The bias of forecasts from a first-order autoregression

Econometric Theory, 7, 1991, 222–235. Printed in the United States of America.

THE BIAS OF FORECASTS FROM A FIRST-ORDER AUTOREGRESSION

JAN R. MAGNUS
London School of Economics and CentER, Tilburg University

BAHRAM PESARAN
Economics Division, Bank of England

The exact finite sample behavior is investigated on the bias of multiperiod least-squares forecasts in the normal autoregressive model $y_t = \alpha + \beta y_{t-1} + u_t$. Necessary and sufficient conditions are given for the existence of the bias and an expression is presented which we use to obtain exact numerical results for finite samples. The unit root and near unit root behavior is studied in detail and some popular preconceptions about the behavior of the bias are shown to be false.

1. INTRODUCTION

In this paper we hope to shed some light on the bias of (multiperiod) least-squares forecasts in dynamic econometric models by studying the simplest example of such a model, namely the first-order autoregressive process $\{y_t\}$ defined by

$$y_t = \alpha + \beta y_{t-1} + u_t, \tag{1}$$

where $\{u_t\}$ is a sequence of independent $N(0, \sigma^2)$ distributed random variables.

Two theoretical results are available regarding the bias of least-squares forecasts based on n observations $y_1, y_2, \ldots, y_n$ generated by (1).

If α is *known* to be zero, so that only β is estimated, then Malinvaud [8, p. 554] showed the forecast bias to be zero. If α is not known, so that both α and β are estimated, then Fuller and Hasza [2] showed that the bias is zero if the process is mean stationary. Both proofs are based on symmetry arguments. (See Cryer, Nankervis and Savin [1] for more general results using symmetry conditions.)

Little is known, however, about the case where both α and β are unknown and the process is not mean stationary. This includes the important fixed start-up case ($y_0 = c$) and the case where the process is covariance stationary (though not mean stationary). No theoretical results are available and the

We are grateful to the editor, coeditor, and two referees for their comments.

only relevant Monte Carlo study we know of is Lahiri's [4]. His results, being based on only 100 replications, are not, however, trustworthy. Our *exact* results fill this gap. The results show, inter alia, that the bias is not, in general, a monotone function of either β, n, or s (the number of periods ahead). The behavior of the bias for values of β close to -1 or 1 is particularly important and is studied in detail.

The plan of the paper is as follows. In Section 2 we present the model and obtain an exact expression for the bias of the (multiperiod) least-squares forecast. Numerical results for the fixed start-up model when α and σ are of the same order of magnitude are presented and discussed in Section 3. In Section 4 we study the unit root and near-unit root behavior of the forecast bias and in Section 5 we discuss sigma asymptotics. An appendix, containing the proofs of the four theorems, concludes the paper.

2. THE FORECAST BIAS: THEORY

As in Magnus and Pesaran [6], we shall be exclusively concerned with the first-order autoregressive process with an intercept term,

$$y_t = \alpha + \beta y_{t-1} + u_t, \qquad t = 2,3,\ldots, \tag{2}$$

where both α and β are unknown and $\{u_2, u_3, \ldots\}$ is a sequence of i.i.d. $N(0,\sigma^2)$ random variables. Regarding the initial observation y_1 we postulate

$$y_1 = \mu_1 + \delta u_1 \tag{3}$$

where $u_1 \sim N(0,\sigma^2)$ is independent of $u_2, u_3, \ldots$, and $\delta > 0$. In finite samples the actual values of μ_1 and δ are important. If $\mu_1 = \alpha/(1-\beta)$, then the series $\{y_1, y_2, \ldots\}$ is mean stationary; if $\delta = (1-\beta^2)^{-1/2}$, then the series is covariance stationary; and if both conditions hold the process is a normal *strictly stationary* time series. On the other hand, if (2) also holds for $t = 1$ and if $\mu_1 = \alpha + \beta c$, then y_0 is distributed symmetrically about the constant c; if $\delta = 1$, then y_0 is a nonrandom constant; and if both $\mu_1 = \alpha + \beta c$ and $\delta = 1$, then $y_0 = c$ and we have the *fixed start-up* model. There are, of course, numerous other possible choices for the initial conditions.

Let $y = (y_1, y_2, \ldots, y_n)'$ be an $n \times 1$ vector of observations generated by (2) and (3). Then y is normally distributed, say $y \sim N(\mu, LL')$ where L can be chosen lower triangular. The least-squares estimators of α and β based on y are given by

$$\hat{\alpha} = \bar{y}_{**} - \hat{\beta}\bar{y}_{*} \tag{4}$$

and

$$\hat{\beta} = \sum_{t=2}^{n} (y_t - \bar{y}_{**})(y_{t-1} - \bar{y}_{*}) \Big/ \sum_{t=2}^{n} (y_{t-1} - \bar{y}_{*})^2, \tag{5}$$

where

$$\bar{y}_* = \frac{1}{n-1}\sum_{t=2}^{n} y_{t-1}, \qquad \bar{y}_{**} = \frac{1}{n-1}\sum_{t=2}^{n} y_t. \tag{6}$$

Clearly $\hat{\alpha}$ is an odd function of y while $\hat{\beta}$ is a ratio of two quadratic forms in y, say $\hat{\beta} = y'Ay/y'By$, and hence an even function of y. The s-periods-ahead forecast is given by

$$\hat{y}_{n+s} = \hat{\alpha}\sum_{j=0}^{s-1}\hat{\beta}^j + \hat{\beta}^s y_n, \qquad s = 1,2,\ldots.$$

The forecast error can thus be written as

$$\hat{y}_{n+s} - y_{n+s} = \hat{\alpha}\sum_{j=0}^{s-1}\hat{\beta}^j + (\hat{\beta}^s - \beta^s)y_n - \alpha\sum_{j=0}^{s-1}\beta^j - \sum_{j=0}^{s-1}\beta^j u_{n+s-j} \tag{7}$$

and the expected value of (7), if it exists, is the bias of the forecast $\hat{y}_{n+s}$.

The following theorem gives an exact expression for this bias and shows that it is finite if and only if $s \leq n-3$.

THEOREM 1. *Let* $n \geq 4$, $\beta \neq 1$ *and let* w_1 *and* w_2 *be the* $n \times 1$ *vectors*

$$w_1 = \frac{1}{n-1}(-1,0,\ldots,0,1)', \qquad w_2 = \frac{1}{n-1}(-1,-1,\ldots,-1,n-1)'.$$

Then the bias of the forecast $\hat{y}_{n+s}$ *exists if and only if* $1 \leq s \leq n-3$, *in which case*

$$E(\hat{y}_{n+s} - y_{n+s}) = \sum_{k=0}^{s-1}\tau_k(w_1) + \tau_s(w_2) + \left(\mu_1 - \frac{\alpha}{1-\beta}\right)\left(\frac{1-\beta^{n-1}}{(n-1)(1-\beta)} - \beta^{n+s-1}\right), \tag{8}$$

where for any $n \times 1$ *vector* $a, \tau_k(a)$ $(1 \leq k \leq n-3)$ *is defined as* $\tau_k(a) = E(\hat{\beta}^k(a'y))$ *and* $\tau_0(a) = E(a'y)$. ■

Proofs of the theorems are in the Mathematical Appendix. Using the theory developed in Magnus [5, Theorem 5], where τ_k is expressed as a univariate integral, we see that (8) allows us to obtain the forecast bias using standard numerical integration techniques. (We used the numerical algorithms group [9] (the so-called NAG) subroutine DO1AMF for the numerical integrations in this paper. This subroutine also gives an estimate of the absolute error in the integration. For all results reported the absolute error was less than 10^{-5}.) The only restriction is that the covariance matrix of y must be nonsingular; thus we require that $\delta \neq 0$.

The only theoretical result which is relevant to our present study is Theorem 1 of Fuller and Hasza [2], which states that if the process is mean stationary then $E(\hat{y}_{n+s} - y_{n+s}) = 0$. As a consequence, we shall be primarily concerned with the fixed start-up model.

3. THE FIXED START-UP MODEL

The bias of the least-squares forecast $\hat{y}_{n+s}$ (BIAS) depends on seven parameters: $\alpha, \beta, \sigma, \mu_1, \delta, n, s$. But in fact it depends on only five parameters. We see this as follows. Defining

$$z_t = \frac{y_t - \bar{y}_*}{\sigma}, \qquad u_t^* = \frac{u_t}{\sigma}, \tag{9}$$

we have (see Magnus and Rothenberg [7])

$$z_t = \varphi_t(\delta(1-\beta)u_1^* - \lambda) + \sum_{j=2}^{t} \beta^{t-j} u_j^* - \frac{1}{n-1} \sum_{j=2}^{n-1} \frac{1-\beta^{n-j}}{1-\beta} u_j^*, \tag{10}$$

where

$$\varphi_t = \frac{\beta^{t-1} - \dfrac{1}{n-1}(1-\beta^{n-1})/(1-\beta)}{1-\beta}, \qquad \lambda = \frac{\alpha - (1-\beta)\mu_1}{\sigma}, \tag{11}$$

and also

$$\hat{\beta} = \beta + \frac{\sum_{t=2}^{n} z_{t-1} u_t^*}{\sum_{t=2}^{n} z_{t-1}^2}. \tag{12}$$

The parameter λ is a mean stationarity parameter: if $\lambda = 0$ the process is mean stationary, otherwise it is not.

Also, using (A.1) in the appendix,

$$\frac{\text{BIAS}}{\sigma} = E(\hat{\beta}^s - \beta^s)z_n + \frac{1}{n-1} \sum_{j=0}^{s-1} E(\hat{\beta}^j - \beta^j)(z_n - z_1) \tag{13}$$

and this depends only on the five parameters β, n, s, $\delta(1-\beta)$, and λ. Of course, BIAS/σ depends also on σ, but only through λ.

In the fixed zero start-up case $y_0 = 0$, we have $\lambda = \alpha\beta/\sigma$ and $\delta = 1$. We shall distinguish between three cases: the case where α and σ are of the same order of magnitude, the case where σ is much smaller than α, and finally the case where σ is much larger than α. Thus, in the present section, we shall deal with the case $\sigma/\alpha = 1$, the limiting behavior of the BIAS when $\sigma/\alpha \to 0$ or

$\sigma/\alpha \to \infty$ being postponed to Section 5. With $\sigma/\alpha = 1$ we have $\lambda = \beta$ and $\delta = 1$ and Table 1 presents the numerical results for this case at selected values of β, n, and s.

We observe first of all that the BIAS vanishes at $\beta = 0$ due to the fact that the process is then mean stationary.

The complete lack of symmetry between positive and negative values of β is illustrated in Figure 1 ($n = 20$). BIAS/σ is negligible unless β is close to plus or minus one. If β is close to -1, then the sign of the BIAS depends on whether $n + s$ is even or odd: if $n + s$ is even then BIAS < 0, otherwise BIAS > 0. If β is close to $+1$ the sign of $n + s$ is not relevant. BIAS/σ is always negative and can be substantial, especially for small n and large s; but even for $n = 30$ and $s = 1$, BIAS/σ is around -0.06 when β is close to $+1$.

We notice also that some preconceptions about the behavior of the BIAS are not necessarily true. First, $|\text{BIAS}|/\sigma$ is not always increasing in $|\beta|$. For example, when $s = 1$ and $n = 10$, BIAS/σ has local minima at $\beta = -0.50$ and $\beta = 0.99$ and a local maximum at $\beta = 0.20$, and this lack of monotonicity persists for n as large as 30. Second, $|\text{BIAS}|/\sigma$ does not always increase with s. This occurs when n is small ($n = 10$ or $n = 15$) and β is around 0.80, but it also occurs for small negative β's even when $n = 30$. This finding strengthens the result obtained in Hoque, Magnus, and Pesaran [3], and explained in Magnus and Pesaran [6], that the mean square forecast error in the autoregressive model without an intercept decreases as s increases for values of β close to zero. Third, $|\text{BIAS}|/\sigma$ does not always decrease with n, for

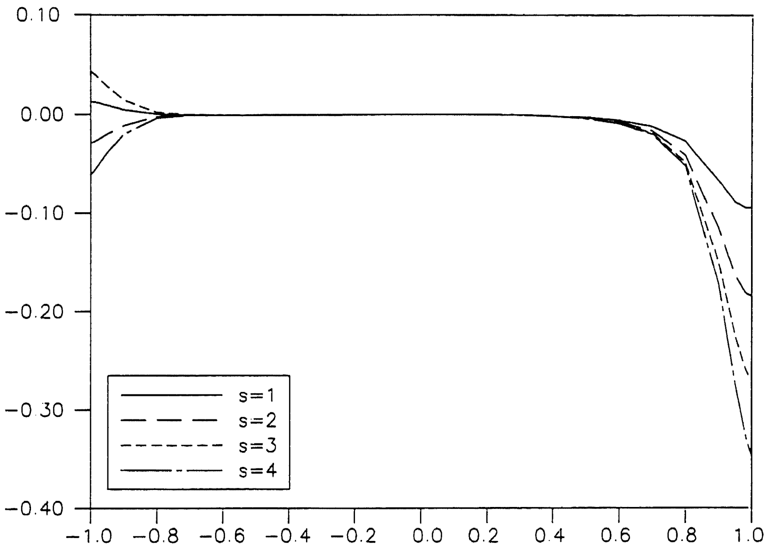

FIGURE 1. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1 = \sigma$ ($\lambda = \beta$), $\delta = 1$, $n = 20$.

TABLE 1. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1 = \sigma$ $(\lambda = \beta)$, $\delta = 1$

		β												
n	s	−1.00	−0.90	−0.80	−0.40	0.00	0.20	0.40	0.60	0.80	0.90	0.95	0.99	1.00
10	1	0.0293	0.0201	0.0094	−0.0016	0.0000	0.0003	−0.0005	−0.0085	−0.0696	−0.1313	−0.1559	−0.1672	−0.1687
10	2	−0.0677	−0.0450	−0.0232	−0.0016	0.0000	0.0016	0.0044	0.0040	−0.0781	−0.1936	−0.2530	−0.2902	−0.2975
10	3	0.1066	0.0637	0.0289	−0.0007	0.0000	0.0010	0.0049	0.0126	−0.0628	−0.2083	−0.2986	−0.3654	−0.3804
10	4	−0.1561	−0.0878	−0.0398	−0.0021	0.0000	0.0019	0.0085	0.0278	−0.0208	−0.1734	−0.2852	−0.3790	−0.4018
15	1	−0.0214	−0.0126	−0.0051	−0.0007	0.0000	−0.0001	−0.0017	−0.0075	−0.0409	−0.0919	−0.1148	−0.1223	−0.1225
15	2	0.0394	0.0198	0.0056	−0.0005	0.0000	0.0000	−0.0012	−0.0074	−0.0571	−0.1524	−0.2051	−0.2309	−0.2343
15	3	−0.0642	−0.0307	−0.0095	−0.0004	0.0000	−0.0002	−0.0016	−0.0074	−0.0623	−0.1899	−0.2731	−0.3237	−0.3325
15	4	0.0860	0.0364	0.0090	−0.0006	0.0000	−0.0001	−0.0016	−0.0068	−0.0602	−0.2082	−0.3192	−0.3979	−0.4137
20	1	0.0131	0.0048	0.0002	−0.0004	0.0000	−0.0002	−0.0013	−0.0059	−0.0269	−0.0665	−0.0886	−0.0948	−0.0944
20	2	−0.0286	−0.0109	−0.0022	−0.0003	0.0000	−0.0001	−0.0014	−0.0075	−0.0407	−0.1145	−0.1628	−0.1831	−0.1845
20	3	0.0436	0.0146	0.0019	−0.0003	0.0000	−0.0002	−0.0017	−0.0086	−0.0482	−0.1485	−0.2238	−0.2640	−0.2690
20	4	−0.0610	−0.0198	−0.0035	−0.0003	0.0000	−0.0002	−0.0019	−0.0092	−0.0518	−0.1712	−0.2722	−0.3362	−0.3464
25	1	−0.0116	−0.0039	−0.0009	−0.0003	0.0000	−0.0001	−0.0010	−0.0045	−0.0197	−0.0500	−0.0712	−0.0772	−0.0765
25	2	0.0220	0.0058	0.0005	−0.0002	0.0000	−0.0001	−0.0012	−0.0062	−0.0313	−0.0878	−0.1325	−0.1506	−0.1509
25	3	−0.0347	−0.0090	−0.0013	−0.0002	0.0000	−0.0002	−0.0014	−0.0072	−0.0386	−0.1161	−0.1848	−0.2196	−0.2225
25	4	0.0463	0.0104	0.0007	−0.0002	0.0000	−0.0002	−0.0015	−0.0080	−0.0433	−0.1370	−0.2285	−0.2837	−0.2904
30	1	0.0085	0.0014	−0.0003	−0.0002	0.0000	−0.0001	−0.0007	−0.0034	−0.0155	−0.0389	−0.0588	−0.0651	−0.0643
30	2	−0.0181	−0.0034	−0.0003	−0.0001	0.0000	−0.0001	−0.0009	−0.0049	−0.0253	−0.0692	−0.1104	−0.1276	−0.1273
30	3	0.0274	0.0044	−0.0001	−0.0001	0.0000	−0.0001	−0.0011	−0.0058	−0.0320	−0.0928	−0.1551	−0.1872	−0.1888
30	4	−0.0377	−0.0061	−0.0005	−0.0001	0.0000	−0.0001	−0.0011	−0.0065	−0.0367	−0.1110	−0.1936	−0.2436	−0.2482

example, when $\beta = 0.60$ and $s = 2$. This phenomenon only occurs, however, when n is small.

So far we have discussed the fixed start-up case $y_0 = 0$ which implies that $\mu_1 = \alpha$ and $\delta = 1$, so that $y_1 \sim N(\alpha, \sigma^2)$. Let us now consider that case where $\mu_1 = \alpha$ and $\delta \neq 1$. Then y_0 is a random variable with a symmetric distribution about zero and $y_1 \sim N(\alpha, \delta^2)$. The choice of δ is somewhat arbitrary, but numerical results not reported here indicate that the BIAS is not very sensitive to small changes in δ. We shall choose $\delta = (1 - \beta^2)^{-1/2}$, which implies that the process is covariance stationary (although not mean stationary) and that $y_1 \sim N(\alpha, \sigma^2/(1 - \beta^2))$.

In Table 2 we see that BIAS/σ is virtually unaffected by this choice of δ even when β is close to +1, unless β is close to −1 in which case BIAS/σ approaches zero. This, again, is the same kind of result as obtained in Hoque, Magnus, and Pesaran [3] who found, in the autoregressive model without an intercept, that the mean square forecast error is similar for $\delta = 1$ and $\delta = (1 - \beta^2)^{-1/2}$ except for values of $|\beta|$ close to 1. For a further explanation and analysis of this phenomenon, see Magnus and Rothenberg [7].

We know, of course, that BIAS → 0 as $n \to \infty$, but we would also like to know how rapid the convergence takes place and from which value of n the convergence is monotonic. For β close to −1 BIAS/σ is an alternating sequence whose absolute value decreases. The behavior for β close to +1 is rather different. Figure 2 gives BIAS/σ in the fixed zero start-up case for $\beta = 0.90$ as a function of n for various values of s. We see that if n is not

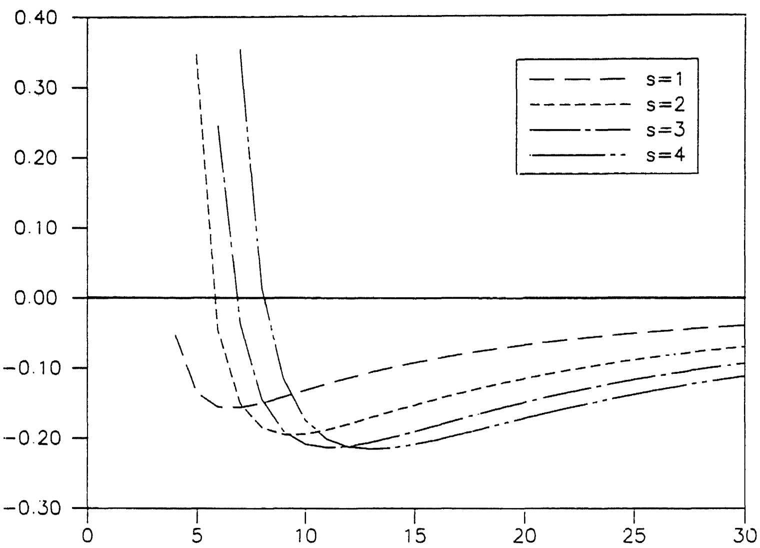

FIGURE 2. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1 = \sigma$, $\beta = 0.9$ ($\lambda = 0.9$), $\delta = 1$.

Table 2. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1 = \sigma$ $(\lambda = \beta)$, $n = 20$[a]

		β												
δ	s	−1.00	−0.90	−0.80	−0.40	0.00	0.20	0.40	0.60	0.80	0.90	0.95	0.99	1.00
1	1	0.0131	0.0048	0.0002	−0.0004	0.0000	−0.0002	−0.0013	−0.0059	−0.0269	−0.0665	−0.0886	−0.0948	−0.0944
1	2	−0.0286	−0.0109	−0.0022	−0.0003	0.0000	−0.0001	−0.0014	−0.0075	−0.0407	−0.1145	−0.1628	−0.1831	−0.1845
1	3	0.0436	0.0146	0.0019	−0.0003	0.0000	−0.0002	−0.0017	−0.0086	−0.0482	−0.1485	−0.2238	−0.2640	−0.2690
1	4	−0.0610	−0.0198	−0.0035	−0.0003	0.0000	−0.0002	−0.0019	−0.0092	−0.0518	−0.1712	−0.2722	−0.3362	−0.3464
*	1	0.0000	0.0038	−0.0002	−0.0004	0.0000	−0.0002	−0.0013	−0.0056	−0.0252	−0.0651	−0.0894	−0.0953	−0.0944
*	2	0.0000	−0.0083	−0.0019	−0.0002	0.0000	−0.0001	−0.0014	−0.0072	−0.0383	−0.1119	−0.1641	−0.1842	−0.1845
*	3	0.0000	0.0111	0.0017	−0.0003	0.0000	−0.0002	−0.0017	−0.0083	−0.0454	−0.1450	−0.2253	−0.2655	−0.2690
*	4	0.0000	−0.0148	−0.0030	−0.0003	0.0000	−0.0002	−0.0018	−0.0090	−0.0490	−0.1673	−0.2740	−0.3380	−0.3464

[a] A star (*) indicates $\delta = 1/\sqrt{1-\beta^2}$.

Table 3. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1 = \sigma$, $\beta = 1.0$ $(\lambda = 1.0)$

	s							
	1		2		3		4	
n	Exact	Approximate	Exact	Approximate	Exact	Approximate	Exact	Approximate
10	−0.1687	−0.2000	−0.2975	−0.4000	−0.3804	−0.6000	−0.4018	−0.8000
15	−0.1225	−0.1333	−0.2343	−0.2667	−0.3325	−0.4000	−0.4137	−0.5333
20	−0.0944	−0.1000	−0.1845	−0.2000	−0.2690	−0.3000	−0.3464	−0.4000
25	−0.0765	−0.0800	−0.1509	−0.1600	−0.2225	−0.2400	−0.2904	−0.3200
30	−0.0643	−0.0667	−0.1273	−0.1333	−0.1888	−0.2000	−0.2482	−0.2667

too small, say $n \geq 15$, BIAS/σ converges smoothly to zero. The closer β is to one, the slower is the convergence of BIAS/σ to zero.

4. UNIT ROOT

The most important range of the β's is near unity, both because of the recent interest in the random walk hypothesis and also because the downward bias can be substantial there (see Figure 1). Hence, in this section we shall study the behavior of the forecast bias when β approaches one.

We observe that BIAS/σ is stable very near the unit root. For example, whether $\beta = 0.99$ or $\beta = 1.0$ makes little difference for BIAS/σ (see Tables 1 and 2); however, there is a substantial increase in BIAS/σ from $\beta = 0.90$ to $\beta = 1.0$.

Our theory allows us to calculate the forecast bias at $\beta = 1$ exactly for any sample size. But since this involves numerical integration, it is important to obtain a large sample approximation. This approximation is given to Theorem 2.

THEOREM 2. *If, as* $\beta \to 1$, $\delta(1-\beta) \to 0$, *and* $\lambda \to \lambda_1$, *then*

$$\lim_{\beta \to 1} \frac{\text{BIAS}}{\sigma} = \begin{cases} \dfrac{-2s}{n\lambda_1} + 0(n^{-3/2}) & \text{if } \lambda_1 \neq 0 \\ 0 & \text{if } \lambda_1 = 0. \end{cases} \qquad \blacksquare$$

It is clear from Theorem 2 that the approximation depends very much on λ_1. On the other hand, the approximation (and indeed the exact limit) does not depend on δ as long as $\delta(1-\beta) \to 0$ as $\beta \to 1$, and this condition holds both in the fixed start-up ($\delta = 1$) and in the covariance stationary case ($\delta = (1-\beta^2)^{-1/2}$). This point was first made by Magnus and Rothenberg [7].

In the fixed start-up case $y_0 = c$ we have $\lambda \to \alpha/\sigma$ when $\beta \to 1$. For $\alpha/\sigma = 1$ we compare the approximation with the exact values in Table 3. The approximation turns out to be very good even for rather small n. If, however, the limiting value λ_1 of λ is small (for example, because σ is large relative to α), then the approximation is likely to be poor. This is so because the approximation becomes infinite when $\lambda_1 \to 0$, whereas BIAS/σ itself tends to zero. Table 4 compares the approximation with the exact values for $0 \leq \lambda_1 \leq 1$. Two facts emerge from Table 4. First, the approximation appears to be good even for $\lambda_1 = 0.5$ but, as expected, poor for smaller λ_1. Second, for λ_1 close to zero the exact value of BIAS/σ turns out to be roughly proportional to λ_1. More formally, we have the following.

THEOREM 3. *For small* λ,

$$\text{BIAS}/\sigma = \lambda 0(1) + 0(\lambda^3). \qquad \blacksquare$$

We remark that Theorem 3 is valid for every value of β, including $\beta \to 1$.

Table 4. BIAS/σ of LS forecast $\hat{y}_{n+s}$: $\beta = 1.0$, $n = 20$

	s							
	1		2		3		4	
λ_1	Exact	Approximate	Exact	Approximate	Exact	Approximate	Exact	Approximate
0.00	0.0000	—	0.0000	—	0.0000	—	0.0000	—
0.01	−0.0089	−10.0000	−0.0165	−20.0000	−0.0234	−30.0000	−0.0298	−40.0000
0.05	−0.0439	−2.0000	−0.0817	−4.0000	−0.1159	−6.0000	−0.1474	−8.0000
0.10	−0.0851	−1.0000	−0.1586	−2.0000	−0.2251	−3.0000	−0.2865	−4.0000
0.20	−0.1508	−0.5000	−0.2823	−1.0000	−0.4016	−1.5000	−0.5119	−2.0000
0.30	−0.1868	−0.3333	−0.3518	−0.6667	−0.5025	−1.0000	−0.6419	−1.3333
0.40	−0.1944	−0.2500	−0.3690	−0.5000	−0.5295	−0.7500	−0.6781	−1.0000
0.50	−0.1827	−0.2000	−0.3497	−0.4000	−0.5043	−0.6000	−0.6473	−0.8000
1.00	−0.0944	−0.1000	−0.1845	−0.2000	−0.2690	−0.3000	−0.3464	−0.4000

TABLE 5. BIAS/α of LS forecast $\hat{y}_{n+s}$: $\alpha = \mu_1$, $\sigma = \infty$, $n = 20$[a]

		β												
δ	s	−1.00	−0.90	−0.80	−0.40	0.00	0.20	0.40	0.60	0.80	0.90	0.95	0.99	1.00
1	1	0.0131	0.0048	0.0002	−0.0004	0.0000	−0.0002	−0.0013	−0.0063	−0.0397	−0.1819	−0.4263	−0.7814	−0.8870
1	2	−0.0286	−0.0110	−0.0022	−0.0003	0.0000	−0.0001	−0.0014	−0.0078	−0.0565	−0.2982	−0.7469	−1.4364	−1.6503
1	3	0.0436	0.0147	0.0019	−0.0003	0.0000	−0.0002	−0.0017	−0.0089	−0.0649	−0.3798	−1.0052	−2.0150	−2.3399
1	4	−0.0610	−0.0199	−0.0035	−0.0003	0.0000	−0.0002	−0.0019	−0.0096	−0.0683	−0.4372	−1.2178	−2.5377	−2.9766
*	1	0.0000	0.0038	0.0002	−0.0004	0.0000	−0.0002	−0.0013	−0.0060	−0.0348	−0.1501	−0.3586	−0.7392	−0.8870
*	2	0.0000	−0.0083	−0.0019	−0.0003	0.0000	−0.0001	−0.0014	−0.0076	−0.0505	−0.2485	−0.6320	−1.3614	−1.6503
*	3	0.0000	0.0111	0.0017	−0.0003	0.0000	−0.0002	−0.0017	−0.0086	−0.0586	−0.3172	−0.8524	−1.9119	−2.3399
*	4	0.0000	−0.0148	−0.0030	−0.0003	0.0000	−0.0002	−0.0019	−0.0093	−0.0622	−0.3647	−1.0326	−2.4093	−2.9766

[a] A star (*) indicates $\delta = 1/\sqrt{1-\beta^2}$.

5. SIGMA ASYMPTOTICS

In the fixed zero start-up case $y_0 = 0$, we have $\lambda = \alpha\beta/\sigma$ and $\delta = 1$. In Section 3 we put $\alpha/\sigma = 1$. In this section we shall comment briefly on the limiting behavior of BIAS/σ when $\alpha/\sigma \to 0$ and when $\alpha/\sigma \to \infty$. The small and large sigma asymptotics are given by Theorem 4.

THEOREM 4. *We have*

$$\frac{\text{BIAS}}{\sigma} = \frac{1}{\sigma}\,0(1) + 0(\sigma^{-3}) \qquad \text{for large } \sigma$$

and

$$\frac{\text{BIAS}}{\sigma} = \sigma 0(1) + 0(\sigma^{3}) \qquad \text{for small } \sigma. \qquad \blacksquare$$

Theorem 4 tells us that BIAS/σ approaches zero when $\sigma \to 0$ and when $\sigma \to \infty$. It also tells us that BIAS/σ is roughly proportional to σ when σ is small and roughly proportional to $1/\sigma$ when σ is large. Thus, it turns out that the BIAS itself remains finite when σ tends to ∞. Indeed, one can show that the BIAS reaches a limit.

Since, when $\mu_1 = \alpha$ and $\sigma = \infty$, the BIAS is proportional to α, we find that BIAS/α depends on n, s, β, and δ. In Table 5 we present the limiting values of BIAS/α for $n = 20$ in the fixed start-up case ($\delta = 1$) and the covariance stationary case ($\delta = (1 - \beta^2)^{-1/2}$). When we compare Table 5 ($\alpha = 1$, $\sigma = \infty$) with Table 2 ($\alpha = 1$, $\sigma = 1$) we find a most remarkable result, namely that the BIAS is virtually unaffected by the value of $\sigma \in (1,\infty)$ unless β is close to $+1$. But for β close to unity the effect of σ is substantial. At $\beta = 1$, for example, BIAS is approximately ten times as large when $\sigma = \infty$ than when $\sigma = 1$. This highlights again the different behavior of the BIAS when β is close to unity.

REFERENCES

1. Cryer, J.D., J.C. Nankervis & N.E. Savin. Forecast error symmetry in ARIMA models. The University of Iowa typescript, 1989.
2. Fuller, W.A. & D.P. Hasza. Predictors for the first-order autoregressive process. *Journal of Econometrics* 13 (1980): 139–157.
3. Hoque, A., J.R. Magnus & B. Pesaran. The exact multi-period mean-square forecast error for the first-order autoregressive model. *Journal of Econometrics* 39 (1988): 327–346.
4. Lahiri, K. Multiperiod predictions in dynamic models. *International Economic Review* 16 (1975): 699–711.
5. Magnus, J.R. On certain moments relating to ratios of quadratic forms in normal variables: further results. *Sankhyā*, Series B (Part 1) 52 (1991): 1–13.
6. Magnus, J.R. & B. Pesaran. The exact multi-period mean-square forecast error for the first-order autoregressive model with an intercept. *Journal of Econometrics* 42 (1989): 157–179.
7. Magnus, J.R. & T.J. Rothenberg. Least-squares autoregression with near-unit root. LSE typescript, 1988.

8. Malinvaud, E. *Statistical Methods of Econometrics*, 2nd Ed. Amsterdam: North-Holland, 1970.
9. Numerical Algorithms Group. NAG Fortran Library Manual, Mark 11. Oxford: NAGLIB, 1984.

MATHEMATICAL APPENDIX

Proof of Theorem 1. Let $\bar{u} = (u_2 + \cdots + u_n)/(n-1)$, so that $\alpha = \bar{y}_{**} - \beta\bar{y}_* - \bar{u}$. Then, using (4), we may rewrite (7) as

$$\hat{y}_{n+s} - y_{n+s} = -\sum_{j=0}^{s-1} \beta^j u_{n+s-j} + \bar{u}\sum_{j=0}^{s-1}\beta^j + (\hat{\beta}^s - \beta^s)(y_n - \bar{y}_*) + \frac{1}{n-1}\sum_{j=0}^{s-1}(\hat{\beta}^j - \beta^j)(y_n - y_1). \quad \textbf{(A.1)}$$

If the expectation exists we get

$$E(\hat{y}_{n+s} - y_{n+s}) = E(\hat{\beta}^s - \beta^s)(w_2'y) + \sum_{j=0}^{s-1} E(\hat{\beta}^j - \beta^j)(w_1'y)$$
$$= \tau_s(w_2) - \beta^s w_2'\mu + \sum_{j=0}^{s-1}\tau_j(w_1) - \left(\sum_{j=0}^{s-1}\beta^j\right)(w_1'\mu)$$

and (8) follows.

To prove the existence of the expectation we note that $\hat{\beta} = y'Ay/y'By$ and that the matrix B has rank $n-2$. Hence, there exists an $n \times 2$ matrix $Q \equiv (q_1, q_2)$ of rank 2 such that $BQ = 0$. One choice for q_1 and q_2 is

$$q_1 = (0,0,\ldots,0,1)', \qquad q_2 = (1,1,\ldots,1,0)'.$$

We have

$$Q'AQ = 0, \qquad AQ \neq 0, \qquad Q'w_1 \neq 0, \qquad Q'w_2 \neq 0,$$

and hence, by Theorem 2(iii) of Magnus [5], $\tau_1, \ldots, \tau_s$ all exist if and only if $s < n-2$. This completes the proof.

Proof of Theorem 2. See (A.10) in the Appendix of Magnus and Rothenberg [7].

Proof of Theorem 3. From (10) we write $z_t = -\lambda\varphi_t + L_t$ where L_t is linear in $u_1^*, \ldots, u_{n-1}^*$ and independent of λ. Then, using (12),

$$\hat{\beta} - \beta = \frac{\sum(-\lambda\varphi_{t-1} + L_{t-1})u_t^*}{\sum(-\lambda\varphi_{t-1} + L_{t-1})^2} = A_0 + \lambda A_1 + \lambda^2 A_2 + 0(\lambda^3)$$

where A_0 and A_2 are even functions of the u^*'s and A_1 is an odd function of the u^*'s. This gives, for $k = 1,2,\ldots,$

$$\hat{\beta}^k - \beta^k = \sum_{j=0}^{k-1} \binom{k}{j} \beta^j A_0^{k-j} + k\lambda(\beta + A_0)^{k-1}A_1 + \lambda^2 A_4 + 0(\lambda^3)$$

where A_4 is some even function of the u^*'s. Hence,

$$E(\hat{\beta}^k - \beta^k)z_t = \lambda\left[-\varphi_t \sum_{j=0}^{k-1} \binom{k}{j} \beta^j E A_0^{k-j} + kE(\beta + A_0)^{k-1}A_1 L_t\right] + 0(\lambda^3)$$

and the result follows from (13).

Proof of Theorem 4. The large σ result is a direct consequence of Theorem 3. To prove the small σ result we define $m = \lambda\sigma$, so that

$$y_t - \bar{y}_* = -m\varphi_t + \sigma L_t$$

and

$$\frac{\hat{\beta} - \beta}{\sigma} = \frac{\sum(-m\varphi_{t-1} + \sigma L_{t-1})u_t^*}{\sum(-m\varphi_{t-1} + \sigma L_{t-1})^2} = B_1 + \sigma B_2 + \sigma^2 B_3 + 0(\sigma^3)$$

where B_1 and B_3 are odd functions and B_2 an even function of the u^*'s. Considering $E(\hat{\beta}^k - \beta^k)(y_t - \bar{y}_*)$ then gives the required result.