

August 1976

Division of Research  
Graduate School of Business Administration  
The University of Michigan

MULTIPERIOD PREDICTIONS FROM AN AUTOREGRESSIVE  
MODEL USING EMPIRICAL BAYES METHODS

Working Paper No. 135

by

R.W. Andrews

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

None of this material is to be quoted or  
reproduced without the express permission  
of the Division of Research.

## ABSTRACT

The ability of an estimated dynamic model to produce accurate predictions depends on the method used to estimate the parameters in the model. This report investigates multiperiod predictions for a first-order autoregressive process. A Bayes predictor is derived using a bivariate normal prior distribution over the parameters of the model. The loss structure used depends on the prediction and the variable to be predicted; therefore this is a statistical decision theory problem. The form of the Bayes predictor lends itself to Empirical Bayes procedures. Two Empirical Bayes procedures are presented which bypass the assessment of a prior distribution. All three procedures are developed so that the estimates of the model parameters depend on the number of steps ahead the forecast is desired.

## Introduction

Since economic time series data is often analyzed using autoregressive models, the methods employed to estimate the parameters of these models is of paramount importance. Using the estimated model to predict is often the purpose of observing time series data. The predictions could be required for one period of time in the future or for many periods ahead. Estimated models which exhibit certain properties for predicting one period ahead may not exhibit these same properties for multiperiod predictions. Therefore, when estimating the model, it is reasonable to consider the number of periods ahead for which the forecast is desired. Consequently, this is a statistical decision theory problem rather than just an exercise in fitting the data to a model.

The dynamic economic structure considered in this report is limited to a first-order autoregressive model. This problem of finding multiperiod predictions from an autoregressive process has been investigated in Klein [3], Chow [1], and Lahiri [4]. The estimator suggested by Klein [3], and referred to by Fair [2] as the DYN estimator, attempts to account for the unknown values of all lagged endogeneous variables in multiperiod predictions. A Bayesian analysis of the problem of estimating an autoregressive model has been reported by Chow [1], who gives a Bayesian solution for the first-order autoregressive process. His analysis assumes that the prior distribution is diffuse. In section three, a Bayesian solution with respect to a normal prior distribution is reported. The loss function is the square of the specific multiperiod prediction error. The use of this loss

structure results in the estimator depending on the number of periods ahead for which the prediction is desired.

In this report, Empirical Bayes procedures are utilized in estimating a first-order autoregressive process. The Empirical Bayes methods of estimating the parameters are developed to depend on the number of periods ahead for which the forecast is required. These procedures employ the same criterion as Chow [1]. However, the Empirical Bayes procedures do not require the assumption of a specific prior distribution on the parameters. The Empirical Bayes estimators presented below use the form of the derived Bayes estimator as given in section three. The artificial assumption of a known prior distribution is circumvented, however.

In section three the Bayes estimator is given as a function of the prior distribution. From this form two Empirical Bayes predictors are presented in section four. The concluding section comments on the applicability of these predictors and suggests essential further research.

### The Model

A first-order autoregressive model can be represented as follows:

$$(1) \quad y_t = ay_{t-1} + c + u_t ; \text{ for } t = 1, 2, \dots$$

$$\text{for which } E[u_t] = 0 \text{ and } E[u_t u_s] = h^{-1} \delta_{ts}$$

Throughout this development it is assumed that the initial value of the series,  $y_0$ , is fixed and known. In addition, the error term is assumed to have a normal distribution with a known precision  $h = 1/\sigma^2$ . The

unknown parameters are therefore the coefficients  $a$  and  $c$ .

It is not the purpose of this report to investigate procedures of directly estimating  $a$  and  $c$ . The problem of interest in this report will be predicting  $y_{n+k}$  based on the observations  $y_1, y_2, \dots, y_n$ . In the next section, Bayes predictors of  $y_{n+2}$  and  $y_{n+3}$  are reported. The following section gives two different forms of the Empirical Bayes predictor of  $y_{n+k}$ .

Since the Bayesian framework will be utilized for both the Bayes and Empirical Bayes predictors, an assumption must be made concerning the realization of the random variables  $a$  and  $c$ . This assumption may be made in two different ways. The augmentation of this assumption to equation (1) results in two different models. This assumption, denoted by either A or B, concerns the function of the prior distribution on  $a$  and  $c$ .

Assumption A: For a specific time series;  $y_0, y_1, y_2, \dots, y_n, \dots$ ;  $(a, c)$  is a realization from a bivariate prior distribution and  $(a, c)$  is fixed throughout the time series. That is, the value of  $(a, c)$  is the same for each time  $t = 1, 2, \dots, n, \dots$ .

Assumption B: For a specific time series;  $y_0, y_1, y_2, \dots, y_n, \dots$  and for each time  $t = 1, 2, \dots, n, \dots$ , a realization from the prior distribution of  $(a, c)$  is experienced. That is, for a given time series, the value of  $(a, c)$  at time  $t$  might be different from the value of  $(a, c)$  at time  $s \neq t$ , even though they are both from the same distribution.

In essence, under Assumption A, the values of  $(a, c)$  are fixed from one step to the next; whereas, under Assumption B, the values of  $(a, c)$  change from step to step. Throughout this report it is assumed that

the prior distribution of  $a$  and  $c$  are independent. (Comments as to how the predictors developed should perform under Assumption A or B will be given in the concluding remarks.) All classical developments of this problem, such as Klein [3], assume that  $a$  and  $c$  are fixed throughout the time series. This is consistent with Assumption A except that classical theory does not assume a prior distribution.

Under Assumption B, equation (1) should be written:

$$(2) \quad y_t = a_t y_{t-1} + c_t + u_t ; \text{ for } t = 1, 2, \dots .$$

#### The Bayes Predictor

This section will develop the form of the Bayes estimator as well as the Bayes predictor of  $y_{n+2}$  and  $y_{n+3}$  for an assumed normal prior. The unknown parameters of the model as given by equation (1) are  $(a, c)$ . Let  $p(a, c)$  denote the prior distribution for these parameters. At the  $n^{\text{th}}$  stage we have observed  $\tilde{y}_n = (y_1, y_2, \dots, y_n)$ , and we want to predict  $y_{n+k}$ . Since the predictor or forecast is a function of  $\tilde{y}_n$ , we write  $\hat{y}_{n+k} = \hat{y}_{n+k}(y_1, y_2, \dots, y_n)$  as the predictor of  $y_{n+k}$ . To find the form of  $\hat{y}_{n+k}$ , the criterion is to minimize the Bayes risk given by

$$(3) \quad R(p) = E_p E\{(\hat{y}_{n+k} - y_{n+k})^2\} .$$

This implicitly assumes a squared error loss on the value of the  $y$  we wish to predict. This also will be the loss function assumed in the development of the Empirical Bayes procedures. The interior expected value of equation (3) is taken with respect to the conditional distri-

bution of  $y_1, y_2, \dots, y_n$  and  $y_{n+k}$  given  $a$  and  $c$ . The other expectation is taken with respect to the prior density  $p(a, c)$  as noted.

Using equation (1) repeatedly we may write,

$$y_{n+k} = a^k y_n + (a^{k-1}c + a^{k-2}c + \dots + c) + (a^{k-1}u_{n+1} + a^{k-2}u_{n+2} + \dots + u_{n+k}) .$$

Then,

$$\begin{aligned} (4) \quad E(\hat{y}_{n+k} - y_{n+k})^2 &= \\ E\{\hat{y}_{n+k} - a^k y_n - (a^{k-1}c + a^{k-2}c + \dots + c) - (a^{k-1}u_{n+1} + a^{k-2}u_{n+2} + \dots + u_{n+k})\}^2 &= \\ E\{\hat{y}_{n+k} - a^k y_n - (a^{k-1}c + a^{k-2}c + \dots + c)\}^2 + & \\ E\{a^{k-1}u_{n+1} + a^{k-2}u_{n+2} + \dots + u_{n+k}\}^2 . & \end{aligned}$$

The last equality holds since  $u_t$  with  $t > n$  are uncorrelated with  $\tilde{y}_n$ .

Since we want to find the  $\hat{y}_{n+k}$  which minimizes equation (3), and since the second term on the right hand side of equation (4) does not contain  $\hat{y}_{n+k}$ , we have reduced the problem to minimizing

$$\begin{aligned} (5) \quad E_p E\{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1}c + \dots + c)\}^2 &= \\ \int E\{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1}c + \dots + c)\}^2 p(a, c) da dc &= \\ \iint \{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1}c + \dots + c)\}^2 p(\tilde{y}_n | a, c) p(a, c) d\tilde{y}_n da dc , & \end{aligned}$$

where  $p(\tilde{y}_n | a, c)$  is the conditional distribution of  $\tilde{y}_n$  given  $(a, c)$ .

The integrals written here, as throughout this entire report, are integrated over the entire appropriate sample space. Since

$$p(\tilde{y}_n | a, c) p(a, c) = p(a, c | \tilde{y}_n) p(\tilde{y}_n) ,$$

we may rewrite equation (5) as

$$(6) \quad E_p E\{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1} c + \dots + c)\}^2 = \iint \{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1} c + \dots + c)\}^2 p(a, c | \tilde{y}_n) p(\tilde{y}_n) d\tilde{y}_n da dc .$$

For an observed  $\tilde{y}_n = (y_1, y_2, \dots, y_n)'$ , equation (6) is minimized if we minimize the integrand

$$\int \{\hat{y}_{n+k} - a^k y_n - (a^k c + a^{k-1} c + \dots + c)\}^2 p(a, c | \tilde{y}_n) da dc ,$$

with respect to  $\hat{y}_{n+k}$ . Setting the derivative equal to zero we obtain

$$(7) \quad \hat{y}_{n+k} = \int a^k y_n + (a^k c + a^{k-1} c + \dots + c) p(a, c | \tilde{y}_n) da dc .$$

We may write that

$$p(a, c | \tilde{y}_n) = \frac{p(\tilde{y}_n | a, c) p(a, c)}{\int p(\tilde{y}_n | a, c) p(a, c) da dc} ,$$

and therefore the form of the Bayes estimator may be written

$$(8) \quad \hat{y}_{n+k} = \frac{\int \{a^k y_n + a^{k-1} c + a^{k-2} c + \dots + c\} p(\tilde{y}_n | a, c) p(a, c) da dc}{\int p(\tilde{y}_n | a, c) p(a, c) da dc} .$$

Equation (7) is the form of the Bayes estimator from which the Bayes predictor for two and three steps ahead will be developed. Equation (8) is the form which will motivate the Empirical Bayes estimator of  $y_{n+k}$ , as given in the next section.

If we assume that  $(a, c)$  are bivariate normal with mean vector  $(\mu_1, \mu_2)'$  and variance-covariance matrix  $\begin{bmatrix} 1/f_1 h & 0 \\ 0 & 1/f_2 h \end{bmatrix}$ ,



then

$$(9) \quad p(a, c) \propto \exp \frac{-h}{2} \left\{ \begin{bmatrix} a-\mu_1 \\ c-\mu_2 \end{bmatrix}' \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix} \begin{bmatrix} a-\mu_1 \\ c-\mu_2 \end{bmatrix} \right\} .$$

The conditional distribution of the data may be represented by

$$(10) \quad p(\tilde{y}_n | a, c) \propto \exp \frac{-h}{2} \left\{ \begin{bmatrix} y_1 - ay_0 - c \\ y_2 - ay_1 - c \\ \vdots \\ y_n - ay_{n-1} - c \end{bmatrix}' \begin{bmatrix} y_1 - ay_0 - c \\ y_2 - ay_1 - c \\ \vdots \\ y_n - ay_{n-1} - c \end{bmatrix} \right\} .$$

Using equations (9) and (10) we derive

$$p(\tilde{y}_n | a, c) \propto \exp \frac{-h}{2} \left\{ \begin{bmatrix} a-\mu_1 \\ c-\mu_2 \end{bmatrix}' \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix} \begin{bmatrix} a-\mu_1 \\ c-\mu_2 \end{bmatrix} + \begin{bmatrix} y_1 - ay_0 - c \\ y_2 - ay_1 - c \\ \vdots \\ y_n - ay_{n-1} - c \end{bmatrix}' \begin{bmatrix} y_1 - ay_0 - c \\ y_2 - ay_1 - c \\ \vdots \\ y_n - ay_{n-1} - c \end{bmatrix} \right\} .$$

In order to simplify this conditional density we will use the following identity as given in Raiffa and Schlaiffer [7; p. 337]:

$$\begin{aligned} & (y-X\beta)'(y-X\beta) + (\beta-b)'V(\beta-b) = \\ & (\beta-[V+X'X]^{-1}[Vb+X'y])'(V+X'X)(\beta-[V+X'X]^{-1}[Vb+X'y]) \\ & + b'Vb + y'y - (Vb+X'y)'(V+X'X)^{-1}(Vb+X'y) . \end{aligned}$$

Notice that  $\beta$  is contained only in the first term of the right hand

side of this identity. If we set

$$\beta = \begin{bmatrix} a \\ c \end{bmatrix}, \quad b = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad v = \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix},$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \text{and} \quad X = \begin{bmatrix} y_0 & 1 \\ y_1 & 1 \\ \vdots & \vdots \\ y_{n-1} & 1 \end{bmatrix},$$

then we can find the posterior distribution of  $(a, c)$  given  $\tilde{y}_n$ . It is obviously normal with the parameters as given below. All summations are from  $i = 1$  to  $n$ . Next, set

$$D^{-1} = (n+f_2) (f_1 + \sum y_{i-1}^2) - (\sum y_{i-1})^2.$$

The posterior mean of  $a$  is

$$\mu_a = D \{ (n+f_2) (\mu_1 f_1 + \sum y_i y_{i-1}) - (\mu_2 f_2 + \sum y_i) \sum y_{i-1} \};$$

the posterior mean of  $c$  is

$$\mu_c = D \{ (f_1 + \sum y_{i-1}^2) (\mu_2 f_2 + \sum y_i) - (\mu_1 f_1 + \sum y_i y_{i-1}) \sum y_{i-1} \};$$

the posterior variance of  $a$  is

$$\sigma_a^2 = D h^{-1} (n+f_2);$$

the posterior variance of  $c$  is

$$\sigma_c^2 = D h^{-1} (f_1 + \sum y_{i-1}^2); \text{ and}$$

the posterior correlation coefficient of  $a$  and  $c$  is

$$\rho_{ac} = (-\sum y_{i-1}) / [(n+f_2)(f_1 + \sum y_{i-1}^2)] .$$

Using these posterior parameters and equation (7) we can find the Bayes predictors for two and three steps ahead. In order to accomplish this, we need the following statements: If (X,Y) are bivariate normal  $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ , then

$$E[Y] = \mu_Y ,$$

$$E[XY] = \rho\sigma_X\sigma_Y + \mu_X\mu_Y ,$$

$$E[X^2Y] = 2\rho\mu_X\sigma_X\sigma_Y + \mu_Y\sigma_X^2 + \mu_Y\mu_X^2 ,$$

$$E[X^3Y] = 3\rho\mu_X^2\sigma_X\sigma_Y + 3\rho\sigma_X^3\sigma_Y + 3\mu_X\mu_Y\sigma_X^2 + \mu_X^3\mu_Y .$$

Using these identities in equation (7) we find that

$$\hat{y}_{n+2} = y_n(\sigma_a^2 + \mu_a^2) + \rho_{ac}\sigma_a\sigma_c + \mu_a\mu_c + \mu_c ,$$

and

$$\begin{aligned} \hat{y}_{n+3} = & y_n(3\mu_a\sigma_a^2 + \mu_a^3) + \mu_c\sigma_a^2 + \mu_c\mu_a^2 + 2\rho_{ac}\mu_a\sigma_a\sigma_c + \rho_{ac}\sigma_a\sigma_c \\ & + \mu_a\mu_c + \mu_c . \end{aligned}$$

### Empirical Bayes Predictors

Using the form of the Bayes predictor given in equation (8) Empirical Bayes procedures will now be investigated. In order to demonstrate the Empirical Bayes procedures as described in Lemon and Krutchkoff [5], we will display the data in the following form:

experience	0	1	2	3	4	.....	j	.....	n
	$y_0$	$y_0$	$y_0$	$y_0$	$y_0$	.....	$y_0$	.....	$y_0$
		$y_1$	$y_1$	$y_1$	$y_1$		$y_1$		$y_1$
			$y_2$	$y_2$	$y_2$		$y_2$		$y_2$
				$y_3$	$y_3$		$y_3$		$y_3$
					$y_4$		$y_4$		$y_4$
					$\vdots$		$\vdots$		$\vdots$
					$\vdots$		$y_j$		$\vdots$
					$\vdots$				$\vdots$
					$\vdots$				$y_n$
					$\vdots$				
					$\vdots$				
					$\hat{a}_4$	.....	$\hat{a}_j$	.....	$\hat{a}_n$
					$\hat{c}_4$		$\hat{c}_j$		$\hat{c}_n$
					$\hat{\sigma}_4^2$	.....	$\hat{\sigma}_j^2$	.....	$\hat{\sigma}_n^2$

If the problem is considered as arrayed above, every new observation will enable us to consider a new experience. For example, after  $j$  observations we are at the  $j^{\text{th}}$  experience and have the data  $y_1, y_2, \dots, y_j$  available. Consider the situation in which we have observed

$\tilde{y}_n = (y_1, y_2, \dots, y_n)'$  and we wish to predict  $y_{n+k}$ . We want to use the form of the Bayes predictor, equation (8); however we do not

have nor do we want to assess a prior distribution. For each experience after the third stage we have been able to estimate  $a, c$ , and  $\sigma^2 = h^{-1}$ .

These estimates could have been developed in numerous ways, but

throughout this report we will use the ordinary least squares estimates and designate them by  $\hat{a}_j, \hat{c}_j$ , and  $\hat{\sigma}_j^2$ ; for  $j = 4, 5, \dots, n$ . If we

give equal weight to each of the estimates of the parameters, i.e.,

let  $p(a, c) = 1/n-3$  for  $(a, c) = (\hat{a}_j, \hat{c}_j)$ , then we can rewrite equation (8)

as

$$(11) \quad \hat{y}_{n+k}^E = \frac{\sum [\hat{a}_j^k y_n + (\hat{a}_j^{k-1} \hat{c}_j + \hat{a}_j^{k-2} \hat{c}_j + \dots + \hat{c}_j)] p(\tilde{y}_n | \hat{a}_j, \hat{c}_j)}{\sum p(\tilde{y}_n | \hat{a}_j, \hat{c}_j)}$$

where the summations are taken from  $j = 4$  to  $n$ . The quantity

$$\hat{a}_j^k y_n + (\hat{a}_j^{k-1} \hat{c}_j + \hat{a}_j^{k-2} \hat{c}_j + \dots + \hat{c}_j)$$

is a prediction of  $y_{n+k}$  based on the estimates obtained at experience  $j$ .

Therefore, equation (11) is a weighted average of estimates of  $y_{n+k}$ .

The weighting function is given by

$$p(\tilde{y}_n | \hat{a}_j, \hat{c}_j) \propto \hat{\sigma}_j^{-n} \exp \left\{ \frac{1}{2\hat{\sigma}_j^2} \sum (y_i - \hat{a}_j y_{i-1} - \hat{c}_j)^2 \right\},$$

with the summation taken from  $i = 1$  to  $n$ . The weight is proportional to the conditional likelihood of the present experience given the estimates of the parameters at the experience to be weighted. In essence we are weighting the prediction developed from each experience by how well the estimated parameters explain the current data,  $\tilde{y}_n$ . Notice that in using this technique we may, without loss of generality, assume that  $(a, c)$  is changing from experience to experience. Also, in using the Empirical Bayes procedure we have eliminated the assumption of a known conditional precision,  $h$ . Using ordinary least squares methods we are able to estimate  $\sigma^2 = 1/h$ .

At the  $j^{\text{th}}$  experience, the ordinary least squares estimates are given by

$$(\hat{a}_j, \hat{c}_j)' = (X_j' X_j)^{-1} X_j' \tilde{y}_j$$

where

$$X_j = \begin{bmatrix} y_0 & 1 \\ y_1 & 1 \\ \vdots & \\ y_{j-1} & 1 \end{bmatrix}, \quad \tilde{y}_j = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_j \end{bmatrix},$$

and

$$\hat{\sigma}_j^2 = \frac{\sum (y_i - \hat{c}_j - \hat{a}_j y_{i-1})^2}{j-2}, \quad \text{with the summation taken from}$$

$i = 1$  to  $j$ . The decision to start after the third observation insures that we have enough degrees of freedom to estimate the conditional variance. The weights assigned to the early experiences are usually, but not always, smaller than the weights assigned to the later experiences.

The properties of Empirical Bayes procedures, such as equation (11), have always been investigated using Monte Carlo experiments. The results of these experiments seem to indicate that the procedure is asymptotically optimal [5]. That is, the risk of using the proposed estimator approaches the risk of the Bayes estimator as the number of experiences approach infinity. Only for some very simple examples have the Empirical Bayes procedures actually been shown to be asymptotically optimal, using analytic, not Monte Carlo, techniques. In all cases heretofore considered, an assumption of the model was that the conditional data at experience  $i$  was independent of the conditional data at experience  $j \neq i$ . This is certainly not the case in our development since the data at experience  $j$  contain all the observations from experience  $j-1$ .

In order to adjust for the dependence between the experiences, a fixed number of observations could be assigned to each experience.

For example, with quarterly data, four observations per experience might be chosen. That situation could be represented as follows:

experience	0	1	2	.....	j	.....	n
	$y_0$	$y_1$	$y_5$	.....	$y_{4j-3}$	.....	$y_{4n-3}$
		$y_2$	$y_6$		$y_{4j-2}$		$y_{4n-2}$
		$y_3$	$y_7$		$y_{4j-1}$		$y_{4n-1}$
		$y_4$	$y_8$		$y_{4j}$		$y_{4n}$
		⋮	⋮				
		⋮	⋮				
		⋮	⋮				
		$\hat{a}_1$	$\hat{a}_2$		$\hat{a}_j$		$\hat{a}_n$
		$\hat{c}_1$	$\hat{c}_2$		$\hat{c}_j$		$\hat{c}_n$
		$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$		$\hat{\sigma}_j^2$		$\hat{\sigma}_n^2$

Since we are assuming an autoregressive model as given by equation (1), the  $j^{\text{th}}$  experience is still not independent of the  $j-1^{\text{st}}$  experience. For example  $y_5$  is not independent of  $y_4$  ; however, the conditional distributions of nonadjacent experiences are independent. In this case the conditioning occurs with respect to the last observation.

The ordinary least squares estimates of  $a$ ,  $c$ , and  $\sigma^2$  are given by

$$(\hat{a}_j, \hat{c}_j)' = (X_j'X_j)^{-1} X_j'y_j$$

where

$$X_j = \begin{bmatrix} y_{3j} & 1 \\ y_{4j-1} & 1 \\ y_{4j-2} & 1 \\ y_{4j-3} & 1 \end{bmatrix} \quad \text{and} \quad y_j = \begin{bmatrix} y_{4j-3} \\ y_{4j-2} \\ y_{4j-1} \\ y_{4j} \end{bmatrix} .$$

Therefore the Empirical Bayes predictor of  $y_{n+k}$ , based on the form of the Bayes predictor in equation (8), is given by

$$(12) \quad \hat{y}_{n+k}^* = \frac{\sum (\hat{a}_j^k y_{4n} + \hat{a}_j^{k-1} \hat{c}_j + \hat{a}_j^{k-2} \hat{c}_j + \dots + \hat{c}_j) p(\tilde{y}_n^* | \hat{a}_j, \hat{c}_j)}{\sum p(\tilde{y}_n^* | \hat{a}_j, \hat{c}_j)}$$

where  $\tilde{y}_n^* = (y_{4n-3}, y_{4n-2}, y_{4n-1}, y_{4n})'$

and

$$p(\tilde{y}_n^* | \hat{a}_j, \hat{c}_j) \propto \hat{\sigma}_j^{-4} \exp \left\{ \frac{1}{2\hat{\sigma}_j^2} \sum (y_{4n-i} - \hat{a}_j y_{4n-i-1} - \hat{c}_j)^2 \right\},$$

with the summation taken from  $i=0$  to 3.

#### Concluding Remarks

The Bayes predictor, as given by equation (7), implicitly incorporates Assumption A into its model. Since the Bayes predictor uses the entire series  $\tilde{y}_n$  as a single entity, it assumes that  $(a,c)$  is a fixed realization from  $p(a,c)$  throughout the series. The Bayes predictor also has the difficult requirement of assessing a prior distribution on  $(a,c)$ . Even though the mean vector of  $(a,c)$  might be easy to assess, it is unlikely that the variance of  $(a,c)$  could be assessed accurately.

The Empirical Bayes predictor,  $\hat{y}_{n+k}^E$ , permits Assumption B to be part of the model. However, the advantage of a more flexible model is probably offset by the disadvantage of strong dependence between experiences. The Empirical Bayes predictor,  $\hat{y}_{n+k}^*$ , implicitly allows the assumption that  $(a,c)$  changes after every fourth observation. For quarterly data, this would be consistent with an annual revision of



model parameters. Also, this Empirical Bayes procedure has the advantage of conditional independence of experiences as explained in the last section.

In order to test the worth of these proposed methods of prediction, a Monte Carlo study, similar to that reported by Orcutt and Winokur [6] and Lahiri [4], should be executed. The study should include observations generated from both models (1) and (2).

REFERENCES

- [1] Chow, G.C. "Multiperiod Predictions from Stochastic Difference Equations by Bayesian Methods." Econometrica 41, No. 1 (Jan. 1973): 109-18.
- [2] Fair, R.C. "A Comparison of Alternative Estimators of Macroeconomic Models." International Economic Review 14, No. 2 (June 1973): 261-77.
- [3] Klein, L.R. "An Essay on the Theory of Economic Prediction." 1968-Yrjö Jahnsson Lectures. Chicago, Ill.: Markham Publishing Company, 1971.
- [4] Lahiri, K. "Multiperiod Predictions in Dynamic Models." International Economic Review 16, No. 3 (Oct. 1975): 699-711.
- [5] Lemon, G.H., and Krutchkoff, R.G. "An Empirical Bayes Smoothing Technique." Biometrika 56, No. 2 (1969): 361-5.
- [6] Orcutt, G.H., and Winokur, H.S., Jr. "First Order Autoregression: Inference, Estimation, and Prediction." Econometrica 37, No. 1 (Jan. 1969): 1-14.
- [7] Raiffa, H., and Schlaiffer, R. Applied Statistical Decision Theory. Boston, Mass.: Harvard University, 1961.