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

Empirical Identification Procedures for Earnings Models

BALA G. DHARAN*

1. Introduction

A large number of studies in the accounting literature have focused on the identification of a stochastic univariate time-series structure for corporate earnings, with the identified models being used primarily to generate an "expected" or forecasted earnings.[1] Most studies have used the iterative procedure of Box and Jenkins [1970] for model identification. This procedure requires the analyst to identify a preliminary model of earnings from the observed characteristics (such as autocorrelations) of the earnings data, estimate the model, and then modify the model based on certain diagnostic tests. The selection of the model is sensitive to the characteristics of the series and so judgment plays an important role in the process.

In this paper, I test two noniterative identification procedures that also provide unique model selections for earnings data. One is the Akaike procedure. It relies on a definition of sample information and selects a model that maximizes the subsequent information measure. The procedure is theoretically based, is easy to implement, and has achieved a wide following in the statistics literature since 1974. In addition, it selects, by design, a unique parsimonious earnings model for a given firm, making it attractive to accounting researchers working with moderate sample sizes. The second procedure uses a predictability criterion and selects the model which achieves the minimum forecast error on a holdout sample.

* Assistant Professor, Rice University. I gratefully acknowledge help and guidance from Robert S. Kaplan. Suggestions from George Foster on an earlier draft improved the focus and methodology of this paper. I also thank the reviewers for their comments. [Accepted for publication August 1982.]

[1] See Abdel-khalik and Thompson [1978] and Foster [1978] for review and references.

In various forms, this procedure is widely used in the economics literature.

The two procedures are tested empirically using quarterly earnings models for 30 firms. The results indicate that the "premier" or sample-wide earnings model selected by the Akaike procedure is the same model selected by using the Box-Jenkins procedure. The results also indicate that the Predictability procedure, despite its obvious myopic focus, results in the selection of models that are comparable in forecasting efficiency to models selected by the conceptually superior Akaike method. The Predictability procedure is thus a surprisingly strong alternative to the more sophisticated Akaike and Box-Jenkins identification procedures.

The Akaike and Predictability procedures are described in section 2 using a unifying framework in which the identification problem is defined. The various design elements of the empirical study to test these procedures are described in section 3. Section 4 presents the results on model selection and forecasting. Conclusions appear in section 5.

2. *The Procedures*

To better understand the Akaike and Predictability procedures, I first define the identification problem addressed by these and other procedures. For a given observation set, X, the identification problem can be stated as the appropriate specification of an unknown density function $g(X)$ using the information from the sample, X. In the studies on accounting earnings, the unknown $g(X)$ is generally assumed to be of the form $f_\theta(X \mid \beta)$, where the function f belongs to the class of mixed autoregressive and moving average (*ARIMA*) models described by Box and Jenkins [1970], θ is a set of parameters that define the exact form of f, and β is a set of model coefficients.

For *ARIMA* models, θ consists of three nonseasonal parameters, p, d, and q; three seasonal parameters P, D, and Q; and a seasonality factor.[2] The latter equals four quarters for quarterly earnings data, and hence the identification problem for earnings models is one of specifying the values for the remaining six parameters. The estimation problem is to obtain the coefficient set β corresponding to the parameter set θ, usually estimated by minimizing a loss criterion such as the sum of squared residuals (given θ and X).

The Box-Jenkins procedure used in accounting studies does not rely on a model or theory to specify θ. Instead, it relies on an iterative search, in which an initial guess for θ is made using X, β is estimated (given θ and X), and then θ is respecified until the model satisfies some diagnostic tests. Note that if we are dealing with earnings data, the information from the series is used first as a substitute for a theory of the firm, in that

[2] The parameter p refers to the number of nonseasonal autoregressive terms used in the model, d refers to the number of nonseasonal differencing performed on the original data, and q refers to the number of nonseasonal moving average terms used in the model. The parameters P, D, and Q are the respective seasonal counterparts of these.

the data are used to specify θ, and then it is used in the traditional role of estimating the coefficients, β. Despite the use of diagnostic tests, both the identification of θ, given X, and the subsequent respecifications require analysts' judgments.

Akaike [1974] introduced the notion that one can treat identification of θ, given X, as an estimation problem. Just as the values of β are found by minimizing a loss criterion (given θ and X), the values of θ can be found by maximizing an information criterion. To implement this idea, consider a large set of models belonging to the model class, f. Among these models, the one which maximizes an information criterion is the model that fits the data best, subject to validation or diagnostic tests on the residuals from the model.

Akaike's information criterion (AIC) is based on the Kullback-Leibler information measure (see Kullback [1959]), which defines the discrimination between the true distribution $g(X)$ and the assumed distribution $f_\theta(X \mid \beta)$ as:

$$I(g, \theta) = \int g(X) \log g(X)\, dX - \int g(X) \log f_\theta(X \mid \beta)\, dX. \tag{1}$$

When $f_\theta(X \mid \beta) = g(X)$, the information measure I is zero. Thus, choosing a model can be formulated as the estimation problem of choosing $\hat{\theta}$ to minimize $I(g, \theta)$. Alternatively, since the first term of $I(g, \theta)$ in (1) is a constant, the second term must be maximized. Akaike shows that the average log-likelihood of the sample can be substituted for the second term. Using second-order approximation to the log-likelihood, he then shows that the problem reduces to minimizing:

$$AIC(\hat{\theta}) = -2 \log \text{(maximum likelihood)} + 2m, \tag{2}$$

where $m(= p + q + P + Q)$ is the total number of coefficients in $\hat{\beta}$ corresponding to $\hat{\theta}$. Under the standard assumption of normally distributed residuals, (2) is equivalent to minimizing:

$$AIC(\hat{\theta}) = N \log s^2 + \frac{N}{N - 4D - d} 2m, \tag{3}$$

where N is the total number of observations, $N - 4D - d$ is the available data for estimation after differencing (assuming a seasonality factor of four quarters), and s^2 is the estimate of the residual variance.[3]

The minimization criterion (3) results in the Akaike procedure favoring parsimonious models, that is, models with small parameter values (usually $p, q \leqslant 2$ and $P, Q, d, D \leqslant 1$), unless an increase in AIC due to an increase in m or the differencing terms is more than offset by a reduction in AIC due to log s^2. Hence, if a model is selected by minimizing AIC over a model set that includes all estimable parsimonious models, that same model should also minimize AIC over an expanded model set that

[3] See Cohen and Peles [1977*a*; 1977*b*].

includes both parsimonious and other models. In this sense, the Akaike procedure is said to produce unique identification of a parsimonious model for a given sample of observations.[4]

Despite its reliance on the Kullback-Leibler measure, the Akaike procedure does not, unfortunately, have the theoretically desirable property of consistency in large samples. As shown by Shibata [1976] for the case of pure autoregressive processes and by Hannan [1980] for general *ARMA* models, the Akaike procedure will tend to identify with a probability greater than zero a model having too many parameters compared to the true model, even in larger samples. However, as will be seen later, estimable models of accounting earnings are usually very parsimonious, and hence the above direction of bias of the Akaike estimates of the number of model parameters may be tolerable. More recently, criteria very similar to *AIC* have been proposed that do have the consistency property. An example is *BIC*, analyzed by Akaike [1977] and Rissanen [1978]. I focus, however, on the *AIC* measure in the rest of the paper because of its more widespread use at the present time.

As indicated earlier, the Akaike procedure uses the entire observation set X for both model identification—i.e., minimization of $AIC(\hat{\theta})$—and model estimation—i.e., estimation of $\hat{\beta}$ corresponding to $\hat{\theta}$. This also holds for the Box-Jenkins procedure, but not for the Predictability procedure. The latter uses only a subset of X for estimation and the remainder of the data for model selection, thus partitioning the observation period from which X was obtained into a model estimation period and a model selection period. Let Y be the set of observed earnings data of the model estimation period, and Z the rest of the observations constituting the selection period. Let $\hat{Z}_\theta$ be the forecasts for the selection period obtained from the assumed model $f_\theta(Y \mid \hat{\beta})$.[5] Then the Predictability procedure selects a model from a given set of models that minimizes a forecast error function $h(Z, \hat{Z}_\theta)$. The assumption behind this procedure is that the forecast error function, h, is a measure of information on how "far" the assumed model is from the "true" model. The procedure assumes, in addition, that the true model has the lowest value of h.

Three problems arise in implementing this procedure: defining the model set, the error function, and the partitioned sets Y and Z.

Regarding the first, the Predictability procedure does not necessarily favor parsimonious models over other models. That is, the model selected by this procedure may change if the initial model set over which h is minimized is expanded to include more nonparsimonious models. In practice, however, this has not been a problem. For typical sample sizes of quarterly earnings data of about 100 observations, it is usually impossible to obtain significant coefficient estimates for a model which is not

---

[4] This procedure has been applied to a variety of data. See Tong [1975; 1976; 1977] and Ooe [1978]. Akaike [1974] applied the procedure to the time-series data of Anderson [1971].

[5] Note that β is estimated using Y, rather than the full observation set X.

parsimonious. Hence, by defining the initial model set to be large enough to include all estimable parsimonious models, one can ensure that the model selection by the Predictability procedure is unique.

As for the second, different specifications of the form of the error function $h(Z, \hat{Z}_\theta)$ can lead to different identified models. In addition, while the assumption that the true model has the lowest value of h holds when h is defined as the mean-squared forecast error,[6] it may not hold for other forms of h.

Finally, the total information set, X, must be allocated to two mutually exclusive sets, Y and Z, for model estimation and selection, respectively. Different partitionings can result in different identified models.

Despite these problems, predictability criteria have often been employed in the accounting and economics literature to compare, evaluate, or select models. For example, Foster [1977] used a predictability criterion to select a premier quarterly earnings model. Similarly, Brown and Rozeff [1978] compared the earnings model implied in *Value Line* forecasts with one selected by the Box-Jenkins procedure using a criterion of predictability. And recently, the economic and econometric literature has elevated to a paradigm the Wiener-Granger concept of causality based on predictability.[7]

In short, the range of past applications of the predictability criterion is quite extensive.[8] Nevertheless, I should emphasize that the Predictability procedure as described here is "myopic." Of the total observation set, X, only the information in the subset Z is used for model selection. As is true in many applications, most of the data in X are retained in set Y for estimation for reasons of estimation efficiency. As a result, only a few observations are usually left in the set Z for model selection. At an extreme, Z could consist of just one observation (the most recent quarter). By contrast, since both the Akaike and Box-Jenkins procedures use the larger information in X for model selection, one would expect the Akaike and Box-Jenkins model selection procedures to be superior to the Predictability procedure in a task such as forecasting beyond the model selection period. The extent of this superiority, however, is an empirical issue. This is addressed in the subsequent sections.

3. *Design of the Empirical Study*

The Akaike and Predictability procedures were tested in an empirical study with a sample of 30 firms. The test had two objectives: (*i*) to identify an *ARIMA* model structure for quarterly earnings using each procedure, and (*ii*) to compare the predictive performance of the models selected by the two procedures. Since the total model set considered here

[6] See Pindyck and Rubinfeld [1976].

[7] See Granger [1969] for the original essay and Sargent [1979] for a discussion of various applications of this use of predictability.

[8] See, however, Simon [1979] and Zellner [1978] for a critique of this criterion.

included the earnings models selected by other researchers using the Box-Jenkins procedure, objective (*ii*) indirectly included the evaluation of the predictive performance of models from the Box-Jenkins procedure also.

The five major elements in the design of the empirical study consisted of the selection of firms, definition of the initial model set, definition of error measures, definition of estimation and model selection periods, and a model identification strategy.

FIRMS

Firms were selected from the Moody's *Handbook of Common Stocks* subject to the following conditions: (*i*) the firm's quarterly earnings data were available continuously for the period 1950/first quarter to 1980/fourth quarter in the Moody's *Handbook* series, (*ii*) the firm's fiscal year ended in December, and (*iii*) the firm's quarterly earnings for the 16 quarters in 1977–80 were positive.

These conditions are similar to the ones used by Brown and Rozeff [1978]. Condition (*iii*) is required so as not to distort the comparisons of the error measures (described below), some of which are based on the ratio of forecast error to actual quarterly earnings. Table 1 lists the 30 selected firms. All but one of the firms in table 1 are in the 1980 "Fortune 500," the smallest of these having a rank of 252. The bias toward large size in the sample selection is due to condition (*i*).

MODEL SET

The Akaike procedure selects a model for a firm from a given set of models by minimizing *AIC*, and the Predictability procedure selects a model from the given set of models by minimizing an error measure. As noted in section 2, all parsimonious models must be included in the model set to ensure that the model selections by the procedures are unique. In this study, 56 parsimonious *ARIMA* models constituted the "given" model set, having $p \leq 2$, $q \leq 2$, $D = 1$, and d, P, and $Q \leq 1$. Models more complex than these were excluded due to the difficulty of obtaining significant coefficient estimates.[9] Furthermore, models with seasonal differencing $D = 0$ were excluded since quarterly earnings of every firm in the sample showed marked seasonality. Note that, except for this omission, there are no "holes" in the parameter space defined above. Researchers should beware of using the *AIC* or the Predictability procedure on just some values of the parameters in the above range and then selecting the optimal parameter values. In other words, it is not advisable to look at nonnested subsets of the parameter space.

For ease of analysis, the 56 models were assigned to seven model

[9] Even with these 56 models, only about 70 percent could be estimated with significant coefficients for most firms. Thus, expansion of the model set to include more models would not have altered much the results on identification.

TABLE 1
List of Sample Firms

No.	Name	1950 Sales ($ millions)	1980 Sales ($ millions)	Industry (major product)
1.	Abbott Laboratories	74	2,038	Pharmaceuticals
2.	Allis-Chalmers Corporation	344	2,064	Industr. machinery
3.	American Cyanamid Company	322	3,454	Chemicals
4.	Armstrong World Industries	187	1,323	Textiles, flooring
5.	Boeing Company	307	9,426	Aerospace
6.	Borg-Warner Corporation	331	2,673	Auto accessories
7.	Bristol-Myers Company	52	3,158	Pharmaceuticals
8.	Caterpillar Tractor Company	337	8,598	Constrn. machinery
9.	Celanese Corporation	233	3,348	Chemicals
10.	Coca-Cola Company	215	5,913	Beverages
11.	Continental Group, Inc.	398	5,119	Glass containers
12.	Corning Glass Works	117	1,530	Glass, concrete
13.	Curtiss-Wright Corporation	136	228	Industr. machinery
14.	Eaton Corporation	148	3,176	Auto accessories
15.	Fruehauf Corporation	128	2,082	Auto accessories
16.	General Electric Corporation	2,233	24,959	Power systems
17.	Gillette Company	99	2,315	Metal products
18.	Gulf Oil Corporation	1,150	26,483	Oil and gas
19.	IBM Corporation	215	26,213	Computers
20.	International Paper Company	498	5,043	Paper, wood
21.	PepsiCo, Incorporated	40	5,975	Beverages
22.	Philip Morris, Incorporated	306	7,328	Tobacco
23.	Philips Petroleum Company	533	13,377	Oil and gas
24.	RCA Corporation	584	8,011	Entertainment
25.	R. J. Reynolds Industries, Inc.	758	8,449	Tobacco
26.	Standard Brands Incorporated	301	3,018	Food
27.	Sterling Drug Incorporated	139	1,701	Pharmaceuticals
28.	The Timken Company	144	1,338	Industr. machinery
29.	Union Carbide Corporation	758	9,994	Chemicals
30.	Westinghouse Electric Corp.	1,020	8,514	Power systems
	Average Sales	404	6,862	

TABLE 2
ARIMA Representations of the Model Classes Examined

Model Class	Models	P	D	Q	d
1	1–8	0	1	0	0
2	9–16	0	1	0	1
3	17–24	0	1	1	0
4	25–32	0	1	1	1
5	33–40	1	1	0	0
6	41–48	1	1	0	1
7	49–56	1	1	1	0

Note: The eight models in each class have the (p, q) values of (0, 1), (0, 2), (1, 0), (2, 0), (1, 1), (1, 2), (2, 1), and (2, 2).

classes, with each model class having a unique set of P, D, Q, and d values.[10] Table 2 shows the model class assignment. The eight models in each class share the P, D, Q, d values, but differ in their (p, q) values.

ERROR MEASURES

Four different error measures were used to implement the Predictability procedure. Two of these, mean-squared percentage error ($MSPE$) and mean absolute percentage error ($MAPE$), have been used often in past accounting studies (see Foster [1978]). The other two error measures used are mean-square error (MSE) and mean percentage error (MPE). To define the four measures, let q_t be the actual earnings in period t in the model selection period, $\tilde{q}_t$ be the earnings predicted from a model, and n be the number of quarters in the model selection period (i.e., the number of elements in the set Z). Then:

$$(a)\quad MSE = \frac{1}{n}\sum_{i=1}^{n}(\tilde{q}_i - q_i)^2,$$

$$(b)\quad MSPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\tilde{q}_i - q_i}{q_i}\right)^2,$$

$$(c)\quad MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\tilde{q}_i - q_i}{q_i}\right|,$$

$$\text{and}\quad (d)\quad MPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\tilde{q}_i - q_i}{q_i}\right).$$

Measures (b), (c), and (d) are based on the ratio (or the percentage) of the forecast error to the actual earnings. These different forms correspond to three different ways of measuring the cost of the forecast errors. Measure (a) is normally used in conventional estimation procedures

[10] Additional justification for the use of model classes is given later in this section.

which require minimization of *MSE* in the estimation period. Here I extended it to the selection period as well.

ESTIMATION AND SELECTION PERIODS

Quarterly earnings data of the period 1950/first quarter to 1978/fourth quarter (116 data) constituted the observation set X described in section 2. As noted there, the Akaike procedure uses the entire set X for both model estimation and model selection. For the Predictability procedure, however, I partitioned the 116 observations into set Y for estimation—1950/first quarter to 1976/fourth quarter—and set Z for selection—1977/first quarter to 1978/fourth quarter.

MODEL IDENTIFICATION STRATEGY

The model selection strategy in the case of the Akaike and Predictability procedures was simple: for each firm in the sample, estimate all 56 *ARIMA* models and select the one with the lowest value of the given selection measure, provided the residuals from the model satisfy the usual diagnostic tests and provided the coefficients are significant. It turns out in practice, however, that if a model in a particular class has the lowest value of a selection measure, say *AIC*, other models belonging to the same class often have the next lowest *AIC*, third lowest *AIC*, etc. In order to avoid letting minor variations in observation set affect the model identification process, the identification strategy was altered to select a model class, rather than a specific model, of the earnings series. The strategy adopted for model class selection was to find the six models with significant coefficients that had the lowest value of a given selection measure for each firm, and in turn find the number of times each model class appeared when this "six best" list is added up for all firms.[11] The identified model class is the one that appears most often in this list. This method yields identification of a "premier" model, that is, a single model applicable to the entire sample.

4. *Results*

I consider first the results on model class identification based on the strategy of compiling model class appearances in "six best" lists. For a given selection criterion, the 30 firms and the "six best" models for each firm together result in 180 model citations. If the identification procedure has no discriminatory power among the models, one would expect each model class in table 2 to appear about 180/7 or 26 times in the compiled "six best" list. On the other hand, if the number of appearances of the

[11] With 56 models, six models represent the top 10 percent category. The choice is arbitrary, but the rankings in table 3, explained below, are not very sensitive to it. The rankings under the *AIC* criterion, for example, were essentially unaltered when "ten best" models were used.

model classes in the "six best" list is not uniformly distributed, the identification procedure has discriminatory power, and the model class that appears most often in the list is the sample-wide representative model class chosen by that identification procedure.

Table 3 presents the results on model class appearances for each of the five selection criteria. Consider first the results using the *AIC* criterion. The distribution of model class appearances in the "six best" list is clearly not "uniform,"[12] and the *AIC* criterion led to a clear selection of model class 3. In contrast, the four predictability criteria unanimously selected model class 2, even though the four criteria behaved differently in discriminating between the model classes. That is, the "uniform distribution" assumption could not be rejected for the *MSPE* criterion, but it could be rejected at the .1 level for the *MSE* and *MAPE* criteria, and .001 level for the *MPE* criterion.[13]

To see whether the Predictability procedure was sensitive to the choice of the error measure, the degree of agreement among the four error criteria in ranking the seven model classes was examined using Kendall's W statistic.[14] Using table 3 data, the computed W was 0.8705, indicating a high degree of association between the four rankings. Using a procedure described by Siegel [1956] to test the significance of this association, the null hypothesis that the four sets of rank orders were independent was clearly rejected at the .01 level.[15] In other words, the rank ordering of the seven model classes by the Predictability procedure was fairly insensitive to the choice of error measure among the four considered.

Next, I considered whether the Akaike and Predictability procedures led to dissimilar ranking of model classes. For this test, I computed Kendall's tau statistic for each of the ten possible pairs of rank orders.[16] The results indicate that the null hypothesis that two given sets of rankings are independent cannot be rejected for the four pairs involving *AIC* and an error criterion, but can be rejected for five of the six pairs involving error criteria at the .05 level.[17] Thus, the tau test indicated that the Akaike and Predictability procedures led to dissimilar choice and ranking of model classes, and in addition confirmed the finding from the

[12] The null hypothesis that the distribution of model class appearances is "uniform" can be tested by computing the Pearson χ^2 goodness-of-fit statistic. For the *AIC* criterion, $\chi^2 = 106.92$, which is significant at the .001 level. Hence the null hypothesis can be rejected.

[13] The computed χ^2 were: *MSE*, 12.27; *MSPE*, 8.53; *MAPE*, 11.02; and *MPE*, 37.99.

[14] Also called the coefficient of concordance. See Hays [1973] or Siegel [1956].

[15] Let R_j be the sum of the ranks of model class j, where $j = 1, \ldots 7$. Let $\bar{R}$ be the mean of R_j. Then the distribution of the variable $s = \sum_{j=1}^{7} (R_j - \bar{R})^2$ under the null hypothesis of independent rank orders is given in table R of Siegel [1956]. From this table, the .01 level cutoff value of s for four error measures and seven model classes is 265. The computed s for the table 3 data was 386.5. Hence the null hypothesis could be rejected.

[16] This statistic is sometimes called the coefficient of disarray. It is believed to be superior to the Spearman rank correlation statistic. See Siegel [1956].

[17] The lone exception was the *MSE*–*MPE* pair. For this pair, the null hypothesis could be rejected at the .06 level.

TABLE 3

Model Class Appearance in "Six Best" under Each Selection Criterion

Model Class	Model Class Representation				*AIC*		*MSE*		*MSPE*		*MAPE*		*MPE*	
	P	*D*	*Q*	*d*	Number	Rank	Number	Rank	Number	Rank	Number	Rank	Number	Rank
1	0	1	0	0	8	7	30	3	32	2	33	2	21	4
2	0	1	0	1	14	4	37	1	34	1	36	1	43	1
3	0	1	1	0	67	1	25	4	23	5	25	4	19	5
4	0	1	1	1	36	2	18	6½	24	4	23	5	35	3
5	1	1	0	0	12	5	18	6½	19	6½	21	6	15	6
6	1	1	0	1	10	6	31	2	29	3	26	3	37	2
7	1	1	1	0	33	3	21	5	19	6½	16	7	10	7

Note: This summary is based on the results of 30 firms. *AIC* was estimated with 1950–78 quarterly data. The other four criteria were computed for the 1977–78 data.

W test that the Predictability procedure was insensitive to the choice among the four error measures.

The sample-wide model class selections by the Akaike and Predictability procedures can be contrasted with those identified in other accounting studies. The four most cited results are summarized below:

Author(s)	Firms in Sample	No. of Quarters	Model $(pdq)(PDQ)_4$	Model Class
Brown and Rozeff [1979]	50	100	$(1, 0, 0)\ (0, 1, 1)_4$	3
Foster [1977]	69	64	$(1, 0, 0)\ (0, 1, 0)_4$	1
Griffin [1977]	94	56	$(0, 1, 1)\ (0, 1, 1)_4$	4
Watts [1975]	175	18–50	$(0, 1, 1)\ (0, 1, 1)_4$	4

Model classes 3 and 4 were ranked respectively 1 and 2 by the Akaike procedure, but model class 1 was ranked highly by the Predictability procedure. Recall that the other studies selected premier model classes using the iterative Box-Jenkins procedure.

The 67 appearances of the model class 3 under the *AIC* criterion were examined further to determine which model within this class appeared in the "six best" list more frequently. This distribution is given below:

Model Number	17	18	19	20	21	22	23	24
Model (p, q)	(0, 1)	(0, 2)	(1, 0)	(2, 0)	(1, 1)	(1, 2)	(2, 1)	(2, 2)
Appearances	3	6	16	16	12	5	7	2

The null hypothesis that the above distribution of model appearances is "uniform" can be rejected at the .001 level.[18] It appears that model 19—the Brown-Rozeff model—and model 20 cause model class 3 to perform best under *AIC*. These two models, then, are the premier models suggested by the Akaike procedure. In contrast, the Watts-Griffin identified model—17—performed poorly in this study.

I now consider the results on forecasting, which was the second objective of this study. For each of the 30 firms, the model with the lowest *AIC* was selected and estimated using the 1950–78 earnings data. Next, the models with the lowest error measure under each of the four error criteria were selected using the 1977–78 earnings data and then reestimated with the 1950–78 data.[19] Using the lowest *AIC* model and the four lowest error models, one-step ahead earnings forecasts for the eight quarters in 1979–80 (the forecast period) were generated and compared to actual quarterly earnings data for this same period. The *MSPE* and *MAPE* measures were calculated for the five models and ranked for the 30 firms. The resulting mean ranks (with 1 being the best) of the five

[18] Pearson χ^2 value is 26.02. See n. 12.

[19] With this reestimation, all models were estimated with the same 1950–78 data, thus avoiding a source of bias for the forecasting tests.

models under each of the two forecast error measures were as follows:

Model	Mean Ranks Based on *MSPE*	*MAPE*
AIC	2.32	2.48
MSE	2.89	2.82
MSPE	3.13	3.15
MPE	3.20	3.22
MAPE	3.47	3.33

The models selected by *AIC* forecasted better than the models selected by the four predictability criteria under both forecast error measures. Moreover, both error measures produced identical rankings for the five models, although the dispersion among the mean ranks was small in each case. To measure the statistical significance of these results, I first tested whether the mean ranks of the four Predictability models were different from one another, using Friedman's S statistic, which is distributed approximately as a χ^2 for large samples. This test could not reject the hypothesis that the mean ranks (under either of the two forecast error measures) of the four Predictability models were equal. I then tested whether the mean rank of the *AIC* model was different from the mean ranks of the four Predictability models, using a statistic related to Friedman's S statistic, as described by Hollander and Wolfe [1973]. This text examined all the ten possible model pairs and found that the null hypothesis that two given mean ranks (under either of the two forecast error measures) were equal could not be rejected for any model pair except the *AIC–MAPE* pair.

These two tests imply that the models selected by the *AIC* procedure and the four versions of the Predictability procedure have similar forecasting efficiency. More important, the paired comparisons indicated that the models selected by the *AIC* procedure did not have significantly superior forecasts compared to models selected by the Predictability procedure. In other words, the mean rank superiority of the *AIC*-based models shown above was not statistically significant. Nevertheless, the Akaike procedure has two distinct advantages over the Predictability procedure: (*i*) it is theoretically based, and (*ii*) it uses a larger information set than the Predictability procedure for model selection.

5. Conclusion

In this study, I compared two noniterative identification procedures which can be used to identify quarterly earnings models. Subject to proper design of the initial model set, both procedures provided unique parsimonious earnings model identifications for a given firm. The first, the Akaike procedure, selects unique parsimonious models by design and, being a noniterative procedure, is easy to implement. Moreover, it is theory-based to the extent that one agrees with its notion of sample information, though it does not have the theoretically desirable property of consistency. These characteristics make the Akaike procedure a serious alternative to the Box-Jenkins procedure normally used by accounting researchers.

The Predictability procedure also has strong proponents (and critics). Though a "myopic" procedure by design, it nevertheless is based on the popular notion that the forecast errors of a model measure how far that model is from the true model. The procedure, however, is more difficult to implement than the Akaike procedure, since the user must make critical decisions on partitioning the available data set into one set for estimation and one for model selection. Still, it usually produces unique model identification, also making it attractive to accountants.

The empirical study presented in the second part of this paper had two major results. One established that the Akaike procedure produced the same sample-wide earnings models as reported by other researchers using the iterative Box-Jenkins procedure. The other result was that the forecast errors of the models selected by the Akaike and Predictability procedures were surprisingly similar, even though the former seemed to exhibit some superiority in forecasting. The latter result parallels the results in management science, where "myopic" or mechanical heuristics are often found to have performance characteristics comparable to those of complex algorithms.

Given a choice between Akaike, Box-Jenkins, and the Predictability procedures, which should one use to forecast earnings? The answer naturally depends on the various cost and benefits of implementing each procedure, which have not been addressed here. One relevant factor in such a decision is the fact that model selection by the Box-Jenkins procedure requires the analyst to make decisions during data analysis about parameter specification. In contrast, the other two procedures require only that the analyst make practical decisions on error measures, selection periods, model set, etc. prior to data analysis. This factor, together with the two empirical results summarized above, suggests that the Akaike and Predictability procedures may be attractive alternatives to accounting researchers for earnings model specification at the firm level, particularly if the objective is to determine unique model representations for corporate earnings for forecasting applications.

REFERENCES

ABDEL-KHALIK, A. R., AND R. B. THOMPSON. "Research on Earnings Forecasts: The State of the Art." *The Accounting Journal* (1978): 180–209.

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

———. "On Entropy Maximization Principle." In *Applications of Statistics*, edited by P. R. Krishnaiah. Amsterdam: North-Holland, 1977.

ANDERSON, T. W. *The Statistical Analysis of Time Series*. New York: Wiley, 1971.

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. "The Superiority of Analyst Forecasts as Measures of Expectations: Evidence from Earnings." *Journal of Finance* (March 1978): 1–16.

———. "Univariate Time Series Models of Quarterly Earnings Per Share: A Proposed Premier Model." *Journal of Accounting Research* (Spring 1979): 179–89.

COHEN, A., AND Y. PELES. "Utilizing Akaike's Criterion for Iterative Fitting of ARIMA Models." Working paper, O.R., Statistics and Economics Mimeograph Series, no. 195, Industrial and Management Engineering, Technion, June 1977*a*.

———. "A Moving ARIMA Model for Better Forecasts—A Case Study." Working paper, O.R., Statistics and Economics Mimeograph Series, no. 198, Industrial and Management Engineering, Technion, June 1977*b*.

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.

GRANGER, C. W. J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* (1969): 424–38.

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

HANNAN, E. J. "The Estimation of the Order of an ARMA Process." *Annals of Statistics* (September 1980): 1071–81.

HAYS, W. L. *Statistics for the Social Sciences*. New York: Holt, Rinehart & Winston, 1973.

HOLLANDER, M., AND D. A. WOLFE. *Nonparametric Statistical Methods*. New York: Wiley, 1973.

KULLBACK, S. *Information Theory and Statistics*. New York: Wiley, 1959.

OOE, M. "An Optimal Complex ARMA Model of the Chandler Wobble." *The Geophysical Journal* (1978): 445–57.

PINDYCK, R. S., AND D. L. RUBINFELD. *Econometric Models and Economic Forecasts*. New York: McGraw-Hill, 1976.

RISSANEN, J. "Modeling by Shortest Data Description." *Automatica* (1978): 465–71.

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

SHIBATA, R. "Selection of the Order of an Autoregressive Model by Akaike's Information Criterion." *Biometrica* (1976): 117–26.

SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.

SIMON, H. A. "Rational Decision Making in Business Organizations." *American Economic Review* (September 1979): 493–513.

TONG, H. "Autoregressive Model Fitting with Noisy Data by Akaike's Information Criterion." *IEEE Transactions on Information Theory* (July 1975): 476–80.

———. "Fitting a Smooth Moving Average to Noisy Data." *IEEE Transactions on Information Theory* (July 1976): 493–96.

———. "More on Autoregressive Model Fitting with Noisy Data by Akaike's Information Criterion." *IEEE Transactions on Information Theory* (May 1977): 409–10.

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

ZELLNER, A. "Causality and Econometrics." Paper presented at the Rochester-Carnegie-Mellon Conference on Public Policy, April 1978.