Journal of Accounting Research
Vol. 20 No. 2 Pt. II Autumn 1982
Printed in U.S.A.

Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation

ROBERT N. FREEMAN*, JAMES A. OHLSON, AND STEPHEN H. PENMAN†

1. Introduction

Over the years, there has developed a fairly substantial body of research on the time series of earnings. As a whole, this literature concludes that changes in (annual) accounting earnings are unpredictable, that is, earnings follow a "random walk."[1] Based on this result, some inferences of economic substance (policy) have been claimed.[2] In this paper we reconsider empirical issues which, at least to some extent, have been obscured by this conclusion. We argue that the above result is only true in a limited sense, since a modest enlargement of the predictive information set should allow for a rejection of the hypothesis that earnings changes are

* Project Manager, FASB; † Professor, and Assistant Professor, University of California, Berkeley. We are indebted to Bill Beaver and members of the Vanderbilt Accounting Workshop, especially Kenneth Gaver, for helpful comments. [Accepted for publication May 1982.]

[1] Precisely, the consensus is that earnings follow a martingale (possibly with drift) so that, conditioned on past earnings, earnings changes (net of drift) are drawn from a distribution with mean zero. This has loosely been referred to as a random walk in earnings. It has been claimed to be descriptive of both undeflated earnings and earnings per share. See Little [1962]; Little and Rayner [1966]; Brealey [1967]; Lintner and Glauber [1967]; Lintner [1969]; Ball and Watts [1972]; Gonedes [1973]; Albrecht, Lookabill, and McKeown [1977]; and Watts and Leftwich [1977].

[2] Ball and Watts [1972] infer that attempts to smooth corporate income cannot be successful and provide interpretations of growth of corporate earnings based on their findings (as do the forerunner papers). For a general discussion of the importance of earnings predictions models, see Foster [1977].

unpredictable. Specifically, we hypothesize that book rate-of-return predicts earnings changes. If this is so, past inferences based on the "random walk hypothesis" are incorrect.

There are two empirical results which together would suggest that this alternative hypothesis would be more descriptive: (*i*) book rates-of-return themselves follow a mean-reverting process; (*ii*) changes in rates-of-return correlate strongly with changes in earnings. Our evidence unambiguously supports the subhypotheses (*i*) and (*ii*), and, to a lesser extent, the major hypothesis. Hence, current book rate-of-return provides a basis for predicting future earnings changes. A relatively low rate-of-return implies that earnings are "temporarily depressed"; similarly, a high rate-of-return implies that earnings are "unusually good." The evidence thus suggests that, while the "random walk hypothesis" is quite robust with respect to past earnings, more successful predictions can be made by expanding the conditioning information set to include book value of net assets.

2. *The Time-Series Behavior of Earnings: Some Preliminary Considerations*

It is commonly suggested in the accounting-finance literature that the time series behavior of earnings are well approximated by a "random walk" model, that is, changes in earnings cannot be predicted. Even though much of the empirical research supports this hypothesis, there are reasons why on an a priori basis we might expect this to be a rather poor approximation. One is simply the existence of vast differences in Price/Earnings ratios (henceforth P/E ratios) across firms at any given point in time. Similarly, the P/E ratio of any given firm typically has substantial time-series variability. This suggests that if future expected earnings are of importance in security valuation, then firms with high P/E ratios would be expected to have relatively high expected increases in earnings compared to those firms with relatively low P/E ratios. The argument is consistent with existing empirical evidence. Beaver and Morse [1978] showed that P/E ratios computed at year-end are positively correlated with subsequent years' growth in earnings.[3] However, as discussed in Beaver, Lambert, and Morse [1980], this result in itself does not mean that the random walk hypothesis can be rejected, since expected earnings differ from current earnings because forecasts of future (expected) earnings are based on an information set which is potentially much larger than that of current and past earnings.[4] Rational investors

[3] It is tempting to suggest that differences in risk explain differences in P/E ratios. However, the explanatory power of such variables appears to be extremely modest. The issue is considered empirically in Beaver and Morse [1978].

[4] There is evidence that past and current earnings are a relatively small part of the information set used to set security prices. Not only do analysts' forecasts of earnings consistently outperform Box-Jenkins identified time-series models (Brown and Rozeff

presumably employ the broadest possible information sets, and these must be reflected in observed prices. Information from accounting statements, other than earnings, may be included in these "global" sets of information. However, if the random walk hypothesis does not hold on the global information sets, the same could be true for a narrower, accounting-based information set. A concrete example which comes to mind is book rate-of-return (i.e., net earnings divided by shareholders' equity, *ROR* henceforth).

In addition to the existence of differential P/E ratios, there is another reason we would expect earnings changes to correlate with *ROR*. Two conditions, both of which are empirically plausible, imply that current *ROR* predicts changes in earnings. First, *ROR* is governed by a mean-reverting process, that is, $ROR_{t+1} - ROR_t$ is negatively correlated with ROR_t. Although it appears that this proposition has not been formally examined, the evidence provided in Beaver [1970] and Lookabill [1976] supports the hypothesis.[5] Second, changes in book rates-of-return are correlated with changes in earnings. (In the extreme case of unchanging common equity—i.e., no sale/purchase of shares and dividends equal to earnings—the correlation is perfect.) If these two conditions are met, we would expect earnings changes to be predictable.

To formalize these ideas, consider the process:

$$ROR_t = \delta + \gamma ROR_{t-1} + \epsilon_t \tag{1}$$

where ROR_t is the book rate-of-return on common equity; ϵ_t, ϵ_s are independent if $s \neq t$; $E(\epsilon_t) = 0$; $\text{Cov}(ROR_{t-1}, \epsilon_t) = 0$; and $0 \leqq \gamma \leqq 1$. If ROR_t is a random walk, $\gamma = 1$; if ROR_t is pure mean reversion, $\gamma = 0$. Prior empirical works suggests that ROR_t is a hybrid with $\delta > 0$ and $\gamma < 1$.

If ROR_t follows the process in (1), what may be implied about the forecasts of change in income in year t given ROR_{t-1}?
Define

$I_t \equiv$ accounting income in year t; $CE_t \equiv$ common equity at the end of year t; and $\Delta I_t \equiv I_t - I_{t-1}$.

[1978]; Crichfield, Dyckman, and Lakonishok [1978]), but earnings announcements account for less than half of the variance in forecast revisions (Brown and Rozeff [1979]).

In this context we might further note that in discussing the random walk hypothesis it is important that the conditioning information set is identified; this has been emphasized by Beaver, Lambert, and Morse [1980]. Thus, the general statement is that:

$$E[\Delta \tilde{l}_{t+1} \mid \phi_t] \quad \text{is independent of } \phi_t$$

where $\Delta l_{t+1} = l_{t+1} - l_t$ is the change in earnings and ϕ_t is the conditioning information set. The "pure" random walk hypothesis then puts ϕ_t equal to $l_t, l_{t-1} \ldots$. The other extreme is the case in which ϕ_t is the total public information set used by the market.

[5] Beaver [1970] divided his sample firms into portfolios of high and low ROR_t. He found that ROR_{t+1} decreases (increases) for the high (low) ROR_t portfolios for $i = 1, \ldots, 8$. Lookabill [1976] shows that shifts in risk (β from the *CAPM*) do not account for Beaver's results.

Then

$$
\begin{aligned}
ROR_t &= I_t/CE_{t-1}, \text{ and} \\
E_{t-1}(\Delta I_t) &= E_{t-1}(ROR_t \cdot CE_{t-1} - I_{t-1}) \\
&= (\delta + \gamma ROR_{t-1}) \cdot CE_{t-1} - ROR_{t-1} \cdot CE_{t-2} \\
&= \delta \cdot CE_{t-1} + (\gamma CE_{t-1} - CE_{t-2}) ROR_{t-1},
\end{aligned}
\tag{2}
$$

assuming that the disturbance term ϵ_t is uncorrelated with CE_{t-1}.

Thus, the expected change in earnings is inversely related to ROR_t if and only if:

$$\gamma CE_{t-1} < CE_{t-2}. \tag{3}$$

Expression (3) will be met in two extreme cases: (*i*) the process of ROR_t is purely mean-reverting, that is, $\gamma = 0$; (*ii*) the process of ROR_t has at least a slight tendency toward mean-reversion, that is, $\gamma < 1$, and, additionally, $CE_{t-1} = CE_{t-2}$. Neither (*i*) nor (*ii*) can be expected to hold exactly. The impact of alternative dividends and equity financing policies is ambiguous, and there appears to be no prior theoretical reason to rule out that some policies may preclude (or at least operate against) prediction of earnings changes on the basis of *ROR*. However, the general idea should be transparent in the sense that there are two factors affecting the hypothesis: the degree of mean-reversion in the ROR_t process, and the stability in the common equity base.[6]

The random walk hypothesis, no matter what its precise conditioning information set, cannot be strictly true. In point of fact, evidence provided by Brooks and Buckmaster [1976; 1979] indicates that the "random walk hypothesis" can be rejected, at least for certain firms under certain conditions.[7] The issue is not whether the hypothesis is fundamentally true in some absolute sense, but rather whether it is a "good" or "poor" (first-order) approximation. It is a "poor" approximation to the extent that the methodology is straightforward but still suffices to reject the null hypothesis. However, prior research suggests that there is a trade-off between methodological simplicity and chances of rejecting the null. Early work which employed pooled data and simple models (such as naive extrapolative models and autoregressive analysis) uniformly supported the random walk hypothesis for almost any earnings numbers (see, for instance, Ball and Watts [1972] and Gonedes [1973]). Later work either performed tests at the individual firm level (thus abandoning the goal of a general approximation) and/or introduced more sophisticated econometric techniques. The latter efforts have led to ambiguous evi-

---

[6] To be sure, note the "noise" which will be present in any empirical analysis. The term δCE_{t-1} does also affect the dependent variable. However, it is readily argued that CE_{t-1} and ROR_{t-1} must be uncorrelated, so the omission of the variable δCE_{t-1} in the econometric analysis should not lead to any biases.

[7] Brooks and Buckmaster [1976] found that "income tends to revert to previous levels in the period subsequent to a substantial deviation from an operationally defined norm."

dence (see, for example, Albrecht, Lookabill, and McKeown [1977] and Watts and Leftwich [1977]). In reviewing the literature, it appears fair to state that no *transparent* modeling and econometric procedure allows the data to reveal that the null ought to be rejected.

It is evident what the basic econometric problem is and why the null is generally not rejected. The variability in the (first) differences of earnings is too large to be compared to the variability in the expected differences in earnings.[8] Thus, the null is difficult to reject because of the substantial noise relative to the amount of data that is typically available. The problem is amplified in that the distribution of differences in earnings is not normal, and the presence of a few outliers can swamp the analysis. It follows that the null will generally not be rejected unless the basic hypothesis is modified. One way is to reduce the variability in the dependent variable (first difference in earnings) through an appropriately selected transform, such as the nonparametric indicator-variable: earnings increase versus decrease. The null hypothesis now is that the probability of an earnings increase (or decrease) is independent of the predictor *ROR*; the alternative is that the probability of an earnings decrease increases as *ROR* increases.[9]

3. A Probabilistic Model of Earnings Changes

In this section we present a probabilistic model of earnings changes that uses an indicator dependent variable. Let $\mathbf{X}_{it}$ denote a vector of descriptors for firm i at time t; let $\boldsymbol{\theta}_i$ be a vector of known parameters; and let $P(\mathbf{X}_{it}, \boldsymbol{\theta}_i)$ denote the probability of an earnings increase for the subsequent period for a given $\mathbf{X}_{it}$ and $\boldsymbol{\theta}_i$. The natural logarithm of the likelihood of any series of earnings increases (set S) and decreases (set S') is given by:

$$L(\boldsymbol{\theta}_i) = \sum_{t \in S} ln[P(\mathbf{X}_{it}, \boldsymbol{\theta}_i)] + \sum_{t \in S'} ln[1 - P(\mathbf{X}_{it}, \boldsymbol{\theta}_i)].$$

Hereafter the firm subscript i is omitted if not needed for clarity. For any function $P(\cdot)$, the maximum likelihood estimates of the elements of $\boldsymbol{\theta}$ are obtained by solving $\text{Max}_{\theta} L(\boldsymbol{\theta})$.

The hypothesis will be tested through use of the logistic distribution-

---

[8] In other words:

$$\text{Var}[E[\Delta l_{t+1} \mid \phi_t]] \quad \text{is small relative to} \quad \text{Var}[\Delta l_{t+1}].$$

[9] The use of the indicator-variable, earnings increase versus decrease, has another methodological advantage. A dollar and cents statement regarding the change in earnings (or *EPS*) means little, if anything, beyond the fact that there has been an increase (or decrease). The problem here is, of course, that one needs some deflation such that the "degree" of increase (decrease) will be reflected. In statistical terms, this translates into the fact that earnings changes are heteroscedastic depending upon some appropriately specified level variable. Unfortunately, this problem has no straightforward or apparent resolution, so the simplest approach is one which considers increases versus decreases.

function, that is:

$$P(y_t) = [1 + \exp\{-y_t\}]^{-1}, y_t \equiv \mathbf{X}_t'\boldsymbol{\theta},$$

the major motivation being its computational simplicity. In the context of the model under consideration one has $\mathbf{X}_t' = (1, ROR_t)$ and $\boldsymbol{\theta}' = (\alpha, \beta)$. Hence:

$$P \equiv Pr\{EPS_t > EPS_{t-1} \mid ROR_{t-1}\} = [1 + \exp\{-\alpha - \beta ROR_{t-1}\}]^{-1}$$

or, equivalently:

$$ln(P/1 - P) = \alpha + \beta ROR_{t-1}.$$

In the Logit framework the null hypothesis is $\beta = 0$, consistent with a random walk, versus the directional alternative $\beta < 0$, consistent with the idea that a low (high) *ROR* predicts an earnings increase (decrease).

The hypothesis is tested using firm-specific (*FS*) and pooled cross-sectional time-series (*PO*) data. The Logit procedure has considerable appeal over models with continuous dependent variables. Extreme changes in earnings have no more influence on the parameter estimates than other observations, and, further, the heteroscedastic disturbance problem associated with a continuous dependent variable technique is not present. This makes pooling of data across firms and time periods a far more plausible procedure than otherwise.

These statistical gains are not obtained without some related losses. As previously indicated, in choosing between a simple metric (increase vs. decrease) and a continuous metric of earnings change, there is an obvious trade-off between "information" and "noise." More extensive and sophisticated models may provide more definitive results, but only if they are not misspecified. If they are, their estimators will tend to be inefficient (and, perhaps, biased), and related t-statistics may be grossly misleading (i.e., overstated). For example, a linear model with the change in *EPS* as a dependent variable would be a more efficient approach than the Logit model if it were correctly specified. But this is likely to require the identification of an autoregressive parameter and a modeling of the heteroscedastic effects. The joint solution to these problems becomes quite difficult because of limited data, and unless it is done reasonably well the statistical output will be somewhat less than credible. Hence, we choose Logit and a binary metric on the assumption that this will lead to (relatively) more robust statistical conclusions.[10]

[10] Discriminant analysis provides another means of testing the hypothesis in the binary metric space, but the important disadvantage here is the requirement that the independent variable must be assumed to follow a normal distribution. The estimations which follow were repeated using *OLS* and continuous dependent variables. As expected, we observed high t-statistics without any improvement in predictive space.

4. Variable Definitions and Sample Selection

Parameter estimates were obtained for two different dependent variables:

Model 1 $Z(\Delta ROR_t)$ where $\Delta ROR_t = ROR_t - ROR_{t-1}$;
Model 2 $Z(\Delta EPS_t)$ where $\Delta EPS_t = EPS_t - EPS_{t-1}$.

$Z(x) = 1$ or 0 as x is positive or negative (nonpositive). Model 1 was estimated because *ROR* must follow a mean-reverting behavior—that is, $\beta < 0$ in Model 1—in order for there to be a basis for suggesting that *ROR* predicts earnings changes in Model 2. Both models were estimated using firm-specific data (*FS*) and data pooled from all firms (*PO*).

The sample was obtained by randomly selecting firms with complete earnings and equity series on the *Compustat Annual Industrial Tape*, 1972 version. All data were traced to *Moody's Industrial Manuals* to assure that post-reporting-date adjustments were excluded, that is, the earnings and equity series are "as reported." *Compustat* tapes and *Moody's Manuals* were used to extend the data series to cover 32 years: 1946–77. Since first differences are used as the argument of the dependent variable, all tests are based on 31 observations per sample firm.

The sample selection proceeded as follows: we randomly selected a firm with complete data on the 1972 *Compustat* file. These data were then extended using *Moody's* and later *Compustat* files. Any firm whose "as reported" data could not be determined was deleted. We continued this process until 30 firms met the sample criteria. Summary statistics for the *ROR* and *EPS* series were analyzed in some detail for all firms. No unreasonable series or summary statistics were noted. *ROR* increases and decreases were approximately equal (470 to 460), and *EPS* increases outnumbered decreases 597 to 333. The former is consistent with a process which is mean-reverting and has no drift; the latter is consistent with any process which has a positive drift-term.

5. Results

The maximum likelihood estimates of the β-parameters for Model 1 are given in table 1 together with supporting statistics. Results are given for each firm and, in the last row of the table, for the pooled model. The signs on the estimates of β are negative for all firms, reinforcing the prior hypothesis that accounting rates-of-return can be identified as being mean-reverting. Although the β-estimates are not necessarily independent across firms,[11] the reported t-values indicate that for a one-tail test 43 percent of the estimates of β's are less than zero at the 0.05 significance

[11] The problem should not be overemphasized. A necessary condition for cross-sectionally correlated t-statistics is that the *ROR* variables are cross-sectionally correlated. The data indicated that these latter correlations were generally quite modest.

TABLE 1

Logit Results for Model 1

$P[Z(\Delta ROR_t)] = [1 + \exp(-\alpha - \beta \cdot ROR_{t-1})]^{-1}$

Firm Number	$\hat{\beta}$		*Significance Tests for Predictions*	
	MLE	$\mid t(\hat{\beta}) \mid$	χ^2	Fisher Exact
1	−19.39	2.47	9.31*	
2	−4.11	1.05	4.01*	
3	−5.36	1.10	.88	
4	−10.24	1.77		0.37
5	−8.36	.91		n.s.
6	−10.21	.88		n.s.
7	−29.72	1.83	2.58	
8	−6.32	1.45		.014
9	−17.11	1.66		n.s.
10	−17.48	1.88	1.57	
11	−25.91	1.76		n.s.
12	−14.27	1.49	.78	
13	−3.79	.83		n.s.
14	−14.41	1.03	.27	
15	−71.18	2.55	8.79*	
16	−2.99	.54	.03	
17	−18.38	1.54	.26	
18	−21.91	1.75		n.s.
19	−6.08	1.06		n.s.
20	−33.60	2.01		n.s.
21	−10.44	1.20		n.s.
22	−33.35	2.34	3.89*	
23	−7.58	1.28		n.s.
24	−33.65	1.99	11.51*	
25	−13.32	1.68		n.s.
26	−29.11	2.30	2.84	
27	−4.42	.68		n.s.
28	−14.58	1.13		n.s.
29	−27.04	2.19	2.64	
30	−48.27	2.21		.003
Pooled	−6.65	6.72	41.37*	

χ^2 omitted when any cell has an expected value less than 5.
* A χ^2 of 3.84 is significant at the .05 level.
n.s. ≡ Fisher Exact Test not significant at .05 level.

level, and 60 percent at the 0.10 level. This must be viewed as compelling evidence against the null hypothesis. Furthermore, the t-value on $\hat{\beta}$ for the pooled model is significant at the .001 level.[12]

[12] The t-statistics are based on the usual kind of large sample asymptotic analysis. This means that there is a possibility that the computed t-statistics overstate the true significance. To investigate this matter, Bayesian analysis was performed. Specifically, assuming a noninformative prior distribution of (α, β), one can calculate the posterior distribution:

$$Pr\{\beta < 0 \mid \text{data}\} = \int_{-\infty}^{0} \int_{-\infty}^{\infty} l(\alpha, \beta \mid \text{data})\, d\alpha\, d\beta\ \text{const}^{-1}$$

where $l(\alpha, \beta \mid \text{data})$ is the likelihood function and "const" is the normalizing constant. These

An alternative method of evaluating the significance of the estimated β-coefficients is to evaluate the predictive performance of the estimated models. These results also appear in table 1. Given the simple metric of increase versus decrease, the analysis was cast in the form of a 2×2 contingency table as follows:

		Prediction: Increase	Prediction: Decrease	Total
Actual	Increase	$C1$	$E1$	$C1 + E1$
	Decrease	$E2$	$C2$	$C2 + E2$

Here $C1$ and $C2$ are correct classifications and $E1$ and $E2$ are incorrect classifications. The predictions are based on a cutoff point of .5; that is, an increase in ROR is predicted if $P_t(1 \mid ROR_{t-1}, \hat{\alpha}, \hat{\beta}) > 0.5$. Predictive ability was assessed to be significantly different from random using a "fixed row" χ^2 test or, when the expected number of observations in a cell was less than five, a Fisher exact probability. These statistics are reported in the last two columns of table 1. Eight of the 30 *FS* models produce statistics which are significant at the 0.05 level, which is greater than the 1.5 expected for random predictions. The eight cases generally correspond to those which have significant $\hat{\beta}$'s. The value of χ^2 for the pooled data is significant at the 0.001 level.[13]

The above prediction tests were performed on the same data used for estimation of the parameters. To validate the results further and to assess the robustness of the models, we repeated the tests using parameter estimates from pooled data, deleting one firm at a time, on the data for the deleted firm. The prediction tests on holdout firms again produced eight significant χ^2 statistics or Fisher exact probabilities, of which seven were in common with the significant statistics for the prediction tests using parameter estimates from individual firm models. Hence the predictive ability in the simple metric appears to be reasonably robust over different sets of procedures.[14]

The conclusions we draw from the results in table 1 are that the

computations were performed for 10 representative *FS* models. In all cases the quantities $Pr\{\beta < 0 \mid \text{data}\}$ were somewhat greater than the corresponding probabilities implied by the asymptotic t-values reported in the table. Hence, it would appear that these t-values are sufficiently conservative for the purposes of this study.

[13] The .5 cutoff implicitly assumes that the loss function is symmetric across the two types of classification errors. The prediction analysis was repeated on the pooled data using cutoffs ranging from .1 to .9. Although total errors ranged from 39.6 percent for the .5 cutoff to 50.4 percent for the .9 cutoff, the results generally indicated that the null hypothesis of statistical independence could be rejected.

[14] In addition, the likelihood ratio index (not reported) for the *PO* model in table 1 is only somewhat less than the index for the typical *FS* model, suggesting that the estimates of α and β are not highly unstationary. (As expected, the likelihood ratio index is less for tests on pooled data because α and β are constrained.) However, a likelihood ratio test did reject the hypothesis that the parameters were the same for each firm.

observed accounting rates-of-return appear consistent with a mean-reverting process, and that the use of a simple metric to describe changes in rates-of-return combined with Logit estimation gives a model which is quite robust in predictive space over different sets of firm data. This mean-reverting behavior of *ROR* appears to have been observed first by Beaver [1970] but only in a rather informal fashion. The tests provided here are, to the best of our knowledge, the first formal ones rejecting the random walk hypothesis when the dependent variable is defined as the accounting rate-of-return.

More important for the purposes of this study, a mean-reverting behavior of *ROR* is essential if earnings-per-share changes are to be predicted from *ROR*. From this perspective, the results are not overwhelmingly positive since the mean-reverting behavior is generally weak, at least for some firms. Nevertheless, we proceed to investigate the empirical relationship between earnings-per-share changes (dependent variable) and rates-of-return (independent variable) by performing similar tests on Model 2, thus addressing the primary hypothesis.

Table 2 provides the results for Model 2. Note that all save one estimate of β are negative. The t-statistics are generally much lower in table 2 (Model 2) as compared to those given in table 1 (Model 1). Only seven (23 percent) are significant at the .05 level, as compared with ten for Model 1. For the *PO* Model, $\hat{\beta}$ is negative and significantly different from zero at the .001 level. Thus, a negative relationship between *EPS* changes and rates-of-return in the previous year is suggested, but the relationship appears to be weaker than that observed for changes in rates-of-return. Taken at face value, the results do not unambiguously support the hypothesis that *ROR* predicts changes in *EPS*. The far-right columns in table 2 provide the prediction results for Model 2. Only two significant χ^2 statistics are observed, and we would expect 1.5 given random predictions (and independence). The prediction success of the *PO* model is also not better than random. Closer scrutiny of the results revealed a particularly striking inability of these models to predict earnings decreases.

These rather negative results can be explained, at least partially, by an "*EPS* bias," making Logit Model 2 an inefficient (or inadequate) representation of the outcome to be predicted. The sign of changes in *EPS* (the dependent variable in Model 2 estimation) is heavily biased toward earnings increases: *EPS* increased in 64 percent of our 930 observations, while *ROR* increased in only 51 percent. Therefore, the Logit procedures may indicate significant coefficients and yet fail the relatively rigorous contingency-table prediction test. This is because the Logit estimation technique "rewards" correct classifications ($C1$ and $C2$ in the 2×2 contingency table) unconditionally, while the prediction test rewards correct classifications only to the extent that they exceed conditional probabilities. In other words, the Logit model essentially rewards on $C1 + C2$ given the total number of observations ($C1 + C2 + E1 + E2$); but,

TABLE 2

Logit Results for Model 2

$P[Z(\Delta EPS_t)] = [1 + \exp(-\alpha - \beta \cdot ROR_{t-1})]^{-1}$

Firm Number	β		Significance Tests for Predictions	
	MLE	$\lvert t(\hat{\beta}) \rvert$	χ^2	Fisher Exact
1	−5.39	1.04	2.59	
2	−4.11	1.05	4.00*	
3	−6.50	1.33		n.s.
4	.31	.08		n.s.
5	−20.89	1.99		n.s.
6	−15.53	.85		n.s.
7	−33.16	2.05	2.34	
8	−2.48	.80		n.s.
9	−12.40	1.36	.79	
10	−7.08	.85		n.s.
11	−17.00	1.23		n.s.
12	−1.88	.25		n.s.
13	−1.28	.27		n.s.
14	−16.43	1.12		n.s.
15	−22.19	1.83	2.34	
16	−.34	.06		n.s.
17	−18.38	1.54	.26	
18	−17.03	1.46	1.55	
19	−2.38	.46		n.s.
20	−13.07	1.25		n.s.
21	−9.50	1.09		n.s.
22	−18.68	1.75		n.s.
23	−7.16	1.21		n.s.
24	−9.93	1.11		n.s.
25	−8.15	1.17		n.s.
26	−34.30	2.59	5.22*	
27	−4.42	.68		n.s.
28	−4.55	.34		n.s.
29	−13.62	1.93		n.s.
30	−14.31	1.75		n.s.
Pooled	−3.09	3.61	1.42	

χ^2 omitted when any cell has an expected value less than 5.
* A χ^2 of 3.84 is significant at the .05 level.
n.s. ≡ Fisher Exact Test not significant at .05 level.

in the prediction test, $C1$ (and $C2$) must be large relative to the exogenous quantity $C1 + E1$ ($C2 + E2$).[15]

Deflating earnings by net-book-value, instead of by number-of-shares-outstanding, apparently deemphasizes earnings trends and enables Model

[15] For example, if all observations were earnings increases, Model 2 estimators would provide perfect predictors yet fail the contingency-table prediction test. To investigate whether the "*EPS* bias" accounts for the failure of Model 2, where Model 1 succeeds, a linking of the two models is necessary. We noted above that two conditions, if met jointly, imply that changes in *EPS* can be predicted by *ROR*. The first condition required that *ROR* itself followed a mean-reverting process; as we saw, this condition was met. The second required that the dependent variables in Models 1 and 2 were correlated. This is

1 efficiently to capture the *ROR*-conditioned mean-reversion process. Hence, the estimated coefficients from Model 1 might also better predict the sign of *EPS* changes.[16] Accordingly, for each firm, the Model 1 parameter estimates from the pooled data, after omitting the 31 observations for that firm, were used to predict the sign of *EPS* changes. The prediction results were much improved: three χ^2 statistics were significant at the .05 level, four additional Fisher exact probabilities were below 10 percent, and the χ^2 statistic for the pooled results were significant at the .001 level.

Finally, we also tested the earnings prediction hypothesis over a holdout sample in order to avoid a potential bias from cross-sectional correlations within each sample year. Parameter estimates were derived from pooled data of the 30 sample firms for the years 1946–63. Predictions tests were conducted for the pooled Model 1 coefficients on all industrial *Compustat* firms with enough data to compute the actual and predicted changes in *ROR* and *EPS* for the years 1964–80. This procedure leads to a total of 31,120 observations. There is no overlap between the estimation and test data periods, thus eliminating potential problems of statistical overfitting.

Results for the holdout sample are reported in table 3, and these strongly support the hypothesis that *ROR* has predictive ability with respect to changes in *EPS* and, even more so, changes in *ROR*. Using a cutoff point of .5 yields a χ^2 statistic of 8.59 for the prediction of earnings changes. However, while statistical dependence is undoubtedly present, it is also quite small: a χ^2 statistic of 8.59 is hardly impressive given the

---

indeed the case. The signs of ΔROR and ΔEPS are the same in roughly 84 percent of all cases, and 94 percent of the differences between signs is due to cases in which $\Delta EPS > 0$ when $\Delta ROR < 0$. If "bias" in *EPS* accounts for the disappointing prediction results of Model 2, then the statistical significance of β in Model 2 is inversely related to the correlation between the dependent variables of Models 1 and 2. Hence the *FS* *t*-statistics of Model 2 should be explained by the *FS* *t*-statistics of Model 1, and a measure of the correlation between $Z(\Delta EPS)$ and $Z(\Delta ROR)$. As a measure of the latter, one can take percentage of disagreements; this leads to the following regression:

Dependent variable: *t*-statistics *FS*-2

Independent variable: *t*-statistics *FS*-1, percent of $Z(\Delta ROR)$ and $Z(\Delta EPS)$ of same sign.

The *t*-statistics for the independent variables were 3.73 and −2.00, respectively. (The R^2 was .34.) In sum, the relationships are statistically significant and in the hypothesized direction.

[16] In other words, the .5 cutoff point used for Model 2 predicts too many earnings increases, and the predictions fail the contingency-table test due to an inability to predict earnings decreases. The use of Model 1 parameters, combined with a .5 cutoff point, is intended to rectify the problem. In this context it might be noted that Models 1 and 2 are isomorphic: their *rankings* of probabilities are equivalent. This is immediate since both models have negative beta coefficients. From such a perspective it is seen that the issue is not how beta coefficients are derived (as long as they are negative); rather, the issue is how one combines a beta with a cutoff point when the prediction evaluation is the χ^2 statistic of a contingency table.

TABLE 3

Summary of Logit Model 1 Predictions of ΔROR and ΔEPS for 1964–80 "Holdout" Years

Probabilities Required for Prediction of		Number of Observations	χ^2 Statistic* for Prediction of the Sign of		Binomial Test t-statistics: Logit Model Predicts Sign of ΔEPS Better (Worse) Than Best Benchmark	
Incr.	Decrs.		ΔROR	ΔEPS	Reversal	Trend
>.9	<.1	661	279.80	109.21	6.93	
>.8	<.2	1,345	505.34	114.44	6.23	
>.7	<.3	2,885	927.87	122.93	6.48	
>.6	<.4	8,433	1,444.20	36.74	10.46	
≧.5	<.5	31,120	1,512.83	8.59		(19.70)

* χ^2 statistic of 3.84 is significant at the .05 level.

number of data points. As a comparison, predictions of changes in *ROR* yield a χ^2 statistic of no less than 1,512.83. Additional analysis reveals that the predictions of earnings changes are no better than random when the probability is close to 50 percent. If observations within the intervals $[.5 - k, .5 + k]$, $k = .1, .2, .3, .4$ are deleted, then the χ^2 statistics are much larger (although they are also based on fewer observations). The evidence thus supports the following conclusion: when *ROR* deviates significantly from its mean (approximately one standard deviation or more), then *ROR* has a definite relationship to the probability of an earnings change. The same cannot be said when *ROR* is reasonably close to its mean.

It is useful to compare Model 1 predictions to those of naive benchmark predictions based on past earnings behavior. Such a comparison bears on the issue of whether *ROR* is a surrogate for information already present in prior periods' earnings, that is, knowledge of a firm's book value adds no information beyond what is captured in past earnings. These kinds of tests are therefore more "demanding" than the one considered so far, that is, the implicit benchmark in earlier predictions tests presumed that earnings (beyond a drift parameter) were unpredictable.

Two naive benchmark models were used: (*i*) the earnings change prediction is the same as prior period's earnings change ("trend model"); (*ii*) the earnings change predicted is opposite that of prior period's earnings change ("reversal model"). Statistical tests match the Logit model against the "better" of the two models. The last two columns in table 3 give the results of the following binomial test: when the Logit-model prediction differs from that of the benchmark model, which model is most often right? The t-statistics are all significant at respectable levels. Note, however, that the trend model does better than the Logit model if the cutoff point is .5. This result suggests that the trend model does capture an element of earnings behavior; in fact, a test (not reported in the table) of the trend model against the alternative of a random walk is significant at a .001 level. From this perspective it is not surprising that the Logit model does relatively poorly. More than two-thirds of all Logit predictions are within the [.4, .6] interval, and these predictions are no better than random. Thus, if one excludes those predictions for which the probabilities of earnings increases are between .6 and .4, or excludes even larger intervals, then the Logit model does significantly better than either benchmark model. The latter is now always the reversal model. In sum, the evidence suggests, rather unambiguously, the *ROR* has predictive content when it deviates from its mean by a significant amount.[17]

6. *Conclusions*

Previous time series of earnings studies generally have failed to reject the "random walk hypothesis." This study demonstrates that a simple expansion of the conditioning information set can succeed where complex model building has failed. The evidence suggests that *ROR* has predictive content with respect to earnings changes, at least when it significantly deviates from its mean. Furthermore, when *ROR* is close to its mean, a simple earnings trend model has predictive power with respect to earnings changes.

Future research could focus on improving these forecasts by augmenting the conditioning information set. One approach might be to employ macroeconomic variables. Recently Chant [1980] found that the money supply added enough information to outperform extrapolative time-series models in terms of predictive power. Other accounting variables besides the book rate-of-return might also add to the predictive model used here. In the latter case, the idea would be to assess the improvement in predictions by widening the accounting information set and eventually to compare the predictions of earnings implicit in security prices (see Beaver, Lambert, and Morse [1980]) with those based on accounting information sets.

REFERENCES

ALBRECHT, W., L. LOOKABILL, AND J. MCKEOWN. "The Time-Series Properties of Annual Earnings". *Journal of Accounting Research* (Autumn 1977): 226–44.

BALL, R., AND R. WATTS. "Some Time Series Properties of Accounting Income." *Journal of Finance* (June 1972): 663–81.

BEAVER, W. "The Time Series Behavior of Earnings." *Empirical Research in Accounting: Selected Studies, 1970.* Supplement to *Journal of Accounting Research* 8: 62–89.

BEAVER, W., R. LAMBERT, AND D. MORSE. "The Information Content of Security Prices." *Journal of Accounting and Economics* (March 1980): 3–28.

BEAVER, W., AND D. MORSE. "What Determines Price Earnings Ratios?" *Financial Analysts Journal* (July/August 1978): 65–76.

BREALEY, R. "Statistical Properties of Successive Changes in Earnings." Working paper, Keystone Custodian Funds, March 1967.

[17] In other words, a "combination model" of the trend and Logit models does extremely well in predicting earnings changes; that is, use the trend model if the Logit model predicts within the [.4, .6] interval; otherwise use the Logit model.

BROOKS, L., AND D. BUCKMASTER. "Further Evidence of the Time Series Properties of Accounting Income." *Journal of Finance* (December 1976): 1359–73.

———. "First Differences Signals and Accounting Income Time Series Properties." Working paper, 1979.

BROWN, L., AND M. ROZEFF. "The Superiority of Analysts Forecasts as Measures of Expectations: Evidence from Earnings." *Journal of Finance* (March 1978): 1–16.

———. "Adaptive Expectations, Time Series Models, and Analyst Forecast Revision." *Journal of Accounting Research* (Autumn 1979): 341–51.

CHANT, P. D. "On the Predictability of Corporate Earnings per Share Behavior." *Journal of Finance* (March 1980): 13–22.

CRICHFIELD, T., T. DYCKMAN AND J. LAKONISHOK. "An Evaluation of Security Analysts' Forecasts". *The Accounting Review* (July 1978): 651–68.

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

GONEDES, N. "Properties of Accounting Numbers: Model and Tests." *Journal of Accounting Research* (Autumn 1973): 212–37.

LINTNER, J. "Further Observations of Higgledy Piggledy Growth." Paper presented at the Seminar on the Analysis of Security Prices, University of Chicago, May 9, 1969.

———, AND R. GLAUBER. "Higgledy Piggledy Growth in America." Paper presented at the Seminar on the Analysis of Security Prices, University of Chicago, May 11–12, 1967.

LITTLE, I. *Higgledy Piggledy Growth.* Oxford: Institute of Statistics, 1962.

———, AND A. RAYNER. *Higgledy Piggledy Growth Again.* Oxford: Basil Blackwell, 1966.

LOOKABILL, L. "Some Additional Evidence on the Time Series Properties of Accounting Earnings." *The Accounting Review* (October 1976): 724–38.

WATTS, R., AND R. LEFTWICH. "The Time Series of Annual Accounting Earnings." *Journal of Accounting Research* (Autumn 1977): 253–71.