Journal of Accounting Research
Vol. 21 No. 1 Spring 1983
Printed in U.S.A.

Identification and Estimation Issues for a Causal Earnings Model

BALA G. DHARAN*

1. Introduction

In this paper, I postulate a causal, or structural, model of corporate earnings and present a theoretical analysis of various empirical issues related to the identification and estimation of time-series earnings models.[1] The postulated model describing corporate earnings is derived endogenously by specifying a model of the firm's production and investment structure and its inventory accounting rule. The main characteristic of this model is that the firm uses linear stochastic decision rules to determine its production, inventory, and capital investment levels. Such decision rules are frequently postulated in the economics literature. Using the theoretical properties of the derived earnings model, I then address estimation and forecasting issues in a general manner without actually doing the estimation.

Starting with Dopuch and Watts [1972], much of the empirical research in accounting on the structural (i.e., time-series) properties of earnings numbers has first postulated a linear, stochastic, time-series model as the underlying earnings-generating process, and then identified and estimated the model using a sample of realized earnings numbers. These accounting studies usually have the goal of estimating an expectation model to generate "expected" or forecasted earnings. This type of end

* Assistant Professor, Rice University. This paper is based in part on my Ph.D. dissertation from the Graduate School of Industrial Administration, Carnegie-Mellon University. I gratefully acknowledge the help from the committee members Robert S. Kaplan (chairman), Dan Givoly, Yuji Ijiri, and Thomas Morton. I am also grateful to the reviewer for providing many substantive and editorial comments. Remaining errors are mine. [Accepted for publication September 1982.]

[1] The model is "causal" in that the factors determining, or causing fluctuations in earnings are specified by the model.

use makes the correct identification of the earnings model and the estimation of the model coefficients relevant empirical issues for the model designer, since any misspecification will increase the forecast error variance.

Estimation problems generally depend on the characteristics of the firm whose earnings are being analyzed and the number of observations available. However, researchers usually know very little about the effects of the firm's characteristics on model estimation because the model of the firm is not explicitly specified. Instead, they assume that earnings are described by a linear, stochastic, univariate model, which in turn is implicitly assumed to be the "reduced" form of the unspecified structural model of the firm. Thus, the coefficients of the reduced form model are unknown functions of the basic model parameters which determine the firm's characteristics. As a result, such studies cannot rely on a priori information on the effects of model parameters on coefficient estimation.

In the causal earnings model of this paper, the firm's decision characteristics are explicitly stated and therefore the coefficients of the earnings model are known functions of the firm's parameters. Moreover, the covariance matrix of the coefficients of the earnings model is also a known function of the parameters and the sample size. Hence, by specifying a set of values for the model parameters and by assuming a sample size, it is possible to compute the ratio of each coefficient to its expected standard error. Based on these likelihood ratios and other related statistics, it is then possible to predict the model that may be actually identified by an empiricist.

The analysis presented here indicates that the derived earnings model may be significantly more complex than the models that would be identified and estimated by a time-series analysis of the observed earnings numbers in sample sizes commonly encountered in accounting research (e.g., 100 observations). In other words, the estimated models will be more parsimonious than the derived model.[2] The effect of this misspecification is to increase the mean-square forecast error by a median value of about 12 percent.

Earnings model estimation issues discussed in the accounting literature are examined here in the light of these results. For example, it is likely that an empiricist comparing the systematic properties of two accounting earnings series will probably be unable to test for differences in the order of the identified models, since limited sample sizes would restrict the analysis to very parsimonious models in each case. Instead, the empiricist is more likely to find differences in coefficient values. Similarly, it is the observed variation in the quarterly earnings model structures identified by various accounting researchers (e.g., Foster [1977], Griffin [1977], etc.)

[2] Time-series models are "simple" or "parsimonious" if they have only one or two autoregressive and moving average terms. "Complex" models are those that are not parsimonious in this sense.

that is likely due to the specification errors caused by the parsimony effect of sample size.

In the next section of this paper, the firm and its decision rules for production and capacity are described, and the earnings model is derived by adding an accounting rule for inventory valuation. Section 3 analyzes the behavior of the coefficients of this model over a wide range of the firm's parameter values. The estimation problem is discussed in section 4 after the covariance matrix of the coefficients is computed. Section 5 examines the effect of misspecification on forecasting. The last section provides some general implications for accounting time-series research.

2. *The Firm and the Earnings Model*

The firm described here is based on Gonedes and Dopuch [1979] in that the firm's stochastic earnings are the product of the joint action of the accounting system and economic events. The accounting system is represented by the firm's inventory accounting rule. The economic environment is characterized by stochastic sales quantities and linear production-investment decisions.[3]

The firm decides on a level of production in order to achieve a desired level of ending inventory, and then decides on a capacity that will accomplish this production. The actual production and capacity, however, are determined by adjustment speeds that are less than one. The stochastic model for the sales quantity is assumed to be a stationary or "mean-reverting" process.[4] This assumption models the unit sales S_t for period t as:

$$S_t = S_0 + e_t, \tag{1}$$

where S_0 is the process mean and e_t is the realization from a white noise process with $E(e_t) = 0$, $E(e_t e_s) = \sigma^2(e_t)$ for $t = s$, and $E(e_t e_s) = 0$ for $t \neq s$. Note that expected sales in period t are constant for all t.

Following Metzler [1941], let the desired ending inventory in period t, I_t^D, be a linear function of expected sales, $\hat{S}_{t+1}$, such that:

$$I_t^D = c\hat{S}_{t+1} = cS_0, \tag{2}$$

where c is a type of stock-sales ratio, $0 < c < 1$. The 0–1 range for c is reasonable in most economies when a period is defined as a quarter or a year.

The production quantity, q_t, is chosen to maintain inventories at the desired level. Let:

$$q_t = \hat{S}_{t+1} + b(I_t^D - I_{t-1}), \tag{3}$$

[3] The main elements of the production-investment model are based on studies on macroeconomic modeling of the U.S. economy. References to this literature are given later in this section.

[4] This assumption facilitates the derivation of model equations. For other sales processes, one may often have to use simulation to obtain model equations. See, however, Dharan [1981] for simple extensions to a random-walk or seasonal sales process.

where I_{t-1} is the beginning inventory in period t, and b is the inventory stock adjustment speed, $0 < b < 1$. Equation (3) is based on the flexible stock adjustment (acceleration) model of Goodwin [1948], also used in studies by Mantell [1977] and Trivedi [1973].

The inventory balancing equation is:

$$I_t = I_{t-1} + q_t - S_t. \tag{4}$$

Using (1) through (4), the decision rules or equations for production and inventory are obtained. Let $b_1 = 1 - b$.[5] Then the decision rules are:

$$I_t = b_1 I_{t-1} + bcS_0 - e_t, \tag{5}$$

and:

$$q_t = b_1 q_{t-1} + bS_0 + be_{t-1}. \tag{6}$$

The desired capital stock in period t, K_t^D is proportional to q_t. This assumption is characteristic of acceleration models as well as the neoclassical model of Hall and Jorgenson [1967].[6] We will let $K_t^D = q_t$. The actual physical capital stock of period $t - 1$, K_{t-1} units, depreciates such that $(1 - d)K_{t-1}$ is the capacity in units available in period t in the absence of new investment, where d is the physical depreciation rate, $0 < d < 1$. Let $d_1 = 1 - d$. The period t capacity, K_t units, is given by:

$$K_t = d_1 K_{t-1} + a(K_t^D - d_1 K_{t-1}), \tag{7}$$

where $a(K_t^D - d_1 K_{t-1})$ is the new investment in period t and a is the capital stock adjustment speed, $0 < a < 1$. The 0–1 range for a is acceptable when one assumes that instantaneous capacity adjustment is infeasible due to physical limits. Let $a_1 = 1 - a$. With mean-reverting sales, and using (6), (7) gives the capacity decision rule:

$$K_t = (b_1 + a_1 d_1)K_{t-1} - b_1 a_1 d_1 K_{t-2} + abS_0 + abe_{t-1}. \tag{8}$$

The firm is assumed to incur production costs in period t, given by q_t^2/S_0 dollars. This implies increasing marginal cost of production.[7] In addition to this out-of-pocket cost, the "cost of goods manufactured" used in the computation of accounting earnings is assumed to include a depreciation expense. With appropriate normalization of the unit cost of capacity, the depreciation expense will be K_t dollars.[8] The cost of goods sold (CGS) in period t is then determined by how the inventory account-

[5] b_1 is a subscripted constant; so are a_1, d_1, and c_1 (used later).

[6] See Visscher [1978].

[7] The marginal out-of-pocket production cost is thus independent of capacity. Note that the fractional adjustment speed, a, can also imply that there are quadratic capacity adjustment costs. These costs are not considered in this model.

[8] A new investment in period t, say X_t, can produce a total of X_t/d units in periods t, $t + 1, \ldots$. If the unit cost of capacity is normalized as $1/d$ dollars, then the investment X_t/d dollars is to be depreciated over the X_t/d total units. Under the units-of-production method, period t depreciation is X_t dollars. By extension, capacity of K_t units leads to a depreciation expense of K_t dollars.

ing rule allocates these two costs to various periods. Pretax earnings then are given by the difference between the sales revenue and the *CGS*.

I assume that the weighted average cost method of inventory valuation is used to compute the *CGS*.[9] If i_t is the weighted average cost per unit of goods available for sale in period t, this method defines i_t as:

$$i_t = (i_{t-1}I_{t-1} + q_t^2/S_0 + K_t)/(I_{t-1} + q_t), \tag{9}$$

where q_t^2/S_0 is the prodution cost and K_t is the depreciation expense. Then, if S_t units are sold, the *CGS* is i_tS_t. Assuming a unit selling price, sales revenue is S_t dollars, and accounting earnings are $Z_t = (1 - i_t)S_t$.

Appendix A summarizes the derivation of the time-series expression for Z_t. Using the widely recognized notational conventions of Box and Jenkins [1970], the resulting autoregressive and moving average (*ARMA*) equation for the mean-adjusted earnings series is written as:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \tilde{e}_t - \theta_1\tilde{e}_{t-1} - \theta_2\tilde{e}_{t-2} - \theta_3\tilde{e}_{t-3}, \tag{10}$$

where the random shock $\tilde{e}_t$ is a function of the variable e_t from the sales process.[10] This is an *ARMA* (3, 3) model, where the numbers (3, 3) denote the number of lags of earnings and the random shock needed to describe current earnings. The coefficients of this model are derived in Appendix A in terms of the firm's parameters.

To summarize, the firm is characterized by four parameters:

a: capacity stock adjustment speed,	$0 < a < 1$,
b: inventory stock adjustment speed,	$0 < b < 1$,
c: desired stock to expected sales ratio,	$0 < c < 1$,
d: capacity depreciation rate,	$0 < d < 1$.

The firm's decision rules for production, inventory, and capacity are given by equations (5), (6), and (8), and its earnings are given by (10).

When considered independently, the production-inventory decision rules and the capacity decision rule can be shown to be optimal under certain assumed cost structures.[11] However, they are not necessarily jointly optimal. Nevertheless, it can be shown that jointly optimal decision rules from a model with rational expectation equilibrium will resemble the decision rules (5), (6), and (8) under certain conditions.[12] It is thus reasonable to assume that these rules, and hence the model in (10), approximate the optimality property.

3. *Properties of the Coefficients*

The values of the coefficients of (10) are determined by the values of the four model parameters. For the "0 to 1" parameter values, and using

[9] See Dharan [1981] for extensions to *FIFO* and *LIFO* methods.

[10] See expression (A.20) in Appendix A.

[11] See, for example, Hall and Jorgenson [1967] for the capacity investment rule and Holt and Modigliani [1961] for the production-inventory rules.

[12] Such a model was derived in an earlier version of this paper and was based on the

(A.11) to (A.13) and (A.22) to (A.24) from Appendix A, it is seen that the coefficients alternate in sign as follows:

$$\phi_1 > 0, \phi_2 < 0, \quad \text{and} \quad \phi_3 > 0, \quad \text{and} \tag{11a}$$

$$\theta_1 > 0, \theta_2 < 0, \quad \text{and} \quad \theta_3 > 0. \tag{11b}$$

Moreover, when the decision parameters a and b are greater than 0.5, we get:

$$\phi_1 > -\phi_2 > \phi_3, \quad \text{and} \tag{12a}$$

$$\theta_1 > -\theta_2 > \theta_3. \tag{12b}$$

Properties (12a) and (12b) seem reasonable for a time-dependent process, since the weights assigned to lagged values of Z and $\tilde{e}$ in (10) are smaller for larger lags. One can show that the following properties are also true:

$$\theta_1 > \phi_1, -\theta_2 > -\phi_2, \quad \text{and} \quad \theta_3 > \phi_3. \tag{12c}$$

The ability to estimate the model coefficients of (10) depends on (i) their relative magnitudes as given by (12), and (ii) on the effect of each model parameter on the magnitudes of the coefficients. This latter information can be obtained by studying the signs of the partial derivatives of each of the six coefficients with respect to each of the decision parameters.

The partial derivatives of the autoregressive coefficients with respect to each of the four model parameters are simple functions of the four parameters. From the signs of the partial derivatives, the effect of the model parameters on the autoregressive coefficients is summarized as follows:

The magnitude of all three autoregressive coefficients always increases if any of a_1, b_1, d_1, and c increases. (13a)

The signs of the partial derivatives of the moving average coefficients are not as easily summarized. The signs often depend on whether the parameters are closer to one or zero. Hence, the effect of the model parameters on the moving average coefficients is summarized in this weaker statement:

When parameter values are greater than 0.5, the magnitude of all three moving average coefficients usually increases if any of a_1, b_1, d_1, and c increases. (13b)

To summarize, the magnitudes of the autoregressive coefficients and, often, the moving average coefficients of the earnings model will increase, if (i) the stock adjustment speeds a and b are reduced (the firm reacts more slowly to catch up to the desired capital or inventory levels); (ii)

equilibrium models of Lucas and Prescott [1971], Lucas [1976], Sargent [1979], Eichenbaum [1980], and Hansen and Sargent [1980].

the depreciation rate d is reduced (the assets are productive for longer periods); (*iii*) the desired stock-to-expected-sales ratio c is increased (the desired inventory level for a given level of anticipated activity is larger).

This summary agrees with one's intuitive notion of time dependence induced by economic actions, since an increase in a_1, b_1, d_1, and c would generally mean that less of the current period shocks in sales and production cost is absorbed in current earnings and more is passed on to future earnings. When current earnings are affected relatively more by the economic shocks of past periods, the magnitudes of the coefficients of the lag terms in the earnings model tend to increase.

The analysis thus far has been based on the "0 to 1" constraints on the model parameters noted in section 2. The requirement that the earnings model (10) be both stationary and invertible may impose additional constraints on the parameters, and these are identified next.[13]

A (second-order) stationary earnings series is one whose mean and variance are time-invariant and finite. For an *ARMA* (3, .) process, the requirement that the variance be finite results in the constraints that the three characteristic roots, G_i, $i = 1, \ldots 3$, of the characteristic equation:

$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2 - \phi_3 x^3 = (1 - G_1 x)(1 - G_2 x)(1 - G_3 x) = 0$$

lie inside the unit circle.[14] In other words, the stationarity condition is that $|G_i| < 1$, where the G_i can be real or complex.

The invertibility condition is required, in practical terms, because of the need to compute the residuals $\tilde{e}_t$, $\tilde{e}_{t-1}$, ... of (10) during model estimation. For a given set of coefficients, and for a given sample of earnings data, recursive computation of $\tilde{e}_t$ in (10) is possible only if it can be expressed as a convergent series in terms of present and past earnings. Thus, the invertibility condition is implicitly assumed by all estimation routines which rely on minimizing the sum of squares of the computed residuals. For an *ARMA* (., 3) model, the invertibility requirement results in the constraint that the three roots, H_i, $i = 1, \ldots 3$, of the characteristic equation $\theta(x) = 1 - \theta_1 x - \theta_2 x^2 - \theta_3 x^3 = (1 - H_1 x)(1 - H_2 x)(1 - H_3 x) = 0$ lie inside the unit circle. Thus, the invertibility condition requires that $|H_i| < 1$, where the H_i can be either real or complex.

These constraints on the characteristic roots G and H imply some corresponding constraints on the coefficients of (10) and hence on the firm's decision parameters. Usually these are not easy to formulate except for parsimonious models such as *AR* (2) or *ARMA* (1, 1). Fortunately, the stationarity conditions can be derived explicitly for the average cost earnings model.

Examining the autoregressive coefficients in Appendix A, and using the constant c_1 from that appendix, the real roots G_i, $i = 1, \ldots 3$, are given by

[13] The stationarity and invertibility conditions are explained in Box and Jenkins [1970], and the definitions given here will be brief.

[14] In this equation, and in the subsequent characteristic equation for moving average terms, x may be any variable.

$G_1 = c_1$, $G_2 = b_1$, and $G_3 = a_1 d_1$. Hence the restriction $|G_i| < 1$ is satisfied for the average cost earnings model, given the "0 to 1" restrictions on the parameter values already noted. Thus, the average cost earnings series is stationary. Since there is no nonstationarity in the sales-generating process (which is mean reverting), this result confirms that the accounting process used to compute the earnings did not introduce any nonstationarity.

Satisfaction of the invertibility conditions is more difficult to verify. The characteristic equation $\theta(x) = 0$ does not have easily identifiable roots for the moving average coefficients given by (A.22) to (A.24) in Appendix A. The solution for the three roots in terms of the three moving average coefficients are complicated. Furthermore, as seen from the expressions (A.14) to (A.16) and (A.22) to (A.24), the expressions for the three roots in terms of the model parameters are also complicated. Moreover, two of the three roots could be complex conjugates.

Under these circumstances, numerical computation of the characteristic roots for various combinations of model parameter values provides one way of verifying whether the invertibility conditions hold. For various combinations of values of the model parameters ranging from 0.3 to 0.9 with increments of 0.2, the moving average coefficients were computed and the three corresponding characteristic roots were found. Of the 256 cases examined, the invertibility condition $|H_i| < 1$ was valid for 152 cases, and for a vast majority of them, the values of a, b, $(1 - c)$, and d were nearer to one than to zero.

Table 1 lists both cases where invertibility is satisfied and cases where it does not hold. From (13), the magnitude of the six model coefficients can be increased by increasing a_1, b_1, c, or d_1. It turns out, however, that these cases of large coefficient magnitudes also often represent noninvertible situations, and hence are not estimable.

Thus, the invertibility condition restricts the allowable parameter values to ranges narrower than the "0 to 1" noted earlier. Specifically, the values of a, b, $(1 - c)$, and d should preferably be greater than 0.5. The assumption that these values are > 0.5 does not guarantee, of course, that the model is invertible (see, for example, cases 5 and 7 in table 1). However, as noted above, the assumption that the model is invertible implies that the parameter values are large. The following sections assume that the earnings model is invertible and that the earnings model coefficients are relatively small.

4. *The Estimation Problem*

Given a sample size, reliable estimation of the coefficients of the earnings model will depend on their theoretical standard errors. If the ratio of a coefficient estimate to its theoretical standard error (the likelihood ratio) is very small, then it is usually assumed in time-series research that the estimate is not significantly different from zero. Such an assumption is, of course, inappropriate when the estimate being tested

TABLE 1

Average Cost Earnings: Some Invertible and Noninvertible Processes

No.	a	b	c	d	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	G_1	G_2	G_3	H_1	H_2	H_3
1*	.3	.3	.3	.3	1.4208	−.6176	.0792	5.0326	−6.2205	2.0702	.2308	.7000	.4900	3.3683	1.1112	.5531
2*	.3	.5	.7	.5	1.2618	−.5250	.0726	2.6009	−2.3818	.5898	.4118	.5000	.3500	.3865	1.1072 ±	.5478 i
3	.5	.9	.3	.9	.3808	−.0396	.0012	.5812	− .2521	.0133	.2308	.1000	.0500	.0601	.2606 ±	.3911 i
4	.5	.9	.7	.9	.5618	−.0668	.0021	.7122	− .2263	.0112	.4118	.1000	.0500	.0595	.3264 ±	.2845 i
5*	.7	.3	.5	.5	1.1833	−.3883	.0350	2.8870	−2.4882	.4311	.3333	.7000	.1500	.2295	1.3288 ±	.3355 i
6	.7	.7	.7	.7	.8018	−.1876	.0111	1.1553	− .5869	.0569	.4118	.3000	.0900	.1238	.5157 ±	.4395 i
7*	.9	.3	.5	.5	1.0833	−.2850	.0117	2.4274	−1.7516	.1342	.3333	.7000	.0500	.0867	1.1703 ±	.4231 i
8	.9	.9	.5	.3	.5033	−.0637	.0023	.6275	− .1980	.0125	.3333	.1000	.0700	.0816	.2729 ±	.2810 i
9	.9	.9	.5	.9	.4433	−.0377	.0003	.5293	− .1251	.0018	.3333	.1000	.0100	.0157	.2568 ±	.2261 i
10	.9	.9	.9	.9	.5837	−.0531	.0005	.6505	− .1211	.0016	.4737	.1000	.0100	.0147	.3179 ±	.1035 i

* Noninvertible processes

Note: H_2 and H_3 are complex conjugates in all cases except case 1. For these cases, the notation for H_2 and H_3 is that one of them is $x + yi$ and the other is $x - yi$, where x is given under the heading H_2 and y is given under the heading H_3.

is highly correlated with other estimates. In such cases, even a low coefficient estimate need not indicate insignificance. Conventional time-series research, however, has used the univariate likelihood ratio tests to evaluate coefficient estimates even in such cases, and hence the discussion here will start with this approach to examine the problem of estimating the earnings model.

When the shock e_t in (1) is assumed to be normally distributed, the likelihood ratio is the time-series analogue of the univariate t-ratio of linear regression. It will be assumed that the model can be "reasonably estimated" if each of the model coefficients is significant as measured by the t-ratio.[15] In any case, the model will be assumed estimable if all coefficients, taken together, are significant as measured by an F-test.

For this analysis, the variance-covariance matrix (or simply, the covariance matrix) of the coefficients must be known. Here, I use the fact that the covariance matrix of the coefficients of (10) is uniquely given once the values of the coefficients and the sample size are specified. For compact description of the procedure to compute the covariance matrix, define the (6×1) vector of coefficients $\beta = (\phi_1, \phi_2, \phi_3, \theta_1, \theta_2, \theta_3)^T$, where T denotes transpose. Then, from Box and Jenkins [1970], the information matrix of (10), $\mathbf{I}_l$, can be written as:

$$\mathbf{I}_l = n \cdot \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \cdot \sigma^{-2}(\tilde{e}), \tag{14}$$

where n is the sample size, and A, B, and D are matrices of the order (3×3). The elements of $\mathbf{I}_l$ are the negative of the second-order partial derivatives of the log-likelihood function of (10) with respect to the coefficients, and are defined below. First, let $\mathbf{V}(\beta)$ be the covariance matrix of β. Box and Jenkins [1970] note that, for moderate to large sample sizes (i.e., $n > 50$), $\mathbf{V}(\beta)$ is obtained by:

$$\mathbf{V}(\beta) = \mathbf{I}_l^{-1} = \frac{\sigma^2(\tilde{e})}{n} \begin{bmatrix} A & B \\ B^T & D \end{bmatrix}^{-1}. \tag{15}$$

The matrix $\mathbf{A}$ in (14) is given by:

$$\mathbf{A} = \begin{bmatrix} V_{uu}(0) & V_{uu}(1) & V_{uu}(2) \\ V_{uu}(1) & V_{uu}(0) & V_{uu}(1) \\ V_{uu}(2) & V_{uu}(1) & V_{uu}(0) \end{bmatrix} \tag{16}$$

where $V_{uu}(0)$ is the variance and $V_{uu}(k)$ is the autocovariance at lag k of a variable u_t described by the process:

$$u_t - \phi_1 u_{t-1} - \phi_2 u_{t-2} - \phi_3 u_{t-3} = \tilde{e}_t. \tag{17}$$

The variable u_t is interpreted by Box and Jenkins [1970, eq. 7.2.9] as the negative of the partial derivative of residuals of (10) with respect to the

[15] In practice, significance can be assumed if the likelihood ratio of a coefficient is greater than two. It is assumed that the model satisfies other diagnostic tests in addition to having significant estimates.

autoregressive coefficients.[16] The matrix $\mathbf{D}$ is given by:

$$\mathbf{D} = \begin{bmatrix} V_{xx}(0) & V_{xx}(1) & V_{xx}(2) \\ V_{xx}(1) & V_{xx}(0) & V_{xx}(1) \\ V_{xx}(2) & V_{xx}(1) & V_{xx}(0) \end{bmatrix} \tag{18}$$

where $V_{xx}(0)$ is the variance and $V_{xx}(k)$ is the autocovariance at lag k of a variable x_t described by the process:

$$x_t - \theta_1 x_{t-1} - \theta_2 x_{t-2} - \theta_3 x_{t-3} = -\tilde{e}_t. \tag{19}$$

The variable x_t is interpreted by Box and Jenkins [1970, eq. 7.2.10] as the negative of the partial derivative of residuals of (10) with respect to moving average coefficients.[17] Finally, the matrix $\mathbf{B}$ is given by cross-covariances between u and x:

$$\mathbf{B} = \begin{bmatrix} V_{ux}(0) & V_{ux}(-1) & V_{ux}(-2) \\ V_{ux}(1) & V_{ux}(0) & V_{ux}(-1) \\ V_{ux}(2) & V_{ux}(1) & V_{ux}(0) \end{bmatrix} \tag{20}$$

where $V_{ux}(k)$ is the covariance between u_t and x_{t+k}.

For time-series processes such as (10), (17), and (19), the variance and autocovariances are all scaled by the residual variance. Hence the elements of $\mathbf{A}$ and $\mathbf{D}$ are scaled by $\sigma^2(\tilde{e})$, and so are the elements of $\mathbf{B}$. It then follows from (14) that the elements of $\mathbf{I}_l$ are independent of $\sigma^2(\tilde{e})$. $\mathbf{I}_l$ and $\mathbf{V}(\boldsymbol{\beta})$ are thus uniquely determined given β and the sample size n. Hence, if theoretical values of the coefficients are known, one can compute the expected standard errors of the coefficients for an assumed sample size and compute the likelihood ratios.

Table 1 gave a sampling of the over 250 parameter combinations of the earnings model that were examined for invertibility. For each of the cases where the model was invertible, the covariance matrix was computed using the procedures described in Dharan [1982*a*] assuming a sample size of 100 observations. Table 2 gives the computed standard errors for case 4 of table 1. It is seen from the top panel of table 2 that all the six coefficients of the *ARMA*(3, 3) model in (10) have theoretical standard errors many times greater than their expected values when a sample size of 100 observations is assumed. A standard univariate t-test cannot reject the hypothesis that the estimates are zero. Based on the univariate t-tests, it seems that in order to get significant estimates of all six model coefficients, one would probably need sample sizes in excess of one million.

A closer examination of the data in table 2 provides some indication that the above result from the univariate t-tests may be understating the significance of at least some of the coefficient estimates. For example, the

[16] $u_{t-j} = -\partial\hat{e}_t/\partial\phi_j$, evaluated at a given β, where $\hat{e}_t$ is the computed residual given β and Z of (10).

[17] $x_{t-j} = \partial\hat{e}_t/\partial\theta_j$.

TABLE 2

Average Cost Earnings Series: Case 4 of Table 1

Model parameters: $a = 0.5$ $b = 0.9$ $c = 0.7$ $d = 0.9$

ARMA(3,3) Model:

Autoregressive coefficients: $\phi_1 = .56176$ $\phi_2 = -.06676$ $\phi_3 = .00206$

Moving average coefficients: $\theta_1 = .71217$ $\theta_2 = -.22626$ $\theta_3 = .01115$

Information matrix of the *ARMA*(3,3) coefficients:

$$n \begin{bmatrix} .13906E{+}01 & .73293E{+}00 & .32040E{+}00 & -.14738E{+}01 & -.76988E{+}00 & -.33590E{+}00 \\ .73293E{+}00 & .13906E{+}01 & .73293E{+}00 & -.87918E{+}00 & -.14738E{+}01 & -.76988E{+}00 \\ .32040E{+}00 & .73293E{+}00 & .13906E{+}01 & -.30124E{+}00 & -.87918E{+}00 & -.14738E{+}01 \\ -.14738E{+}01 & -.87918E{+}00 & -.30124E{+}00 & .15896E{+}01 & .92603E{+}00 & .31015E{+}00 \\ -.76988E{+}00 & -.14738E{+}01 & -.87918E{+}00 & .92603E{+}00 & .15896E{+}01 & .92603E{+}00 \\ -.33590E{+}00 & -.76988E{+}00 & -.14738E{+}01 & .31015E{+}00 & .92603E{+}00 & .15896E{+}01 \end{bmatrix}$$

Variance-covariance matrix of the *ARMA*(3,3) coefficients:

$$\frac{1}{n} \begin{bmatrix} .20509E{+}10 & -.10284E{+}10 & .74868E{+}08 & .20509E{+}10 & -.13369E{+}10 & .38335E{+}09 \\ -.10284E{+}10 & .51569E{+}09 & -.37542E{+}08 & -.10284E{+}10 & .67037E{+}09 & -.19223E{+}09 \\ .74868E{+}08 & -.37542E{+}08 & .27331E{+}07 & .74868E{+}08 & -.48803E{+}08 & .13994E{+}08 \\ .20509E{+}10 & -.10284E{+}10 & .74868E{+}08 & .20509E{+}10 & -.13369E{+}10 & .38335E{+}09 \\ -.13369E{+}10 & .67037E{+}09 & -.48803E{+}08 & -.13369E{+}10 & .87144E{+}09 & -.24989E{+}09 \\ .38335E{+}09 & -.19223E{+}09 & .13994E{+}08 & .38335E{+}09 & -.24989E{+}09 & .71657E{+}08 \end{bmatrix}$$

T-ratios of the *ARMA*(3,3) coefficients (with $n = 100$):

AR coefficients:	.124046*E*−03	−.294003*E*−04	.124534*E*−04
MA coefficients:	.157258*E*−03	−.766445*E*−04	.131668*E*−04

Assuming that an *ARMA*(2,2) is estimated,
Variance-covariance matrix of the *ARMA*(2,2) coefficients:

$$\frac{1}{n} \begin{bmatrix} .48498E{+}02 & -.31224E{+}01 & .47133E{+}02 & -.68291E{+}01 \\ -.31224E{+}01 & .36639E{+}02 & -.23591E{+}01 & .33769E{+}02 \\ .47133E{+}02 & -.23591E{+}01 & .46767E{+}02 & -.65701E{+}01 \\ -.68291E{+}01 & .33769E{+}02 & -.65701E{+}01 & .32397E{+}02 \end{bmatrix}$$

T-ratios of the *ARMA*(2,2) coefficients (with $n = 100$):

AR coefficients:	.806662*E*+00	−.110301*E*+00
MA coefficients:	.104139*E*+01	−.397511*E*+00

correlation coefficient between the estimates of ϕ_1 and ϕ_3 is given by:

$$\rho(\phi_1, \phi_3) = \frac{\text{cov}(\phi_1, \phi_3)}{[\text{var}(\phi_1).\text{var}(\phi_3)]^{1/2}} = \frac{.74868}{[(.20509)(2.7331)]^{1/2}} = .99999.$$

Similarly, the correlation coefficient between any of (ϕ_1, ϕ_2, θ_1, and θ_2) and (ϕ_3 and θ_3) is close to 1 or − 1. These high correlations indicate that the confidence regions for pairs of estimates such as (ϕ_1, ϕ_3) will be attenuated ellipses along one of the diagonals. In other words, while a coefficient estimate such as ϕ_3 may be insignificant when viewed alone, the estimate, conditional on other coefficient estimates such as ϕ_1 and ϕ_2, may be acceptable as significant.[18] Thus, the univariate t−tests, which are routinely employed by researchers, are not appropriate for cases like the one in table 2.

A more acceptable procedure to test for the significance of any given coefficient (or combination of coefficients) is to derive the conditional t

[18] I am grateful to James Patell for pointing this out.

distribution for the coefficient estimate(s) given all other estimates and perform the t-test based on the conditional distribution. The procedure for deriving the conditional univariate or multivariate t distributions is well known and is also summarized in Box and Jenkins [1970, appendix A7.1].

In addition to the above procedure, the joint multiple t distribution of β can be used to test the joint hypothesis H_0: $\beta_1 = \beta_2 = ..\beta_6 = 0$. This procedure is adopted here, since it is equivalent to the conventional F-test. For the above hypothesis, the probability contours of the joint six-variate t distribution of the estimates are ellipsoids defined by $\beta^T \cdot \mathbf{V}(\boldsymbol{\beta})^{-1} . \beta =$ constant, where $\mathbf{V}(\boldsymbol{\beta})$ is the covariance matrix of β and T denotes transpose. The probability mass outside the density contour of the distribution $\beta^T \cdot \mathbf{V}(\boldsymbol{\beta})^{-1} \cdot \beta = 6F_0$ is given by $\Pr\{F > F_0\}$, where the F distribution has $(6, n - 6)$ degrees of freedom.[19] Thus, the joint significance of β can be tested by computing the statistic F_0 using the covariance matrix of β, and by comparing the F_0 value against an appropriate cutoff value from the F distribution with (6, 94) degrees of freedom. For the $ARMA(3, 3)$ case in table 2, F_0 is computed as 0.03984, which is clearly not large enough to reject H_0.

Hence, with 100 observations, the null hypothesis $\beta_1 = \beta_2 = ..\beta_6 = 0$ cannot be rejected. Reported quarterly earnings data of firms are available for only about 30 years, and empirical research on earnings typically relies on about 100 or fewer observations. For this sample size, estimating an $ARMA(3, 3)$ model for the average cost earnings data of the firm would lead to disappointing results.

The univariate t-statistics and the F-statistic are very small and the elements of the $\mathbf{V}(\boldsymbol{\beta})$ matrix in the top panel of table 2 are very large because the $\mathbf{I}_l$ matrix of the $ARMA(3, 3)$ model is very close to singularity. Rows 3 and 6 (and equivalently columns 3 and 6) of the information matrix in the top panel of table 2 are nearly identical, but with opposite sign. The magnitude of the determinant of $\mathbf{I}_l$ could thus be increased when these rows and columns are dropped. In other words, if a lower-order $ARMA$ model is estimated, the coefficients of that model will have smaller standard errors. In the example given in table 2, this would mean attempting to estimate an $ARMA(2, 2)$ model by dropping the coefficients ϕ_3 and θ_3 from consideration.

The bottom panel of table 2 presents the theoretical (4×4) covariance matrix of an $ARMA(2, 2)$ model, computed using the procedures in Dharan [1982a], when the firm's parameters remain the same as in the top panel. It is seen that all the elements of the new covariance matrix have considerably smaller values compared to the top panel. The magnitudes of the univariate t-ratios of the ϕ_1, ϕ_2, θ_1, and θ_2 coefficients now are larger. Still, with a sample size of 100, the t-test cannot reject the hypothesis that the estimates of the coefficients are zero. To test the

[19] See Box and Jenkins [1970, appendix A7.1].

hypothesis that all four coefficients are zero, an F-statistic can be computed, as was done for the $ARMA(3, 3)$ case. For the $ARMA(2, 2)$ model, the joint four-variate t distribution is given by $\gamma^T \cdot \mathbf{V}(\gamma)^{-1} \cdot \gamma =$ constant, where $\gamma = (\beta_1, \beta_2, \beta_4, \beta_5)^T$ and $\mathbf{V}(\gamma)$ is the covariance matrix in the bottom panel. Setting the constant equal to $4F_0$ gives the required F-statistic with (4, 96) degrees of freedom. For the sample size of 100, the computed $F_0(4, 96)$ is 0.9165, which is not large enough to reject the null at acceptable significance levels.

Thus, even an $ARMA(2, 2)$ model does not give satisfactory estimates of the coefficients when the sample size is 100. For such sample sizes, it is very likely that a researcher analyzing the earnings data of the table 2 case would identify an $ARMA(1, 2)$ or an $ARMA(1, 1)$ model as the appropriate earnings model.

These observations hold for all the invertible processes examined. Table 3 shows case 6 from table 1. All six autoregressive and moving average coefficients for this case are larger (in magnitude) than those of table 2. However, the magnitudes of the expected univariate t-ratios for an $ARMA(3, 3)$ model are still very small. In addition, the hypothesis "all

TABLE 3

Average Cost Earnings Series: Case 6 of Table 1

Model parameters: $a = 0.7 \quad b = 0.7 \quad c = 0.7 \quad d = 0.7$
$ARMA(3,3)$ Model:
Autoregressive coefficients: $\phi_1 = .80176 \quad \phi_2 = -.18759 \quad \phi_3 = .01112$
Moving average coefficients: $\theta_1 = 1.15530 \quad \theta_2 = -.58686 \quad \theta_3 = .05686$
Information matrix of the $ARMA(3,3)$ coefficients:

$$n \begin{bmatrix} .19146E{+}01 & .12991E{+}01 & .69686E{+}00 & -.22981E{+}01 & -.15116E{+}01 & -.80104E{+}00 \\ .12991E{+}01 & .19146E{+}01 & .12991E{+}01 & -.18135E{+}01 & -.22981E{+}01 & -.15116E{+}01 \\ .69686E{+}00 & .12991E{+}01 & .19146E{+}01 & -.83235E{+}00 & -.18135E{+}01 & -.22981E{+}01 \\ -.22981E{+}01 & -.18135E{+}01 & -.83235E{+}00 & .30432E{+}01 & .22492E{+}01 & .94052E{+}00 \\ -.15116E{+}01 & -.22981E{+}01 & -.18135E{+}01 & .22492E{+}01 & .30432E{+}01 & .22492E{+}01 \\ -.80104E{+}00 & -.15116E{+}01 & -.22981E{+}01 & .94052E{+}00 & .22492E{+}01 & .30432E{+}01 \end{bmatrix}$$

Variance-covariance matrix of the $ARMA(3,3)$ coefficients:

$$\frac{1}{n} \begin{bmatrix} .17822E{+}07 & -.12015E{+}07 & .18221E{+}06 & .17822E{+}07 & -.18314E{+}07 & .81274E{+}06 \\ -.12015E{+}07 & .80998E{+}06 & -.12284E{+}06 & -.12015E{+}07 & .12347E{+}07 & -.54791E{+}06 \\ .18221E{+}06 & -.12284E{+}06 & .18636E{+}05 & .18221E{+}06 & -.18724E{+}06 & .83098E{+}05 \\ .17822E{+}07 & -.12015E{+}07 & .18221E{+}06 & .17822E{+}07 & -.18314E{+}07 & .81274E{+}06 \\ -.18314E{+}07 & .12347E{+}07 & -.18724E{+}06 & -.18314E{+}07 & .18820E{+}07 & -.83519E{+}06 \\ .81274E{+}06 & -.54791E{+}06 & .83098E{+}05 & .81274E{+}06 & -.83519E{+}06 & .37064E{+}06 \end{bmatrix}$$

T-ratios of the $ARMA(3,3)$ coefficients (with $n = 100$):
AR coefficients: $.600571E{-}02 \quad -.208434E{-}02 \quad .814401E{-}03$
MA coefficients: $.865394E{-}02 \quad -.427784E{-}02 \quad .933906E{-}03$

Assuming that an $ARMA(2,2)$ is estimated,
Variance-covariance matrix of the $ARMA(2,2)$ coefficients:

$$\frac{1}{n} \begin{bmatrix} .48766E{+}01 & -.26354E{+}01 & .36278E{+}01 & -.22414E{+}01 \\ -.26354E{+}01 & .46995E{+}01 & -.17582E{+}01 & .33589E{+}01 \\ .36278E{+}01 & -.17582E{+}01 & .33669E{+}01 & -.20121E{+}01 \\ -.22414E{+}01 & .33589E{+}01 & -.20121E{+}01 & .30939E{+}01 \end{bmatrix}$$

T-ratios of the $ARMA(2,2)$ coefficients (with $n = 100$):
AR coefficients: $.363068E{+}01 \quad -.865322E{+}00$
MA coefficients: $.629618E{+}01 \quad -.333642E{+}01$

coefficients are zero" still cannot be rejected using an F-test based on the joint t distribution of the coefficients. From the bottom panel of table 3, it seems likely that an $ARMA(1, 2)$ model will be identified for this case.

To summarize the estimation problem, it is very likely that a researcher analyzing the reported time series of earnings data of the firm would not identify or estimate the derived $ARMA(3, 3)$ model and instead would identify a more parsimonious model such as $ARMA(1, 1)$ or $ARMA(1, 2)$ as the appropriate earnings model. Less parsimonious models might be identified if model coefficients are tested for significance using conditional univariate or multivariate t distributions instead of marginal univariate t distributions. Conventional time-series research has not, however, used such tests widely even when estimation results required them. Note also that one can expect some variation on model identification depending on the identification procedures adopted. However, the inability to identify the underlying model in the case of the earnings model is caused by the effect of sample size and coefficient magnitudes on the coefficient variances rather than solely by the identification procedure.

5. *Effect on Forecasts*

It was noted in the introduction that time-series earnings models have been used in various empirical studies primarily to generate "expected" or forecasted earnings. The effect of the estimation problem on one-step-ahead forecasting is examined in this section. Other implications of the estimation problem are addressed in the next section.

A measure of the effectiveness of a one-step-ahead forecast is the "mean-squared forecast error" ($MSFE$), defined as $E(\hat{Z}_t - Z_t)^2$, where Z_t is the true earnings and $\hat{Z}_t$ is the forecast. When the coefficients are known for the $ARMA(3, 3)$ model in (10), the minimum $MSFE$ forecast $\hat{Z}_t$ is obtained by setting $\tilde{e}_t$ equal to zero. For this forecast, $E(\hat{Z}_t - Z_t)^2 = \sigma^2(\tilde{e})$. In other words, the residual variance is the smallest forecast error variance any forecast could have, even when the coefficients are known.

I shall examine the effect of misspecification on forecasts by measuring the minimum forecast error variance for the misspecified earnings models. Naturally, the minimum $MSFE$ of a misspecified model such as $ARMA(1, 1)$ or $ARMA(1, 2)$ cannot be less than $\sigma^2(\tilde{e})$. However, if the minimum $MSFE$ is not very different from $\sigma^2(\tilde{e})$, then it can be assumed that the misspecification does not adversely affect the value of forecasts.

This comparison of the minimum $MSFE$ of the true and the misspecified models is based, as noted, on the assumption that the coefficients of the true model are known exactly. Clearly, the estimated $MSFE$ of these models (based on a given sample of observations) will be higher than the minimum values when the true coefficients are replaced by their estimates. For example, if an $ARMA(3, 3)$ model is fitted to a given sample of realizations from the earnings model in (10), the $MSFE$ of the estimated true model will be greater than $\sigma^2(\tilde{e})$. If n is the sample size, it can be shown that the increase in the *MSFE* of the true model is of the order n^{-1}. In other words, as the sample size increases, the discrepancy between the *MSFE* of the estimated true model and its minimum *MSFE* approaches zero. Intuitively, this can be seen by observing that the variances of the coefficient estimates are a function of n^{-1}, as seen from (15).[20] Similarly, the *MSFE* of an estimated misspecified model such as *ARMA*(1, 2) will differ from the minimum *MSFE* for that model by a factor which is of the order n^{-1}. Thus, for a given sample size, the ratio of the estimated *MSFE* of the true earnings model to that of a misspecified model can be expected to vary around the value given by the ratio of the minimum *MSFE*s of the two models. For this reason, the latter ratio is a reasonable benchmark to measure the effect of misspecification on forecasting.

The minimum *MSFE* for the misspecified model will be computed by first assuming that the researcher identifies an *ARMA*(1, 2) model for the observed earnings data which are in fact generated by (10). Since only the ratio of the minimum *MSFE* of the misspecified model to the minimum *MSFE* of the true model is of interest, let $\sigma^2(\tilde{e}) = 1$. Let $\tilde{\phi}_1$, $\tilde{\theta}_1$, and $\tilde{\theta}_2$ be the coefficients of the *ARMA*(1, 2) misspecified model. The earnings forecast computed from the misspecified model is:

$$\tilde{Z}_t = \tilde{\phi}_1 Z_{t-1} - \tilde{\theta}_1 \tilde{e}_{t-1} - \tilde{\theta}_2 \tilde{e}_{t-2}. \tag{21}$$

Since Z_t is generated by (10), and since $\sigma^2(\tilde{e}) = 1$, the minimum *MSFE* of the *ARMA*(1, 2) model is given by:

$$\begin{aligned} E(\tilde{Z}_t - Z_t)^2 = 1 &+ \gamma_2{}^2 + \gamma_3{}^2 + \theta_3{}^2 + V_0(\gamma_1 + \phi_2 + \phi_3) \\ &- 2V_1(\gamma_1\phi_2 - \phi_2\phi_3) \\ &- 2V_2(\gamma_1\phi_3) - 2(\gamma_1\gamma_2 - \phi_2\gamma_3 + \phi_3\theta_3) \\ &+ 2V_{e,1}(\gamma_1\gamma_3 + \phi_2\theta_3) \\ &+ 2V_{e,2}(\gamma_1\theta_3), \end{aligned} \tag{22}$$

where:

$$\gamma_1 = \tilde{\phi}_1 - \phi_1,\ \gamma_2 = \tilde{\theta}_1 - \theta_1,\ \gamma_3 = \tilde{\theta}_2 - \theta_2,\ V_0 = E(Z_t)^2,$$

$$V_1 = E(Z_t Z_{t-1}),\ V_2 = E(Z_t Z_{t-2}),\ V_{e,1} = E(Z_t e_{t-1}) = \phi_1 - \theta_1, \text{ and}$$

$$V_{e,2} = E(Z_t e_{t-2}) = \phi_1 V_{e,1} + \phi_2 - \theta_2.$$

To compute the minimum *MSFE* from (22), the variance V_0 and the autocovariances V_1 and V_2 of the process in (10) must first be computed. From their definitions, and using (10), it is readily seen that:

$$V_0 = \phi_1 V_1 + \phi_2 V_2 + \phi_3 V_3 + 1 - \theta_1 V_{e,1} - \theta_2 V_{e,2} - \theta_3 V_{e,3}, \tag{23}$$

$$V_1 = \phi_1 V_0 + \phi_2 V_1 + \phi_3 V_2 - \theta_1 - \theta_2 V_{e,1} - \theta_3 V_{e,2}, \tag{24}$$

[20] See also Box and Jenkins [1970, appendix A7.3].

$$V_2 = \phi_1 V_1 + \phi_2 V_0 + \phi_3 V_1 - \theta_2 - \theta_3 V_{e,1}, \text{ and} \tag{25}$$

$$V_3 = \phi_1 V_2 + \phi_2 V_1 + \phi_3 V_0 - \theta_3, \tag{26}$$

where $V_3 = E(Z_t Z_{t-3})$, and $V_{e,3} = E(Z_t e_{t-3}) = \phi_1 V_{e,2} + \phi_2 V_{e,1} + \phi_3 - \theta_3$. These form a set of simultaneous equations in V_0, V_1, V_2, and V_3. Thus, given β, one can compute the variance and autocovariances of Z_t in (10).

To compute the minimum *MSFE* from (22), we also need to know the coefficients $\tilde{\phi}_1$, $\tilde{\theta}_1$, and $\tilde{\theta}_2$ of the misspecified *ARMA*(1, 2) model. A reasonable procedure to compute these is to assume that the empiricist knows V_0, V_1, V_2, etc., since s/he observes the realizations from the true process (10), and that s/he will use these values to compute the estimated coefficients of his assumed *ARMA*(1, 2) model. The variance and covariances of an *ARMA*(1, 2) process can be described by equations similar to (23) to (26). These are:

$$L_1 = V_0 - V_1\tilde{\phi}_1, \tag{27}$$

$$L_2 = -V_0\tilde{\phi}_1 + V_1, \text{ and} \tag{28}$$

$$\tilde{\theta}_2 = \tilde{\phi}_1 V_1 - V_2, \tag{29}$$

where:

$$L_1 = 1 + \tilde{\theta}_2{}^2 - (\tilde{\phi}_1 - \tilde{\theta}_1)(\tilde{\theta}_1 + \tilde{\phi}_1\tilde{\theta}_2), \text{ and} \tag{30}$$

$$L_2 = -\tilde{\theta}_1 - \theta_2(\tilde{\phi}_1 - \tilde{\theta}_1). \tag{31}$$

Since V_0, V_1, and V_2 are now assumed known, it is possible to compute $\tilde{\phi}_1$, $\tilde{\theta}_1$, and $\tilde{\theta}_2$ from these equations. Starting with a guess $\tilde{\phi}_1{}^*$, one can compute L_2 using (28), $\tilde{\theta}_2$ using (29), and then $\tilde{\theta}_1$ using (31) and $\tilde{L}_1$ using (30). When this L_1 is substituted in (27), one gets a new guess for $\tilde{\phi}_1$, say $\tilde{\phi}_1^{**}$. When the two guesses are equal, a solution has been found. Otherwise, one must iterate again by modifying the initial guess.[21]

To summarize, the minimum *MSFE* of the misspecified *ARMA*(1, 2) model can be computed from (22) by first computing the variance and autocovariances using (23) to (26) and then computing the *ARMA*(1, 2) coefficients using (27) to (31). If the misspecified model is assumed to be an *ARMA*(1, 1), the minimum *MSFE* can be computed from (22) by letting $\tilde{\theta}_2 = 0$. In addition, the coefficients $\tilde{\phi}_1$ and $\tilde{\theta}$ of this model can be computed by directly solving (27), (28), (30), and (31) once V_0 and V_1 are known.[22]

In section 3, 152 invertible cases were generated by various parameter combinations of the earnings model. For each of these cases, minimum mean-squared forecast errors were computed for *ARMA*(1, 2) and *ARMA*(1, 1) approximations to the underlying *ARMA*(3, 3) process.

[21] For an *ARMA*(1, 2) model, (26) yields $V_k = \tilde{\phi}_1 V_{k-1}$ for $k \geq 2$. The solution $\tilde{\phi}_1$ from the iterative search may not satisfy this equation since the iterative search ignores the available data on V_2, V_3, etc. This is the effect of aproximating (10) by *ARMA*(1, 2).

[22] Note 21 is applicable to the case of *ARMA*(1, 1) also. In other words, in solving for *ARMA*(1, 1) coefficients, equation (29) and the data on V_2, V_3, etc., are ignored.

Table 4 summarizes the data on *MSFE*s. Since $\sigma^2(\tilde{e})$ was assumed equal to one, the *MSFE* in table 4 represents the ratio of the minimum *MSFE* of the assumed model to the minimum *MSFE* of the true model. From the ratios in table 4, it is seen that the two approximations result in median increases in minimum *MSFE* of about 12 percent. The mean increases, though, are different for the two approximations. The *ARMA*(1, 1) approximation has a higher mean increase in *MSFE* because it sometimes has increases of more than 100 percent in the forecast error variance. (By contrast, the maximum increase in *MSFE* for the *ARMA*(1, 2) approximation is only 78 percent.) However, the extreme increases in *MSFE* for the *ARMA*(1, 1) approximation mostly occurred when the dropped θ_2 value (in magnitude) was relatively large in relation to the retained θ_1 value. One would expect in these cases that the probability that an *ARMA*(1, 1) approximation would be accepted instead of *ARMA*(1, 2) is also relatively small. Hence, the equal median increases in the *MSFE* of the two misspecified models are perhaps better practical indicators of the potential effect of the two misspecifications on forecasting.

Whether these median increases in earnings forecast error variance are acceptable to a forecaster naturally depends on the cost (to the forecaster) of developing a theoretical model to identify the true underlying process and the cost of forecast errors. It is clear from table 4 that the increase in forecast error variance due to the misspecification can be negligible for some cases. For firms that correspond to these cases of parameter value combinations, the effect of model misspecification may be considered tolerable by a forecaster. Of course, a modeling approach to identifying the structure of earnings process would still be necessary to recognize such cases.

TABLE 4

Minimum Mean-Squared Forecast Errors of Misspecified Models: Summary of 152 Invertible Cases

		Minimum *MSFE* of *ARMA*(1,2)	Minimum *MSFE* of *ARMA*(1,1)
Percentiles.	10	1.0090	1.0067
	20	1.0242	1.0188
	30	1.0490	1.0409
	40	1.0722	1.0701
	(Median)50	1.1186	1.1176
	60	1.1524	1.1844
	70	1.2289	1.2867
	80	1.3272	1.4954
	90	1.5034	1.9691
Minimum		1.0025	1.0017
Mean		1.1842	1.2899
Maximum		1.7808	2.6650

Note: The minimum *MSFE* of the true *ARMA*(3,3) process is normalized to 1, and hence the above data represent ratios of the minimum *MSFE* of the misspecified model to the minimum *MSFE* of the true model.

6. Conclusion

Based on the analysis above, it appears that accounting earnings of real firms follow a more complex process than those identified by prior empirical work. Regardless of the identification method used, earnings models identified and estimated without reference to a theory of the firm will likely be more parsimonious than the model postulated by that theory. In this sense, the estimated models can be said to be misspecified.

My analysis was based on a wide range of parameter values of a firm, and, when some of the assumptions of this model about the firm and its accounting system are relaxed, the resulting earnings model can only be expected to have more autoregressive and moving average terms than the $ARMA(3, 3)$ model identified here. For example, Dharan [1981] shows that replacing (1) with a random-walk sales process or a sales process with seasonality would lead to differencing terms and possibly more moving average terms in the decision rules. Similarly, using a straight-line depreciation method instead of the units-of-production method used here would lead to more autoregressive terms in the capacity decision rule. Hence, when the model assumptions are relaxed, one can expect the estimation problems to remain.

The parsimony of the estimated models is due to the large theoretical variances of the coefficients when sample sizes are of the order encountered in accounting research and when some of the coefficients (usually of higher lags) have near-zero values. Of course, parsimony is also sometimes deliberately sought by identification procedures, as in Akaike [1974]. Similarly, the subjective or iterative nature of the identification procedure used may lead the empiricist to give preference to parsimonious models, as in Box and Jenkins [1970].

The fact that small sample sizes and small coefficients will limit identification to misspecified and parsimonious models suggests that, depending on the identification procedure used, different researchers analyzing the same earnings data may identify different parsimonious (but misspecified) models. This is particularly likely when the identification procedure requires subjective inputs from the researcher in model selection, as in the case of the Box-Jenkins and (to a lesser extent) the Akaike procedures. This could explain the variation in the earnings model structures observed in the quarterly earnings models identified by different accounting researchers.[23]

This parsimony effect also has some implications for empirical studies that test for changes in the earnings model structure caused by, say, an accounting change. For example, suppose that a researcher hypothesizes a change in the earnings model structure as a result of an accounting change or a decision rule change and tests this hypothesis by estimating

[23] See Watts [1975], Griffin [1977], Foster [1977], Brown and Rozeff [1979], and Dharan [1982*b*]. Foster [1978] provides a review.

separate models for the prechange and postchange earnings data. The researcher would likely fail to recognize differences in the order of the identified models, since limited sample sizes would preclude estimation of anything but the most parsimonious models for both earnings series. On the other hand, hypotheses focusing on the differences in the estimated coefficients of each series (e.g., differences in the value of ϕ_1 of each series) are more likely to be testable despite small sample sizes. Such a hypothesis test was employed by Dopuch and Watts [1972].

I should emphasize that my main objective was to document a rigorous causal earnings modeling effort following the framework of Gonedes and Dopuch [1979] so the observations outlined above are by-products of this effort. Other researchers who wish to pursue further the above results on estimation and forecasting or, more likely, the general modeling methods outlined in this paper should realize that extending this modeling effort to other environments will lead to analytical derivations which are likely to become much more complicated when more realistic model assumptions are introduced. Hence, it may be more feasible to rely on simulating the model elements, such as the sales process and the inventory decision, and then studying the simulated earnings data. The simulation model could include real-world data (e.g., actual sales data to describe the stochastic process for sales) and optimal decision rules. With these additions, the technique of causally linking accounting earnings to managerial decision rules via an accounting system becomes powerful, and accounting researchers may apply this technique to areas such as forecasting efficiency, firm valuation, and income-smoothing—areas that require knowledge of the effect of firms' managerial decisions on earnings.

APPENDIX A

Average Cost Earnings Model

This appendix gives a brief derivation of the coefficients of the $ARMA(3, 3)$ earnings model given in (10) and is based on Dharan [1981].

As noted in section 4, the time-series expression for the earnings, Z_t, is obtained by first getting a time-series expression for the unit cost i_t. To get the latter, one must substitute the decision rules for production, inventory, and capacity—expressions (5), (6), and (8)—into the expression for unit cost (9). Define a lag operator L such that $L^x Z_t$ gives the lagged variable Z_{t-x} for integer $x \geq 0$. Then the three decision rules can be written as:

$$(1 - b_1 L) I_t = bcS_0 - e_t, \tag{A.1}$$

$$(1 - b_1 L) q_t = bS_0 + be_{t-1}, \text{ and} \tag{A.2}$$

$$(1 - b_1 L)(1 - a_1 d_1 L) K_t = abS_0 + abe_{t-1}, \tag{A.3}$$

where a is the capital stock adjustment factor, b is the inventory stock adjustment factor, c is the stock to sales ratio, d is the physical depreciation rate, $a_1 = 1 - a$, $b_1 = 1 - b$, and $d_1 = 1 - d$. S_0 is the mean of the sales process.

In the sales process described by (1), the shocks e_t, e_{t-1}, ... are generated independently by a white noise process and then added to the value S_0 to get S_t, S_{t-1}, ..., etc. Viewed this way, one can see that S_0 can be scaled arbitrarily for any given set of shocks to get a correspondingly scaled series of sales. The variance of the sales process is unaltered. Scaling S_0 does not affect the stationarity or invertibility of the resulting earnings series. Furthermore, scaling of S_0 does not require altering the firm's four parameters a, b, c, and d. In order to get tractable linear expressions for earnings under average cost method, I assume that S_0 is scaled such that e_t/S_0 is treated as a small fraction. Then nonlinear terms involving e_te_s/S_0^2 for all t and s can be ignored as too small. With this approximation, (A.2) results in:

$$(1 - b_1L)^2q_t^2 = S_0^2b^2 + S_0b^2e_{t-1}. \tag{A.4}$$

From (A.1) and (A.2):

$$(1 - b_1L)(I_{t-1} + q_t) = bS_0 + bcS_0 - b_1e_{t-1}. \tag{A.5}$$

From (A.1), (A.3), and (A.4):

$$\begin{aligned}(1 - b_1L)&(i_{t-1}I_{t-1} + q_t^2/S_0 + K_t)\\ &= S_0\left(\frac{b^2}{1 - b_1L} + \frac{ab}{1 - a_1\,d_1L}\right) + S_0bci_{t-1} \\ &+ \left(\frac{b^2}{1 - b_1L} + \frac{ab}{1 - a_1\,d_1L}\right)e_{t-1} - i_{t-1}e_{t-1}.\end{aligned} \tag{A.6}$$

To derive a linear time-series expression for i_t in terms of the shocks, first assume that this expression will be of the form:

$$i_t = \theta + \sum_{j=1}^{\infty} \beta_je_{t-j}, \tag{A.7}$$

where θ is a constant and β_j are coefficients of the infinite moving average terms. θ can be viewed as the mean value of i_t. Then the terms $i_{t-1}e_{t-1}$ in (A.6) can be substituted by θe_{t-1} given the earlier assumption about e_t/S_0. Now the time-series expression for i_t can be derived by substituting (A.5) and (A.6) into equation (9) defining i_t. To facilitate these substitutions, it is useful to use (A.5) to write the term $(I_{t-1} + q_t)^{-1}$ in (9) as:

$$(I_{t-1} + q_t)^{-1} = \left(\frac{S_o}{(1 - b_1L)}\right)^{-1}(b + bc)^{-1}\left(1 - \frac{b_1}{b + bc}u_{t-1}\right)^{-1},$$

where $u_{t-1} = e_{t-1}/S_0$. Since e_t/S_o is assumed to be a small fraction, the term $\left(1 - \frac{b_1}{b + bc}u_{t-1}\right)^{-1}$ can be approximated as $\left(1 + \frac{b_1}{b + bc}u_{t-1}\right)$. With this substitution, the rest of the derivation merely involves accumulating similar terms. The resulting expression for i_t is given by:

$$i_t = \frac{1}{b + bc}\left(b + \frac{ab}{1 - a_1 d_1}\right) + \frac{c}{1 + c} i_{t-1} + \frac{1 + bc}{(b + bc)^2}\left(\frac{b^2}{1 - b_1 L} + \frac{ab}{1 - a_1 d_1 L} - b\theta\right)u_{t-1}. \tag{A.8}$$

Let $c_1 = c/(1 + c)$. Since c is a stock-sales ratio, c_1 is the ratio of stocks to stocks plus sales. When the term $c_1 i_{t-1}$ is taken to the left side and when both sides are multiplied by $(1 - b_1 L)(1 - a_1 d_1 L)$, one gets:

$$(1 - c_1 L)(1 - b_1 L)(1 - a_1 d_1 L)i_t = \lambda_0 + (\lambda_1 + \lambda_2 L + \lambda_3 L^2)u_{t-1}, \tag{A.9}$$

where λ_0, λ_1, λ_2, and λ_3 are constants defined by the four parameters a, b, c, and d. These are defined below. The left side of (A.9) can be expanded to get:

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3)i_t = \lambda_0 + (\lambda_1 + \lambda_2 L + \lambda_3 L^2)u_{t-1}. \tag{A.10}$$

Expression (A.10) shows that i_t is described by an $ARMA(3, 2)$ process in terms of the shock $\lambda_1 u_{t-1}$, and this process has three autoregressive coefficients, ϕ_1, ϕ_2, and ϕ_3, and two moving average coefficients. In addition, (A.10) confirms that i_t indeed has the form assumed in (A.7), since any stationary $ARMA$ process can be written equivalently as an infinite order moving average process.

The autoregressive coefficients of (A.10) are given by:

$$\phi_1 = c_1 + b_1 + a_1 d_1, \tag{A.11}$$

$$\phi_2 = - c_1 b_1 - b_1 a_1 d_1 - a_1 d_1 c_1, \text{ and} \tag{A.12}$$

$$\phi_3 = c_1 b_1 a_1 d_1. \tag{A.13}$$

The moving average coefficients are given by:

$$\lambda_1 = -\frac{1 + bc}{b(1 + c)^2}(\theta - a - b), \tag{A.14}$$

$$\lambda_2 = \frac{1 + bc}{b(1 + c)^2}(b_1(\theta - a) + a_1 d_1(\theta - b)), \text{ and} \tag{A.15}$$

$$\lambda_3 = -\frac{1 + bc}{b(1 + c)^2} b_1 a_1 d_1 \theta. \tag{A.16}$$

The moving average constant is:

$$\lambda_0 = \frac{b(1 + a - a_1 d_1)}{1 + c}. \tag{A.17}$$

From (A.9), the mean of the i_t process is seen as:

$$\theta = \frac{\lambda_0}{(1 - c_1)(1 - b_1)(1 - a_1 d_1)} = 1 + \frac{a}{1 - a_1 d_1}. \tag{A.18}$$

The time-series expression for the earnings Z_t is obtained by substituting (A.10) into $Z_t = (1 - i_t)S_t$. This gives:

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3)Z_t = \lambda_0' + (1 - \theta_1 L - \theta_2 L^2 - \theta_3 L^3)\tilde{e}_t, \quad \text{(A.19)}$$

where ϕ_1, ϕ_2, and ϕ_3 were defined above:

$$\tilde{e}_t = -(\theta - 1)e_t, \text{ and} \quad \text{(A.20)}$$

$$\lambda_0' = [(1 - c_1)b(1 - a_1\, d_1) = \lambda_0]S_0 = \frac{abS_0}{1 + c}. \quad \text{(A.21)}$$

The three moving average coefficients of (A.19) are given by:

$$\theta_1 = \phi_1 - \frac{\lambda_1}{\theta - 1}, \quad \text{(A.22)}$$

$$\theta_2 = \phi_2 - \frac{\lambda_2}{\theta - 1}, \text{ and} \quad \text{(A.23)}$$

$$\theta_3 = \phi_3 - \frac{\lambda_3}{\theta - 1}. \quad \text{(A.24)}$$

Expression (A.19) describes the earnings under average cost method as an *ARMA*(3, 3) process in terms of the shock $\tilde{e}_t$. Since e_t is from a white noise process, the shock $\tilde{e}_t$ is also from a white noise process, with zero mean and a variance of $\sigma^2(\tilde{e}_t) = (\theta - 1)^2\sigma^2(e_t)$. The three autoregressive coefficients of Z_t are given by (A.11) to (A.13), and the three moving average coefficients are given by (A.22) to (A.24). Expression (A.19) is identical to (10) except that the moving average constant λ_0' in (A.19) has been dropped in (10), which implies that the mean of the realizations $Z_t, Z_{t-1}, \ldots$, etc. has been subtracted from the $Z_t, Z_{t-1}, \ldots$, etc. in (10).

In conclusion, it should be noted that the finite lags of Z and $\tilde{e}$ observed in (A.19) do not mean that only three past periods affect current earnings. When the earnings are to be written exclusively in terms of the shocks, the stationary *ARMA* model in (A.19) is equivalently written in an infinite moving average form similar to (A.7). Such a form makes clear the dependence of Z_t on all present and past shocks. Also, it should be noted that the linear stochastic form in (A.19) was derived by ignoring all nonlinear terms in the derivation. Thus, (A.19) is a linear approximation to a possibly more complex nonlinear stochastic equation.

REFERENCES

AKAIKE, H. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* (December 1974): 716–23.

BOX, G. E. P., AND G. M. JENKINS. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day, 1970.

BROWN, L. D., AND M. S. ROZEFF. "Univariate Time-Series Models of Quarterly Earnings per Share: A Proposed Premier Model." *Journal of Accounting Research* (Spring 1979): 179–89.

DHARAN, B. G. "Time Series Analysis of Inventory Decisions and Accounting Earnings." Ph.D. dissertation, Carnegie-Mellon University, 1981.

———. "Theoretical Evaluation of Sample Size for ARMA Model Estimation." Working paper, Northwestern University, February 1982*a*.

———. "Empirical Identification Procedures for Earnings Models." Working paper, Rice University, July 1982*b*.

DOPUCH, N., AND R. WATTS. "Using Time Series Models to Assess the Significance of Accounting Changes." *Journal of Accounting Research* (Spring 1972): 180–94.

EICHENBAUM, M. S. "A Rational Expectations Equilibrium Model of the Cyclical Behavior of Inventories and Employment." Working paper, University of Minnesota, January 1980.

FOSTER, G. "Quarterly Accounting Data: Time-Series Properties and Predictive Ability Results." *The Accounting Review* (January 1977): 1–21.

———. *Financial Statement Analysis*. Englewood Cliffs, N.J.: Prentice-Hall, 1978.

GONEDES, N. J., AND N. DOPUCH. "Economic Analysis and Accounting Techniques: Perspective and Proposals." *Journal of Accounting Research* (Autumn 1979): 384–410.

GOODWIN, R. M. "Secular and Cyclical Aspects of the Multiplier and Accelerator." In *Income, Employment, and Public Policy*, edited by A. H. Hansen. New York: W. W. Norton, 1948.

GRIFFIN, P. A. "The Time-Series Behavior of Quarterly Earnings—Preliminary Evidence." *Journal of Accounting Research* (Spring 1977): 71–83.

HALL, R. E., AND D. W. JORGENSON. "Tax Policy and Investment Behavior." *American Economic Review* (June 1967): 391–414.

HANSEN, L. P., AND T. J. SARGENT. "Formulating and Estimating Dynamic Linear Rational Expectations Models: I." *Journal of Economic Dynamics and Control* (1980): 7–46.

HOLT, C. C., AND F. MODIGLIANI. "Firm Cost Structures and the Dynamic Response of Inventories, Production, Work Force, and Orders to Sales Fluctuations." In *Inventory Fluctuations and Economic Stabilization, Part II*, Joint Economic Committee, U.S. 87th Congress, 1961.

LUCAS, R. E. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and the Labor Market*, edited by K. Brunner and A. Meltzer. Amsterdam: North-Holland, 1976.

———, AND E. C. PRESCOTT. "Investment under Uncertainty." *Econometrica* (1971): 659–81.

MANTELL, E. H. "Forecasting Aggregate Changes in U.S. Manufacturing and Nonmanufacturing Inventory Investment." *Journal of Business* (January 1977): 40–49.

METZLER, L. A. "The Nature and Stability of Inventory Cycles." *Review of Economic Statistics* (August 1941): 113–29.

SARGENT, T. J. *Macroeconomic Theory*. New York: Academic Press, 1979.

TRIVEDI, P. K. "Retail Inventory Investment Behavior." *Journal of Econometrics* (March 1973): 61–80.

VISSCHER, M. "Tax Policies of R&D and Technological Innovation." Working paper, Carnegie-Mellon University, 1978.

WATTS, R. "The Time Series Behavior of Quarterly Earnings." Working paper, University of New Castle, April 1975.