



American Finance Association

Autoregressive Modeling of Earnings-Investment Causality

Author(s): Sasson Bar-Yosef, Jeffrey L. Callen, Joshua Livnat

Source: *The Journal of Finance*, Vol. 42, No. 1 (Mar., 1987), pp. 11-28

Published by: Blackwell Publishing for the American Finance Association

Stable URL: <http://www.jstor.org/stable/2328416>

Accessed: 26/11/2009 08:06

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Blackwell Publishing and American Finance Association are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance*.

<http://www.jstor.org>

Autoregressive Modeling of Earnings-Investment Causality

SASSON BAR-YOSEF, JEFFREY L. CALLEN, and JOSHUA LIVNAT*

ABSTRACT

The purpose of this paper is to empirically test the relationships between corporate earnings and investment. In particular, the study investigates whether knowledge of past investments improves the prediction of future earnings beyond predictions that are based on past earnings alone. Similarly, it investigates whether knowledge of past earnings improves the prediction of future investments beyond knowledge of past investments alone. This is the empirical definition of Granger causality. The empirical results show that the bivariate past series of earnings and investments is superior to the univariate series in predicting future investments but not in predicting future earnings.

FUTURE CORPORATE EARNINGS ARE an important parameter in almost all stock-valuation models. The dividend-stock model, perhaps the most widely employed quantitative tool by professional investors, is a case in point.¹ It is therefore not surprising to find that an enormous amount of intellectual capital has been expended studying corporate earnings patterns and the extent to which corporate earnings can be accurately forecast. There are, in fact, a number of related bodies of literature on these issues.

The time-series literature on annual corporate earnings concludes that annual earnings follow a random walk.² This means essentially that annual earnings cannot be forecast, so that the best estimate of tomorrow's earnings is today's earnings. On the other hand, the time-series literature on interim corporate earnings concludes that interim earnings contain seasonality patterns, so that they are forecastable, at least in theory.³ However, the problem is that no time-series model seems to predominate, and each firm appears to have its own unique applicable model.

Another related body of literature deals with the ability of financial analysts to forecast future earnings and earnings growth.⁴ The results are not conclusive. Cragg and Malkiel [8] and Elton and Gruber [9] find no significant difference between analysts' forecast accuracy and their best naive earnings-prediction

* Bar-Yosef is from Hebrew University, Jerusalem; Callen is from Hebrew University and the Faculty of Business, McMaster University; Livnat is from Hebrew University and Owen Graduate School of Management, Vanderbilt University.

¹ See Gordon [15].

² See, for example, Ball and Watts [4], Albrecht, Lookabill, and McKeown [3], and Watts and Leftwich [33].

³ See, for example, Foster [11], Griffin [17], Lorek [27], and Brown and Rozeff [6].

⁴ For a review of this literature, see Givoly and Lakonishok [14].

model. More recently, Collins and Hopwood [7], Brown and Rozeff [6], and Fried and Givoly [13] find that financial analysts, on average, significantly outperform naive and Box-Jenkins earnings models. Nevertheless, financial analysts' prediction errors are rather large. Indeed, as Elton, Gruber, and Gultekin [10] emphasize, although financial analysts are reasonably accurate in forecasting earnings growth at the aggregate level, they are quite inaccurate in forecasting individual corporate-level earnings growth.

In general, earnings prediction is a function of the underlying information set upon which the prediction is based. One potentially important element of this information set is the current and past investment record of the firm. At least in theory, the firm's current and past investment activity—together with its financing activity—is one of the most important determinants of its future earnings.⁵ Therefore, we should expect that the additional information contained in the firm's investment record should enhance earnings predictability. Nevertheless, none of the literature cited above has investigated whether financial analysts utilize the corporation's investment record to form their earnings predictions, nor has anyone investigated whether adding the time series of past investments to the time series of past earnings enhances earnings prediction.⁶

Although we have emphasized the importance of earnings prediction to the finance literature, predicting future corporate investment is also of interest, especially to macrofinance specialists and macroeconomists concerned with predicting aggregate corporate investment behavior.⁷ In fact, if there is a linkage between corporate investment and corporate earnings, the linkage may be two-sided; past investment may be useful in predicting future earnings, while past earnings may be useful in predicting future investment.

Despite the potential linkage between corporate earnings and investment, the empirical literature to date is ambiguous. Little [26], using U.K. data, and Lintner and Glauber [25], employing U.S. data, find that corporate earnings are not necessarily affected by past investments. Baumol et al. [5], using U.S. data, conclude that "the rate of return to firms relying on ploughback for their new investment is typically uncomfortably small" relative to the rate of return to external financing.⁸ This implies that the reinvestment of corporate earnings may have little impact on future corporate earnings. In contradistinction, using Canadian data, McFetridge [29] concludes that the "difference between the marginal rate of return to retained earnings and the marginal rate of return to equity does not itself differ significantly from zero." Shapiro et al. [32] reconfirm

⁵ On theoretical grounds, one would expect on average a positive relationship between past investments and current earnings. This is because investments are undertaken only if their expected yield (adjusted for risk) is greater than their opportunity cost. If expectations are realized on average, then current earnings should be an increasing function of past investments.

⁶ Only one study to date has attempted to relate investment to earnings prediction. Freeman, Ohlson, and Penman [12] find that book rates of return can be used to predict annual earnings changes better than a random walk for some firms. Specifically, they show that firms with large book returns in one year are more likely to exhibit lower earnings the next year. Similarly, firms with small book returns are more likely to exhibit larger earnings in the following year.

⁷ See the literature cited below.

⁸ See also Whittington's [34] results on U.K. data and his penetrating remarks about the Baumol et al. study.

McFetridge's results and suggest that these contradictory results may be due to international differences in managerial behavior.

The above empirical evidence concerning the impact of past investment on current earnings is rather vague and contradictory. Similarly, the evidence concerning the impact of past earnings on current investment, while less vague, is also contradictory, as is evident from the literature on the determinants of corporate investment behavior.⁹ This literature suggests, on a priori grounds, that there should be an empirical linkage between past earnings and current investment.¹⁰ Specifically, corporate investment is a function of future desired levels of capital, which in turn are a function of expected future profits. Although expected future profits are unobservable, they can be estimated from the time series of past earnings. Therefore, it is claimed, past earnings are useful in estimating future corporate investment. However, to repeat, the empirical evidence is contradictory. While some studies, notably those based on accelerator models of investment, found past earnings to be a relatively insignificant determinant of investment, studies based on optimal capital accumulation models found past earnings to be moderately significant.

The purpose of this paper is to use a different empirical approach, based on Granger [16] causality, to examine the linkage, if any, between corporate earnings and corporate investment.¹¹ This approach is predictive in nature in that it tries to determine whether the time series of past investment provides additional information, beyond the earnings time series itself, that can be used to predict earnings. However, unlike lead-lag analysis or standard regression analysis, causality analysis recognizes that the investment-earnings relationship could well run in both directions. Thus, we shall also investigate whether the time series of past earnings provides additional information, beyond the investment time series, that is useful in predicting future corporate investment.

In what follows, Section I defines Granger causality more fully and formalizes the model relating earnings and investment. Section II describes the data and discusses the empirical results. Section III concludes the paper.

I. Modeling Earnings-Investment Causality

A. Defining Granger Causality

Consider the bivariate stationary stochastic process $Z_t = \{E_t, I_t\}$, where E_t denotes earnings at time t and I_t is the investment at time t . Suppose that, at

⁹ See Jorgenson [22] and Jorgenson and Siebert [23].

¹⁰ In a perfect capital market, past earnings should have no *theoretical* impact on current investment. However, in an imperfect market, past earnings may be a useful signal to the market. Thus, a firm having a good earnings track record may find that it can raise additional capital at a lower cost, thereby increasing the incentive for additional investments.

¹¹ Also, unlike the literature on the determinants of corporate investment behavior, we use a *time series* of past annual earnings to model the impact of earnings on investment.

Granger [16] causality is a statistical definition and does not correspond necessarily to the common usage of the term causality. For an in-depth characterization of Granger causality and some reservations about it on philosophical grounds, see Newbold [31]. As a counterweight to Newbold's reservations, see Nelson and Schwert [30].

time $t - 1$, we try to predict next-period earnings, E_t . If E_t is better predicted by adding the past investment time series to the past earnings time series than by using the past earnings time series alone, then investment is said to cause earnings. Similarly, earnings are said to cause investment if next-period investment, I_t , is better predicted by the bivariate time series than by the univariate investment series alone.

More formally, let $\bar{E}_{t-1}$, $\bar{I}_{t-1}$, and $\bar{Z}_{t-1}$ denote the set of past values of the earnings, investment, and bivariate time series, respectively, so that

$$\bar{E}_{t-1} = \{E_{t-1}, E_{t-2}, \dots\},$$

$$\bar{I}_{t-1} = \{I_{t-1}, I_{t-2}, \dots\},$$

and

$$\bar{Z}_{t-1} = \{E_{t-1}, E_{t-2}, \dots, I_{t-1}, I_{t-2}, \dots\}.$$

Let $\sigma^2(E_t | \bar{Z}_{t-1})$ denote the error in predicting period- t earnings given that the analyst bases predictions on the information set that includes both the past earnings and the past investment time series.¹² By contrast, let $\sigma^2(E_t | \bar{Z}_{t-1} - \bar{I}_{t-1})$ be the error in predicting period- t earnings given that the analyst chooses to exclude the past investment time series for prediction purposes—in other words, period- t earnings are predicted on the basis of the past earnings time series alone.

A.1. Definition of Granger Causality

If $\sigma^2(E_t | \bar{Z}_{t-1}) < \sigma^2(E_t | \bar{Z}_{t-1} - \bar{I}_{t-1})$, then investment is said to cause earnings. In words: investment causes earnings if the information set that includes past investment data, $\bar{Z}_{t-1}$, yields a more accurate prediction of future earnings than does the same information set without past investment data, $\bar{Z}_{t-1} - \bar{I}_{t-1}$. If the inequality in the definition does not hold, then investment is said not to cause earnings. In a similar fashion, we can define $\sigma^2(I_t | \bar{Z}_{t-1})$ and $\sigma^2(I_t | \bar{Z}_{t-1} - \bar{E}_{t-1})$ to be the errors in predicting future investment, where the prediction is based on the bivariate information set or the information set excluding past earnings, respectively. Earnings cause investment if $\sigma^2(I_t | \bar{Z}_{t-1}) < \sigma^2(I_t | \bar{Z}_{t-1} - \bar{E}_{t-1})$; otherwise, earnings do not cause investment.

There is nothing in the definition of causality to exclude the possibility that earnings and investment cause each other. This is called feedback.

A.2. Definition of Feedback

If $\sigma^2(E_t | \bar{Z}_{t-1}) < \sigma^2(E_t | \bar{Z}_{t-1} - \bar{I}_{t-1})$ and $\sigma^2(I_t | \bar{Z}_{t-1}) < \sigma^2(I_t | \bar{Z}_{t-1} - \bar{E}_{t-1})$, then feedback is said to occur between investment and earnings.

These causality definitions will be adapted further below to an empirical model of earnings and investment, which we now describe.

¹² The reason for the use of the variance symbol σ^2 to denote the prediction error is made clearer below. Basically, $\sigma^2(E_t | \bar{Z}_{t-1})$ is the minimum mean-square linear prediction error of E_t , given the information set $\bar{Z}_{t-1}$.

B. The Model

It is well known (e.g., Masani [28]) that a regular full-rank stationary stochastic process, $Z_t = \{E_t, I_t\}$, can be modeled, under fairly general conditions, by the autoregressive representation:

$$I_t = \psi_{111}I_{t-1} + \psi_{112}I_{t-2} + \cdots + \psi_{121}E_{t-1} + \psi_{122}E_{t-2} + \cdots + v_t, \quad (1)$$

$$E_t = \psi_{211}I_{t-1} + \psi_{212}I_{t-2} + \cdots + \psi_{221}E_{t-1} + \psi_{222}E_{t-2} + \cdots + u_t, \quad (2)$$

where the ψ_{ijk} are parameters and $\{v_t, u_t\}$ are conventional zero-mean error terms with constant variance-covariance matrix. Equations (1) and (2) state that the firm's future investment and future earnings (in period t) are a linear function of the past investment record and the past earnings record of the firm. These equations can be written more compactly in the polynomial form:

$$I_t = \psi_{11}(L)I_t + \psi_{12}(L)E_t + v_t, \quad (1')$$

$$E_t = \psi_{21}(L)I_t + \psi_{22}(L)E_t + u_t. \quad (2')$$

Here, $\psi_{ij}(L)$ is the lag polynomial $\sum_{k=1} \psi_{ijk}L^k$, where L^k denotes the lag operator ($L^k I_t = I_{t-k}$).

Causal relationships would appear to enter this model in a very natural way. If $\psi_{12}(L) = 0$ (i.e., $\psi_{12k} = 0$ for all k), then it is clear from equation (1') that past earnings have no effect on future investment; that is, earnings do not cause investment. Similarly, if $\psi_{21}(L) = 0$, investment does not cause earnings. Thus, in theory, one could determine the causal relationships between earnings and investment by first fitting equations (1') and (2') by least squares—yielding estimates that are consistent and asymptotically normally distributed—and then testing to see whether $\psi_{ij}(L) = 0$, $i \neq j$.

This approach is potentially problematic, however, because the test of $\psi_{ij}(L) = 0$ is quite sensitive to the order of the lags of $\psi_{ij}(L)$. (See Hsiao [19, 20].) If, as is normally the case, the lag structure is prespecified, the test results may simply be a result of the imposed lag specification rather than of the data showing causality. One way out of the problem, as suggested by Hsiao, is to let the data determine the lag structure rather than imposing some arbitrary lag structure on the model. His suggested procedure for obtaining the optimal order of the lags for each of the $\psi_{ij}(L)$ in each of equations (1') and (2') is to employ the Final Prediction Error (FPE) criterion developed originally by Akaike [1, 2]. Consider first the investment variable I_t . The FPE of I_t is defined to be the mean squared prediction error:

$$E(I_t - \hat{I}_t)^2, \quad (3)$$

where I_t is the actual investment in period t , $\hat{I}_t$ is the predicted investment in period t , and $E(\cdot)$ is the expectations operator. The predicted investment value $\hat{I}_t$ is determined by regressing equation (1') using ordinary least squares for a given lag structure of orders m and n on $\psi_{11}(L)$ and $\psi_{12}(L)$, respectively. In other words, $\hat{I}_t$ is the least-squares estimate:

$$\hat{I}_t = \hat{b} + \hat{\psi}_{11}^m(L)I_t + \hat{\psi}_{12}^n(L)E_t, \quad (4)$$

where $\hat{\delta}$ is an estimated constant, $\hat{\psi}_{11}^n(L)$ are the estimated parameters of $\psi_{11}(L)$ assuming a lag structure of order m , and $\hat{\psi}_{12}^n$ are the estimated parameters of ψ_{12} assuming a lag structure of order n . Since the population mean of the FPE's is unknown, Akaike estimates the Final Prediction Error by

$$\text{FPE}_I(m, n) = \left(\frac{T + m + n + 1}{T - m - n - 1} \right) (\sum_{t=1}^T (I_t - \hat{I}_t)^2 / T), \quad (5)$$

where T is the total number of data points. The first bracketed term of the product on the right-hand side of equation (5) is a measure of estimation error, while the second bracketed term represents a measure of the average modeling error. Akaike's FPE criterion is to choose the lag structure (m, n) that *minimizes* the Final Prediction Error given by equation (5).¹³ In a similar fashion, one can estimate equation (2') by ordinary least squares to obtain an estimated value $\hat{E}_t$ for earnings. The estimated final prediction error for earnings would then be

$$\text{FPE}_E(m, n) = \left(\frac{T + m + n + 1}{T - m - n - 1} \right) (\sum_{t=1}^T (E_t - \hat{E}_t)^2 / T). \quad (6)$$

Minimizing equations (5) and (6) with respect to the different lag structures on the earnings and investment time series can be computationally prohibitive for a large sample of firms. To minimize the computational difficulties, Hsiao [19] has suggested the following sequential procedure. First, an upper bound on the maximal lag order is specified, say Q . Second, the FPE criterion is used to determine the optimal order of the one-dimensional autoregressive process for investment alone. Call this order q ($\leq Q$), so that the resulting FPE is $\text{FPE}_I(q, 0)$. Third, fix the lag structure for investment at q and use the FPE criterion to specify equation (1'). Let $r > 0$ denote the potential lag order for the earnings lag operator $\psi_{12}(L)$, so that the FPE is $\text{FPE}_I(q, r)$. To select r , equation (1') is estimated with earnings lags of 1 through the maximum lag Q . Thus, r minimizes $\text{FPE}_I(q, r)$ for all positive earnings lags up to Q . Fourth, holding the order of the lag operator $\psi_{12}(L)$ at r , let the order of lag operator $\psi_{11}(L)$ vary from 1 to Q . Choose the order of $\psi_{11}(L)$ that gives the smallest FPE, say s (s being not necessarily equal to q), thereby yielding $\text{FPE}_I(s, r)$. This fourth step is a check to see whether the lag structure for investment (which was originally q) is sensitive to the lag structure of earnings. Finally, if $\text{FPE}_I(q, 0) \leq \text{FPE}_I(s, r)$, then investment is best represented by a one-dimensional autoregressive process, so that earnings do not cause investment. Conversely, if $\text{FPE}_I(q, 0) > \text{FPE}_I(s, r)$, then earnings cause investment. A similar approach can be used on equation (2') to see whether investment causes earnings. This procedure is repeated for each firm in the sample.

It is worth emphasizing that the FPE approach for determining causality yields a number of distinct benefits in terms of identifying the model. First, as we have already pointed out, the data are used to determine the lag structure of the model rather than imposing some arbitrary lag-order specification. Second, the FPE criterion does not constrain the lag structure of each variable to be identical; i.e.,

¹³ Essentially, Akaike's criterion balances the bias from choosing too small a lag order against the increased variance of a higher lag-order specification.

m is not constrained to be necessarily equal to n . Third, it can be shown that the FPE criterion is equivalent to choosing the model specification on the basis of an F -test with varying significance levels.¹⁴

II. The Empirical Results

A. The Data

Annual data from the Compustat tapes were collected for a sample of U.S. manufacturing firms. The sample was limited to firms reporting Net Income Before Extraordinary Items (NIBEI) and Capital Expenditure data for the years 1960 through 1981.¹⁵ Six hundred forty-four firms satisfied these sample selection criteria. A further subsample was generated by requiring each firm to have positive NIBEI over the entire data period. Four hundred ninety of the original 644 firms satisfied this latter requirement.

The data were divided into two nonoverlapping time periods. The data from 1960 through 1979 were used to estimate the causality relationships. The data for 1980 and 1981 were used as a holdout sample to test the predictive validity of the estimated causality models.

To reduce the problem of serial correlation in the time series, the causality models were estimated after first-order differencing of the *logarithms* of the earnings and investment data.¹⁶ This necessitated employing the 490-firm subsample of positive-earnings firms. Since limiting the sample to positive-earnings firms could potentially bias the results, the models were also estimated on the larger sample of 644 firms, after first-differencing the earnings and investment data. However, since the results turned out to be substantially the same for both samples, we report the model based on the subsample, i.e., after first-differencing the logarithms of the data.

B. Estimating the Causality Model

The model (equations (1') and (2')) was estimated for *each* firm separately using five different earnings definitions (two of these being cash flow surrogates: Operating Income Before Depreciation and Net Income Plus Depreciation) and two different investment definitions. Investment was defined to be Capital Expenditures and Capital Expenditures plus Investment in Unconsolidated Subsidiaries. Since the results for both investment definitions were virtually identical, we report the Capital Expenditures outcomes only.

The univariate and bivariate time-series models were estimated as described above. In particular, the bivariate model was estimated using all four steps of the Hsiao sequential time-series procedure. In all cases, five was set to be the

¹⁴ See Hsiao [21].

¹⁵ See Larcker [24] for the definition and limitations of the Capital Expenditure Item.

¹⁶ In addition, the data are thereby transformed into yields that are essentially independent of firm size. It is, of course, important that the test statistic (the FPE's) not be a function of firm size.

Although earnings and investment are correlated, multicollinearity is not an issue in our analysis because the FPE is an overall-fit statistic (like an R^2 , for example) and is insensitive to multicollinearity between the independent variables.

maximum number of potential lags for each variable (i.e., $Q = 5$). Table I presents summary statistics for earnings as the dependent variable (equation (2')). The column headed $FPE_E(0, r)$ lists the median Final Prediction Error—the median being taken over the number of firms in the sample—where future earnings are predicted by past earnings alone. Similarly, the column headed $FPE_E(q, r)$ (where $q > 0$) provides the median FPE for earnings where earnings are predicted by the bivariate time series of earnings and investment.¹⁷ For earnings definitions other than NIBEI, there were differing amounts of data reported on the Compustat tapes. Column (3) gives the number of data points (i.e., firms) for each income definition.

On the basis of the median FPE, the results would appear to be relatively unambiguous. For all income definitions except for Net Income, investment does not cause earnings since $FPE_E(0, r) \leq FPE_E(q, r)$. In the case of Net Income, the direction of causality is indeterminate. The results are similar using more rigorous tools. In particular, columns (4) and (5) in Table I list the sum of rank differences of $FPE_E(0, r) - FPE_E(q, r)$. Column (4) gives the positive sums and column (5) the negative sums. For most earnings definitions, the sum of the negative ranks is greater than the sum of the positive ranks and, to a significant degree, for the cash flow surrogates. Thus, on the basis of the Wilcoxon signed-rank test, we also conclude that investment does not cause earnings.¹⁸ In other words, expanding the earnings information set to include investment data does not improve the predictability of future earnings.

Table II presents summary statistics in the same fashion as Table I, except now investment is the dependent variable. (See equation (1').) In the case of predicting investment, the results are unambiguous. For all earnings definitions, the median FPE using the bivariate time series is less than the median FPE using the investment series alone. This result is also confirmed by the Wilcoxon signed-rank test. Therefore, we conclude that earnings significantly cause investment. In other words, past earnings data are useful in predicting future firm investment.

In addition to comparing FPE's and rank sums, we thought that it might be illuminating to plot comparative cumulative distributions of the (square root of the) FPE's resulting from the univariate and bivariate models.¹⁹ Figures 1 and 2 show the cumulative distributions of the FPE's for the earnings variable, defined as Net Income Before Extraordinary Items (NIBEI) and Net Income Plus Depreciation, respectively. In the case of the cash flow definition, the univariate model clearly dominates the bivariate model, in the sense of yielding smaller FPE's, by second-order stochastic dominance.²⁰ In the case of NIBEI, on the

¹⁷ $FPE_E(q, r)$ in the tables is determined by searching over $q > 0$ (i.e., $q = 1, \dots, 5$). $FPE_E(q, r)$ is then compared with $FPE_E(0, r)$ to see which yields the true minimum. Similarly, $FPE_I(q, r)$ is determined for $r > 0$. Since we employed all steps of the Hsiao procedure to compute $FPE_E(q, r)$ (and $FPE_I(q, r)$), the r 's of $FPE_E(q, r)$ and $FPE_E(0, r)$ in the tables need not be identical (similarly for the q 's of $FPE_I(q, r)$ and $FPE_I(0, r)$).

¹⁸ The nonparametric Wilcoxon signed-rank test is appropriate because we do not know the distribution of the FPE's. Also, this test attenuates the problem of outliers in the data. See also the discussion further below.

¹⁹ We show the square root of the FPE's simply to enhance the clarity of the plots.

²⁰ On stochastic dominance, see, for example, Hanoch and Levy [18].

Table I
Results of the FPE Test for Earnings

Income Definition ^a	Median FPE _E (0,r)	Median FPE _E (q,r)	No. of Cases	Sum of Ranks of Positive Differences FPE _E (0,r) – FPE _E (q,r)	Sum of Ranks of Negative Differences FPE _E (0,r) – FPE _E (q,r)	Significance Level Less Than
	(1)	(2)	(3)	(4)	(5)	(6)
1. Net income before extraordinary items	.090	.122	490	56,816	63,479	.14404
2. Net income	.111	.111	455	56,006	47,734	.07029
3. Operating income after depreciation	.088	.125	524	64,875	72,675	.13036
4. Operating income before deprecia- tion	.061	.125	580	57,830	110,660	.00001
5. Net income plus depreciation	.048	.134	574	49,873	115,152	.00001

^a These definitions correspond to the following Compustat items:

Income Definition	Compustat Item
1	18
2	18 – 48
3	13 – 14
4	13
5	18 + 14

Table II
Results of the FPE Test for Investment

Income Definition	Median $FPE_t(q,0)$	Median $FPE_t(q,r)$	No. of Cases	Sum of Ranks of Positive Differences $FPE_t(q,0) -$ $FPE_t(q,r)$	Sum of Ranks of Negative Differences $FPE_t(q,0) -$ $FPE_t(q,r)$	Significance Level Less Than
	(1)	(2)	(3)	(4)	(5)	(6)
1. Net income before extraordinary items	.256	.028	490	114,895	5,400	.00001
2. Net income	.252	.030	455	96,615	7,125	.00001
3. Operating income after depreciation	.269	.031	524	131,423	6,127	.00001
4. Operating income before deprecia- tion	.278	.021	580	166,472	2,018	.00001
5. Net income plus depreciation	.282	.015	574	163,723	1,302	.00001

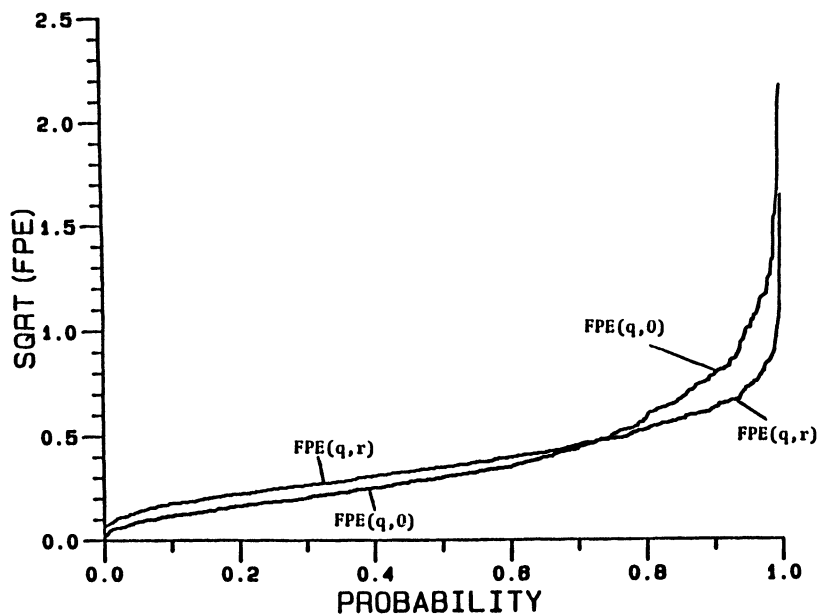


Figure 1. FPE Earnings—Income Definition 1

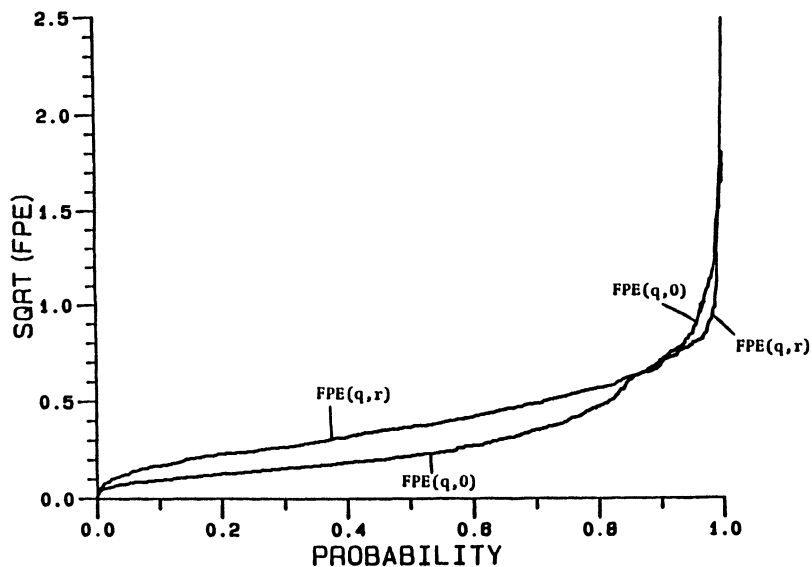


Figure 2. FPE Earnings—Income Definition 5

other hand, dominance is not obvious. This actually corresponds to the results of the Wilcoxon signed-rank test in Table I. There the univariate model yielded significantly smaller FPE's than the bivariate model for the cash flow definition, whereas the result was far less clear-cut for the NIBEI definition.

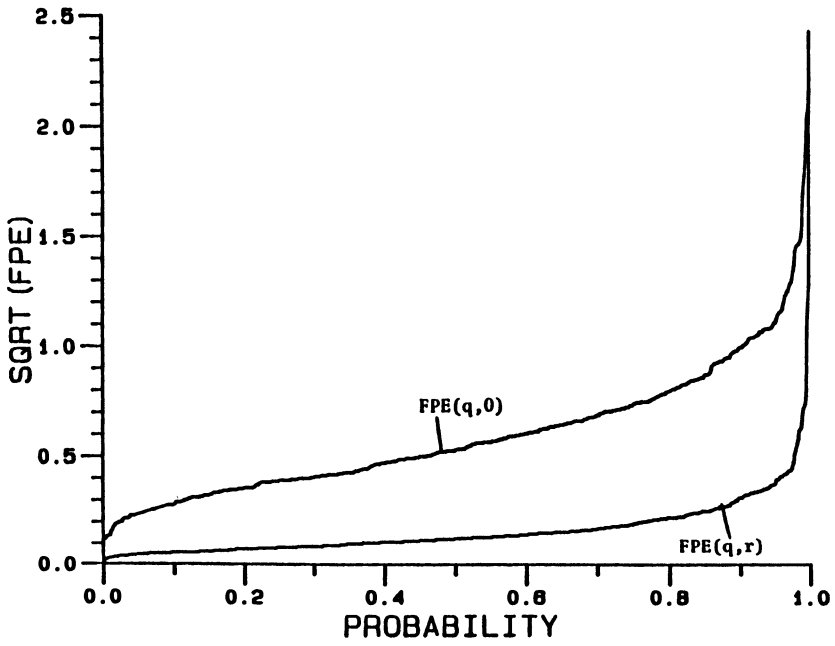


Figure 3. FPE Investment—Income Definition 1

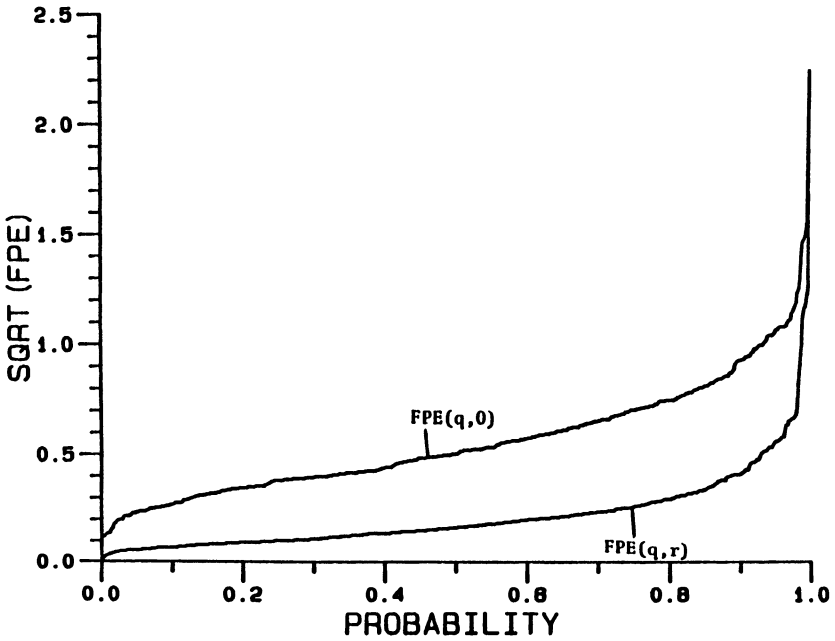


Figure 4. FPE Investment—Income Definition 5

The results for the investment variable are even more impressive. Figures 3 and 4 plot the cumulative distributions of the FPE's for investment, with earnings for the bivariate model defined by NIBEI and Net Income Plus Depreciation, respectively. For both earnings definitions, the bivariate model dominates the univariate model (in the sense of yielding smaller FPE's) by *first-order stochastic dominance*. Therefore, we are once more led to conclude that investment does not cause earnings but that earnings cause investment in the Granger sense.

Since it is always possible that the preceding empirical results are sensitive to industry variation, we broke the data down by two-digit S.I.C. industry codes with the requirement that each industry be comprised of at least twenty firms. In this way, we were able to obtain six two-digit industries comprising about forty percent of the sample. The results were then summarized and analyzed on the basis of each industry. We found substantially the same results as for all industries together. Therefore, we conclude tentatively that industry variation does not affect the causality relationship between investment and earnings.

In addition to industry variation, we also tested the robustness of our results to alternative perturbations of the investment and earnings variables. Specifically, each firm in the sample was ranked on the basis of its average growth rate in capital expenditures. The ranked sample was partitioned by quartiles and the FPE tests replicated for the lower and upper quartiles separately. In an alternative perturbation, the FPE tests were replicated, with investment defined as the ratio of capital expenditures to total assets. In a further trial, firms were ranked in terms of the correlation between earnings and past investment.²¹ The FPE tests were again replicated for the lower and upper quartiles separately. The final perturbation involved repeating the FPE tests for those firms with an optimal lag structure of four or less in each of the two variables, earnings and investment.²² For each of these perturbations, the results were substantially the same as, if not identical to, those in Tables I and II.²³

C. Validating the Causality Model—Prediction

To validate our conclusions, we also tested the predictive ability of our estimated causality models on the holdout data, namely, the 1980 and 1981 observations. In addition, we compared the causality model's predictive ability with a naive model, which assumes that the growth in earnings (or investment) this year is the same as the growth in earnings (or investment) last year. Table

²¹ For each firm, earnings in period t were regressed against past investment lagged over the previous five periods (i.e., $t - 1$ to $t - 5$). The resulting R^2 from the regression represents the correlation between the firm's earnings and its past investment. Firms were then ranked on the basis of this correlation. This was done to see whether firms with a high correlation demonstrate greater predictability of earnings from past capital-expenditure data.

²² In carrying out the FPE tests, five was set to be the maximum lag structure ($Q = 5$) for each variable. The number five, we felt, mitigated against the potential problems of insufficient degrees of freedom and yet allowed for a sufficient number of lags to undertake a meaningful study. Since five is an arbitrary upper limit, it is possible that the true optimal lag structure is larger than five, especially for those firms that were found to have an optimal lag structure of five in at least one of the variables. To see whether our results are sensitive to this upper limit, the FPE tests were repeated for those firms with an optimal lag structure of less than five in each variable.

²³ The results are available from the authors.

Table III
Results of Predictions for Earnings

Income Definition	1980 Observation					1981 Observation				
	Significance Level					Significance Level				
	FPE _E (0, <i>r</i>)	FPE _E (<i>q</i> , <i>r</i>)	Naive	(2 - 1)	(3 - 1)	FPE _E (0, <i>r</i>)	FPE _E (<i>q</i> , <i>r</i>)	Naive	(7 - 6)	(8 - 6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	.026	.114	.034	.00001	.00001	.032	.156	.033	.00001	.04798
2	.029	.104	.037	.00001	.00057	.033	.131	.034	.00001	.01146
3	.037	.090	.032	.00001	.23689	.039	.143	.045	.00001	.00113
4	.025	.105	.021	.00001	.41943	.024	.156	.030	.00001	.00594
5	.018	.111	.020	.00001	.00341	.020	.165	.017	.00001	.52452

III gives summary statistics for the earnings' prediction errors where the prediction error is defined to be the predicted minus the actual squared.²⁴ Columns (1), (2), and (3) list the median prediction errors (across all firms) for the optimal univariate, optimal bivariate, and naive earnings models, respectively, for the 1980 observation. Using the Wilcoxon signed-rank test, column (4) reports the significance level at which the univariate model is a better predictor than the bivariate model. Similarly, column (5) provides the significance level at which the univariate model is a better predictor than the naive model.

On the basis of simple medians, the univariate earnings model predicts 1980 earnings better than the bivariate model for all earnings definitions. Also, on the basis of the Wilcoxon signed-rank test, the univariate model predicts significantly better than the bivariate model, confirming our earlier conclusion that investment does not cause earnings. Table III also shows that, for the NIBEI, Net Income, and Net Income Plus Depreciation definitions, the univariate model predicts significantly better for 1980 than does the naive model. On the other hand, for the other income definitions, the univariate model does not predict significantly better than the naive model. Table III also presents summary statistics for the 1981 prediction errors where the causality models were not updated for the 1980 observation. In all cases, the univariate earnings model predicted earnings significantly better than the bivariate model. With one exception, the univariate model also predicts significantly better than the naive model.

Table IV is the same as Table III, except now (future) investment is being predicted rather than earnings. On the basis of median prediction errors, the bivariate model appears to predict 1980 investment better than the alternatives, which is what we would expect from our conclusion that earnings cause investment. Using a Wilcoxon signed-rank test, the results are somewhat more ambiguous. While the bivariate model predicts significantly better than the naive model for all earnings definitions, it predicts better than the univariate investment model only for the two cash flow definitions. In the case of the other earnings definitions, the bivariate model does not predict significantly better than the univariate model. The reason for this result may be due to the hypothesis that investment, which itself is a cash outflow, is more closely related to the firm's cash inflows than to the firm's income. Therefore, Operating Income After Depreciation and the cash flow definitions, which are proxies for the firm's cash inflows, are better able to predict investment than is NIBEI or Net Income.

The 1981 investment variable results are very similar to the 1980 results except that the bivariate model is not significantly better than the naive model for the Net Income definition of earnings.

III. Conclusion

The purpose of this paper has been to test empirically the causality relationships between annual corporate earnings and investment. Using Akaike's Final Prediction Error criterion and the Hsiao sequential time-series procedure, we esti-

²⁴ This error metric is most consistent with the FPE criterion. See equation (3). In fact, since this error metric is essentially a cross-sectional FPE, we so label it in Tables III and IV.

Table IV
Results of Predictions for Investment

Income Definition	1980 Observation					1981 Observation				
	Significance Level					Significance Level				
	$FPE_I(q,0)$	$FPE_I(q,r)$	Naive	(1 - 2)	(3 - 2)	$FPE_I(q,0)$	$FPE_I(q,r)$	Naive	(6 - 7)	(8 - 7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	.069	.064	.093	.59482	.00001	.084	.076	.073	.96216	.01931
2	.068	.064	.091	.77204	.00120	.083	.083	.074	.99164	.21594
3	.074	.066	.102	.33505	.00001	.084	.075	.092	.33120	.00141
4	.079	.061	.105	.02260	.00001	.086	.065	.100	.00722	.00001
5	.078	.061	.110	.05157	.00001	.086	.073	.096	.13425	.00001

mated autoregressive causality models for both earnings and investment. On the basis of these models, we showed that expanding the earnings information set to include investment data does not increase the predictability of earnings. On the other hand, corporate earnings were shown to have predictive content when used in conjunction with past investment data to predict future investment behavior. Thus, we concluded that investment does not cause earnings but that earnings cause investment in the sense defined by Granger.

We further attempted to validate our conclusions by employing the estimated causality models and, also, a naive model to predict earnings and investment in a holdout sample. On the whole, the holdout sample confirmed our earlier results. Thus, with some minor exceptions, the univariate earnings model predicted earnings better than both the bivariate and naive models. Furthermore, the bivariate model predicted investment better in general than both the univariate and naive models for both the 1980 and 1981 observations, especially for the cash flow definitions of earnings.

Perhaps the most important conclusion of our study is that corporate earnings are a determinant of corporate investment. There are two potential reasons for this. In an imperfect capital market, the firm's wealth may constrain its investment opportunities. Thus, if the time series of past earnings is a proxy for wealth, one would expect such a time series to help predict corporate investment. Also, the earnings time series may provide a signal about the firm's ability to find and exploit truly profitable investment opportunities. Thus, suppliers of capital may be willing to reduce their required yield to firms with a good earnings track record (and vice versa), which, in turn, increases (decreases) the firm's incentive to undertake new investments.

REFERENCES

1. H. Akaike. "Statistical Predictor Identification." *Annals of the Institute of Statistical Mathematics* (1969), 203-17.
2. ———. "Fitting Autoregressions for Prediction." *Annals of the Institute of Statistical Mathematics* (1969), 243-47.
3. W. Steve Albrecht, Larry L. Lookabill, and James C. McKeown. "The Time-Series Properties of Annual Earnings." *Journal of Accounting Research* 15 (Autumn 1977), 226-44.
4. Ray Ball and Ross Watts. "Some Time Series Properties of Accounting Income." *Journal of Finance* 27 (June 1972), 663-82.
5. William J. Baumol, Peggy Heim, Burton G. Malkiel, and Richard E. Quandt. "Earnings Retention, New Capital and the Growth of the Firm." *Review of Economics and Statistics* 52 (November 1970), 345-55.
6. Lawrence D. Brown and Michael S. Rozeff. "Univariate Time-Series Models of Quarterly Accounting Earnings Per Share: A Proposed Model." *Journal of Accounting Research* 17 (Spring 1979), 179-89.
7. William A. Collins and William S. Hopwood. "A Multivariate Analysis of Annual Earnings Forecasts Generated from Quarterly Forecasts of Financial Analysts and Univariate Time-Series Models." *Journal of Accounting Research* 18 (Autumn 1980), 390-406.
8. John G. Cragg and Burton G. Malkiel. "The Consensus and Accuracy of Some Predictions of the Growth of Corporate Earnings." *Journal of Finance* 23 (March 1968), 67-84.
9. Edwin J. Elton and Martin J. Gruber. "Earnings Estimates and the Accuracy of Expectational Data." *Management Science* 18 (April 1972), 409-24.
10. ——— and Mustafa N. Gultekin. "Profession Expectations: Accuracy and Diagnosis of Errors." *Journal of Financial and Quantitative Analysis* 19 (December 1984), 351-63.

11. George Foster. "Quarterly Accounting Data: Time-Series Properties and Predictive Ability Results." *The Accounting Review* 52 (January 1977), 1-21.
12. Robert N. Freeman, James A. Ohlson, and Steven H. Penman. "Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation." *Journal of Accounting Research* 20 (Autumn 1982), 639-53.
13. Dov Fried and Dan Givoly. "Financial Analysts' Forecasts of Earnings: A Better Surrogate for Market Expectations." *Journal of Accounting and Economics* 4 (October 1982), 85-107.
14. Dan Givoly and Joseph Lakonishok. "The Quality of Analysts' Forecasts of Earnings." *Financial Analysts Journal* (September-October 1984), 40-47.
15. Myron J. Gordon. *The Cost of Capital to a Public Utility*. Michigan State University Public Utilities Studies, 1974.
16. Clive W. J. Granger. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (June 1969), 424-38.
17. Paul Griffin. "The Time-Series Behavior of Quarterly Earnings: Preliminary Evidence." *Journal of Accounting Research* 15 (Spring 1977), 71-83.
18. G. Hanoch and H. Levy. "The Efficiency Analysis of Choices Involving Risk." *The Review of Economic Studies* 36 (July 1969), 335-46.
19. Cheng Hsiao. "Autoregressive Modeling of Canadian Money and Income Data." *Journal of the American Statistical Association* 74 (September 1979), 553-60.
20. ———. "Causality Tests in Econometrics." *Journal of Economic Dynamics and Control* 1 (November 1979), 321-46.
21. ———. "Autoregressive Modeling and Money-Income Causality Detection." *Journal of Monetary Economics* 7 (January 1981), 85-106.
22. Dale W. Jorgenson. "Econometric Studies of Investment Behavior." *Journal of Economic Literature* 9 (December 1971), 1111-47.
23. ——— and Calvin D. Siebert. "A Comparison of Alternative Theories of Corporate Investment Behavior." *American Economic Review* 58 (September 1968), 681-712.
24. David F. Larcker. "The Association between Performance Plan Adoption and Corporate Capital Investment." *Journal of Accounting and Economics* 5 (April 1983), 3-30.
25. J. Lintner and R. Glauber. "Higgledy Piggledy Growth in America." In James Lorie and Richard Brealy (eds.), *Modern Developments in Investment Management*. New York: Praeger, 1972, 645-62.
26. I. M. D. Little. "Higgledy Piggledy Growth." *Bulletin of the Oxford University Institute of Statistics* 24 (November 1962), 387-412.
27. Kenneth S. Lorek. "Predicting Annual Net Earnings with Quarterly Earnings Time-Series Models." *Journal of Accounting Research* 17 (Spring 1979), 190-204.
28. P. Masani. "Recent Trends in Multivariate Prediction Theory." In P. R. Krishnaiah (ed.), *Multivariate Analysis I*. New York: Academic Press, 1966, 282-351.
29. Donald G. McFetridge. "The Efficiency Implications of Earnings Retentions." *The Review of Economics and Statistics* 60 (May 1978), 218-24.
30. Charles R. Nelson and G. William Schwert. "Tests for Predictive Relationships between Time Series Variables: A Monte Carlo Investigation." *Journal of the American Statistical Association* 77 (March 1982), 11-18.
31. P. Newbold. "Causality Testing in Economics." In O. D. Anderson (ed.), *Time Series Analysis: Theory and Practice I*. Amsterdam: North Holland, 1982, 701-16.
32. Daniel M. Shapiro, William A. Sims, and Gwenn Hughes. "The Efficiency Implications of Earnings Retentions: An Extension." *Review of Economics and Statistics* 65 (May 1983), 327-31.
33. Ross L. Watts and Richard W. Leftwich. "The Time Series of Annual Accounting Earnings." *Journal of Accounting Research* 15 (Autumn 1977), 253-71.
34. G. Whittington. "The Profitability of Retained Earnings." *Review of Economics and Statistics* 54 (May 1972), 152-60.