte Carlo study show that the usual $t$-test will be very misleading. The percentiles of the "t-statistic" using the least squares estimator were in poor agreement with the percentiles of Student's $t$ with $k(n-2)-1$ degrees of freedom. The "t-statistic" using $\hat{\rho}_{MP}$ can be used for inference about the parameter.
of a stationary normal first-order seasonal autoregressive process unless the true value of $\rho$ is close to one. For the test $H_0: \rho = 1$ versus $H_A: \rho < 1$, the tables of Dickey and Fuller (1979) can be used.

The stationary normal second-order seasonal autoregressive process defined by

$$Y_{ij} = \theta_1 + \alpha_1 Y_{i-1, j-1} + \alpha_2 Y_{i-2, j-2} + \varepsilon_{ij},$$

where the $\varepsilon_{ij}$ are independent normal $(0, \sigma^2)$ random variables, was considered in Chapter IV. It is assumed that the roots of the polynomial equation, $m^{2k} - \alpha_1 m^k - \alpha_2 = 0$, are less than unity in modulus. Approximations to the least squares biases for the second-order seasonal autoregressive process were derived, but the complexity of the expressions did not allow for a convenient method of correcting for the biases.

The bias correction procedure proposed for the first-order case does not extend to higher order processes. Three bias correction procedures which remove the bias due to the estimated means were proposed for constructing estimators of the second-order seasonal autoregressive coefficients. The least squares estimator of $\alpha = (\alpha_1, \alpha_2)$ is

$$\hat{\alpha}_{1,OLS}, \hat{\alpha}_{2,OLS} = \hat{H}^{-1}_{\alpha},$$

where

$$\hat{H} = [k(n-2)]^{-1} \sum_{i=0}^{k-1} \sum_{j=3}^{n} X_{ij} X_{ij}' ,$$

$$\hat{N} = [k(n-2)]^{-1} \sum_{i=0}^{k-1} \sum_{j=3}^{n} X_{ij} (Y_{ij} - \bar{Y}_{ij} - \bar{Y}_{10}) ,$$
\[ X'_{ij} = (Y_{i,j-1} - \bar{Y}_{i1}, Y_{i,j-2} - \bar{Y}_{i2}) , \]

and

\[ \bar{Y}_{i\ell} = (n-2)^{-1} \sum_{j=3}^{n} Y_{i,j-\ell} , \quad \ell = 0,1,2 . \]

Let \( H = \text{var}(\{Y_{i,j-1}, Y_{i,j-2}\}) \) and \( N = \text{cov}(\{Y_{i,j-1}, Y_{i,j-2}\}, Y_{ij}) \).

When the least squares estimator of \( \alpha \) lie outside the stationary region, the estimators \( \hat{\alpha}_{OLS} \) are truncated to lie on the boundaries of the stationary region and no further modification is considered. If the least squares estimators obey the stationary conditions, the two modifications of \( \hat{\alpha}_{OLS} \) were:

1. A linear transformation of the least squares estimator. Let

\[ E = H^{-1}[E(\hat{N} - \hat{N}) - (H-H)\alpha] . \]

Then \( \hat{\alpha} = \hat{\alpha}_{OLS} - \hat{E} \), where \( \hat{E} \) is the estimator of \( E \) obtained by substituting \( \hat{\alpha}_{OLS} \) for \( \hat{\alpha} \).

In the case of the second-order seasonal autoregressive process, the estimators were

\[ \hat{\alpha}_2 = \left[ \hat{\alpha}_{2,OLS}(n-2) + 1 \right](n-3)^{-1} , \quad \hat{\alpha}_{2,OLS} \in (-1, 1/2(n-2)^{-1}) \]

\[ = 1 , \quad \hat{\alpha}_{2,OLS} = 1 \]

\[ = -1 , \quad \hat{\alpha}_{2,OLS} = -1 \]

and

\[ \hat{\alpha}_{2,OLS} \leq 1/2(n-2)^{-1} \]

\[ \hat{\alpha}_{2,OLS} \leq -1 \]
\[ \tilde{\alpha}_1 = \hat{\alpha}_{1,\text{OLS}} + (n-2)^{-1}(1+\tilde{\alpha}_2). \]

2. The modified least squares estimator which generalizes the method that produced the estimator, \( \hat{\rho}_1 \), in the first-order case. Let \( \hat{\gamma}(0) = [k(n-2)]^{-1} \sum_{i=0}^{k-1} \sum_{j=3}^{n} (Y_{ij} - \overline{Y}_{i0})^2 \) and \( V = \text{var}(\overline{Y}_{i0}) \). Then \( \hat{\alpha} = \frac{\hat{H}}{\hat{N}} \), where

\[
\hat{H} = \hat{H} + VJ' \\
\hat{N} = \hat{N} + VJ \\
J' = (1, 1),
\]

and \( \hat{V} \) is the estimator of \( V \) obtained by substituting \( \hat{\gamma}(0) \) and \( \hat{\alpha}_{\text{OLS}} \) for \( \gamma(0) \) and \( \alpha \), respectively.

The third bias correction procedure was given by Bora-Senta and Kounias (1980) and was described in detail in Chapter II. The three procedures are expected to have similar small sample properties and are asymptotically equivalent to the least squares estimators. Although it is seldom necessary to consider seasonal autoregressive processes of order greater than two, the three bias correction procedures can be extended to include higher order processes.
CHAPTER VII. EXAMPLES

To illustrate the proposed modified estimation procedure, the electrical usage in Iowa is investigated. The data are monthly observations on peak electrical load covering the period January 1974 to November 1979 from the Iowa Power and Light Company. The data are given in Table 25. It is necessary to use an inverse transformation to stabilize the variances. Let \( Y_t = 1000 \cdot X_t^{-1} \), where \( X_t \) denotes the original series.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>603</td>
<td>660</td>
<td>723</td>
<td>777</td>
<td>791</td>
<td>760</td>
</tr>
<tr>
<td>Feb.</td>
<td>602</td>
<td>625</td>
<td>634</td>
<td>732</td>
<td>787</td>
<td>741</td>
</tr>
<tr>
<td>Mar.</td>
<td>556</td>
<td>620</td>
<td>594</td>
<td>684</td>
<td>688</td>
<td>686</td>
</tr>
<tr>
<td>Apr.</td>
<td>544</td>
<td>554</td>
<td>591</td>
<td>616</td>
<td>670</td>
<td>654</td>
</tr>
<tr>
<td>May</td>
<td>784</td>
<td>630</td>
<td>796</td>
<td>795</td>
<td>701</td>
<td>769</td>
</tr>
<tr>
<td>June</td>
<td>879</td>
<td>780</td>
<td>982</td>
<td>1008</td>
<td>862</td>
<td>1074</td>
</tr>
<tr>
<td>July</td>
<td>927</td>
<td>944</td>
<td>1064</td>
<td>1057</td>
<td>993</td>
<td>1150</td>
</tr>
<tr>
<td>Aug.</td>
<td>955</td>
<td>921</td>
<td>834</td>
<td>1058</td>
<td>1119</td>
<td>1110</td>
</tr>
<tr>
<td>Sept.</td>
<td>933</td>
<td>803</td>
<td>798</td>
<td>1064</td>
<td>1002</td>
<td>1063</td>
</tr>
<tr>
<td>Oct.</td>
<td>629</td>
<td>604</td>
<td>586</td>
<td>625</td>
<td>642</td>
<td>667</td>
</tr>
<tr>
<td>Nov.</td>
<td>633</td>
<td>658</td>
<td>661</td>
<td>741</td>
<td>748</td>
<td>668</td>
</tr>
<tr>
<td>Dec.</td>
<td>648</td>
<td>693</td>
<td>736</td>
<td>777</td>
<td>763</td>
<td>---</td>
</tr>
</tbody>
</table>

*Source: Iowa Power and Light Co., 666 Grand Avenue, Des Moines, Iowa 50309.*
An automatic time series package identified and estimated the model

\[(1 + 0.47B^{1.2})[(1-B^{1.2})Y_t + 0.045] = (1 + 0.54B)e_t, \quad (7.1)\]

\[(0.13) \quad (0.014) \quad (0.12)\]

where \( B \) is the usual backshift operator, \( B^jY_t = Y_{t-j} \). The numbers in parentheses are the estimated standard errors obtained from the automatic time series program. The residual mean square error for model (7.1) is equal to 0.0097. Inspection of the residuals indicated that the model is not totally adequate and that an alternative model should be considered. A multiplicative seasonal autoregressive process of order \((1, 0) \times (1, 0)\) with seasonal means was fitted to the time series using the least squares procedure. The regression equation is

\[\hat{Y}_t = 0.35 D_{1t} + 0.49 D_{2t} + 0.53 D_{3t} + 0.56 D_{4t}\]

\[(0.15) \quad (0.15) \quad (0.16) \quad (0.17)\]

\[+ 0.18 D_{5t} + 0.09 D_{6t} + 0.19 D_{7t} + 0.33 D_{8t}\]

\[(0.18) \quad (0.15) \quad (0.12) \quad (0.11)\]

\[+ 0.37 D_{9t} + 0.90 D_{10t} + 0.29 D_{11t} + 0.33 D_{12t}\]

\[(0.12) \quad (0.15) \quad (0.17) \quad (0.16)\]

\[+ 0.74 Y_{t-1} - 0.11 (Y_{t-12} - 0.74 Y_{t-13}), \quad (7.2)\]

\[(0.10) \quad (0.14) \quad (0.10)\]

where \( D_{1t}, D_{2t}, \ldots, D_{12t} \) are the seasonal dummy variables. The residual mean square for this model is \( s^2 = 0.0067 \). Assuming \( Y_t - 0.74 Y_{t-1} \) follows a first-order seasonal autoregressive process, the coefficient of \( Y_{t-12} - 0.74 Y_{t-13} \) is estimated using the modified
least squares estimator given in (3.28). Using \( n = 6 \) and \( k = 12 \), the modified least squares estimator, \( \hat{\rho}_{MP} \), is 
\[
\left[ 1 - 14(12 \cdot 5)^{-1} \right]^{-1}
\]
\[
[1-14(12+5)^{-1}] = 0.12
\]
To test \( H_0: \rho = 0 \) versus \( H_A: \rho \neq 0 \), the "t-statistic" using \( \hat{\rho}_{MP} \) is computed as 
\[
(0.12)(0.14)^{-1} = 0.89
\]
From the normal tables, the "t-statistic" is not significant at the 0.05 level.

The autocorrelation function of the residuals from model (7.2) suggested that a multiplicative seasonal autoregressive process of order \((1, 0) \times (2, 0)\) might give a better fit. Suppressing the seasonal intercepts, the least squares procedure estimated the model

\[
\hat{Y}_t = \sum_{i=1}^{12} D_{it} \hat{\theta}_i + 0.71 Y_{t-1} - 0.10 (Y_{t-12} - 0.71 Y_{t-13})
\]
\[
- 0.29 (Y_{t-24} - 0.71 Y_{t-25}),
\]
(7.3)

where \( s^2 = 0.0068 \). Assuming \( Y_t - 0.71 Y_{t-1} \) satisfies a second-order seasonal autoregressive process, the modified least squares estimators of \( \alpha_{24} \) and \( \alpha_{12} \) defined in equations (4.30) and (4.31) are calculated as 
\(-0.06\) and \(0.13\), respectively. To test \( H_0: \alpha_{24} = 0 \) versus \( H_A: \alpha_{24} \neq 0 \), the "t-statistic" using \( \hat{\alpha}_{24} \) is 
\[
(-0.06)(0.17)^{-1} = -0.33
\]
which does not exceed 1.645 in absolute value. The extra parameter does not significantly improve the fit and so model (7.2) using 0.12 as the estimated coefficient of \( Y_{t-12} - 0.74 Y_{t-13} \) is judged an adequate model. In fact, the model with no seasonal autoregression is accepted by the data.
Using the first 60 observations to reestimate the parameters of model (7.2), forecasts for the next 11 months were obtained. The least squares predictors using $\hat{\rho}_{MP}$ defined in Chapter V were calculated. The predictors were transformed back to the original scale and the results are given in Table 26. Both methods of prediction gave fairly accurate forecasts with prediction error never greater than 16 percent of the observed value. The mean average percentage error of the least squares predictor is 7.8 percent while the mean average percentage error of the predictors using $\hat{\rho}_{MP}$ is 7.3 percent.

The estimated standard errors of the least squares predictors of the transformed series, $Y_t$, are 0.10, 0.12, 0.13, 0.14, 0.14, 0.14, 0.14, 0.14, 0.14, 0.14, and 0.14. The two methods of prediction gave predictions which were always less than two standard errors from the observed values and in most cases the prediction errors were less than one standard error.

Table 26. Forecasts of peak loads for the year 1979

<table>
<thead>
<tr>
<th>Month</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
<th>Month</th>
<th>$\hat{\rho}_{OLS}$</th>
<th>$\hat{\rho}_{MP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>784</td>
<td>775</td>
<td>July</td>
<td>1033</td>
<td>1011</td>
</tr>
<tr>
<td>Feb.</td>
<td>707</td>
<td>721</td>
<td>Aug.</td>
<td>953</td>
<td>994</td>
</tr>
<tr>
<td>Mar.</td>
<td>658</td>
<td>655</td>
<td>Sept.</td>
<td>891</td>
<td>953</td>
</tr>
<tr>
<td>Apr.</td>
<td>606</td>
<td>615</td>
<td>Oct.</td>
<td>611</td>
<td>631</td>
</tr>
<tr>
<td>May</td>
<td>752</td>
<td>751</td>
<td>Nov.</td>
<td>688</td>
<td>701</td>
</tr>
<tr>
<td>June</td>
<td>929</td>
<td>911</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To avoid the possibility of having considered the wrong transformation, the logarithmic transformation was considered and the results were similar to those obtained for the inverse transformation.

The second example considers the unemployment rate of civilian workers in the United States. Monthly observations of seasonally unadjusted unemployed civilian workers and seasonally unadjusted total labor force for the period January 1967 to November 1979 were taken from various issues of Survey of Current Business. The unemployment rate was computed as the ratio of unemployed civilian workers to total labor force. The unemployment rate displays nonstationary behavior which may require taking differences. To test for a unit root in the autoregressive process, the procedure outlined in Fuller (1976, pp. 366-382) is used. Let $Y_t$ denote the unemployment rate at time $t$.

An autoregressive model was fit to the data using ordinary least squares. Suppressing the seasonal intercepts, the regression equation is

$$
\hat{Y}_t = 12 \sum_{i=1}^{12} D_i \hat{Y}_{t-i} + 0.00112 t + 0.946 Y_{t-1} - 0.107 \left( Y_{t-2} - Y_{t-1} \right) \\
- 0.375 \left( Y_{t-3} - Y_{t-2} \right) - 0.028 \left( Y_{t-12} - Y_{t-3} \right) \\
- 0.189 \left( Y_{t-13} - Y_{t-12} \right) - 0.042 \left( Y_{t-14} - Y_{t-13} \right) \\
+ 0.100 \left( Y_{t-15} - Y_{t-14} \right), \\
$$

(7.4)
where \( s^2 = 0.042 \). Let \( a_1 \) and \( \beta \) denote the coefficients of \( Y_{t-1} \) and \( t \). To test \( H_0: a_1 = 1, \beta = 0 \) versus \( H_A: a_1 < 1 \), the statistic denoted by \( \hat{\tau}_\tau \) is computed as \((0.946 - 1)(0.024)^{-1} = -2.28\). Using the third part of Table 8.5.2 of Fuller (1976, p. 373), \( \hat{\tau}_\tau \) is compared with \(-3.43\) when testing at the .05 level. Since \( \hat{\tau}_\tau = -2.28 > -3.43 \), the null hypothesis is not rejected.

A multiplicative seasonal autoregressive process of order \((2, 0) \times (1, 0)\) was fit to the first difference of the original series, \( W_t = Y_t - Y_{t-1} \). The least squares procedure estimated the model as

\[
\hat{W}_t = 0.80 D_{1t} - 0.30 D_{2t} - 0.60 D_{3t} - 0.48 D_{4t} - 0.12 D_{5t} \\
(0.11) (0.09) (0.11) (0.08) (0.07)
+ 1.14 D_{6t} - 0.23 D_{7t} - 0.58 D_{8t} + 0.08 D_{9t} - 0.10 D_{10t} \\
(0.14) (0.11) (0.12) (0.07) (0.06)
+ 0.17 D_{11t} + 0.05 D_{12t} + 0.115 W_{t-1} + 0.377 W_{t-2} \\
(0.06) (0.07) (0.083) (0.083) (0.083)
+ 0.165(W_{t-12} - 0.115 W_{t-13} - 0.377 W_{t-14}) \\
(0.089) (0.083) (0.083)
\]

where \( s^2 = 0.043 \). Assuming \( W_t - 0.115 W_{t-1} - 0.377 W_{t-2} \) satisfies a first-order seasonal autoregressive process, the Marriott and Pope modified least squares estimator of the coefficient of \( W_{t-12} - 0.115 W_{t-13} - 0.377 W_{t-14} \) is calculated as \( 0.275 \). To test \( H_0: \rho = 0 \) versus \( H_A: \rho \neq 0 \), the "t-statistic" using \( \hat{\rho}_{MP} \) is equal to \( 3.09 \) which exceeds the 97.5-th percentile of a Student's t with 125 degrees of freedom. The null hypothesis is rejected at the .05 level.
Table 27. Seasonally unadjusted unemployment rate of civilian workers in U.S.\(^a\)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>4.20</td>
<td>4.21</td>
<td>3.91</td>
<td>3.50</td>
<td>3.23</td>
<td>4.59</td>
<td>4.09</td>
<td>3.72</td>
<td>3.73</td>
<td>3.78</td>
<td>3.71</td>
<td>3.48</td>
</tr>
<tr>
<td>1968</td>
<td>4.03</td>
<td>4.25</td>
<td>3.78</td>
<td>3.21</td>
<td>2.94</td>
<td>4.47</td>
<td>3.97</td>
<td>3.46</td>
<td>3.32</td>
<td>3.18</td>
<td>3.25</td>
<td>3.06</td>
</tr>
<tr>
<td>1969</td>
<td>3.68</td>
<td>3.70</td>
<td>3.46</td>
<td>3.19</td>
<td>2.89</td>
<td>4.13</td>
<td>3.84</td>
<td>3.48</td>
<td>3.65</td>
<td>3.48</td>
<td>3.33</td>
<td>3.45</td>
</tr>
<tr>
<td>1970</td>
<td>4.22</td>
<td>4.67</td>
<td>4.57</td>
<td>4.33</td>
<td>4.14</td>
<td>5.56</td>
<td>5.32</td>
<td>5.02</td>
<td>5.20</td>
<td>5.12</td>
<td>5.53</td>
<td>5.58</td>
</tr>
<tr>
<td>1971</td>
<td>6.55</td>
<td>6.58</td>
<td>6.26</td>
<td>5.66</td>
<td>5.29</td>
<td>6.46</td>
<td>6.20</td>
<td>5.91</td>
<td>5.75</td>
<td>5.40</td>
<td>5.66</td>
<td>5.53</td>
</tr>
<tr>
<td>1972</td>
<td>5.44</td>
<td>6.38</td>
<td>6.11</td>
<td>5.51</td>
<td>5.08</td>
<td>6.16</td>
<td>5.84</td>
<td>5.50</td>
<td>5.37</td>
<td>5.13</td>
<td>4.91</td>
<td>4.73</td>
</tr>
<tr>
<td>1973</td>
<td>5.45</td>
<td>5.59</td>
<td>5.17</td>
<td>4.77</td>
<td>4.34</td>
<td>5.36</td>
<td>5.01</td>
<td>4.67</td>
<td>4.68</td>
<td>4.19</td>
<td>4.51</td>
<td>4.52</td>
</tr>
<tr>
<td>1974</td>
<td>5.62</td>
<td>5.75</td>
<td>5.31</td>
<td>4.81</td>
<td>4.61</td>
<td>5.81</td>
<td>5.64</td>
<td>5.28</td>
<td>5.69</td>
<td>5.49</td>
<td>6.21</td>
<td>6.69</td>
</tr>
<tr>
<td>1976</td>
<td>8.82</td>
<td>8.66</td>
<td>8.08</td>
<td>7.37</td>
<td>6.74</td>
<td>7.07</td>
<td>7.80</td>
<td>7.57</td>
<td>7.40</td>
<td>7.15</td>
<td>7.42</td>
<td>7.35</td>
</tr>
<tr>
<td>1978</td>
<td>7.04</td>
<td>6.88</td>
<td>6.58</td>
<td>5.75</td>
<td>5.50</td>
<td>6.19</td>
<td>6.27</td>
<td>5.81</td>
<td>5.75</td>
<td>5.38</td>
<td>5.54</td>
<td>5.63</td>
</tr>
<tr>
<td>1979</td>
<td>6.38</td>
<td>6.40</td>
<td>6.06</td>
<td>5.49</td>
<td>5.18</td>
<td>5.99</td>
<td>5.81</td>
<td>5.88</td>
<td>5.61</td>
<td>5.56</td>
<td>5.57</td>
<td>--</td>
</tr>
</tbody>
</table>

A multiplicative seasonal autoregressive process of order \((2, 0) \times (2, 0)\) was also considered but no significant improvements were detected. It was concluded that model (7.5) is a fairly good approximation and can be used for predicting future values. Model (7.5) was reestimated using the first 144 observations and predictions were made for the next 11 months. The two methods of prediction considered in the first example are compared in Table 28. Both methods tended to overestimate the unemployment rate and the predictors using \(\hat{\rho}_{MP}\) were slightly larger than the least squares predictors. The prediction error never exceeded 16 percent of the observed value for either predictor. The mean average percentage error of the least squares predictor is 8.4 percent as compared with 10.5 percent for the predictors using \(\hat{\rho}_{MP}\).

Table 28. Forecasts of unemployment rates for 1979

<table>
<thead>
<tr>
<th>Month</th>
<th>(\hat{\rho}_{OLS})</th>
<th>(\hat{\rho}_{MP})</th>
<th>Month</th>
<th>(\hat{\rho}_{OLS})</th>
<th>(\hat{\rho}_{MP})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan.</td>
<td>6.71</td>
<td>6.69</td>
<td>July</td>
<td>6.49</td>
<td>6.70</td>
</tr>
<tr>
<td>Feb.</td>
<td>6.83</td>
<td>6.82</td>
<td>Aug.</td>
<td>6.13</td>
<td>6.34</td>
</tr>
<tr>
<td>Mar.</td>
<td>6.53</td>
<td>6.63</td>
<td>Sept.</td>
<td>6.11</td>
<td>6.33</td>
</tr>
<tr>
<td>Apr.</td>
<td>5.95</td>
<td>6.07</td>
<td>Oct.</td>
<td>5.83</td>
<td>6.07</td>
</tr>
<tr>
<td>May</td>
<td>5.63</td>
<td>5.81</td>
<td>Nov.</td>
<td>6.01</td>
<td>6.23</td>
</tr>
<tr>
<td>June</td>
<td>6.71</td>
<td>6.88</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The estimated standard errors of the least squares predictors are 0.22, 0.33, 0.47, 0.59, 0.72, 0.83, 0.93, 1.03, 1.12, 1.21, and 1.29. The prediction errors of the two methods of prediction were always less than two standard errors.
REFERENCES


De Gooijer, J. G. 1980. Exact moments of the sample autocorrelations from series generated by general ARIMA processes of order \((p, d, q)\), \(d = 0\) or 1. J. of Econometrics 14:365-379.


Sawa, T. 1978. The exact moments of the least squares estimator for the autoregressive model. J. of Econometrics 8:159-172.


Slutsky, E. 1937. The summation of random causes as the source of cyclic processes. Econometrica 5:105-146.


Sawa (1978) gives a general procedure for evaluating moments of a ratio of two quadratic forms in a normally distributed random vector. The method is useful for studying the exact moments of serial correlation coefficients in normal autoregressive moving average time series. However, exact formulae for the moments of r(k) are not possible except in the simplest case of independent normal series. In the case of the first-order autoregressive process, Sawa numerically evaluated the first two moments of the least squares estimator of \( \rho \) for various values of \( \rho \) and differing sample sizes. Kendall's (1954) approximation for the mean and Bartlett's (1946) variance approximation were found adequate for samples as small as 20.

Let \( x \) be an \( n \times 1 \) random vector which is distributed as \( N(0, \Sigma) \). Denote the ratio of two quadratic forms by

\[
Q_1 = x'Ax, \quad Q_2 = x'Bx, \quad r = \frac{Q_1}{Q_2}
\]

where A and B are \( n \times n \) symmetric matrices. The moments of \( r \) are derived by straightforward application of the following lemmas given by Sawa.

**Lemma A.1.** Let \( \Phi(\theta_1, \theta_2) \) be the joint moment generating function of \( Q_1 \) and \( Q_2 \). Then the \( h \)-th moment of \( r \) is given by
Lemma A.2. The joint moment generating function of $Q_1$ and $Q_2$ is given by

$$\phi(\theta_1, \theta_2) = |I - 2\theta_1 C - 2\theta_2 A|^{-\frac{1}{2}}, \quad (A.3)$$

where $A$ is a $n \times n$ diagonal matrix containing the eigenvalues of a positive definite matrix $\Sigma$ and $C = P' \Sigma' A P$, where $P$ is the corresponding $n \times n$ matrix of normalized eigenvectors of $\Sigma$. Some results in matrix algebra which are useful in deriving the derivatives of $\phi(\theta_1, \theta_2)$ are presented in the following lemma.

Lemma A.3. Let $A$ be a symmetric matrix whose elements are differentiable functions of $\theta$. Then

$$\frac{\partial |A(\theta)|}{\partial \theta} = |A| \cdot \text{tr} \ A^{-1} \left( \frac{\partial A}{\partial \theta} \right),$$

and

$$\frac{\partial A^{-1}}{\partial \theta} = A^{-1} \left( \frac{\partial A}{\partial \theta} \right) A^{-1}.$$

Let $\phi(\theta_1, \theta_2) = |R(\theta_1, \theta_2)|^{-\frac{1}{2}}$, where $R = I - 2\theta_1 C - 2\theta_2 A$. Then
\[
\frac{\partial^r}{\partial \theta_1^r} \phi(\theta_1, \theta_2) = \frac{\partial^{r-1}}{\partial \theta_1^{r-1}} \left[ \phi(\theta_1, \theta_2) \text{tr}(R^{-1}C) \right],
\]

and

\[
\frac{\partial^s}{\partial \theta_1^s} \text{tr}(R^{-1}C) = s! \ 2^s \text{tr}(R^{-1}C)^{s+1}.
\]

Using Leibnitz's formula for the n-th derivative of the product of two functions, the r-th derivative of \( \phi(\theta_1, \theta_2) \) is given as

\[
\frac{\partial^r}{\partial \theta_1^r} \phi(\theta_1, \theta_2) = \frac{(r-1)!}{2} \sum_{j=0}^{r-1} \frac{1}{j!} \frac{\partial^j}{\partial \theta_1^j} \phi(\theta_1, \theta_2) [2^{r-j} \text{tr}(R^{-1}C)^{r-j}].
\]  

(A.4)

In particular, the first and second derivatives of \( \phi(\theta_1, \theta_2) \) with respect to \( \theta_1 \) evaluated at \( \theta_1 = 0 \), are

\[
\frac{\partial \phi}{\partial \theta_1} \bigg|_{\theta_1=0} = \left[ I - 2\theta_2A \right]^{-1/2} \text{tr} \left[ (I-2\theta_2A)^{-1}C \right]
\]

\[
= \frac{1}{n} \sum_{k=1}^{\infty} \frac{C_{kk}}{(1-2\lambda_k \theta_2)^{1/2}}
\]

(A.5)

and
where $C = (c_{ij})$ and $A = \text{diag}(\lambda_i)$. Using (A.5) and (A.6) in Lemma A.1, the first two moments of $r$ can be compared. These results are presented in the following theorem.

**Theorem A.1.** The first and second moments of $r$ are given by

$$E(r) = \sum_{j=1}^{n} c_{jj} \int_{0}^{\infty} \frac{dx}{(1+2\lambda_j x) \prod_{\ell=1}^{l} (1+2\lambda_{\ell} x)^{l_\ell}} , \quad (A.7)$$

and

$$E(r^2) = \sum_{i=1}^{n} \sum_{j=1}^{n} (c_{ii} c_{jj} + 2c_{ij}^2) \int_{0}^{\infty} \frac{x dx}{(1+2\lambda_i x)(1+2\lambda_j x) \prod_{\ell=1}^{l} (1+2\lambda_{\ell} x)^{l_\ell}} . \quad (A.8)$$

The least squares estimator of $\rho$ in the first-order normal autoregressive process with unknown mean can be written as $\hat{\rho} = Q_1/Q_2$, where
Q₁ = y'Ay and Q₂ = y'B y. The matrices A and B are n x n symmetric matrices whose rows are orthogonal to the (1,1,...,1)' vector and hence Q₁ and Q₂ are distributed independently of μ. The two quadratic forms are equal to (y-μ)'A(y-μ) and (y-μ)'B(y-μ), respectively and have joint moment generating function given by Φ(θ₁, θ₂).

In the case of the first-order seasonal autoregressive process of period k, the least squares estimator has the representation

\[ \rho = Q_3/Q_4 \],

where \( Q_3 = \sum_{i=0}^{k-1} y_i' A y_i \) and \( Q_2 = \sum_{i=0}^{k-1} y_i' B y_i \); the matrices A and B are as defined in the first-order autoregressive case and

\[ y_i = (y_{i1}, y_{i2}, \ldots, y_{in}) \], i = 0, 1, ..., k-1 . The joint moment generating function of \( Q_3 \) and \( Q_4 \), say \( \phi_k(θ_1, θ_2) \), is related to the joint moment generating function of \( Q_1 \) and \( Q_2 \) by

\[ \phi_k(θ_1, θ_2) = \phi^k(θ_1, θ_2) . \] (A.9)

The first two derivatives of \( \phi_k(θ_1, θ_2) \), evaluated at \( θ_1 = 0 \), are given by

\[
\frac{\partial^2 \phi_k}{\partial \theta_1^2} \bigg|_{\theta_1=0} = k \phi^{k-1} \left. \frac{\partial \phi}{\partial \theta_1} \right|_{\theta_1=0} = \frac{2}{n} \sum_{j=1}^{n} \frac{C_{ij}}{(1-2λ_j \theta_2)^{k/2}} \left( \sum_{i=1}^{k} \frac{1}{(1-2λ_i \theta_2)^{k/2}} \right),
\] (A.10)
The first two moments of the least squares estimator of $\rho$ for the first-order seasonal autoregressive process are formulated in the following theorem.

**Theorem A.2.** The first and second moments of the least squares estimator of $\rho$ in the first-order seasonal autoregressive process are given by

$$E(\hat{\rho}) = k \sum_{j=1}^{n} C_{jj} \int_{0}^{\infty} \frac{dx}{(1+2\lambda_j x)(1+2\lambda_j x)^{k/2}} ,$$ \tag{A.12}

and

$$E(\hat{\rho}^2) = k \prod_{i=1}^{n} \prod_{j=1}^{n} (kC_{ii}C_{jj} + 2C_{ij}^2) \int_{0}^{\infty} \frac{xdx}{(1+2\lambda_i x)(1+2\lambda_j x)(1+2\lambda_k x)^{k/2}} .$$ \tag{A.13}

For a general stationary normal autoregressive moving average process, the sample autocorrelations are expressible as ratios of two
quadratic forms given in (A.1). Denote the \( n \times n \) matrix \( J_n \) as the \( n \times n \) matrix with all its elements equal to one and \( L_n \) as the \( n \times n \) matrix with elements equal to one along the first upper diagonal and zeroes elsewhere. Using the definition of the sample autocovariances given by

\[
C_n(h) = n^{-1} \sum_{t=1}^{n-h} (Y_t - \overline{Y})(Y_{t+h} - \overline{Y}),
\]

an alternative representation of \( C_n(h) \) is given as \( n^{-1} Y'VU_yVY \), where

\[ V = I - n^{-1} J_n \text{ and } U_h = (L^h + L'^h)2^{-1}. \]

The sample autocorrelations \( r_n(h) \) are then given by

\[
r_n(h) = \frac{Y'VU_yVY}{\sum_{\hat{h}} Y'VY}, \quad h = 1, 2, \ldots, n-1,
\]  

(A.14)

which are of the form given in (A.1).

Consider the least squares estimator of the first-order seasonal autoregressive parameter defined in (3.3). The numerator and denominator of \( \rho_{OLS} \) are given as \( Q_3 = \sum_{i=0}^{k-1} y_i'A_iy_i \) and \( Q_4 = \sum_{i=0}^{k-1} y_i'B_iy_i \), where \( y_i = (y_{i1}, y_{i2}, \ldots, y_{in}) \) and \( A \) and \( B \) are \( n \times n \) symmetric matrices given as
When the true value of \( \rho \) is equal to zero, the covariance matrix of \( y_i' \) is equal to \( I \sigma^2 \). The value of \( \sigma^2 \) can be taken to be one with no loss of generality since the least squares estimator of \( \rho \) is distributed independently of \( \sigma^2 \). The eigenvalues of \( \Sigma B \Sigma^{-1} = B \) are obtained by noting that \( B \) is an idempotent matrix of rank \( n - 2 \). From matrix theory, a \( n \times n \) idempotent matrix of rank \( p \) has \( p \) eigenvalues equal to one and zeroes elsewhere. The normalized eigenvectors of \( B \) associated with a unit root are given as \( P_j \), \( j = 1, 2, \ldots, n-2 \), where the elements of \( P_j \) are equal to

\[
P_{ij} = [i(i+1)]^{-1/2} \quad i \leq j
\]  

(A.17)
The two normalized eigenvectors of $B$ associated with the zero roots are equal to $\mathbf{v}_1 = (n-1)^{-\frac{1}{2}}(1 1 1 \ldots 1 0)$ and $\mathbf{v}_2 = (0 0 0 \ldots 0 1)$. Let $P$ be the $n \times n$ matrix of normalized eigenvectors of $B$. Then the matrix $C = P' \Sigma \Sigma P' = P'AP$ has elements $C_{ij} = \sum_{k=1}^{n} \sum_{m=1}^{n} P_{ik} \Sigma_{km} P_{mj}$ which are given for $i \leq j$, as

\begin{align*}
C_{ij} &= -[i(i+1)]^{-\frac{1}{2}} \\
&= 0 \\
&= 2^{-1}(i+1)^{-1}[i(i+2)]^{-\frac{1}{2}}(i+2) \\
&= -2^{-1}[i(i+1)](j+1))^{-\frac{1}{2}} \\
&= 2^{-1}(n-1)^{-\frac{1}{2}}[n(n-1)]^{-\frac{1}{2}} - [i(i+1)]^{-\frac{1}{2}} \\
&= 2^{-1}(n-1)^{-1}(n-2)^{\frac{1}{2}} \\
&= 0
\end{align*}

and $C_{ji} = C_{ij}$ by symmetry.

Using equation (A.12) of Theorem A.2, the mean of $\hat{\rho}_{OLS}$ when the true value of $\rho$ is zero is obtained by
The second moment of \( \hat{\rho}_{\text{OLS}} \) can be derived exactly using equation (A.13) of Theorem A.2. However, the form of \( \text{E}(\rho_{\text{OLS}}^2) \) involves the sums of the form \( \sum_{i=1}^{n-2} i(i+1) \) and \( \sum_{i=1}^{n-2} i^{-1} \), which have no convenient expressions.

For values of \( \rho \) other than zero, the exact moments of \( \hat{\rho}_{\text{OLS}} \) are algebraically intractable to derive but can be numerically computed with the aid of a computer. Sawa (1978) has shown that the approximate mean and variance of \( \hat{\rho}_{\text{OLS}} \) based on the work of Marriott and Pope (1954), Kendall (1954) and Bartlett (1946) are adequate for moderately small samples.
APPENDIX B. THE COVARIANCES OF ESTIMATED AUTOCOVARIANCES FOR STATIONARY AUTOREGRESSIVE PROCESSES

Consider the \( p \)-th order stationary autoregressive process with zero mean

\[
Y_t - \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} - \ldots - \alpha_p Y_{t-p} = \epsilon_t,
\]

where \( \epsilon_t \) is a sequence of independent normal \( (0, \sigma^2) \) random variables and the roots of the characteristic equation

\[
m^p - \alpha_1 m^{p-1} - \alpha_2 m^{p-2} - \ldots - \alpha_p = 0
\]

lie within the unit circle. The sample autocovariance of lag \( h \) is defined to be

\[
\gamma(h) = (n-h)^{-1} \sum_{t=1}^{n-h} Y_t Y_{t+h}, \quad h = 0, 1, 2, \ldots, n-1
\]

which is unbiased for \( \gamma(h) \). Bartlett (1946) derived the covariance of \( \gamma(h) \) and \( \gamma(q) \) to order \( n^{-1} \) as

\[
\text{cov}(\gamma(h), \gamma(q)) = n^{-1} \sum_{p=0}^{\infty} \{ \gamma(p) \gamma(p-h+q) + \gamma(p+q) \gamma(p-h) \}
\]

\[
+ O(n^{-2}) . \tag{B.1}
\]
In deriving expressions for the bias of the least squares estimators for
the autoregressive parameters, relevant formulas of (B.1) are needed. A
general procedure of obtaining such expressions is given in this section.

Associated with the p-th order autoregressive process, the Yule-
Walker equations are

$$\gamma(h) - a_1\gamma(h-1) - a_2\gamma(h-2) - \ldots - a_p\gamma(h-p) = \sigma^2, \ h = 0$$

$$= 0, \ h > 0. \ (B.2)$$

The first p + 1 equations corresponding to h = 0, 1, 2, ..., p form a
system of p + 1 simultaneous equations which can be written as

$$B\gamma = (\sigma^2, 0, 0, \ldots, 0)' \ , \ \text{where} \ \gamma' = (\gamma(0), \gamma(1), \ldots, \gamma(p)) \ \text{and} \ B \ \text{is a}
(p+1)x(p+1) \ \text{matrix whose elements are functions of} \ (a_1, a_2, \ldots, a_p).$$

In the case of the second-order autoregressive process, the system of
equations is

$$\begin{align*}
1 & -a_1 & -a_2 \\
-a_1 & 1-a_2 & 0 \\
-a_2 & -a_1 & 1
\end{align*} \begin{bmatrix}
\gamma(0) \\
\gamma(1) \\
\gamma(2)
\end{bmatrix} = \begin{bmatrix}
\sigma^2 \\
0 \\
0
\end{bmatrix}. \quad (B.3)
$$

Since B is nonsingular, \( \gamma = B^{-1}(\sigma^2, 0, 0, \ldots, 0)' \).

Denoting \( A(h) = \sum_{j=0}^{\infty} \gamma(j)\gamma(j+h) \), \( h = 0, 1, 2, \ldots \), the Yule-Walker
equations can be used to form a system of p + 1 equations in the p + 1
unknowns \( A(h) \), \( h = 0, 1, 2, \ldots, p \). The first equation in A(h) is
obtained by multiplying (B.2) by \( \gamma(h) \) which gives
\[ \gamma^2(h) - \alpha_1\gamma(h)\gamma(h-1) - \ldots - \alpha_p\gamma(h)\gamma(h-p) = 0, \quad h > 0. \quad (B.4) \]

For a stationary autoregressive process, the autocovariance function is absolutely summable. Since \[ \sum_{j=0}^{\infty} |\gamma(j)\gamma(j+h)| \leq \left( \sum_{j=0}^{\infty} |\gamma(j)| \right)^2 < \infty, \] the function \( A(h) \) is well defined and (B.4) can be summed over \( h \) giving

\[ \sum_{h=1}^{\infty} \gamma^2(h) - \alpha_1 \sum_{h=1}^{\infty} \gamma(h)\gamma(h-1) - \ldots - \alpha_p \sum_{h=1}^{\infty} \gamma(h)\gamma(h-p) = 0. \quad (B.5) \]

Since \[ \sum_{h=1}^{\infty} \gamma^2(h) = A(0) - \gamma^2(0), \quad \sum_{h=1}^{\infty} \gamma(h)\gamma(h-1) = A(1) \quad \text{and} \]

\[ \sum_{h=1}^{\infty} \gamma(h)\gamma(h-j) = A(j) + \sum_{h=1}^{j-1} \gamma(h)\gamma(j-h) \quad \text{for} \quad j \geq 2, \quad \text{equation (B.5) is equal to} \]

\[ A(0) - \alpha_1 A(1) - \ldots - \alpha_p A(p) = C(1), \quad (B.6) \]

where \( C(1) = \gamma^2(0) + \sum_{j=2}^{p} \sum_{h=1}^{j-1} \alpha_j \gamma(h)\gamma(j-h) \). The other \( p \) equations are obtained by multiplying (B.2) by \( \gamma(h-1), \gamma(h-2), \ldots, \gamma(h-p) \), respectively, and summing over \( h = 1 \) to \( \infty \). The system of equations is seen to have the form \( BA = C \), where \( B \) is the nonsingular \((p+1)x(p+1)\) matrix given earlier, \( A' = [A(0), A(1), \ldots, A(p)] \) and \( C \) is a \( p+1 \) vector whose elements are functions of \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_p) \).
and \( \Gamma \). In the case of the second-order autoregressive process, the system of equations is given as

\[
\begin{bmatrix}
1 & -\alpha_1 & -\alpha_2 \\
-\alpha_1 & 1-\alpha_2 & 0 \\
-\alpha_2 & -\alpha_1 & 1
\end{bmatrix}
\begin{bmatrix}
A(0) \\
A(1) \\
A(2)
\end{bmatrix} =
\begin{bmatrix}
\gamma^2(0) + \alpha_2\gamma^2(1) \\
\alpha_2\gamma(0)\gamma(1) \\
0
\end{bmatrix} .
\] (B.7)

The system of equations can be solved directly for \( \mathbf{A} \) by \( \mathbf{A} = \mathbf{B}^{-1} \mathbf{C} \).

It is noted that the function \( A(h) \) satisfies the homogeneous difference equation defining the process.

\[
A(h) - \alpha_1 A(h-1) - \alpha_2 A(h-2) - \ldots - \alpha_p A(h-p) = 0 , \ h \geq p .
\] (B.8)

This result is stated as a lemma.

**Lemma B.1.** Let the autocovariance function of a \( p \)-th order stationary autoregressive process be defined as \( \gamma(h) = \text{cov}\{Y_t, Y_{t+h}\} \). Let

\[
A(h) = \sum_{j=0}^{\infty} \gamma(j)\gamma(j+h) , \ h = 0,1,2,\ldots,
\]

then

\[
A(h) - \alpha_1 A(h-1) - \alpha_2 A(h-2) - \ldots - \alpha_p A(h-p) = 0 , \ h \geq p .
\]

**Proof.** Multiplying the Yule-Walker equation corresponding to \( h = j \) by \( \gamma(j-k) \), gives
\[ \gamma(j-k)\gamma(j) - \alpha_1 \gamma(j-k)\gamma(j-1) - \ldots - \alpha_p \gamma(j-k)\gamma(j-p) = 0, \quad j \geq 1. \]

Noting that the autocovariance function \( \gamma(h) \) is absolutely summable, the system of equations can be summed over \( j = k \) to \( \infty \), i.e.

\[ \sum_{j=k}^{\infty} \gamma(j-k)\gamma(j) - \alpha_1 \sum_{j=k}^{\infty} \gamma(j-k)\gamma(j-1) - \ldots - \alpha_p \sum_{j=k}^{\infty} \gamma(j-k)\gamma(j-p) = 0. \]

For \( k > p \), \( \sum_{j=k}^{\infty} \gamma(j-k)\gamma(j-1) = A(k-1) \), \( i = 0, 1, 2, \ldots, p \), and the equation is given as

\[ A(k) - \alpha_1 A(k-1) - \alpha_2 A(k-2) - \ldots - \alpha_p A(k-p) = 0, \quad k \geq p. \]

In terms of the second-order autoregressive model, the first two values of \( A(h) \) corresponding to \( h = 0, 1 \) are computed as

\[
\begin{bmatrix}
A(0) \\
A(1)
\end{bmatrix} = \frac{1}{(1-\alpha_2)^2 - \alpha_1^2 (1+\alpha_2)} \begin{bmatrix}
(1-\alpha_2) \gamma^2(0) + \alpha_2 \gamma^2(1) + \alpha_1 (1+\alpha_2) \gamma(0) \gamma(1) \\
\alpha_1 \gamma^2(0) + \alpha_2 \gamma^2(1) + (1-\alpha_2) \alpha_2 \gamma(0) \gamma(1)
\end{bmatrix}
\]
\[
\text{and } A(2) = a_1A(1) + a_2A(0) .
\]

Using the results of Bartlett (1946), the variance of \(\hat{\gamma}(0)\) is computed as

\[
\text{var}\{\hat{\gamma}(0)\} = 2n^{-1} \sum_{j=-\infty}^{\infty} \gamma^2(j) + O(n^{-2})
\]

\[
= 2n^{-1}[2A(0) - \gamma^2(0)] + O(n^{-2})
\]

\[
= 2\frac{\gamma^2(0)}{n} \left[ \frac{1+\alpha_2^2}{1-\alpha_2^2} + \frac{2\alpha_1^2(1+\alpha_2)}{[1-\alpha_2^2-a_1^2](1-\alpha_2)} \right] + O(n^{-2}) . \tag{B.9}
\]

Similarly,

\[
\text{var}\{\hat{\gamma}(1)\} = n^{-1} \left[ \sum_{j=-\infty}^{\infty} \gamma^2(j) + \sum_{j=-\infty}^{\infty} \gamma(j+1)\gamma(j-1) \right] + O(n^{-2})
\]

\[
= 2^{-1} \text{var}\{\hat{\gamma}(0)\} + n^{-1}[2A(2) + \gamma^2(1)] + O(n^{-2})
\]

\[
= \frac{\gamma^2(0)}{n} \left[ \frac{1+\alpha_2^2}{1-\alpha_2^2} + \frac{\alpha_1^2(1+2\alpha_2)}{(1-\alpha_2)(1-\alpha_2^2)} \right.
\]

\[+ \left. \frac{\alpha_1^2(4+6\alpha_2^2-5\alpha_2^4)}{[1-\alpha_2^2-a_1^2](1-\alpha_2^2)} \right] + O(n^{-2}) , \tag{B.10}
\]
\[
\text{cov}\{\hat{\gamma}(0), \hat{\gamma}(1)\} = n^{-1} \sum_{j=-\infty}^{\infty} \gamma(j)\gamma(j-1) + O(n^{-2}) \\
= n^{-1} \cdot 4 \cdot A(1) + (n^{-2}) \\
= \frac{4\gamma^2(0)\alpha_1}{n[(1-\alpha_2)^2-\alpha_1^2]} \left[ 1 + \frac{\alpha_1^2\alpha_2}{(1-\alpha_2^2)(1-\alpha_2)} \right] + O(n^{-2}) .
\]
Five other estimators of the parameter of the first-order seasonal autoregressive process were included in the Monte Carlo study and the results are presented in this section. The estimators in this group consist of the following: the least squares estimator \( \hat{\rho}_{\text{OLS}} \) defined in (3.5), the maximum likelihood estimator \( \hat{\rho}_{\text{MLE}} \) given in (3.35), the approximate maximum likelihood estimator \( \hat{\rho}_{\text{MLE}} \) considered in Lemma (3.3), the modified least squares estimator \( \hat{\rho}_2 \) given in (3.33) and the modified symmetric estimator \( \hat{\rho}_{\text{SYM},\text{MP}} \) given in (5.3). The estimators \( \hat{\rho}_{\text{OLS}} \) and \( \hat{\rho}_{\text{MLE}} \) assume the seasonal means are known and equal to zero. More descriptions of the estimators can be found in Chapter III. The empirical means, variances and mean square errors of the various estimators are presented in Tables C.1, C.2 and C.3.

One and three period predictors using \( \hat{\rho}_2 \) and \( \hat{\rho}_{\text{SYM}} \) were generated by the equations (5.5) and (5.6), where the seasonal intercepts are estimated by the approximate generalized least squares estimators of \( \theta_i \) given in (5.11) using \( \hat{\rho}_2 \) and \( \hat{\rho}_{\text{SYM}} \), respectively. The one and three period predictors using \( \hat{\rho}_2 \) are denoted by \( \hat{Y}_{n+1,\text{RH02}} \) and \( \hat{Y}_{n+3,\text{RH02}} \). The one and three period predictors using \( \hat{\rho}_{\text{SYM}} \) are denoted by \( \hat{Y}_{n+1,\text{SYM}} \) and \( \hat{Y}_{n+3,\text{SYM}} \). Table C.4 compares the Monte Carlo variances of the one and three period predictors of \( \hat{\rho}_2 \) and \( \hat{\rho}_{\text{SYM}} \).
### Table C.1. Empirical biases of various estimators of \( \rho \)

<table>
<thead>
<tr>
<th>True value of ( \rho )</th>
<th>n</th>
<th>( \hat{\rho}_{\text{MLE}} )</th>
<th>( \hat{\rho}_{\text{SYM,MP}} )</th>
<th>( \hat{\rho}_2 )</th>
<th>( \hat{\rho}_{\text{OLS}} )</th>
<th>( \hat{\rho}_{\text{MLE}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>5</td>
<td>0.212</td>
<td>0.239</td>
<td>0.027</td>
<td>0.045</td>
<td>0.179</td>
</tr>
<tr>
<td>-0.99</td>
<td>5</td>
<td>0.101</td>
<td>0.095</td>
<td>0.008</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>-0.90</td>
<td>5</td>
<td>0.072</td>
<td>0.083</td>
<td>0.015</td>
<td>0.012</td>
<td>0.013</td>
</tr>
<tr>
<td>-0.70</td>
<td>5</td>
<td>-0.002</td>
<td>0.047</td>
<td>0.019</td>
<td>0.015</td>
<td>0.017</td>
</tr>
<tr>
<td>-0.30</td>
<td>5</td>
<td>-0.137</td>
<td>-0.017</td>
<td>0.010</td>
<td>-0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>0.00</td>
<td>5</td>
<td>-0.248</td>
<td>-0.086</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>0.30</td>
<td>5</td>
<td>-0.371</td>
<td>-0.181</td>
<td>-0.030</td>
<td>-0.004</td>
<td>-0.005</td>
</tr>
<tr>
<td>0.70</td>
<td>5</td>
<td>-0.553</td>
<td>-0.357</td>
<td>-0.104</td>
<td>-0.020</td>
<td>-0.021</td>
</tr>
<tr>
<td>0.90</td>
<td>5</td>
<td>-0.644</td>
<td>-0.460</td>
<td>-0.155</td>
<td>-0.011</td>
<td>-0.013</td>
</tr>
<tr>
<td>0.99</td>
<td>5</td>
<td>-0.683</td>
<td>-0.508</td>
<td>-0.190</td>
<td>-0.006</td>
<td>-0.002</td>
</tr>
<tr>
<td>1.00</td>
<td>5</td>
<td>-0.694</td>
<td>-0.518</td>
<td>-0.194</td>
<td>-0.050</td>
<td>-0.184</td>
</tr>
<tr>
<td>-1.00</td>
<td>10</td>
<td>0.099</td>
<td>0.106</td>
<td>0.020</td>
<td>-0.025</td>
<td>0.101</td>
</tr>
<tr>
<td>-0.99</td>
<td>10</td>
<td>0.001</td>
<td>-0.003</td>
<td>0.004</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>-0.90</td>
<td>10</td>
<td>0.006</td>
<td>0.016</td>
<td>0.008</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>-0.70</td>
<td>10</td>
<td>-0.010</td>
<td>0.017</td>
<td>0.010</td>
<td>0.008</td>
<td>0.009</td>
</tr>
<tr>
<td>-0.30</td>
<td>10</td>
<td>-0.061</td>
<td>0.005</td>
<td>0.007</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>0.00</td>
<td>10</td>
<td>-0.116</td>
<td>-0.026</td>
<td>-0.001</td>
<td>-0.003</td>
<td>-0.003</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
<td>-0.163</td>
<td>-0.055</td>
<td>-0.020</td>
<td>-0.012</td>
<td>-0.012</td>
</tr>
<tr>
<td>0.70</td>
<td>10</td>
<td>-0.226</td>
<td>-0.119</td>
<td>-0.050</td>
<td>-0.017</td>
<td>-0.016</td>
</tr>
<tr>
<td>0.90</td>
<td>10</td>
<td>-0.267</td>
<td>-0.181</td>
<td>-0.082</td>
<td>-0.009</td>
<td>-0.010</td>
</tr>
<tr>
<td>0.99</td>
<td>10</td>
<td>-0.291</td>
<td>-0.222</td>
<td>-0.106</td>
<td>-0.004</td>
<td>-0.002</td>
</tr>
<tr>
<td>1.00</td>
<td>10</td>
<td>-0.295</td>
<td>-0.227</td>
<td>-0.106</td>
<td>-0.024</td>
<td>-0.099</td>
</tr>
<tr>
<td>-1.00</td>
<td>20</td>
<td>0.054</td>
<td>0.055</td>
<td>0.011</td>
<td>0.012</td>
<td>0.054</td>
</tr>
<tr>
<td>-0.99</td>
<td>20</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>-0.90</td>
<td>20</td>
<td>0.003</td>
<td>0.008</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>-0.70</td>
<td>20</td>
<td>-0.008</td>
<td>0.006</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>-0.30</td>
<td>20</td>
<td>-0.031</td>
<td>0.003</td>
<td>0.004</td>
<td>0.003</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Table C.1 (continued)

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>n</th>
<th>$\hat{\rho}_{\text{MLE}}$</th>
<th>$\hat{\rho}_{\text{SYM,MP}}$</th>
<th>$\hat{\rho}_2$</th>
<th>$\rho_{\text{OLS}}^*$</th>
<th>$\rho_{\text{MLE}}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>20</td>
<td>-0.053</td>
<td>-0.006</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
<td>-0.074</td>
<td>-0.015</td>
<td>-0.004</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>0.70</td>
<td>20</td>
<td>-0.099</td>
<td>-0.037</td>
<td>-0.016</td>
<td>-0.003</td>
<td>-0.004</td>
</tr>
<tr>
<td>0.90</td>
<td>20</td>
<td>-0.118</td>
<td>-0.070</td>
<td>-0.035</td>
<td>-0.006</td>
<td>-0.005</td>
</tr>
<tr>
<td>0.99</td>
<td>20</td>
<td>-0.135</td>
<td>-0.103</td>
<td>-0.055</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>1.00</td>
<td>20</td>
<td>-0.137</td>
<td>-0.107</td>
<td>-0.058</td>
<td>-0.012</td>
<td>-0.053</td>
</tr>
</tbody>
</table>

Table C.2. Empirical variances of various estimators of $\rho$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>n</th>
<th>$\hat{\rho}_{\text{MLE}}$</th>
<th>$\hat{\rho}_{\text{SYM,MP}}$</th>
<th>$\hat{\rho}_2$</th>
<th>$\rho_{\text{OLS}}^*$</th>
<th>$\rho_{\text{MLE}}^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.00</td>
<td>5</td>
<td>0.199</td>
<td>0.355</td>
<td>0.340</td>
<td>0.447</td>
<td>0.338</td>
</tr>
<tr>
<td>-0.99</td>
<td>5</td>
<td>0.002</td>
<td>0.004</td>
<td>0.045</td>
<td>0.029</td>
<td>0.003</td>
</tr>
<tr>
<td>-0.90</td>
<td>5</td>
<td>0.091</td>
<td>0.173</td>
<td>0.689</td>
<td>0.480</td>
<td>0.234</td>
</tr>
<tr>
<td>-0.70</td>
<td>5</td>
<td>0.417</td>
<td>0.757</td>
<td>1.738</td>
<td>1.155</td>
<td>0.923</td>
</tr>
<tr>
<td>-0.30</td>
<td>5</td>
<td>1.244</td>
<td>1.961</td>
<td>2.887</td>
<td>1.891</td>
<td>1.774</td>
</tr>
<tr>
<td>0.00</td>
<td>5</td>
<td>1.639</td>
<td>2.275</td>
<td>3.403</td>
<td>2.065</td>
<td>2.065</td>
</tr>
<tr>
<td>0.30</td>
<td>5</td>
<td>1.845</td>
<td>2.235</td>
<td>3.792</td>
<td>1.880</td>
<td>1.791</td>
</tr>
<tr>
<td>0.70</td>
<td>5</td>
<td>2.030</td>
<td>1.918</td>
<td>4.670</td>
<td>1.251</td>
<td>0.930</td>
</tr>
<tr>
<td>0.90</td>
<td>5</td>
<td>2.179</td>
<td>1.752</td>
<td>4.024</td>
<td>0.471</td>
<td>0.198</td>
</tr>
<tr>
<td>0.99</td>
<td>5</td>
<td>2.197</td>
<td>1.719</td>
<td>3.748</td>
<td>0.030</td>
<td>0.003</td>
</tr>
<tr>
<td>1.00</td>
<td>5</td>
<td>2.200</td>
<td>1.709</td>
<td>3.575</td>
<td>0.494</td>
<td>0.364</td>
</tr>
<tr>
<td>-1.00</td>
<td>10</td>
<td>0.086</td>
<td>0.102</td>
<td>0.117</td>
<td>0.131</td>
<td>0.124</td>
</tr>
<tr>
<td>-0.99</td>
<td>10</td>
<td>0.003</td>
<td>0.004</td>
<td>0.020</td>
<td>0.016</td>
<td>0.003</td>
</tr>
<tr>
<td>-0.90</td>
<td>10</td>
<td>0.117</td>
<td>0.160</td>
<td>0.255</td>
<td>0.222</td>
<td>0.125</td>
</tr>
<tr>
<td>-0.70</td>
<td>10</td>
<td>0.357</td>
<td>0.457</td>
<td>0.547</td>
<td>0.502</td>
<td>0.430</td>
</tr>
<tr>
<td>-0.30</td>
<td>10</td>
<td>0.701</td>
<td>0.850</td>
<td>0.950</td>
<td>0.820</td>
<td>0.810</td>
</tr>
</tbody>
</table>
Table C.2 (continued)

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>Variance multiplied by one hundred</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$\hat{\rho}_{MLE}$</td>
<td>$\hat{\rho}_{SYM,MP}$</td>
<td>$\hat{\rho}_2$</td>
<td>$\hat{\rho}^*$</td>
<td>$\hat{\rho}^*_{MLE}$</td>
</tr>
<tr>
<td>0.00</td>
<td>10</td>
<td>0.853</td>
<td>1.007</td>
<td>1.184</td>
<td>0.947</td>
<td>0.954</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
<td>0.929</td>
<td>1.033</td>
<td>1.130</td>
<td>0.858</td>
<td>0.841</td>
</tr>
<tr>
<td>0.70</td>
<td>10</td>
<td>0.800</td>
<td>0.754</td>
<td>1.063</td>
<td>0.515</td>
<td>0.438</td>
</tr>
<tr>
<td>0.90</td>
<td>10</td>
<td>0.727</td>
<td>0.580</td>
<td>0.939</td>
<td>0.197</td>
<td>0.121</td>
</tr>
<tr>
<td>0.99</td>
<td>10</td>
<td>0.692</td>
<td>0.518</td>
<td>0.873</td>
<td>0.017</td>
<td>0.004</td>
</tr>
<tr>
<td>1.00</td>
<td>10</td>
<td>0.682</td>
<td>0.495</td>
<td>0.783</td>
<td>0.130</td>
<td>0.124</td>
</tr>
</tbody>
</table>

Table C.3. Empirical mean square error of various estimators of $\rho$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>Mean square errors multiplied by ten</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$\hat{\rho}_{MLE}$</td>
<td>$\hat{\rho}_{SYM,MP}$</td>
<td>$\hat{\rho}_2$</td>
<td>$\hat{\rho}^*$</td>
<td>$\hat{\rho}^*_{MLE}$</td>
</tr>
<tr>
<td>-1.00</td>
<td>5</td>
<td>0.469</td>
<td>0.608</td>
<td>0.041</td>
<td>0.065</td>
<td>0.354</td>
</tr>
<tr>
<td>-0.99</td>
<td>5</td>
<td>0.101</td>
<td>0.091</td>
<td>0.001</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>-0.90</td>
<td>5</td>
<td>0.061</td>
<td>0.086</td>
<td>0.071</td>
<td>0.050</td>
<td>0.025</td>
</tr>
<tr>
<td>-0.70</td>
<td>5</td>
<td>0.042</td>
<td>0.098</td>
<td>0.177</td>
<td>0.118</td>
<td>0.095</td>
</tr>
</tbody>
</table>
Table C.3 (continued)

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>Mean square errors multiplied by ten</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
</tr>
<tr>
<td>-0.30</td>
<td>5</td>
</tr>
<tr>
<td>0.00</td>
<td>5</td>
</tr>
<tr>
<td>0.30</td>
<td>5</td>
</tr>
<tr>
<td>0.70</td>
<td>5</td>
</tr>
<tr>
<td>0.90</td>
<td>5</td>
</tr>
<tr>
<td>0.99</td>
<td>5</td>
</tr>
<tr>
<td>1.00</td>
<td>5</td>
</tr>
<tr>
<td>-1.00</td>
<td>10</td>
</tr>
<tr>
<td>-0.99</td>
<td>10</td>
</tr>
<tr>
<td>-0.90</td>
<td>10</td>
</tr>
<tr>
<td>-0.70</td>
<td>10</td>
</tr>
<tr>
<td>-0.30</td>
<td>10</td>
</tr>
<tr>
<td>0.00</td>
<td>10</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
</tr>
<tr>
<td>0.70</td>
<td>10</td>
</tr>
<tr>
<td>0.90</td>
<td>10</td>
</tr>
<tr>
<td>0.99</td>
<td>10</td>
</tr>
<tr>
<td>1.00</td>
<td>10</td>
</tr>
<tr>
<td>-1.00</td>
<td>20</td>
</tr>
<tr>
<td>-0.99</td>
<td>20</td>
</tr>
<tr>
<td>-0.90</td>
<td>20</td>
</tr>
<tr>
<td>-0.70</td>
<td>20</td>
</tr>
<tr>
<td>-0.30</td>
<td>20</td>
</tr>
<tr>
<td>0.00</td>
<td>20</td>
</tr>
<tr>
<td>0.30</td>
<td>20</td>
</tr>
<tr>
<td>0.70</td>
<td>20</td>
</tr>
</tbody>
</table>
Table C.3 (continued)

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>Mean square errors multiplied by ten</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
</tr>
<tr>
<td>0.90</td>
<td>20</td>
</tr>
<tr>
<td>0.99</td>
<td>20</td>
</tr>
<tr>
<td>1.00</td>
<td>20</td>
</tr>
</tbody>
</table>
Table C.4. Empirical variances of various predictors of $Y_{i,n+1}$ and $Y_{i,n+3}$

<table>
<thead>
<tr>
<th>True value of $\rho$</th>
<th>$\hat{Y}_{n+1,RH02}$</th>
<th>$\hat{Y}_{n+3,RH02}$</th>
<th>$\hat{Y}_{n+1,SYM}$</th>
<th>$\hat{Y}_{n+3,SYM}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.275 1.143 1.056</td>
<td>3.354 3.084 3.023</td>
<td>1.886 1.112 1.054</td>
<td>6.667 3.070 3.001</td>
</tr>
<tr>
<td>20</td>
<td>1.291 1.122 1.057</td>
<td>2.858 2.629 2.537</td>
<td>1.284 1.115 1.055</td>
<td>2.802 2.587 2.528</td>
</tr>
<tr>
<td>-0.70</td>
<td>1.280 1.119 1.056</td>
<td>1.970 1.821 1.772</td>
<td>1.251 1.116 1.056</td>
<td>1.903 1.816 1.772</td>
</tr>
<tr>
<td>-0.30</td>
<td>1.260 1.118 1.056</td>
<td>1.248 1.216 1.131</td>
<td>1.263 1.119 1.058</td>
<td>1.250 1.217 1.131</td>
</tr>
<tr>
<td>0.00</td>
<td>1.237 1.111 1.056</td>
<td>1.202 1.099 1.051</td>
<td>1.278 1.122 1.059</td>
<td>1.204 1.099 1.051</td>
</tr>
<tr>
<td>0.30</td>
<td>1.214 1.105 1.054</td>
<td>1.444 1.277 1.193</td>
<td>1.327 1.130 1.059</td>
<td>1.444 1.277 1.193</td>
</tr>
<tr>
<td>0.70</td>
<td>1.155 1.084 1.048</td>
<td>2.316 2.080 1.941</td>
<td>1.466 1.170 1.067</td>
<td>2.404 2.160 1.967</td>
</tr>
<tr>
<td>0.90</td>
<td>1.091 1.059 1.037</td>
<td>2.911 2.827 2.713</td>
<td>1.561 1.234 1.094</td>
<td>3.383 3.286 2.950</td>
</tr>
<tr>
<td>0.99</td>
<td>1.059 1.036 1.021</td>
<td>3.176 3.162 3.098</td>
<td>1.618 1.293 1.136</td>
<td>4.010 4.148 3.773</td>
</tr>
<tr>
<td>1.00</td>
<td>1.053 1.031 1.019</td>
<td>3.198 3.189 3.135</td>
<td>1.621 1.301 1.144</td>
<td>4.075 4.271 3.905</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I wish to express my appreciation and gratitude to Professor Wayne A. Fuller for sharing his enthusiasm, experience and insight and for providing the support, both mentally and financially, during the course of this study. The past four years have been a great learning experience and I am grateful to Professor Fuller for making them bearable and enjoyable.

I thank my parents for their constant encouragement and moral support. To my brother, Albert, I am grateful for setting an example to follow.

A special thanks to Professor Roy D. Hickman and Helen Nelson for their constant help and for making the days just a little bit brighter.

To my fiancee, Martha, I am especially thankful for putting up with me and for providing the support and companionship during the past year.

Thanks also to Jane Stowe for the expert typing of this manuscript and the many drafts required to bring it to final form.

This research was partly supported by Joint Statistical Agreements J.S.A. 81-14 with the U.S. Bureau of the Census.