

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

December 1973

Computer Simulation in Varying  
Parameter Regression Models

Working Paper No. 88

by

Archer McWhorter

The University of Houston

and

W. Allen Spivey

William J. Wroblewski

The University of Michigan

© The University of Michigan

FOR DISCUSSION PURPOSES ONLY

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

## Introduction

This paper is concerned with the  $p$ -by-1 time series  $y(t)$  which, for time periods  $t = 1, \dots, n$  has the linear regression structure

$$(1) \quad y(t) = X(t)\beta(t) + e(t),$$

where  $X(t)$  is a known  $p$ -by- $r$  matrix of fixed regressors, and  $e(t)$  is a  $p$ -by-1 vector of equation errors having mean 0 and known  $p$ -by- $p$  variance-covariance matrix  $V_e(t)$ . The equation errors  $e(1), \dots, e(n)$  are serially uncorrelated.

In conventional regression theory the  $r$ -by-1 vector of regression coefficients  $\beta(t)$  is regarded as a fixed, nonrandom vector. In this paper, however, we investigate varying parameter regression models in which  $\beta(t)$  is taken to be a random coefficient vector formed for time periods  $t = 1, \dots, n$  from the transition equation

$$(2) \quad \beta(t) = T(t)\beta(t-1) + u(t),$$

where  $T(t)$  is a known coefficient updating matrix of size  $r$ -by- $r$ , and the  $r$ -by-1 random coefficient error vector  $u(t)$  has mean 0 and known  $r$ -by- $r$  variance-covariance matrix  $V_u(t)$ . The coefficient errors  $u(1), \dots, u(n)$  are also serially uncorrelated. In addition, the

random coefficient errors  $u(t)$  and the random equation errors  $e(t)$  are mutually uncorrelated both within and between time periods. Finally, for the first time period ( $t = 1$ ) we have

$$(3) \quad \beta(1) = T(1)\beta(0) + u(1),$$

where the  $r$ -by-1 vector  $\beta(0)$ , called the starting value, is taken to be a known, fixed, nonrandom vector.

This class of models was first introduced by Kalman [4] and by Kalman and Bucy [5] and it forms a part of the theory of optimal control. It has been applied to many engineering problems and there is an extensive literature on the theory and on these applications (good survey papers are [2] and [9]).

Application of the Kalman-Bucy model requires estimation of the varying regression coefficients  $\beta(t)$  from the observed data (the  $y(t)$  and  $X(t)$ ). Using an approach involving wide sense conditional probability distributions and expectations, Kalman obtained recursive updating equations as orthogonal projections on linear manifolds for the minimum mean squared linear estimators  $\hat{\beta}(t)$  of  $\beta(t)$  based on the use of all the data through time period  $t$ . The estimators  $\hat{\beta}(t)$  are frequently called filtered estimators. These equations can be written as follows:

$$(4) \quad \hat{\beta}(t) = \hat{\beta}(t|t-1) + S(t|t-1)X'(t)D^{-1}(t)(y(t) - X(t)\hat{\beta}(t|t-1))$$

where

$$(5) \quad \hat{\beta}(1|0) = T(1)\beta(0)$$

$$(6) \quad S(1|0) = V_u(1)$$

$$(7) \quad \hat{\beta}(t|t-1) = T(t)\hat{\beta}(t-1) \quad t = 2, 3, \dots, n$$

$$(8) \quad S(t|t-1) = T(t)S(t-1)T'(t) + V_u(t) \quad t = 2, 3, \dots, n$$

$$(9) \quad S(t-1) = S(t-1|t-2) - S(t-1|t-2)X'(t-1)D^{-1}(t-1)X(t-1)S(t-1|t-2) \\ t = 2, 3, \dots, n$$

$$(10) \quad D(t) = V_e(t) + X(t)S(t|t-1)X'(t) \quad t = 1, 2, \dots, n$$

and where the symbol ' denotes matrix transposition and the notation  $t|t-1$  means that the associated variable is an estimate for time  $t$  based on all the data through time  $t-1$ .

It is also shown that the differences  $\beta(t) - \hat{\beta}(t)$  have means 0 and variance-covariance matrices  $S(t)$  and are uncorrelated with the observations through time  $t$ . When the equation errors  $e(t)$  and coefficient errors  $u(t)$  have a Gaussian distribution the Kalman estimator  $\hat{\beta}(t)$  of  $\beta(t)$  will be the minimum mean squared estimator rather than merely the minimum mean squared linear estimator.

The relevance of Kalman-Bucy models for regression theory has been examined recently by Duncan and Horn [3] who presented a wide-sense random coefficient regression interpretation in which, as

a consequence of an extended Gauss-Markov theorem, the Kalman estimators  $\hat{\beta}(t)$  are identified as minimum variance linear (unconditionally) unbiased estimators of  $\beta(t)$ . Under the Gaussian assumption about the distributions of the errors  $e(t)$  and  $u(t)$  the Kalman estimator will be a minimum variance (unconditionally) unbiased estimator of  $\beta(t)$ .

Kalman-Bucy modelling has been applied in econometrics (in [7] and [8]), and the present authors have used the approach in the econometric modelling and forecasting of short-term interest rates [6]. In the latter work a Kalman-Bucy varying parameter adaptation of the reduced form of a standard, fixed coefficient econometric equation system is studied.

It must be emphasized that applying the Kalman-Bucy model to economic time series offers some interesting challenges because of the problem of specifying  $T(t)$ ,  $\beta(0)$ ,  $V_e(t)$  and  $V_u(t)$ . In engineering applications  $T(t)$  and  $\beta(0)$  are typically given by known physical properties of some real world system, so most of the attention in this literature is given to an examination of  $V_e(t)$  and  $V_u(t)$ . In economics, however, none of these parameter sets can be considered to be known, and a tractable and generally applicable method of estimation is not yet available. For example, maximum likelihood estimation of these parameter sets jointly appears to be extremely complex even in the identifiable case. Moreover, in simpler cases in which maximum

likelihood estimates are available, their determination requires simultaneous solution of systems of nonlinear equations. This in turn requires extensive numerical analysis even for small systems (see [1]).

These complexities have led us to use simulation studies to examine the influence of various specifications of the parameter sets in the single equation case. Specifically, we explore by means of simulation the question of how sensitive the Kalman-Bucy model is to the specification of the four parameter sets  $T(t)$ ,  $\beta(0)$ ,  $V_e(t)$  and  $V_u(t)$ . In other words, are the properties of the Kalman estimators  $\hat{\beta}(t)$  of  $\beta(t)$  and  $\hat{y}(t)$  of  $y(t)$  materially influenced by misspecification of one or more of these parameter sets?

We have developed a simulation model with which we can (1) generate simulated observations on  $\beta(t)$  and  $y(t)$  from a known (specified) Kalman-Bucy model, (2) obtain filtered estimates of  $\beta(t)$  and  $y(t)$  from the simulated data using parameter sets which have been both correctly specified and then misspecified, and (3) compare the estimates obtained by correctly specifying the model with those obtained by misspecifying it. In every case but one we have restricted our experiments to single-equation Kalman-Bucy models in which  $V_e(t)$ ,  $V_u(t)$ , and  $T(t)$  do not vary with time, so in order to simplify notation we will write these parameter sets as  $V_e$ ,  $V_u$ , and  $T$  whenever appropriate. Finally, we have restricted our simulation experiments to the

misspecification of one parameter set at a time over what we feel is a reasonably wide variety of underlying models.

#### Description of the Simulation Program

The computer program used to generate the results described in the next section requires the following inputs:

1. A set of  $n$  observations on  $k-1$  independent variables. Rather than randomly generate such a set of independent variables, we have used, for all of the experiments discussed here, thirty observations on three economic time series (thus,  $n = 30$ ,  $k = 4$ ). The means of  $X_1(t)$ ,  $X_2(t)$ , and  $X_3(t)$  are 4.94,  $-.0014$ , and  $-.0116$ , and the corresponding standard deviations are  $.694$ ,  $.0018$ , and  $.0138$ . The correlation between  $X_1(t)$  and  $X_2(t)$  is  $-.811$ , between  $X_1(t)$  and  $X_3(t)$  is  $.269$ , and between  $X_2(t)$  and  $X_3(t)$  is  $-.524$ .
2. The parameters  $\beta(0)$ ,  $V_e$  (note that  $V_e$  is a scalar in the single equation case,  $V_u$ , and  $T$  (to be called the underlying parameters) of the Kalman-Bucy model that are to be used to generate simulated observations on  $\beta(t)$  and  $y(t)$ ,  $t = 1, \dots, 30$ .
3. The parameters  $\beta^*(0)$ ,  $V_e^*$ ,  $V_u^*$ , and  $T^*$  (to be called the assigned parameters) that will be used to obtain filtered estimates of  $\beta(t)$  and  $y(t)$ ,  $t = 1, \dots, 30$ , from the three independent

variables and the simulated  $y(t)$  values. For a given experiment at most one of the assigned parameters will differ from the underlying parameters. In most cases, exactly one will be different. When the assigned and underlying parameters are identical, all parameters are then correctly specified and the simulation results conform to the properties of the Kalman estimators described above.

When the above information has been inputted into the program, it proceeds as follows:

1.  $\beta(t)$  and  $y(t)$  are generated from (1) and (2) for  $t = 1, \dots, 30$  using the underlying parameters. In all cases,<sup>1</sup> the errors  $e(t)$  and  $u(t)$  are generated internally as pseudorandom selections from independent normal distributions with means of 0 and variances as specified by  $V_e$  and  $V_u$ .
2. Given the resulting  $30 \times 5$  data matrix  $[y : X]$ , where the first column of  $X$  is a column of 1's, filtered estimates of  $\beta(t)$  and  $y(t)$ ,  $t = 1, \dots, 30$ , are then obtained using the assigned parameters.

---

<sup>1</sup>In the case of a nondiagonal  $V_u$ ,  $k$  independent  $N(0, 1)$  variates are selected; the resulting vector of four components  $u(t)$  is then premultiplied by an appropriate matrix  $P$ , chosen such that  $PP' = V_u$ , giving  $u^*(t)$  which is distributed as  $N(0, V_u)$ .



3. Steps 1 and 2 are repeated as many times as desired, and various statistics (described below) are accumulated. In each of the experiments reported on here, we generated 30 such replications (i. e., for a given set of input, steps 1 and 2 were repeated 30 times) and then printed a variety of summary statistics.

#### Description of Summary Statistics

The large amount of output generated by our simulation experiments compels us to condense and summarize the results and limits the summary statistics that can be discussed. Of the many statistics that were accumulated for each experiment, only two will be referred to here: the average sum of squares of error per trial, and the mean error per trial (and a related t-statistic).

For trial  $j$  of any given experiment ( $j = 1, \dots, 30$ ), we define the following:

$y_j(t)$  = observed value (simulated using the underlying parameters) of the dependent variable at time  $t$  (a scalar).

$x(t)$  =  $1 \times 4$  vector of observations on four independent variables at time  $t$  (the first element in  $x(t)$  is equal to 1 for all  $t$ ) -- note that  $x(t)$  does not vary over trials and so is the same for all  $j$ .

$\beta_j(t)$  =  $4 \times 1$  vector of observed coefficients (simulated using the underlying parameters) at time  $t$ . The elements of  $\beta_j(t)$  are denoted  $b_{ij}(t)$ ,  $i = 0, 1, 2, 3$ .

$\hat{y}_j(t)$  = the Kalman-Bucy filtered estimate of  $y_j(t)$  derived using the assigned parameters.

$\hat{\beta}_j(t)$  = the Kalman-Bucy filtered estimate of  $\beta_j(t)$  derived using the assigned parameters; the elements of  $\hat{\beta}_j(t)$  are denoted  $\hat{b}_{ij}(t)$ .

The summary statistics on which we will base much of the discussion in the next section are described below in terms of the  $y_j(t)$  and  $\hat{y}_j(t)$ . Statistics similar to those described below are developed based on  $b_{ij}(t)$  and  $\hat{b}_{ij}(t)$  in place of  $y_j(t)$  and  $\hat{y}_j(t)$  for  $i = 0, 1, 2, 3$ .

Table 1 illustrates some typical output from several experiments related to the misspecification of  $V_e$ . All three lines of statistics apply to data generated using the same underlying parameters, but each line represents estimation results for a different set of assigned parameters. For these three experiments, the data were generated with  $V_e$  equal to 1.25 and the estimates were derived by assigning  $V_e^*$  the values of 1.25 (Experiment 1), .125 (Experiment 2), and 5.00 (Experiment 3). The values of  $\beta^*(0)$ ,  $V_u^*$  and  $T^*$  were assigned correctly (i. e., equal to the underlying parameters) for all three experiments.

TABLE 1  
ESTIMATING  $y(t)$  FOR VARIOUS ASSIGNMENTS OF  $V_e^*$

Experiment	Estimation of $y(t)$	
	Avg. SSE/Trial	Mean Error/Trial (t-statistic)
1	18.342	-0.0172 (-0.568)
2	1.843	-0.0026 (-0.482)
3	39.161	-0.0004 (-0.537)

Avg. SSE/Trial. If we define the sum of squares of error for trial j to be

$$SSE_j = \sum_{t=1}^{30} (y_j(t) - \hat{y}_j(t))^2,$$

then the average sum of squares of error per trial is given by

$$\overline{SSE} = \left( \sum_{j=1}^{30} SSE_j \right) / 30$$

Mean Error/Trial. A more accurate title for this statistic would be "mean average error/trial." If we define the average error for trial j to be

$$A_j = \left( \sum_{t=1}^{30} (y_j(t) - \hat{y}_j(t)) \right) / 30$$

then the mean error per trial and its associated t-statistic (which provides a test of the hypothesis that the expected value of the mean error per trial is zero) are given by

$$M = \left( \sum_{j=1}^{30} A_j \right) / 30$$

and

$$t = M \sqrt{30} / \left( \sum_{j=1}^{30} (A_j - M)^2 / 29 \right)^{1/2}$$

### Discussion of Simulation Results

Since we restricted our experiments to the misspecification of one parameter set at a time, it will be convenient to break this

discussion of the sensitivity results into four parts, treating in turn the effects of misspecifying  $V_e$ ,  $\beta(0)$ ,  $V_u$ , and  $T$ .

#### Misspecification of $V_e$

In general, the size of the equation disturbance variance relative to the size of the variance contributed by the coefficient disturbances determines how well the  $y(t)$  are estimated by the  $\hat{y}(t)$ . In fact one can show that  $\hat{y}(t)$  will be identical to  $y(t)$  if  $V_e(t)$  is identically zero for all  $t$ . It also appears from the simulation results that the smaller is the scalar  $V_e$  for a given assignment of  $V_u$  -- whatever is the specification of  $T$  and  $\beta(0)$  -- the closer the  $\hat{y}(t)$  will be to  $y(t)$ . However, this result may induce considerable fluctuations in the estimates  $\hat{\beta}(t)$  of  $\beta(t)$  from one time period to the next. Moreover, these fluctuations may be greater still if other parameters have been misspecified also.

In all of our experiments concerning  $V_e$  (as well as those for  $V_u$  and  $T$ ),  $\beta(0)$  and  $\beta^*(0)$  were set equal to  $(.67, .85, -220., 9.92)'$  and  $V_u$  and  $V_u^*$  were of the form  $\sigma_u^2 I$ . In the first twelve experiments,  $T$  and  $T^*$  were set equal to the identity matrix and the relative magnitudes of  $V_e$  and  $V_u$  were varied. The final twelve experiments were identical to the first twelve except that  $T$  and  $T^*$  were specified to be the matrix

$$\begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 1.3 & 0 & 0 & 0 \\ 0 & 30.0 & 0 & -25.0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

The general conclusion that can be drawn from the first twelve experiments is that, regardless of the relative sizes of the variances of the equation and coefficient disturbances, misspecifying  $V_e$  has considerable effect on the estimation of  $y(t)$  and comparatively little effect on the estimation of  $\beta(t)$  (given that  $V_u$ ,  $\beta(0)$ , and  $T$  are correctly specified). To illustrate, Table 2 contains partial results of Experiments 1, 2, and 3, in which the data were generated with  $V_e = .125$  and  $V_u = (.025)I$ , and the estimates were derived by assigning  $V_e^*$  the values .125, .0125, and 1.25. The results for  $b_2(t)$  and  $b_3(t)$  are similar and have been omitted. The differences in the Avg. SSE/Trial and Mean Error/Trial figures for estimating the  $y(t)$ 's are dramatic, particularly in light of the comparatively stable analogous figures for estimating the coefficients. None of the t-statistics associated with the above mean error figures had an absolute value larger than 0.6, so those figures were deleted from Table 2.

Experiments 13 through 24 were identical to the first twelve except that  $T$  and  $T^*$  were not the identity matrix. The overall results for these experiments were similar to those described for the first twelve experiments, although using a non-identity  $T$  matrix greatly

TABLE 2

SELECTED ESTIMATION RESULTS FOR VARIOUS ASSIGNMENTS OF  $V_e^*$ 

Experiment	Estimation of $y(t)$			Estimation of $b_0(t)$			Estimation of $b_1(t)$		
	Avg. SSE	Mean Error	Avg. SSE	Mean Error	Avg. SSE	Mean Error	Avg. SSE	Mean Error	
1 ( $V_e = .125$ )	0.5320	-.0020	10.868	.0346	0.5776	-.0080			
2 ( $V_e = .0125$ )	0.0097	-.0002	11.018	.0362	0.5931	-.0080			
3 ( $V_e = 1.25$ )	7.2258	-.0152	10.693	.0303	0.7842	-.0100			

reduced the average sum of squares of error per trial and the mean error per trial for estimating  $b_0(t)$  and  $b_1(t)$  as compared to the results presented in Table 2 above. The similar values for estimating  $y(t)$  changed very little from those in the first twelve experiments.

In none of the experiments concerning  $V_e$  did any of the misspecifications of  $V_e$  appear to bias the estimators when the other parameter sets were assigned correctly.

#### Misspecification of $\beta(0)$

We ran 56 experiments representing a variety of underlying models to examine the effect of incorrectly specifying  $\beta(0)$ . In general, we found the Kalman-Bucy model to be very sensitive to the choice of the starting value  $\beta(0)$ . In some of the experiments, each element of  $\beta(0)$  was misspecified with an error of approximately 10 percent of the true value. In other experiments, the error was closer to 50 percent. In still other experiments, only one element of  $\beta(0)$  was misspecified by varying amounts. The relative sizes of  $V_e$  and  $V_u$  were varied over a number of the experiments and some experiments used a T matrix which was not the identity matrix.

The most significant results can be better understood by rewriting (4) in the form

$$(11) \quad \hat{\beta}(t) = \hat{\beta}(t|t-1) + K(t)v(t) \quad (t = 1, \dots, n)$$

where  $K(t) = S(t|t-1)X'(t)D^{-1}(t)$  is the so-called Kalman gain and  $v(t)$  is a scalar equal to  $y(t) - x(t)\hat{\beta}(t|t-1)$ . If  $T$  is the identity matrix, (11) can be rewritten as

$$(12) \quad \hat{\beta}(t) = \hat{\beta}(t-1) + K(t)v(t) \quad (t = 1, \dots, n).$$

Obviously, then, particularly in the  $T = I$  case, the amount of correction that  $\hat{\beta}(t)$  will realize from one time period to the next depends on the elements  $k_i(t)$  of  $K(t)$ . The magnitude of  $k_i(t)$  is considerably affected by the magnitude of  $x_i(t)$  and the relative sizes of  $V_e$  and  $V_u$ .

For example, in many of our experiments concerning  $\beta(0)$ , we specified  $V_u$  and  $V_u^*$  to be of the form  $\sigma_u^2 I$ . The mean of  $x_1(t)$ ,  $t = 1, \dots, 30$ , is approximately 5.0, but the mean of  $x_3(t)$ ,  $t = 1, \dots, 30$ , is approximately  $10^{-3}$ . Consequently, if  $T = I$  and  $V_u = \sigma_u^2 I$ ,  $b_1(t)$  is much more likely to correct toward its true value if misspecified than is  $b_3(t)$ .

In general, then, if  $T = I$  and  $V_u$  is diagonal, the effect of misspecifying  $\beta(0)$  on the estimation of  $\beta(t)$  and  $y(t)$  depends on the magnitude of the specification error of  $b_i(0)$  relative to the magnitude of  $k_i(t)$ . In addition, misspecifying only one of the elements of  $\beta(0)$  may introduce bias in the estimation of the other coefficients that were correctly specified. If  $T$  is not a diagonal matrix (i. e., if there are interrelationships between the coefficients), it is easy to see that



incorrectly specifying one coefficient may lead to biased results for several coefficients. This was observed in several of our experiments.

For the most part, misspecifying  $\beta(0)$ , in whatever way, had very little effect on the estimation of  $y(t)$  unless the variances of the coefficient disturbances were very small compared to the variance of the equation disturbance. In this case none of the coefficients was able to adjust itself if misspecified. For example, in one pair of experiments, the underlying parameters included  $T = I$ ,  $V_e = .125I$ , and  $V_u = .025I$ . The average sum of squares of error for  $y(t)$  was 0.532 with  $\beta^*(0) = \beta(0)$  and 0.771 with  $\beta^*(0)$  approximately equal to  $.5\beta(0)$ . Furthermore, the  $t$ -statistics associated with the mean error per trial were -0.38 in the correctly specified case and -3.33 in the misspecified case. In contrast to those results, another pair of experiments was run in which the conditions were identical except that  $V_u$  was set equal to  $(.00025)I$ . The average sum of squares of error for  $y(t)$  was 3.03 for the correctly specified case and 24.35 in the misspecified case. The corresponding  $t$ -statistics for the mean errors per trial were -0.43 and 43.35.

#### Misspecification of $V_u$

It seems reasonable to believe that in a typical application,  $V_u$  will be assumed to be a diagonal matrix. Thus, we limited these

experiments to cases in which  $V_u^*$  was specified to be a diagonal matrix whether the underlying  $V_u$  was diagonal or not.

In the first nine  $V_u$  experiments, we set  $T = I$  and  $V_u = \sigma_u^2 I$  and looked at the effects of misspecifying  $\sigma_u^2$  for various relative sizes of  $\sigma_u^2$  and  $V_e$ . In all cases (regardless of the size of  $\sigma_u^2$  relative to  $V_e$ , at least in our experiments), misspecifying  $\sigma_u^2$  either on the high or low side had virtually no effect on the estimation of  $\beta(t)$ . On the other hand, underestimating  $\sigma_u^2$  effectively increases the relative size of  $V_e$  and the model then fits  $y(t)$  less well. Conversely, overestimating  $\sigma_u^2$  reduces the relative size of  $V_e$  and the model fits  $y(t)$  more closely than if  $\sigma_u^2$  is specified correctly.

The next nine experiments were identical to the first nine except that instead of setting  $T$  equal to  $I$ , we specified

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 30 & -25 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This transition matrix, if specified correctly as it was here, provides closer estimates of  $y(t)$  than were realized by correctly specifying an underlying identity  $T$  matrix. Nevertheless, misspecifying  $\sigma_u^2$  again had little noticeable effect on the estimation of  $\beta(t)$ , so the results were essentially the same as those observed for the first nine experiments.

The last eight experiments were designed to test the effect of specifying a diagonal  $V_u$  with correct diagonal elements when the underlying  $V_u$  had numerous off-diagonal elements. In other words, what happens if there is correlation among the coefficient disturbances but we fail to recognize that correlation when estimating  $y(t)$  and  $\beta(t)$ ?

Our results, all derived with  $T = I$ , indicate that as the correlations among the coefficient disturbances grow larger, so does the error of estimation if those correlations are ignored, particularly with respect to estimating  $\beta(t)$ . For example, in one pair of experiments we specified  $V_u$  to be a matrix with each diagonal element equal to .025 and each off-diagonal element equal to .005. This gave a correlation of .20 for each pair  $(u_i(t), u_j(t))$ ,  $i \neq j$ . In the first of those two experiments, we assigned  $V_u^*$  correctly, and in the second we set  $V_u^* = (.025)I$ . In another pair of experiments, the same approach was taken, but in this case the off-diagonal elements of  $V_u$  were set equal to .20. This gave a correlation of .80. Table 3 presents the results for estimating  $y(t)$  and  $b_0(t)$  only; note how much the results for estimating  $b_0(t)$  were affected, particularly when the larger correlation was ignored in the estimation (the  $t$ -statistics associated with the mean error per trial were virtually the same whether the correlation was correctly specified or not, so they are not included in Table 3).

TABLE 3

SELECTED RESULTS FOR VARIOUS SPECIFICATIONS OF  $V_u^*$

Experiment	Estimation of $y(t)$		Estimation of $b_0(t)$	
	Avg. SSE	Mean Error	Avg. SSE	Mean Error
1 ( $r=.20, V_u^*=V_u$ )	0.339	0.00367	1.071	0.0625
2 ( $r=.20, V_u^*\neq V_u$ )	0.379	0.00394	1.877	0.0830
3 ( $r=.80, V_u^*=V_u$ )	0.423	0.00449	0.185	0.0204
4 ( $r=.80, V_u^*\neq V_u$ )	0.642	0.00578	4.010	0.1234

Misspecification of T.

Our experiments with the transition matrix involved four matrices,  $T_1, T_2, T_3,$  and  $T_4$ .  $T_1$  was chosen to be the identity matrix and the remaining three matrices were chosen so that each matrix represented more complex interrelationships among the coefficients than the preceding one. Thus,  $T_1$  was the least structured transition matrix and  $T_4$  was the "most structured," forcing the largest number of interrelationships among the coefficients.  $T_2, T_3,$  and  $T_4$  all were chosen in such a way that the elements of  $\beta(0)$  (which were the same for all of these T experiments) conformed to the relationships imposed by those matrices.

We conducted 16 experiments to examine the effects of misspecifying T. In the first four of these experiments, the underlying