Volume 3, Section 1 – McShane, Cost of Capital

Q. (Appendix D, page 4)

a. Please provide copies of all the articles referenced on page 4 of Appendix D.

b. As all of the articles referenced on page 4 of Appendix D were published 18 or more years ago, and in light of the conflicted allegiances of many Wall Street securities analysts that were revealed as result of the investigations into the collapse of Enron, WorldCom, Global Crossing, and other corporations during the early 2000s, is Ms. McShane aware of any studies published during the last 5 years that support the notion that "empirical studies … conclude that investment analysts' growth forecasts serve as a better surrogate for investors expectations than historic growth rates" (Appendix D, page 4, top 2 lines)? If so, please provide copies.

A. a) Please see attached:

Attachment A - Lawrence D. Brown and Michael S. Rozeff "The Superiority of Analyst Forecasts as Measures of Expectations: Evidence from Earnings", *The Journal of Finance* Volume XXXII, No. 1 March 1978;

Attachment B - Dov Fried and Dan Givoly "Financial Analysts Forecasts of Earnings, A Better Surrogate for Market Expectations' *Journal of Accounting and Economics*, Vol. 4 1982;

Attachment C - R. Charles Moyer, Robert E. Chatfield, Gary D. Kelly "The Accuracy of Long-Term Earnings Forecasts in the Electric Utility Industry" *International Journal of Forecasting* Vol. 1 1985;

Attachment D - Robert S. Harris "Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return" *Financial Management* Spring 1986;

Attachment E - James H. Vander Weide and William T. Carleton "Investor Growth Expectations: Analysts vs. History: *The Journal of Portfolio Management* Spring 1988; and,

Attachment F - David Gordon, Myron Gordon and Lawrence Gould "Choice Among Methods of Estimating Share Yield" *The Journal of Portfolio Management* Spring 1989.

(b) Ms. McShane is not aware of newer studies that have focused on this specific attribute of analysts' forecasts.

Lawrence D. Brown and Michael S. Rozeff
"The Superiority of Analyst Forecasts as Measures of Expectations: Evidence from Earnings", *The Journal of Finance*, Vol. XXXII, No. 1
March 1978

The Journal of FINANCE

VOLUME XXXIII MARCH 1978 NO.

The Journal of THE AMERICAN FINANCE ASSOCIATION

The Journal of FINANCE

VOL. XXXIII MARCH 1978 NO.

THE SUPERIORITY OF ANALYST FORECASTS AS MEASURES OF EXPECTATIONS: EVIDENCE FROM EARNINGS

LAWRENCE D. BROWN AND MICHAEL S. ROZEFF*

ACCURATE MEASUREMENT OF EARNINGS expectations is essential for studies of firm valuation, cost of capital and the relationship between unanticipated earnings and stock price changes. Under the rational expectations hypothesis [23], market earnings expectations should be measured by the best available earnings forecasts. Univariate time series forecasts are often used for this purpose ([1], [3], [4], [5], [12], [13], [14], [16], [18], [20]) instead of direct measures of earnings expectations such as security analysts' forecasts. Univariate time series forecasts neglect potentially useful information in other time series and therefore do not generally provide the most accurate possible forecasts [24]. Since security analysts process substantially more data than the time series of past earnings, their earnings forecasts *should* be superior to time series forecasts and provide better measures of market earnings expectations.

However, the mere existence of analysts as an employed factor in long run equilibrium means that analysts *must* make forecasts superior to those of time series models. To reach this conclusion, one need only assume that participants in the market for forecasts act in their own best interests and that both forecast producers and consumers demand forecasts solely on the basis of their predictive ability.[1] Since analysts' forecasts cost more than time series forecasts, the continued employment of analysts by profit-maximizing firms implies that analysts' forecasts must be superior to those of the lower cost factor, time series models.

Past comparisons of analysts' forecasts to sophisticated time series models conclude that analysts' forecasts are not more accurate than time series forecasts (Cragg and Malkiel (CM) [9]; Elton and Gruber (EG) [11]). This evidence plainly conflicts with basic economic theory. Hence, the predictive accuracy of analysts' forecasts is re-examined in this paper. In contrast with other studies, the results overwhelmingly favor the superiority of analysts over time series models.

Part I considers statistical tests and experimental design. Part II contains the empirical results. Summary and implications appear in Part III.

* College of Business Administration, The University of Iowa, Iowa City.

1. We assume that forecast purchasers do not derive nonmonetary benefits from forecasts.

I. Experimental Design

A. *Statistical Evaluation of Forecast Methods*

Without direct information on the costs of imperfect forecasts to forecast users, comparative forecast accuracy is usually evaluated by comparing the error distributions of different forecast methods statistically. However, statistical comparisons in past studies ([9], [11]) utilize test statistics improperly, particularly Theil's U [25] and Student's t. In this section, after discussing the defects of these statistics for evaluating two or more forecast methods, the alternative statistical methods used in this study are introduced.[2]

Theil's U-statistic (applied to earnings) is the square root of

$$U_{ij}^2 = \frac{\sum_{t=1}^{T} (\dot{P}_{ijt} - \dot{A}_{it})^2}{\sum_{t=1}^{T} \dot{A}_{it}^2}$$

where $\dot{A}_{it}$ = change in actual earnings per share of firm i from $t-1$ to t,
$\dot{P}_{ijt}$ = predicted change in earnings per share of firm i from $t-1$ to t by forecast method j, and
T = total number of time series observations.

For its computation, it requires *time series* data on a firm's earnings *changes*.[3] Given forecast method j and earnings time series data on firm i, Theil's U compares the forecast accuracy of method j to that of a naive, no change, earnings forecast model.[4,5] Since analysts' earnings forecasts are currently available only in short time series, use of Theil's U for comparative forecast evaluation necessarily relies on small samples.[6] Larger sample sizes are possible by testing forecast methods on a cross-section of firms. Finally, no procedure is available with tests of significance which uses Theil's U to compare two forecast methods when neither is a no-change method. Direct hypothesis tests are preferable to inferences drawn from ranking the U statistics of different forecast methods.

For hypothesis tests of two forecast methods, an appropriate design is a one-sample or matched pairs case with self-pairing by firm. The members of each pair

2. Past studies also contain experimental biases: CM compare analysts' five-year forecasts with realizations over three and four-year horizons; EG compare analysts' forecasts with the "best" of nine time series models selected from the same time period in which comparisons with analysts' forecasts are made. This procedure introduces *ex post* selection bias.

3. EG computed "Theil's U" using earnings *levels* rather than *changes*. This statistic has unknown sampling properties.

4. $\dot{P}_{ijt} = \dot{A}_{it}$ and $U_{ij} = 0$ if prediction is perfect in every period. If no change is predicted in each period (i.e., $\dot{P}_{ijt} = 0$), $U_{ij} = 1$; $0 < U_{ij} < 1$ if prediction is less than perfect but better than the no-change prediction and $U_{ij} > 1$ if forecast method j is less accurate than the no-change prediction.

5. CM used *cross-sectional* rather than temporal data. This "Theil's U" statistic has unknown sampling properties because each error is drawn from a different error distribution, one for each firm.

6. EG's sample size in computing Theil's U varied between two and six.

are the errors from the two methods; the matched pair is reduced to a single observation by taking the difference in the errors. The usual parametric test of the mean difference is the paired t-test [17]. An alternative non-parametric test of the median difference is the Wilcoxon Signed Ranks test [8].

The parametric paired t-test is inappropriate for testing mean error differences of forecast methods applied to cross-section earnings data. If applied to error measures stated in level form (e.g., $|P_{ijt} - A_{it}|$, where P_{ijt} = firm i's forecasted earnings per share for period t by method j and A_{it} = firm i's actual earnings per share in period t), the test's assumption that paired differences are drawn from the same population is violated since each error difference depends upon each firm's earnings per share level. If applied to error measures stated in ratio form (e.g., $|P_{ijt} - A_{it}|/|A_{it}|$), the distributional assumptions of the paired t-test are also unlikely to be fulfilled since ratio measures applied to earnings per share data are dominated by outliers because actual earnings per share are often close to zero.[7]

Meaningful pairwise comparisons require test statistics which are insensitive to error definition and outliers. We adopt the Wilcoxon Signed Ranks test which meets these requirements and has power comparable to the parametric paired t-test [8, p. 213].

For tests of several forecast methods, the generalization of the paired t-test, two-way analysis of variance, is inapplicable.[8] The Friedman test [8], which is based on two-way analysis of variance by ranks and is independent of error definition, is used instead.

For an error measure, we choose relative error ignoring sign, $|P_{ijt} - A_{it}|/|A_{it}|$, a metric which is likely to be of interest to forecast purchasers.[9] In any event, the Wilcoxon test statistic is insensitive to error definition (see fn. 16).

B. *Forecast Horizon*

Because economic theory provides no guidance concerning the association of analyst superiority with a particular forecast horizon, several horizons should be investigated.[10] Our choice of horizons reflects the following considerations: (i) micro-level information obtained by analysts often concerns earnings of the following several quarters or fiscal year; (ii) current fiscal and monetary policies affect earnings of the subsequent one to five quarters; (iii) published forecasts are available mainly for short horizons. We thus investigate point estimates of quarterly earnings per share for forecast horizons of one to five quarters. We also examine annual earnings forecasts. The basic time series data are quarterly primary

7. EG's cross-section parametric t-test is inappropriate. Their use of an error measure stated in terms of levels squared (mean square error) appears to compound the inherent difficulty in applying the paired t-test to cross-section earnings data (see fn. 16).

8. Preliminary tests indicated serious violation of the homogeneity of variances and additivity assumptions, basically because of error outliers. Violation of the ANOVA assumptions also prevents application below of a factorial design with sample year and forecast horizon as factors, forecast method as treatment and firm as replication.

9. For a discussion of the deficiencies of using $|P_{ijt}|$ or $|P_{ijt} + A_{it}|/2$ in the denominator see [25].

10. The forecast horizons studied in the past have been five years (CM) and one year (EG).

earnings per share before extraordinary items, adjusted for stock splits, stock dividends and other capitalization changes for the years 1951–1975.

Ex ante conditional predictions of all forecast methods are determined as follows for a sample of 50 firms for each of the four years 1972–1975. Starting with third quarter 1971 earnings (III/1971), conditional earnings per share predictions for the ith firm by the jth method are obtained for the individual quarters of 1972. The forecasts of 1972 quarterly earnings, conditional on III/1971, are denoted P_{ij}(I/1972 | III/1971), P_{ij}(II/1972 | III/1971), P_{ij}(III/1972 | III/1971) and P_{ij}(IV/1972 | III/1971). Moving ahead one quarter, predictions are again obtained for each of the four quarters of 1972 made conditional upon IV/1971 earnings data. Again moving ahead one quarter, predictions are obtained for the last three quarters of 1972 conditional upon knowledge of I/1972 earnings, etc. Table 1 shows the set of 1972 predictions so obtained. With these conditional predictions, relative forecast errors ignoring sign are computed for each forecast method j over five distinct quarterly forecast horizons for use in the quarterly error comparisons. Annual earnings forecasts for 1972 are the sum of the forecasts P_{ij}(I/1972 | IV/1971), P_{ij}(II/1972 | IV/1971), P_{ij}(III/1972 | IV/1971), and P_{ij}(IV/1972 | IV/1971), that is, the one to four period ahead point forecasts made conditional upon knowledge of the prior year's fiscal earnings.[11] After obtaining analogous forecasts for the years 1973, 1974 and 1975, quarterly and annual comparisons are repeated for these years.

TABLE 1

SUMMARY OF PREDICTIONS BY FORECAST HORIZON FOR 1972[a,b]

1 Quarter Ahead	2 Quarters Ahead	3 Quarters Ahead	4 Quarters Ahead	5 Quarters Ahead[c]
P_{ij}(I/1972|IV/1971)	P_{ij}(I/1972|III/1971)			
P_{ij}(II/1972|I/1972)	P_{ij}(II/1972|IV/1971)	P_{ij}(II/1972|III/1971)		
P_{ij}(III/1972|II/1972)	P_{ij}(III/1972|I/1972)	P_{ij}(III/1972|IV/1971)	P_{ij}(III/1972|III/1971)	
P_{ij}(IV/1972|III/1972)	P_{ij}(IV/1972|II/1972)	P_{ij}(IV/1972|I/1972)	P_{ij}(IV/1972|IV/1971)	P_{ij}(IV/1972|III/1971)

[a] Predictions missing from the table (e.g., P_{ij}(I/1972|II/1971), P_{ij}(II/1972|II/1971) are absent because our source of analyst data does not contain these forecasts.

[b] i and j refer to firm i and method j, respectively.

[c] Five quarter ahead are available for BJ and V only.

C. *Time Series Models and Analysts' Forecasts*

Within the class of univariate time series models, Box and Jenkins (BJ) [6] models are highly regarded for their ability to make the most efficient use of the time series data. The BJ modelling technique enables one to select the most appropriate time series model consistent with the process generating each firm's time series of quarterly earnings per share data. BJ models, by not making *a priori* assumptions about the processes generating the data, subsume autoregressive,

11. Beaver [1] concludes that a quarterly approach to predicting annual earnings is at least as good as an annual approach to predicting annual earnings. Also see [7], [19] and [22] for other aspects of the usefulness of quarterly earnings per share data.

moving average and mixed models as special cases.[12] Forecasts of individually fitted BJ models should, therefore, perform better than forecasts of a particular class of time series models applied to all firms' time series data. We adopt the BJ modelling technique in this paper. Two other time series models are also included, a "seasonal martingale" (denoted M) and a "seasonal submartingale" (S). These models have been used as standards of comparison in the earnings forecast literature and are available for forecast producers and users at minimal cost.

As a source of analysts' forecasts we choose the Value Line Investment Survey since it contains one to five quarter ahead earnings forecasts which can be accurately dated and measured. Value Line makes earnings forecasts for 1,600 firms in contrast with institutional research firms which provide fewer, more expensive forecasts. Our hypothesis test thus compares a relatively sophisticated time series model with an "average" source of analysts' forecasts.

BJ conditional forecasts are obtained by standard methods after identifying and estimating each firm's appropriate model [6].[13] Value Line's conditional forecasts are taken directly from individual issues of the Value Line Investment Survey. The Survey, published weekly, makes quarterly earnings predictions four times a year for each firm included.

To define conditional forecasts of the naive models for each firm i, let A_{it} denote the tth actual quarterly earnings per share for firm i, where $t=1,\ldots,96$ (I/1951–IV/1974).

Seasonal submartingale (S) conditional one to four quarter ahead forecasts at time t are

one quarter ahead	$A_{it-3}+(A_{it}-A_{it-4})$
two quarters ahead	$A_{it-2}+(A_{it}-A_{it-4})$
three quarters ahead	$A_{it-1}+(A_{it}-A_{it-4})$
four quarters ahead	$A_{it}\quad+(A_{it}-A_{it-4})$.

Seasonal martingale (M) conditional one to four quarter ahead forecasts made in period t are A_{it-3}, A_{it-2}, A_{it-1}, and A_{it}. M's forecasts for a given quarter do not change as actual earnings per share data become available. S modifies M's forecasts with the change of the latest period's quarter over that of the previous year.

Actual quarterly earnings data are announced for most firms approximately five to six weeks into the subsequent quarter. Time series forecasts then become

12. The *ad hoc* time series models used in previous studies at a time when BJ techniques were unavailable are special cases of BJ models.

13. Recent research by Froeschle [15] and diagnostic tests of Dent and Swanson [10] were helpful in identifying the BJ models in addition to the standard diagnostic tests. As an aid to identifying the BJ models, most of which had multiplicative seasonal components, theoretical autocorrelation and partial autocorrelation functions for many quarterly multiplicative seasonal models were obtained. The coefficients of the BJ models, estimated with data through IV/1974, were not re-estimated with less data for earlier periods or more data for later periods. Foster [13] has shown that coefficient re-estimation of BJ quarterly earnings models is unnecessary due to its negligible effect on forecast errors. In any event, our procedure (no re-estimation) favors BJ in nearly all comparisons with Value Line.

possible and Value Line forecasts are published, on average, forty to fifty days later.[14]

The pattern of forecasts for all models is summarized in Table 1. Note that models M and S are not used to generate five quarter ahead forecasts.

II Empirical Results

A. *Sample Selection*

Fifty firms were randomly selected from Moody's Handbook of Common Stocks. Each firm has complete quarterly earnings data available from 1951, is included in the Value Line Investment Survey since 1971 and has a December fiscal year. The resulting sample (Appendix A) is representative of the New York Stock Exchange firms included in Moody's and Value Line. Utilities were excluded due to insufficient quarterly earnings data. Sample sizes are reduced in those rare instances when the Value Line conditional forecasts are unavailable.

B. *Annual Comparisons*

The error distributions of relative annual forecast errors are shown in Table 2 for each of the years 1972–75 using the four forecast methods, seasonal martingale (M), seasonal submartingale (S), Box-Jenkins (BJ) and Value Line (V). Table 2 also contains Friedman test statistics (Chi-square with 3 degrees of freedom) and Wilcoxon test statistics (Student's t with $N-1$ degrees of freedom where N is sample size). The Friedman test statistic examines the null hypothesis that *all four* error distributions are identically distributed; the Wilcoxon statistic tests the null hypothesis that the median error difference of *two* methods being compared exceeds zero.

Using the Friedman test, the null hypothesis is rejected at the 1% level in 1972, 1973 and 1975. In the 12 pairwise hypothesis tests of V's errors against those of M, S, and BJ, the sign of the Wilcoxon test statistic favors Value Line in every instance. Statistical significance occurs 8 times; 6 times at the 1% level and twice at the 5% level. Thus, V generally produces smaller annual errors than the three time series models suggesting that Value Line annual earnings forecasts are superior to those of time series models.

As argued earlier, BJ forecasts should be superior to forecasts of *ad hoc* time series models. The annual comparisons show that the BJ models generally yield smaller forecast errors than the other time series models studied. In 8 comparisons with M and S, the Wilcoxon test favors BJ 7 times with statistical significance 3 times. These findings suggest that BJ's forecasts are superior to those of *ad hoc* naive time series models.

While the annual results provide strong support for the hypothesis of analyst superiority, they use only a fraction of the data. More powerful tests are achieved using the larger sample sizes of the quarterly data and many more comparative tests can be performed with these data. We turn next to quarterly comparisons.

14. The time interval from announcement to forecast varies from approximately 7 to 70 days for our sample firms. The fact that the Investment Survey, published in 13 installments, makes forecasts for different firms each week accounts for the variation.

e. forty to fifty days

ı Table 1. Note that
forecasts.

ıdbook of Common
ıilable from 1951, is
nd has a December
ve of the New York
tilities were excluded
educed in those rare
ıilable.

shown in Table 2 for
seasonal martingale
e Line (V). Table 2
ees of freedom) and
reedom where N is
pothesis that *all four*
atistic tests the null
ds being compared

he 1% level in 1972,
against those of M,
alue Line in every
% level and twice at
than the three time
asts are superior to

asts of *ad hoc* time
dels generally yield
d. In 8 comparisons
stical significance 3
to those of *ad hoc*

pothesis of analyst
l tests are achieved
more comparative
rly comparisons.

ely 7 to 70 days for our
its, makes forecasts for

TABLE 2

WILCOXON AND FRIEDMAN TEST STATISTICS AND ERROR DISTRIBUTIONS, ANNUAL COMPARISONS OF VALUE LINE AND TIME SERIES MODEL PREDICTION ERRORS, 1972–1975[c]

1972

Error Distribution[d]

	<.05	.05–.10	.10–.25	.25–.50	.50–.75	.75–1.00	>1.00
M	3	7	14	17	4	3	2
S	11	6	12	10	3	1	7
BJ	10	6	12	12	4	1	5
V	13	7	17	12	0	0	1

SAMPLE SIZE = 50
Friedman Statistic = 27.10[a]
Wilcoxon Statistics[e]

	S	BJ	V
M	−.55	.24	4.46[a]
S		.46	3.50[a]
BJ			3.45[a]

1973

Error Distribution[d]

	<.05	.05–.10	.10–.25	.25–.50	.50–.75	.75–1.00	>1.00
M	2	6	16	18	6	0	2
S	11	8	14	9	4	1	3
BJ	8	6	15	16	3	0	2
V	10	9	13	16	0	0	2

SAMPLE SIZE = 50
Friedman Statistic = 33.19[a]
Wilcoxon Statistics[e]

	S	BJ	V
M	3.15[a]	2.51[a]	4.61[a]
S		−1.89[b]	0.34
BJ			2.17[b]

1974

Error Distribution[d]

	<.05	.05–.10	.10–.25	.25–.50	.50–.75	.75–1.00	>1.00
M	8	6	12	15	4	1	4
S	12	3	11	12	6	2	4
BJ	5	8	16	13	4	0	4
V	6	7	15	13	5	0	4

SAMPLE SIZE = 50
Friedman Statistic = 4.68
Wilcoxon Statistics[e]

	S	BJ	V
M	−.21	2.37[a]	2.23[b]
S		1.24	1.44
BJ			0.61

TABLE 2 (*continued*)

1975

Error Distribution[d]

	<.05	.05–.10	.10–.25	.25–.50	.50–.75	.75–1.00	>1.00
M	4	7	13	10	2	3	11
S	3	5	12	7	9	4	10
BJ	7	3	13	12	2	3	10
V	7	5	18	5	3	3	9

SAMPLE SIZE = 50

Friedman Statistics = 12.84[a]

Wilcoxon Statistics[e]

	S	BJ	*V*
M	−1.77[b]	0.86	3.29[a]
S		2.99[a]	3.11[a]
BJ			1.28

[a] Significant at the 1% level, one-tailed test.

[b] Significant at the 5% level, one-tailed test.

[c] *V* = Value Line, *M* = Seasonal Martingale, *S* = Seasonal Submartingale, BJ = Box-Jenkins.

[d] Each entry below designates the number of observations for a given model whose relative error ignoring sign is within the stated fractiles.

[e] Each Wilcoxon test statistic below results from comparing the method at the top with the method on the side. Thus, positive Wilcoxon statistics indicate superiority of model on top.

C. *Quarterly Comparisons*

In each year, 1972 to 1975, quarterly forecasts are obtained for the forecast methods in the manner shown in Table 1. Relative forecast errors of all four methods are compared over 1–4 quarter forecast horizons; BJ and *V* are also compared over 5 quarter horizons. In each of the four years, sample sizes are approximately 200 for the 1 and 2 quarter ahead comparisons, 150 for the 3 quarter ahead comparisons, and 100 for the 4 quarter ahead comparisons. Test results over all horizons appear in Table 3 and are summarized in Table 4.

With minor exceptions (3 and 4 quarter horizons in 1974), the Friedman statistics are highly significant when the four methods are tested as a group; the null hypothesis of identically distributed distributions is rejected in 14 of the 16 Friedman tests. Using Wilcoxon test statistics, *V*'s errors are tested pairwise against *M*'s and *S*'s errors 16 times each and against BJ's errors 20 times. The resulting 52 hypothesis tests of *V* against *M*, *S* and BJ are summarized in Table 4A. In the 34 instances of significant Wilcoxon test statistics, *V* is statistically superior 33 times. In the remaining 18 tests, the sign of the t-statistic favors *V* 12 times. In total, *V* is favored 45 times out of 52, revealing an overwhelming dominance of *V* over the time series models.

The data are also summarized in Table 4 by the mean Wilcoxon t-value ($\bar{t}$), the estimated standard deviation of the mean t-value ($s(\bar{t})$) and the ratio $\bar{t}/s(\bar{t})$. The latter ratio is itself a t-statistic only if each t-value being averaged is drawn from the same distribution. Since the distribution of t-values is likely to depend upon the horizon, model and/or year that the experiment is conducted, we refrain from

TABLE 3

WILCOXON AND FRIEDMAN TEST STATISTICS, QUARTERLY COMPARISONS OF VALUE LINE AND TIME SERIES MODEL PREDICTION ERRORS, 1972–1975[c,d]

		Forecast Horizon												
		One Quarter			Two Quarter						Four Quarter			Five Quarter
		S	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*V*
	M	2.14[b]	6.87[a]	8.15[a]	0.79	5.41[a]	6.87[a]	−1.09	2.50[a]	5.77[a]	−3.09[a]	1.41	5.22[a]	—
1972	*S*	—	4.62[a]	5.25[a]	—	4.62[a]	5.57[a]	—	3.03[a]	5.42[a]	—	3.38[a]	5.30[a]	—
	BJ	—	—	1.75[b]	—	—	2.51[a]	—	—	4.09[a]	—	—	3.93[a]	3.11[a]
		Sample Size = 200			Sample Size = 200			Sample Size = 150			Sample Size = 100			Sample Size = 50
		Friedman Stat. = 73.45[a]			Friedman Stat. = 60.54[a]			Friedman Stat. = 41.14[a]			Friedman Stat. = 43.43[a]			
		S	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*V*
	M	8.02[a]	8.98[a]	10.66[a]	5.81[a]	6.41[a]	8.70[a]	4.81[a]	3.52[a]	6.31[a]	2.55[a]	1.69[b]	4.63[a]	—
1973	*S*	—	−0.60	1.62	—	−1.83[b]	1.04	—	−3.57[a]	−0.02	—	−1.59	1.04	—
	BJ	—	—	2.48[a]	—	—	3.47[a]	—	—	3.34[a]	—	—	2.79[a]	1.66
		Sample Size = 199			Sample Size = 200			Sample Size = 150			Sample Size = 100			Sample Size = 50
		Friedman Stat. = 173.51[a]			Friedman Stat. = 119.91[a]			Friedman Stat. = 75.22[a]			Friedman Stat. = 29.12[a]			
		S	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*V*
	M	3.35[a]	6.29[a]	6.19[a]	0.84	4.88[a]	3.78[a]	−0.25	2.59[a]	1.29	−2.69[a]	1.41	0.29	—
1974	*S*	—	2.34[a]	2.95[a]	—	2.31[b]	1.50	—	1.53	0.97	—	2.67[a]	2.80[a]	—
	BJ	—	—	1.16	—	—	−1.45	—	—	−1.04	—	—	−0.92	−2.20[b]
		Sample Size = 199			Sample Size = 199			Sample Size = 149			Sample Size = 100			Sample Size = 50
		Friedman Stat. = 47.57[a]			Friedman Stat. = 22.63[a]			Friedman Stat. = 5.40			Friedman Stat. = 2.92			
		S	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*S*	BJ	*V*	*V*
	M	2.07[b]	5.76[a]	8.22[a]	−2.64[a]	3.63[a]	5.29[a]	−4.49[a]	2.93[a]	2.95[a]	4.89[a]	−0.78	−0.05	—
1975	*S*	—	4.70[a]	6.36[a]	—	6.02[a]	6.14[a]	—	6.13[a]	5.14[a]	—	3.62[a]	3.28[a]	—
	BJ	—	—	3.51[a]	—	—	1.62	—	—	−0.22	—	—	0.08	0.45
		Sample Size = 199			Sample Size = 199			Sample Size = 149			Sample Size = 100			Sample Size = 50
		Friedman Stat. = 80.32[a]			Friedman Stat. = 44.49[a]			Friedman Stat. = 33.25[a]			Friedman Stat. = 15.66[b]			

[a] Significant at the 1% level, one-tailed test.

[b] Significant at the 5% level, one-tailed test.

[c] *V* = Value Line, *M* = Seasonal Martingale, *S* = Seasonal Submartingale, BJ = Box-Jenkins.

[d] Each Wilcoxon test statistic entered in the table results from comparing method at the top with method on the side. Thus, positive Wilcoxon statistics indicate superiority of model on top.

TABLE 4

SUMMARY OF WILCOXON TEST COMPARISONS

A: Value Line vs. Time Series Models[a]

		Forecast Horizon					Forecast Model			Year			
	Total	1Q	2Q	3Q	4Q	5Q	M	S	BJ	1972	1973	1974	1975
Number of Comparisons	52	12	12	12	12	4	16	16	20	13	13	13	13
Comparisons Favorable to V[b]	45	12	11	9	10	3	15	15	15	13	12	9	11
Comparisons Statistically Favorable to V[c]	33	10	8	7	7		13	10	10	13	8	4	8
Comparisons Statistically Unfavorable to V	1	0	0	0	0	1	0	0		0	0		0
Mean Wilcoxon Test Statistic ($\bar{t}$)	3.25	4.86	3.75	2.83	2.37	.76	5.27	3.40	1.51	4.84	3.67	1.18	3.29
$\bar{t}/s(\bar{t})$[d]	8.27	5.45	4.51	3.81	3.72	.67	5.65	6.24	3.48	9.98	4.18	1.81	4.24

B: BJ vs. Naive Time Series Models

		Forecast Horizon				Forecast Model		Year			
	Total	1Q	2Q	3Q	4Q	M	S	1972	1973	1974	1975
Number of Comparisons	32	8	8	8	8	16	16	8	8	8	8
Comparisons Favorable to BJ[b]	27	7	7	7	6	15	12	8	4	8	7
Comparisons Statistically Favorable to BJ[c]	24	7	7	6	4	13	11	7	4	6	7
Comparisons Statistically Unfavorable to BJ	2	0	1	1	0	0	2	0	2	0	0
Mean Wilcoxon Test Statistic ($\bar{t}$)	3.15	4.87	3.93	2.33	1.48	3.97	2.34	3.98	1.63	3.00	4.00
$\bar{t}/s(\bar{t})$[d]	6.37	4.70	4.16	2.41	2.25	6.23	3.25	6.46	1.05	4.99	4.96

[a] V = Value Line, M = Seasonal Martingale, S = Seasonal Submartingale, BJ = Box-Jenkins.

[b] Comparisons are favorable if Wilcoxon statistic in Table 3 is positive.

[c] Comparisons are statistically favorable if Wilcoxon statistic in Table 3 is positive and significant at the 5% level or better.

[d] Both $\bar{t}$ and $s(\bar{t})$ are computed using the number of comparisons in each column of the Table.

Mean Wilcoxon Statistic ($\bar{t}$)	3.15	4.87	3.93	2.33	1.48	3.97	2.34	3.98	1.63	3.00	4.00
$\bar{t}/s(\bar{t})^d$	6.37	4.70	4.16	2.41	2.25	6.23	3.25	6.46	1.05	4.99	4.96

[a] V = Value Line, M = Seasonal Martingale, S = Seasonal Submartingale, BJ = Box-J...

[b] Comparisons are f... if Wilcoxon statistic in Table 3 is positive.

[c] Comparisons are statistically favorable if Wilcoxon statistic in Table 3 is positive and significant at the 5% level or better.

[d] Both $\bar{t}$ and $s(\bar{t})$ are computed using the number of comparisons in each column of the Table.

hypothesis tests on $\bar{t}$ and present $\bar{t}$ and $\bar{t}/s(\bar{t})$ without formal tests of significance. For the 52 comparisons involving V, the mean Wilcoxon test statistic is 3.25 and $\bar{t}/s(\bar{t})$ is 8.27.

Table 4A also decomposes the 52 comparisons of V with the time series models by forecast horizon, model and year.[15] The data show that Value Line's forecast superiority holds over all horizons studied with a tendency for its superiority to decline as horizon lengthens. V's predominance model-by-model is, as hypothesized, quite evident with somewhat less superiority over BJ than over M and S. Turning our attention to the 20 comparisons between V and BJ, V is superior in 10 of 11 cases in which the test statistic is significant. In 5 of the remaining 9 comparisons, the sign of the Wilcoxon test statistic favors V. For completeness, Table 4A summarizes Wilcoxon tests by year. Again we expect V to be superior, on average, but have no hypothesis concerning particular years. Comparisons unfavorable to V tend to be confined to 1974, but even in this year, 4 of the 5 statistically significant comparisons favor Value Line.

In summary, the evidence strongly supports the hypothesis that Value Line consistently makes significantly better predictions than time series models. The statistically significant experiments overwhelmingly favor Value Line. In the remaining experiments the majority of the Wilcoxon tests also favor Value Line, providing additional support for the hypothesis of analyst superiority.

Table 4B summarizes the 32 comparisons of BJ with the naive time series models. The mean Wilcoxon test statistic is 3.15 and $\bar{t}/s(\bar{t})$ equals 6.37. In 26 cases, there are significant differences with BJ statistically superior 24 times. BJ is superior to M and S in 3 of the remaining 6 comparisons. Hence, BJ is favored in 27 of 32 comparisons, providing strong support for the hypothesis that BJ predicts earnings better than *ad hoc* time series models.

Table 4B also summarizes comparisons involving BJ by horizon, model and year. BJ's superiority over the naive models is clearly evident over each forecast horizon with a tendency for its superiority to decline as horizon lengthens. In comparison to individual models, BJ outperforms both M and S with somewhat less dominance over S. Turning to comparisons by year, the superiority of BJ is consistent over time, with most of the comparisons unfavorable to BJ occurring in 1973. Even in this year, the mean Wilcoxon test statistic is 1.63 and 4 of the 6 significant comparisons favor BJ.[16]

In conclusion, the quarterly and the annual comparisons provide convincing evidence both of Value Line's superiority over each of the three time series models and BJ's superiority over the naive models. The quarterly results also show that V's superiority over the time series models and BJ's superiority over the naive models

15. The decomposition is an alternative to analysis of variance which is inapplicable to the error distribution (see fn. 8).

16. As noted earlier, the Wilcoxon tests should be insensitive to error definition. Wilcoxon test statistics were recomputed on annual and selected quarterly comparisons using three additional error measures, mean square error, root mean square error and relative error squared. The small changes in the test statistics left the results virtually unchanged. Parametric t-tests were also applied to the four error measures. Both the sign and magnitude of these test statistics were highly sensitive to error definition. The hypothesis tests using the parametric t-test most often gave results in disagreement with the Wilcoxon test when mean square error was chosen as the error definition. This may account for EG's results differing from ours.

are not confined to particular models, horizons, or years. The very general character of Value Line's superiority in predicting earnings, evidenced over all models, horizons, and years in 64 separate hypothesis tests involving sample sizes averaging 125, lends extraordinary support to the hypothesis of analyst superiority.

D. *Further Analysis*

The superiority of Value Line over time series models follows from the rational behavior of forecast producers and consumers and should be generalizable to other sources of analyst forecasts and other time periods. As a preliminary test of the sensitivity of our results to choice of analyst, we obtained predictions of 1975 annual earnings per share made by the Standard and Poor's Earnings Forecaster (SP) for each firm included in the 1975 annual earnings sample.[17] Wilcoxon tests of SP against M, S, and BJ favored SP, yielding t-statistics of 3.18, 2.85 and 1.45 respectively. These results are remarkably similar to those using Value Line.[18] This evidence suggests that Value Line's forecast superiority over time series models is not unique.

To ascertain whether the sample period posed unusual difficulties for time series earnings forecasting, a BJ model was fitted to the Quarterly Earnings Index of the Dow Jones Industrial Average over the 1951–1975 time period.[19] Average quarterly percentage errors ignoring sign produced by the BJ model for 1972–1975 were 7.31%, 6.61%, 9.99%, and 15.47% respectively. Since the mean and standard deviation of average percentage forecast errors over the 1951–1975 period were 10.14% and 4.38%, it appears that the 1972–1975 period was not a particularly difficult one in which to predict earnings. Indeed, from this standpoint, the 1972–1975 period is comparable to the "stable" years of the sixties, 1962–1967, studied by CM and EG.[20]

These results indicate that if appropriate hypothesis tests are applied to other analysts and time periods, the results are likely to parallel those using Value Line and the 1972–1975 time period.

E. *A Brief Investigation of Value Line Superiority*

To produce forecasts superior to time series models, Value Line must utilize information not contained in the time series of quarterly earnings. During the period between the most recent quarterly earnings announcement and the subsequent Value Line prediction, Value Line acquires incremental information which, if an important part of its total information set, may explain Value Line's

17. SP, published weekly, contains annual predictions made by Standard and Poor's and other investment firms. The SP prediction for each firm is that made by Standard and Poor's on the date closest to the Value Line prediction date.

18. V's t-statistics versus M, S, and BJ were 3.29, 3.11, and 1.28 respectively (See Table 2). A direct Wilcoxon test between V and SP favored $V(t=.77)$.

19. The sample period, 1972–1975, may appear "unusual" since it includes peacetime wage and price controls, high inflation and inventory profits, large changes in employment and new accounting requirements. If events arising during the sample period caused the earnings generating process to change, the forecast ability of the BJ modelling technique may be hampered, unintentionally favoring the analyst.

20. The average percentage errors were 12.67%, 10.71%, 7.03%, 4.93%, 6.08% and 5.26%, respectively for 1962–1967.

superiority. Information arising during this interval is likely to be most important for predicting next quarter's earnings. Assuming that the generation of this incremental information is positively related to the passage of time, earnings should be relatively easier to predict the further Value Line's prediction date is from the most recent earnings announcement date, and one quarter horizon forecast errors should be negatively related to the corresponding intervals.

To test this hypothesis, we obtained for the firms in the 1975 one quarter horizon sample their Value Line errors and the time intervals (7–70 days) since their most recent earnings announcements. A rank correlation was applied to these variables. The insignificantly negative Spearman rho which was obtained suggests that information obtained by Value Line during this interval has a negligible effect on its ability to predict next quarter's earnings.[21] This evidence is consistent with the hypothesis that Value Line's superiority can be attributed to its use of the information set available to it on the quarterly earnings announcement date, and not to the acquisition of information arising after the quarterly earnings announcement date.

III. Summary and Implications

Basic economic theory and the equilibrium employment of analysts, a higher cost factor than time series models, imply that analysts must produce better forecasts than time series models. Past studies ([9], [11]) of comparative earnings forecast accuracy have concluded otherwise but use inappropriate parametric tests and contain experimental biases. Using nonparametric statistics which provide proper yet powerful tests, we find that (1) BJ models consistently produce significantly better earnings forecasts than martingale and submartingale models; (2) Value Line Investment Survey consistently makes significantly better earnings forecasts than the BJ and naive time series models. The findings are in accord with rationality in the market for forecasts and the long-run equilibrium employment of analysts.

[redacted]able earnings forecasts should be used to measure market earnings expectations. Given rational market expectations, our evidence of analyst superiority over time series models means that analysts' forecasts should be used in studies of firm valuation, cost of capital and the relationship between unanticipated earnings and stock price ch[redacted] [21]) [redacted] levels are significantly better explained by analysts' earnings

21. The lack of a significant negative correlation between prediction error and time since last announcement date may occur if the interval is intentionally lengthened by Value Line in order to acquire more information about the firms whose earnings are more difficult to predict. To test this possibility, we measured each firm's prediction "difficulty" by its average one quarter horizon percentage error ignoring sign yielded by its BJ model. No significant correlation was found between this variable and the time interval between the most recent quarterly earnings announcement and the Value Line prediction date.

22. In examining the relationship between unanticipated earnings and stock price changes, for example, the sign of the forecast error from a time series is often used ([7], [12], [13]) as a device for classifying unanticipated earnings into "favorable" or "unfavorable" categories. With this methodology, BJ and V classify earnings differently 213 times out of the 797 one quarter ahead forecasts in our sample.

forecasts than by those of time series models are consistent with our evidence and with market rationality.

The hypothesis of analyst superiority versus univariate time series models is derived from basic economic theory and is not limited to the case of earnings. It is therefore applicable to all types of forecasts subject to the market test. There is no presumption that other non-market forecasts such as those made by corporate executives or government agencies should be better (or worse) than those generated by univariate time series models.

Appendix A

Sample Firms

Abbott Laboratories
Allegheny Ludlum Industries, Inc.
American Airlines, Inc.
Anaconda Company
Boeing Company
Borg-Warner Corporation
Braniff International Corporation
Caterpillar Tractor Company
Champion International Corporation
Chrysler Corporation
Clark Equipment Company
Colgate-Palmolive Company
Continental Can Company, Inc.
Curtiss-Wright Corporation
Cutler-Hammer, Inc.
Eastern Airlines, Incorporated
Eastman Kodak Company
Flintkote Company
Freeport Minerals Company
Fruehauf Corporation
GATX Corporation
General Electric Company
Goodrich (B. F.) Company
Gulf Oil Corporation
Homestake Mining Company
International Business Machines Corporation
International Paper Co.
Kennecott Copper Corporation
Leheigh Portland Cement Co.
Ligget Group Inc.
Lowenstein (M.) & Sons, Inc.
Nabisco, Inc.
National Distillers & Chemical Corporation
National Steel Corporation

with our evidence and

time series models is
case of earnings. It is
arket test. There is no
e made by corporate
) than those generated

Pan American World Airways, Inc.
Pepsico, Inc.
Phelps Dodge Corporation
Phillips Petroleum Co.
Pullman, Incorporated
Raybestos-Manhattan, Inc.
Republic Steel Corporation
Standard Brands, Inc.
Standard Oil Company of Indiana
Sterling Drug, Incorporated
St. Regis Paper Company
Timken Company
United States Gypsum Company
United States Steel Corporation
United Technologies Corp.
Wrigley (W. M.) Jr. Company

REFERENCES

1. W. H. Beaver. "The Information Content of the Magnitude of Unexpected Earnings," Unpublished Manuscript, 1974.
2. F. W. Bell. "The Relation of the Structure of Common Stock Prices to Historical, Expectational and Industrial Variables," *Journal of Finance* (March 1974), pp. 187–197.
3. H. Benishay. "Variability in Earnings-Price Ratios of Corporate Equities," *American Economic Review* (March 1961), pp. 81–94.
4. G. Benston. "Published Corporate Accounting Data and Stock Prices," *Empirical Research in Accounting: Selected Studies, 1967*. Supplement to Vol. 5, *Journal of Accounting Research*, pp. 1–54.
5. R. S. Bower and D. H. Bower. "Risk and the Valuation of Common Stock," *Journal of Political Economy* (May/June 1969), pp. 349–362.
6. G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control* (San Francisco: Holden-Day, 1970).
7. P. Brown and J. Kennelly. "The Informational Content of Quarterly Earnings: An Extension and Some Further Evidence," *Journal of Business* (July 1972), pp. 403–415.
8. W. J. Conover. *Practical Nonparametric Statistics* (New York: John Wiley & Sons, Inc., 1971).
9. J. G. Cragg and B. G. Malkiel. "The Consensus and Accuracy of Some Predictions of the Growth of Corporate Earnings," *Journal of Finance* (March 1968), pp. 67–84.
10. W. T. Dent and J. A. Swanson. "On Forecast Error in Autoregressive Integrated Moving Average Models," Working Paper No. 75-3, College of Business Administration, The University of Iowa, February 1975.
11. E. J. Elton and M. J. Gruber. "Earnings Estimates and the Accuracy of Expectational Data," *Management Science* (April 1972), pp. 409–424.
12. G. Foster. "Stock Market Reaction to Estimates of Earnings per Share by Company Officials," *Journal of Accounting Research* (Spring 1973), pp. 25–37.
13. ———. "Quarterly Accounting Data: Time Series Properties and Predictive-Ability Results," *Accounting Review*, (January 1977), pp. 1–21.
14. I. Friend and M. Puckett. "Dividends and Stock Prices," *American Economic Review* (September 1964), pp. 656–682.
15. J. R. Froeschle. "The Analysis of Seasonal Time Series," Doctoral Dissertation, The University of Iowa, 1975.
16. N. Gonedes. "Properties of Accounting Numbers: Models and Tests," *Journal of Accounting Research* (Autumn 1973), pp. 212–237.

17. W. L. Hays and R. L. Winkler. *Statistics: Probability, Inference and Decision* (New York: Holt, Rinehart and Winston, Inc., 1971).
18. R. C. Higgins. "Growth, Dividend Policy and Capital Costs in the Electric Utility Industry," *Journal of Finance* (September 1974), pp. 1189–1201.
19. J. E. Kiger. "An Empirical Investigation of NYSE Volume and Price Reactions to the Announcement of Quarterly Earnings," *Journal of Accounting Research* (Spring 1972), pp. 113–128.
20. R. H. Litzenberger and C. U. Rao. "Estimates of the Marginal Time Preference and Average Risk Aversion of Inventors in Electric Utility Shares 1960–66," *The Bell Journal of Economics and Management Science* (Spring 1971), pp. 265–277.
21. B. G. Malkiel and J. G. Cragg. "Expectations and the Structure of Share Prices," *American Economic Review* (September 1970), pp. 601–617.
22. R. G. May, "The Influence of Quarterly Earnings Announcements in Investor Decisions as Reflected in Common Stock Price Changes," *Empirical Research in Accounting: Selected Studies, 1971*. Supplement to Vol. 9, *Journal of Accounting Research*, pp. 119–163.
23. J. F. Muth. "Rational Expectations and the Theory of Price Movements," *Econometrica* (July 1961), pp. 315–335.
24. C. R. Nelson. "Rational Expectations and the Predictive Efficiency of Economic Models," *The Journal of Business* (July 1975), pp. 331–343.
25. H. Theil, *Applied Economic Forecasting* (Amsterdam: North-Holland Publishing Company, 1966).

Dov Fried and Dan Givoly
"Financial Analysts Forecasts of Earnings
A Better Surrogate for Market Expectations"
***Journal of Accounting and Economics*, Vol. 4**
1982

Kelly Per
your request.

Journal of Accounting and Economics 4 (1982) 85–107. North-Holland Publishing Company

FINANCIAL ANALYSTS' FORECASTS OF EARNINGS

A Better Surrogate for Market Expectations*

Dov FRIED

New York University, New York, NY 10006, USA

Dan GIVOLY

Carnegie-Mellon University, Pittsburgh, PA 15213, USA
Tel-Aviv University, Ramat-Aviv, Tel-Aviv, Israel

Received July 1981, final version received June 1982

The specification of the market expectation of accounting numbers is a common feature of many empirical studies in accounting and finance. Givoly and Lakonishok (1979) found that financial analysts' forecasts have information content. This study evaluates the quality of analysts' forecasts as surrogates for the market expectation of earnings and compares it with that of prediction models commonly used in research. Results indicate that prediction errors of analysts are more closely associated with security price movements, suggesting that analysts' forecasts provide a better surrogate for market expectations than forecasts generated by time-series models. The study also identifies factors that might contribute to the performance of the financial analysts' forecasts. The broadness of the information set employed by analysts and, to a lesser extent, their reliance on information released after the end of the fiscal year appear to be important contributors to their performance.

1. Introduction

The specification of market expectations of stock returns and of accounting numbers is a common feature of empirical studies in accounting and finance. While expected returns in these studies have been derived customarily by the theoretically founded and empirically supported market model, no such underlying theory exists for the specification of a surrogate for market expectation of earnings. To a great extent, the expectation models selected by researchers relied exclusively on past time-series behavior of the variable.[1] Since no established theory could guide the selection of the earnings expectations models, many researchers used a wide set of time-series

*The authors wish to thank Robert Kaplan, Ross Watts, Jerold Zimmerman, and an anonymous referee for their helpful comments. The financial support of the Deloitte Haskins and Sells Foundation is gratefully acknowledged.

[1]A short list of such studies, which is by no means exhaustive, includes Ball and Brown (1968), Barnea et al. (1976), Beaver and Dukes (1972), Brown and Kennelly (1972), Foster (1977), and Watts (1978).

Fax:4078233047 Feb 19 1999 12:03 P.04

models so that some assessment of the robustness of the results to model selection could be made.

The selection of a time-series model as a surrogate for market expectations is further impaired by the underlying assumptions that the earnings generating processes are stationary with stable parameters and that the model characteristics are applicable to all firms. There is evidence suggesting that models applicable to one period are not necessarily relevant for other periods. Brooks and Buckmaster (1976), for example, showed that while the martingale process might describe the earnings changes in normal years, earnings behavior in periods following unusual fluctuations in earnings may best be described by a mean-reverting process. The use of such models as a proxy for market expectations of earnings thus may limit the validity and the scope of any conclusions.[2]

The purpose of this paper is to examine the performance of an alternative surrogate for market expectations, earnings forecasts made by financial analysts. These forecasts were obtained from the *Earnings Forecaster*, a weekly publication by Standard and Poor that first appeared in 1967. The *Earnings Forecaster* lists the outstanding EPS forecasts for about 1500 companies. The forecasts are those made by S & P and by about 70 other security analysts and brokerage houses who agreed to submit their forecasts, upon release, to the publication.

~~Givoly and Lakonishok (1979) showed that financial analysts' forecasts of earnings have information content. Their study found a significant price reaction to the disclosure of revisions in FAF. The wide dissemination of~~ FAF ~~in the financial community further~~ reinforces the notion that FAF might ~~proxy for market expectations~~.

Given the above evidence, tests on the information content of earnings that use FAF as a surrogate for market expectations are likely to be better specified than those based on time-series models. The first objective of this study is to evaluate FAF as a surrogate for market expectation of earnings, and to compare them with prediction models widely used in the literature. The findings show that FAF are a better surrogate for market expectation of earnings and suggest that the use of other prediction models may have weakened the tests employed by previous research.

The tests of the association between the *API* and the prediction errors, to be described later, follow those employed by Ball and Brown (1968) and Beaver et al. (1979) and rely on the correlation between *API* and forecasts made about a year before the release of the earnings report. Tests on the

[2]This limitation was recognized in the literature [see, for example, Beaver and Dukes (1972) and Collins (1975)]. As Beaver and Dukes conclude: '...any inferences are conditioned upon the prediction models used to test the accounting measures tested...any findings are the joint results of prediction models and accounting method and only appropriately specified joint statements are warranted' (p. 332).

[3]See, for example, the report of the SEC Advisory Committee on Corporate Disclosure (1977).

information content of earnings, however, are best carried out by examining the association between prediction errors from forecasts based on the most up-to-date accounting information available, and *API* calculated on a daily basis in the immediate period surrounding the earnings release date. Nonetheless, the results herein are useful in that they suggest that FAF may serve as a better proxy if used in such studies.

The finding that FAF are a better surrogate for earnings expectation of the market is important for other reasons. Stock valuation models as well as P/E studies often rely on expected earnings or derivation thereof, as a basic parameter. The results of this study would thus offer valuable input to these studies in providing better identification of earnings expectations used by investors.

The existence of an empirical surrogate for earnings expectations will enable researchers to examine more thoroughly the formation of earnings expectations. Questions concerning the rationality of earnings expectations, the extent to which they employ accounting information and their consistency with the observed time-series behavior of earnings might be addressed. Some interesting work on the time-series behavior of FAF has been done by Abdel-Khalik and Espejo (1978) and by Brown et al. (1978, 1979, 1980). Establishing that FAF provide a satisfactory surrogate for market expectations would underscore the relevance of these studies and provides a motivation for further research.

The second objective of this study is to analyze the factors that contribute to FAF having information content. While forecasts of earnings based solely on past accounting data are revisable only in certain time intervals (annual or quarter), FAF incorporate presumably all publicly available (firm-specific, industry, and market) information, and can be continuously updated with the arrival of any new information. These characteristics suggest two factors which explain FAF superiority, and which will come under examination in this study: one is the broadness of the information set available to them, and the other is their timing advantage, in that they employ information that becomes available only after the last accounting report.

The paper is organized as follows. Section 2 describes the data and discusses the statistical tests concerning identification of the best surrogate for market expectations. Section 3 explores the broadness of information and timing issues and provides evidence on their effect on the performance of FAF. Concluding remarks are made and implications for future research are suggested in the final section.

2. FAF vs. time-series models as surrogates for market expectations

The model evaluation methodology follows the one used by Beaver and

Dukes (1972), Collins (1975) and Patell (1976), and which was articulated by Patell (1979). The presumption is made that accounting earnings possess information content. Alternative models are then evaluated by their ability to correctly classify the signal produced by the accounting number and hence by their usefulness in developing profitable trading strategies for which the buy/sell decisions are determined by this signal classification.

The association between the signals (e.g., the prediction error) produced by each expectation model (time-series or FAF) and abnormal stock return is analyzed. The expectation model whose signals (concerning future earnings) are the most strongly associated with stock price behavior is considered the best surrogate for the true, unobservable, market expectation.

This section is divided into four subsections. In the first we describe our data and the forecasting models, and present some results on the forecasting accuracy of the models. The next subsection describes the measure used to gauge stock market reaction. The third subsection discusses the tests to be used for the evaluation of the models; results of these tests are presented and discussed in the fourth subsection.

2.1. Data and forecasting models

Financial analysts' forecasts of earnings of a sample of companies listed in the *Earnings Forecaster* were evaluated in each of the eleven years 1969 to 1979. Considered each year were the FAF of that year's earnings outstanding at the beginning of April. These forecasts were first issued to the public typically in early March. The time of the forecast is between the release of the annual report for the previous year [which is made on average, in February — see Givoly and Palmon (1981)] and the release of the first quarterly report (typically late April).

Included in the sample each year were companies which satisfied these criteria:

(1) fiscal year ending December 31,
(2) N.Y.S.E. listing,
(3) existence of at least four forecasts (by different forecasters) of the current year's earnings,
(4) availability of monthly return data for the forecast year, the following year and the preceding four years,
(5) availability of actual earnings numbers for the forecast year and the preceding nine years.

The third criterion was introduced to allow the derivation of a reliable measure for the average or 'consensus' forecast.

All the contemporaneous company forecasts were for primary EPS before extraordinary items. To ensure that the comparison between the forecast and the actual EPS was not unduly affected by changes in capitalization not incorporated in the forecasts, we adjusted any earnings forecasts announced prior to the disclosure of the change in capitalization.

The final sample consists of 1247 cases (company-years) with a total of 6020 forecasts. The number of cases in each year differs and varies from 95 (1972) to 173 (1969). This sample represents 424 distinct companies. The FAF for each company-year are represented by their simple average.

Two alternative models of earnings expectation were employed to define the news content of earnings announcements:

(a) $P_t = f(A_{t-1}, A_{t-2}, \ldots)$,

(b) $P_t = A_{t-1} + \gamma_t + \delta_t E(\Delta A_{mt})$,

where A_t is the realized earnings. The earnings variable was the primary earnings per share before extraordinary items (EPS) of year t adjusted for capitalization, P_t is the expected (predicted) value of A_t, γ_t and δ_t are regression parameters,[4] and $E(\Delta A_{mt})$ is the expected change in market earnings. A_m is represented by the average EPS of the S & P's Composite 500. The expected change in market earnings is derived from a submartingale model using the (arithmetic) average growth over years $t-6$ to $t-1$ as an estimate of the drift term.[5] The regression parameters are re-estimated each year from the available past annual EPS data (the first available year is 1958).

The first model is a univariate time-series model derived from the results of Brooks and Buckmaster (1976). For most of our observations, the submartingale model of the form

$$P_t = A_{t-1} + C_t$$

was used, where C_t is the (arithmetic) average growth in EPS computed over the years $t-6$ to $t-1$.

This model was found by recent studies to represent quite adequately the time-series behavior of earnings [see Albrecht et al. (1977) and Watts and Leftwich (1977)]. Furthermore, as a general representative firm model, the martingale with drift was found to perform as well as the firm-specific Box–Jenkins models in describing the time-series characteristics of annual earnings

[4]These regression parameters were estimated over the first differences series of ΔA_t and ΔM_t.

[5]The expectation is formed consistent with the model used to predict individual firm's earnings. We also used in all tests a version of the IM in which the *realized* market index is employed. The two versions yielded essentially the same results.

(see also Albrecht et al.). However, periods that follow extreme earnings fluctuations were found by Brooks and Buckmaster — B & B — (1976) to behave in a way more consistent with a mean reverting process. To provide better specification of the earnings time-series, the sample was stratified each year according to the size of the deviation of previous year's earnings from some 'norm'. The model used for the extreme strata was, in accordance with B & B's findings, an exponential smoothing rather than the martingale with trend.[f] About 23% of the cases (company-years) in our sample fell in these extreme strata. For the stratification procedure and the specification of the exponential smoothing models, see the appendix. We shall refer to the univariate time-series model used as the modified submartingale (MSM).

The use of Model b, the index model (IM), is supported by the relationship that was found between the first differences in individual company earnings and an economy-wide index of earnings such as the differences in earnings across all firms [see Ball and Brown (1968) and Gonedes (1973)].

The relative prediction error was defined as

$$e^k_{it} = (A_{it} - P^k_{it})/|A_{it}|. \tag{1}$$

where k denotes the expectation model, i the observation index ($i = 1, \ldots, N$), and t the year.

In the few cases (3–4% of the cases, depending on the model) where $|e^k_{it}| > 1.00$, the error measure was equated to ± 1.0. This truncation of the distribution of e_{it} was introduced to avoid the distortive effect of a small denominator and to suppress the effect of possible data and measurement errors.

One measure of accuracy of model k in period t is the mean absolute relative error,

$$|e^k_t| = (1/N)\sum_i |e^k_{it}|. \tag{2}$$

The corresponding measure of bias in model k in period t is the mean relative error,

$$e^k_t = (1/N)\sum_i e^k_{it}. \tag{3}$$

The relative accuracy of the forecasts is presented in table 1. The table reveals that in almost all years the accuracy of FAF measured by the mean relative error is greater than that of the competing models both for cases of

[f] The smoothing parameter, a, used for each strata was the one found by B & B to be the best smoothing constant (see table 8).

Table 1

Mean relative earnings prediction errors — annual models (percentages).[a]

	All years[b]	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Cases of positive errors												
FAF (608)	10.4	4.8	5.2	7.5	8.9	14.5	19.4	8.9	10.2	10.1	10.9	14.2
MSM (767)	13.8	8.1	8.2	10.9	14.6	20.1	19.7	9.7	17.5	11.5	14.7	16.9
IM (801)	16.3	9.8	3.6	16.8	15.1	18.8	20.6	17.0	22.0	16.4	16.3	17.9
Cases of negative errors												
FAF (639)	−20.4	−18.3	−31.1	−23.8	−12.3	−17.1	−26.0	−25.0	−11.3	−22.0	−19.2	−17.9
MSM (480)	−24.2	−17.8	−34.1	−29.3	−13.4	−17.3	−36.4	−37.5	−13.9	−24.1	−24.9	−17.0
IM (446)	−24.3	−17.7	−35.6	−26.3	−15.3	−23.5	−33.4	−39.4	−15.8	−19.1	−23.4	−17.4
All cases (accuracy results)[c]												
FAF (1247)	16.4	14.0	25.9	17.6	10.4	15.4	22.0	18.6	10.7	17.1	13.7	15.0
MSM (1247)	19.3	13.1	26.2	18.3	14.4	19.6	24.9	28.0	16.8	16.7	17.4	16.9
IM (1247)	20.3	13.6	26.0	20.3	15.1	19.4	24.5	30.1	20.9	17.6	18.1	17.7
All cases (bias results)[d]												
FAF (1247)	−5.3	−10.9	−23.9	−11.8	−0.5	4.3	1.6	−11.6	1.2	−8.7	0.6	1.4
MSM (1247)	−1.2	−5.2	−21.5	−5.2	8.1	12.5	2.4	−21.3	11.2	−3.2	4.2	4.9
IM (1247)	1.4	−3.2	−19.8	1.1	8.7	12.9	4.4	−15.9	15.3	0.1	6.5	5.5

[a] FAF – Financial Analysts' Forecasts of Earnings, MSM – Modified Submartingale, and IM – Index Model. Number of all cases given in parentheses.
[b] Simple average of the 11 years.
[c] See expression (2) in the text.
[d] See expression (3) in the text.

positive prediction error (i.e., actual earnings are above expectation) and for cases of negative prediction error. The average prediction error of FAF is significantly lower than that of the other models for both types of cases. For the positive errors, the t values (computed from the 11 observations) are 5.27 and 6.37 for the comparison with MSM and IM, respectively. For the negative errors, the values are 5.02 and 3.04, and for all cases 3.14 and 3.37. The critical t-value for one-tail test with 10 degrees of freedom and 1% significance level is 2.76.

The bias of each model is provided by the fourth (bottom) panel in the table which shows the mean relative error measured over all cases. The results indicate some tendency for FAF to overestimate next year's earnings.[7] Yet, the bias of FAF is present only in 6 of the 11 years and, except for the first three years, appears to be quite small. The finding of some bias conforms to the persistent optimism of FAF reported by previous studies [Barefield and Comiskey (1971) and McDonald (1973)].[8]

Any comparison between the performance of the models is, however, incomplete if it ignores the potential for improvement inherent in each. The existence of a systematic behavior of the model's errors may allow forecast users to improve upon (increase accuracy and eliminate the bias of) the original forecast. To the extent that stationarity of the prediction and realization processes is assumed, forecast users will rely for that improvement on all available past information.[9]

To examine the potential improvement of each model, we employed the linear correction procedure suggested by Mincer and Zarnowitz [see Mincer (1969)] and Theil (1966). The results reveal that all three models offer very little in terms of potential reduction in error through a linear correction of the forecasts. The tests conducted for the corrected forecasts yielded results similar to those obtained for the raw forecasts; therefore, we report only the latter.

[7]Given the general increase over time in the EPS of all firms (the average annual increase in the average EPS, adjusted for capitalization, of S&P's 500 firms over the 20-year period, 1958 to 1977, was 12.4%), the upward bias in the prediction of earnings' *levels* by FAF implies also an overestimation of the change in earnings. This finding contrasts with the observed tendency of economic forecasters to underestimate changes in variables such as GNP and Personal Consumption [see Theil (1966, ch. V) and Mincer (1969, ch. I)]. Two explanations might be offered for the finding: first, time-series behavior of earnings is apparently less regular and monotonic than that of economic variables leading to less reliance of earnings forecasts on past levels. Second, financial analysts who, as part of the 'establishment' of the investor community and unlike most economic forecasters have a direct stake in the prosperity of the stock market, are perhaps more likely to issue an optimistic outlook than a dim one.

[8]Since only aggregate results are produced, the findings are not comparable neither to those reported by Brown and Rozeff (1979), which show that analysis predict in an adaptive manner — changing the forecasts in a direction opposite to last period's error — nor to those of Elton, Gruber and Gultekin (1981), which suggest persistence of error in consecutive years.

[9]Whether users actually employ corrected forecasts depends on the cost of adjustment and on the degree of stationarity in the systematic behavior of the forecast.

The accuracy of FAF is not necessarily related to the adequacy of their use as a surrogate for market expectations. It is conceivable that FAF are superior to other prediction models in terms of ex-post accuracy tests, but inferior in terms of association with stock price movements. In the next subsection, we describe the metric to be used to measure stock price movements.

2.2. Market reaction measure

Stock price movements are measured in this study by the abnormal return where the expected return was defined according to the familiar market model,

$$E(R_{it}) = \alpha_i + \beta_i R_{mt}, \tag{4}$$

where R_{it} denotes the return of security i for period t, α_i and β_i are parameters and R_{mt} is the actual market rate of return for period t. The market rate of return is represented by the value-weighted rate of return of New York Stock Exchange stocks. Monthly abnormal returns were measured by the difference

$$\hat{\varepsilon}_{it\tau} = R_{it} - (\hat{\alpha}_{it} + \hat{\beta}_{it} R_{m\tau}), \tag{5}$$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ were estimated from the 48 months preceding the test period, t is the year index, and τ is the month index.

The average β in the pooled sample (1247 observations) is 1.133. The slightly higher than one β's apparently reflect the simple averaging of β's which are computed from the value-weighted index.[10]

The test period for evaluating the models' predictions consisted of the 12-month period from April of year t ($\tau = 1$) to March of year $t+1$ ($\tau = 12$) and was designed to cover the period of approximately 11 months preceding the release of the annual report and the month that follows it.

Cumulative abnormal returns were computed as

$$CAR_{it} = \sum_{\tau=1}^{12} \hat{\varepsilon}_{it\tau}, \tag{6}$$

and the Abnormal Performance Index (API) was derived as

$$API_{it}^{k} = \text{sign}(e_{it}^{k}) \cdot CAR_{it}. \tag{7}$$

[10]For randomly selected securities, an unweighted average β greater than one would be expected if securities with low value weights have relatively high β and vice versa. Higher β's for small firms is suggested by the results of Foster (1978) and Reinganum (1981).

2.3. Tests

The models will be tested according to the association of their errors with stock price movements. In examining prediction error and stock price behavior, the magnitude of the prediction error, in addition to its sign, will be considered. As shown by Beaver et al. (1979), the inclusion of the magnitude of the prediction error makes the association tests more powerful. In addition, using only the sign of the prediction error results in a serious limitation of the tests since they rely exclusively on those cases where the models disagree as to the sign of the prediction error. Thus, the only relevant observations would belong to a group which might be a very small subset of the total sample. The following two tests, which incorporate the magnitude of the prediction error, alleviate this problem by exploiting the entire sample.

(a) *Correlation test:* The correlation between the magnitude of the prediction error (e^k_{it}) and the stock price movement (CAR_{it}) is computed. The model which yields the highest correlation is considered to be superior. This association test was employed recently by Beaver et al. (1979) in measuring the relationship between abnormal returns and prediction errors of earnings expectation models.

(b) *Weighted API test:* The second test (magnitude of API) involves the evaluation of an 'investment strategy' under which long or short positions in a portfolio are taken in accordance with the direction and magnitude of the prediction error produced by each model. Previous research which looked at the sign of the prediction error implicitly assumed that the same amount is invested (or disinvested) regardless of the magnitude of the error. It is plausible that the amount invested will be in direct proportion to the magnitude of the error. Indeed, if the 'unexpected' earnings (conveyed by the error) are expected to be permanent (consistent with the random-walk behavior of earnings over time) and the security risk is unaltered, the abnormal return will be proportional to the error. This test, therefore, evaluates an investment strategy under which the cross-sectional prediction errors of a given model k served each year to determine the weight of each security in that year's portfolio k. The API of the portfolio was computed as the weighted average across individual securities. Specifically, the weight assigned to each security i in year t of portfolio k is

$$a^k_{it} = |e^k_{it}| / \sum_i |e^k_{it}|, \tag{8}$$

where e is the relative error from (1) and the portfolio's API is[11]

[11] The model assumes realistically that the proceeds from short sales are not collected at the time of sale and that, in addition, collateral in the amount of the sale is required. Other weighing schemes were also employed but led to essentially the same results.

$$API_{k,t} = \sum_i a^k_{i,t} \cdot API_{k,t}. \tag{9}$$

In designing the statistical tests, one should be aware of the potential existence of cross-sectional dependence between contemporaneous residuals (or abnormal returns). The dependence, which could stem from various sources (e.g., nonlinearity of the return generating function or the omission of common factors, such as industry, from the index model), makes it likely that the sample estimate of the variance of the residuals will be biased in an unknown manner. Cross-sectional dependence is likely to exist also between contemporaneous prediction errors due to the common factors underlying the generation of earnings (e.g., GNP; the use of the 'index model' of earnings may have removed this source of dependence). For these reasons the t-tests to be reported here employ an estimate of the variance taken from a time-series in which the serial correlation is not expected to be significantly different from zero. Specifically, the mean of the variable of interest was computed each year from the cross-section of observations. The 11 mean values were treated as a sample of independent observations. Similar procedures have been used by Beaver et al. (1979), Jaffe (1974), and Mandelker (1974).

2.4. Results and discussion

Table 2 presents the frequency of cases in which the signs of the prediction error and the price movement (measured by CAR) during the test period were consistent, that is, in the same direction. Overall, the models produced errors whose sign was consistent with the sign of CAR. Of the 1247 cases (company-years), the sign of the FAF prediction error was consistent with the sign of the CAR in 743 cases (60%). This is somewhat superior to the performance of the MSM and the index model which experienced prediction errors' signs that were consistent with that of the CAR in 670 cases (55%) and 679 cases (54%), respectively.

A closer examination of the table reveals that FAF perform about equally well when the CAR is positive as when the CAR is negative (i.e., $337/561 \approx 406/686 \approx 743/1247 \approx 60\%$). However, the time-series models do very well in times of positive CAR (MSM yields 69% and IM yields 71% consistent classifications), but rather poorly when the CAR is negative (MSM = 44%, IM = 41% consistent classifications).

The comparison between the models is more meaningful when only the disagreement cases are considered. Panel (b) of the table shows that FAF do poorer than the other models in periods of positive CAR (only $34/113 = 30\%$ of the cases) but do extremely well in periods of negative CAR ($125/152 = 82\%$ of the cases).

The results of table 2 serve to highlight the limitation inherent in constructing API tests based on the sign (but not the magnitude) of the

Table 2

Frequency of cases in which the sign of the prediction error is consistent with the sign of the corresponding cumulative abnormal return (*CAR*).[a]

(a) *All cases*

CAR realization	FAF errors		MSM errors		IM errors	
Positive	337	224	386	175	398	163
Negative	406	280	304	382	281	405
Total	743	504	690	557	679	568

(b) *Cases in which competing models disagree*

	FAF vs. MSM		FAF vs. IM		MSM vs. IM	
CAR realization	FAF errors consistent	MSM errors consistent	FAF errors consistent	IM errors consistent	FAF errors consistent	MSM errors consistent
Positive	34	79	32	92	26	35
Negative	125	27	148	27	51	28
Total	159	106	180	119	77	73

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM = Index Model. Consistent sign is said to exist when the difference between the realized value and the predicted value ($A-P$) has the same sign as the *CAR*.

prediction errors. The difference between the *API*'s of two competing models, which is based on the sign of the prediction error will reflect only those cases in which the models disagree with respect to the sign of the prediction error (in all other cases the models will produce the same *API*). These cases, however, represent only a small percentage of all cases. Indeed, as table 2 reveals, the proportion of disagreement cases in our sample is low (for instance, out of the 1247 predictions made by both models, FAF and MSM produced prediction errors of opposite signs in only 263, or 21% of the cases). Results which utilize information on both the sign and magnitude of the prediction error (and therefore on the entire sample of 1247 observations) are presented in tables 3 and 4.

Table 3 presents the average cross-sectional correlation coefficients for all years, between the prediction error of each model and the corresponding *CAR*. The first three columns (under 'All cases') present the correlation coefficient calculated over the entire sample for each of the models. The next six columns show the correlation coefficient calculated for each model separately for cases with positive prediction errors and cases with negative

Table 3

Correlation coefficients between *CAR* and the earnings prediction error.[a]

	All cases			Cases of positive errors			Cases of negative errors		
		MSM	IM	FAF	MSM	IM			
All years[b]	0.33	0.27	0.27	0.23	0.18	0.18	0.17	0.18	0.17
1969	0.47	0.38	0.39	0.10	0.04	0.08	0.37	0.40	0.38
1970	0.41	0.33	0.31	−0.03	−0.02	0.09	0.37	0.32	0.43
1971	0.39	0.23	0.27	0.26	−0.02	0.02	0.38	0.36	0.32
1972	0.22	0.07	−0.01	−0.05	0.03	−0.13	0.08	0.08	−0.01
1973	0.35	0.46	0.44	0.43	0.45	0.47	0.19	0.21	0.15
1974	0.33	0.28	0.29	0.07	−0.02	0.00	0.30	0.23	0.30
1975	0.12	−0.20	−0.02	0.20	0.32	0.45	−0.04	−0.20	−0.20
1976	0.41	0.32	0.28	0.47	0.34	0.27	0.17	0.34	0.21
1977	0.37	0.43	0.36	0.33	0.19	0.17	0.24	0.21	0.26
1978	0.25	0.27	0.24	0.43	0.42	0.33	−0.17	−0.05	−0.07
1979	0.24	0.36	0.37	0.36	0.18	0.24	0.01	0.10	0.14

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM = Index Model. The critical value for the correlation coefficient at the 5% significance level for $H_0: \rho = 0$ and one-tail test using the z-statistic [see Freund (1962, pp. 310–311)] is 0.13 for most cells ($n \gtrsim 40$).

[b]Simple average of the 11 years.

Table 4

Mean *API* over the test period of a portfolio weighted by the magnitude of the earnings prediction errors of its members (percentages).[a]

	All cases			Cases of positive errors			Cases of negative errors		
				FAF	MSM	IM	FAF	MSM	IM
All years[b]	14.12	9.97	9.45	7.45	2.73	3.32	17.48	17.03	17.63
1969	22.67	19.06	19.02	10.76	3.72	5.08	24.11	25.63	27.53
1970	17.51	15.17	14.57	7.51	2.47	−0.47	17.92	16.52	16.59
1971	18.55	10.53	9.27	1.40	−7.33	−4.11	21.92	20.43	24.25
1972	9.96	−1.45	−4.26	−0.19	−4.68	−8.48	19.22	10.18	11.40
1973	16.98	18.26	16.90	16.64	17.49	16.25	17.67	21.74	20.15
1974	13.56	11.00	11.11	3.76	0.59	0.67	25.02	23.57	26.51
1975	7.33	−3.05	1.82	6.63	−2.61	9.68	7.50	−3.11	−0.60
1976	13.92	5.89	5.39	13.37	3.98	4.00	14.61	15.29	14.30
1977	16.90	17.85	14.09	10.92	9.18	6.13	18.84	23.71	22.12
1978	7.34	6.97	6.06	9.20	7.12	6.33	5.32	6.74	5.49
1979	10.63	9.47	9.99	1.93	0.06	1.41	20.14	26.58	26.23

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM = Index Model.

[b]Simple average of the 11 years.

prediction errors. Generally, errors of all models show positive and, in most cases, significant correlation with *CAR*. Overall, FAF prediction errors are more strongly associated with *CAR* than the prediction errors of the other two models: the average coefficient of correlation over the 11 years between *CAR* and separately FAF, MSM and IM's errors are 0.33, 0.27 and 0.27, respectively. The t-test results for the differences between the correlations produced by FAF and the MSM and IM models are significant at the 10% and the 5% level, respectively. Looking at the positive error and negative error cases separately, the superiority of FAF is evident for positive error cases ($0.23 > 0.18$), but disappears for negative error cases.

Table 4 provides *API* values [calculated using eqs. (8) and (9)] for a portfolio based on both sign and magnitude of the signal (prediction error). The table reveals that FAF errors appear to be more strongly associated with stock price movement than the other models. All models yielded significant average *API*'s for all cases and for negative error cases (the t-test was used over the 11 years). However, only FAF produced significant *API*'s for the positive error cases. The *API* average yielded by FAF (14.12%) is higher than that produced by the MSM (9.97%) and the IM (9.45) models. For all cases FAF performed better than each of the other two models in nine of the 11 years. For positive error cases FAF performed better (i.e., the *API* was higher) than MSM in 10 years and better than IM in all 11 years. The corresponding differences between the *API*'s are significant (at the 5% significance level) for all cases and for the positive error cases. There is no significant difference between the models for the negative error cases.[12]

The foregoing results are consistent with the hypothesis that FAF, or information closely correlated with FAF, serve as an input to investment decisions by market participants. Furthermore, the findings suggest that FAF, or at least those outstanding in early April, might be more representative of market expectation of earnings than some time-series models widely used in the financial literature.

[12]We also derived *API* based only on the sign of the prediction error. The *API* based on FAF predictions calculated over the 11 years was on average 6.94%, while those based upon MSM and the IM yielded 3.79% and 3.42%, respectively. The difference between the FAF's *API* and the other model's *API* is significant at the 5% significance level.

The *API*'s in this study are lower than those reported by Ball and Brown (1968). Note, however, that the survey periods are different. Also, the models are not exactly identical: we use a modified submartingale and an ex-ante index model. Finally, Ball and Brown averaged the *API*'s cross-sectionally and over years giving an equal weight to each company-year. In our analysis, we first find the simple average for each year and then the average across years, giving each year an equal weight. So, for example, 1969, which has the highest average *API*, is given the same weight as any other year, despite the fact that it is represented in the sample by the largest number of cases.

3. Causes of FAF superiority

It can be argued that financial analysts' forecasts may have an edge over time-series prediction models for two main reasons:

(a) They use a broader information set which includes non-accounting information on the firm, its industry and the general economy.
(b) They have a timing advantage in that they are issued some time within the year being forecasted. As such, they can use more recent information about the firm's earnings which becomes available only after the end of the fiscal year.

In this study we provide an analysis of the contribution of each of the ingredients (broadness and timeliness of the information) to the performance of FAF.

3.1. Broadness of the information set

The FAF presumably utilize all publicly available (and occasionally unpublished) information while the time-series models examined rely exclusively on past earnings. There are several interesting questions in this context: the extent to which FAF are a product of a simple extrapolative procedure; the extent to which they incorporate other, autonomous information, unrelated to the time-series of earnings; and the degree to which they efficiently utilize all available extrapolative information.

In our analysis, the MSM and the index model of earnings serve as representatives of the family of extrapolative models.[13] The contribution of each component to the predictive power of FAF is measured by the partial correlation between actual earnings and FAF, given the time-series model's prediction or $r_{AP.X}$, where A denotes the realized value, P is the FAF, and X is the prediction of the time-series model.[14] The extent to which FAF exploit the extrapolative potential of past earnings series (offered by the examined models) is measured by the partial correlation $r_{AX.P}$.

Values of $r_{AP.X} > 0$ suggest that FAF contain predictive power based not only on extrapolation but also on an autonomous component. In addition, the magnitude of $r_{AX.P}$ indicates the extent of underutilization of available extrapolative information by FAF, since $r_{AX.P} > 0$ means that the time-series model contains some amount of predictive power that was not used in FAF.

The partial correlation results are presented in table 5. The average coefficient of the partial correlation between realization and FAF, given the

[13]Other, more efficient extrapolative models probably exist. Thus, the conclusions from our analysis are expected to overstate the weight of the autonomous component and perhaps also the success of FAF in exploiting the available extrapolative information.

[14]The notations A, P, and X, as well as the results presented, are stated in terms of earnings levels. Similar results were obtained for earnings changes.

Table 5

Partial correlations between realization and predictions of different models.[a]

	Correlation coefficient between realization and the prediction by				
	FAF given MSM ($r_{AP.x_1}$)	FAF given IM ($r_{AP.x_2}$)	FAF given MSM and IM ($r_{AP.x_1x_2}$)	MSM given FAF ($r_{Ax_1.P}$)	IM given FAF ($r_{Ax_2.P}$)
All years[b]	0.55	0.56	0.51	−0.04	0.01
1969	0.43	0.45	0.43	−0.08	−0.07
1970	0.38	0.23	0.26	0.02	0.15
1971	0.53	0.80	0.53	0.06	−0.04
1972	0.63	0.60	0.55	0.00	0.13
1973	0.56	0.40	0.40	−0.12	0.01
1974	0.73	0.63	0.61	−0.38	−0.28
1975	0.63	0.64	0.60	0.01	−0.02
1976	0.67	0.79	0.67	0.10	−0.03
1977	0.50	0.52	0.56	−0.20	0.03
1978	0.53	0.59	0.53	0.09	0.08
1979	0.49	0.52	0.49	0.01	0.05

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM-Index Model.
[b]A simple average of the 11 years.

time-series predictions, $r_{AP.X}$, is 0.55 and 0.56 for the comparison with MSM and IM, respectively. The values remain high, 0.51, when the correlation was conditional on the predictions of both of the other models. These values are significantly different from zero. Since $r_{AP.X}$ is a measure of the net contribution of the autonomous component, it appears that FAF utilize a considerable amount of information which is independent of the time series and cross sectional properties of the series as captured by our extrapolative models.

The coefficients of the partial correlation between realization and time-series predictions, given FAF ($r_{AX.P}$), are generally very small and close to zero (the hypothesis that their mean is zero could not be rejected at the 5% significance level). This means that, in addition to the utilization of autonomous information, analysts also fully exploit the time-series and cross-sectional properties of the earnings series that are captured by the MSM and IM models of earnings.

The apparent reliance of FAF on extrapolations is also evident in the association between the performance of FAF and that of the other models: the mean error of each model in each of the 11 years (see table 1) was ranked (from 1 to 11); the Spearman coefficients of rank correlation between the mean error of FAF and those of MSM and IM are 0.77 and 0.85, respectively. Both values are significant at the 5% level. These results suggest that periods which are characterized by unusual deviations of earnings from their past pattern present forecasting difficulties not only to time-series models but also to FAF.

3.2. Timing of information

Analysts presumably make use also of information that becomes available only after the end of the previous fiscal year. To gauge the effect of the use of a more recent information by analysts, it would be desirable to compare the performance of forecasts released at different points of time. For this aim, we collected from the *Earnings Forecaster* the release month of each forecast; this information was not available for 1969, the first year in our sample.

The distribution of forecasts for the remaining 10 years was as follows: 253 issued before January, 435 in January, 1219 in February, 1988 in March and 1299 in early April. We expect forecasts with a later release date to incorporate more (accounting and non-accounting) information and therefore to be superior to earlier forecasts.

To examine whether this is so, we divided our sample into two groups of FAF: one, denoted as 'early' forecasts, consists of forecasts released in January and February, and the other, denoted as 'late' forecasts, consists of those released in March and early April. This particular grouping results in forecasts that were released on average, about six weeks apart. Only companies for which both early and late forecasts were available in a given year were considered.[15] The number of companies considered each year differs and varies from 56 (1979) to 111 (1973).[16]

The research design for this investigation is essentially the one described in section 2, except that we concentrate on comparing early with late FAF. Table 6 exhibits the results of the *API* tests for the early and late FAF. The main findings are that a timing advantage does exist but has no significant impact on the comparative performance of the models considered. The average *API* over the 11 years is 12.78% and 13.15% for the early and late forecasts, respectively. The difference, although in the expected direction, when subjected to a *t*-test proved insignificant. The mean *API*'s for the time-

[15]We also used another version of the test under which this restriction was not imposed. Under this version, however, the composition of the company sample of the early forecasts was not identical to that of the company sample of the late forecasts. The results were essentially similar.

[16]This particular definition of early and late forecasts allowed us to get a large sample size in each group. Looking at January's forecasts alone and comparing them to those made in March and April, although might theoretically accentuate the timing difference between the forecasts, resulted in a large drop in the sample size: in two of the years the number of available companies was less than six. The examination of the other eight years did not in fact show a larger difference between early and late FAF.

Table 6

Mean *API* over the test period of a portfolio weighted by the magnitude of the earnings prediction errors of its members — all cases (percentages).[a]

	Early FAF	Late FAF	MSM[b]	IM[b]
All years	12.78	13.15	8.85	8.60
1970	16.92	18.29	15.42	15.00
1971	17.29	17.88	9.01	7.71
1972	11.78	9.09	0.97	−1.36
1973	17.80	17.05	18.45	17.35
1974	8.34	8.56	17.77	7.04
1975	6.95	9.67	−2.83	2.19
1976	14.15	16.88	7.93	7.75
1977	16.72	16.45	17.39	13.65
1978	7.01	7.00	6.88	6.48
1979	10.83	10.66	7.47	10.22

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM = Index Model. Averages calculated each year only for companies for which both early and late forecasts exist.
[b]The results for the MSM and IM do not correspond to those reported in table 4 since the sample now covers only the years 1970–1979 and consists of companies for which both early and late forecasts were available.

series models computed over the same sample are lower than both FAF groups, 8.85% for the MSM and 8.60% for IM.

Table 7 provides other summary statistics pertinent to the comparison between early and late forecasts (the mean *API* results reported in table 6 are repeated here). The degree of correlation between the CAR and the earnings prediction error for the early forecasts is indistinguishable from that for the late forecasts (0.31 vs. 0.32).

The findings so far suggest that the timing advantage does not result in a significant improvement in the association of FAF with stock price movements. Another relevant consideration is the amount of information incorporated in the late vs. the early FAF. The partial correlation results reveal that late forecasts appear to rely somewhat less on extrapolation of past earnings data and more on autonomous information than early forecasts: the partial correlation between realization and prediction, given the predictions of both the MSM and IM is 0.46 and 0.51 for early and late forecasts, respectively. The timing advantage is more pronounced when we pit early and late forecasts against each other. While the partial correlation between realization and late forecasts, given the early forecasts is 0.26 (suggesting utilization of incremental information by late forecasts) the

Table 7

Comparative performance results for early and late FAF (over years averages).[a]

Performance measure	Predictor: Early FAF	Late FAF	MSM	IM
Correlation of prediction error with *CAR*	0.31	0.34	0.25	0.26
Mean *API*, considering magnitude of error (%)	12.78	13.15	8.85	8.60
Partial correlation of realization and prediction, given both MSM and IM	0.46	0.51		
Partial correlation of realization and prediction, given early FAF		0.26	0.01	0.07
Partial correlation of realization and prediction, given early FAF, MSM and IM		0.23		

[a]FAF = Financial Analysts' Forecasts of Earnings, MSM = Modified Submartingale, and IM = Index Model. The averages are calculated over the company-years for which both early and late FAF existed.

partial correlation between realization and early forecasts (not presented in the table) was practically zero.

The findings indicate that the timing advantage of two months that late forecasts have over early forecasts affect their relative performance. Late forecasts employ a greater amount of autonomous information and their performance is somewhat better than that of early forecasts. Both early and late forecasts outperform the time-series models,[17] and it appears that the main factor behind the better performance of FAF is the broader information set used by them.

4. Concluding remarks

The study provides evidence which indicates that, overall, analyst forecasts are a better surrogate for market expectation of earnings than time-series

[17]It should be noted that the comparison between the early FAF and the naive models is 'unfair' to the former: naive models utilize the most recent earnings numbers and have an advantage over FAF that do not incorporate these yet undisclosed audited results for the year.

P. 13

models customarily used in the literature. This finding does not invalidate the results of studies which use time-series models to find an association between unexpected earnings and share price changes. In fact, it reinforces the results by indicating that the association is even stronger. This paper's results provide added motivation for the study of other important properties of FAF such as time-series behavior and cross-section dispersion.

The study also analyzes the cause of the superior performance of FAF. The results point to the existence of some timing advantage to forecasts that are made well after the end of the fiscal year and which presumably incorporate more recent information. However, the main contributor to the better performance of FAF is their ability to utilize a much broader set of information than that used by the univariate time-series models. The findings further suggest that analysts efficiently exploit the extrapolative power of the earnings series itself.

The findings of the study should be analyzed cautiously. Only two extrapolation models were considered — the submartingale (or MSM) and the index model. It should be noted, however, that these models were found by previous research to perform well when compared to other, sometimes more complex, models.

The representative of FAF was the mean forecast. Even if FAF are associated with the true market expectations, the mean might not be the proper variable. A case can be made for other measures such as the median forecast. To the extent that the mean forecast is not the measure most strongly associated with market expectations, our results underestimate the superiority of FAF as an expectation surrogate.

Another potential source of a bias, possibly against FAF, is the sample selection criterion whereby only firms with at least four contemporaneous forecasts were considered. The criterion, which was introduced to assure a meaningful measure of 'consensus' forecast led inevitably to the exclusion of many small firms which do not attract considerable attention by analysts.[18] If the remaining firms, which are larger, experience smaller earnings variability, the performance of the extrapolative models in the sample is expected to be better than the entire population.

Further research might address the interesting issue of the relationship between the independent, or autonomous, component in analysts' forecasts, which may serve as a measure of the research efforts, and possibly of their costs, and stock characteristics such as risk and marketability.

[18] Indeed, size and earnings variability are negatively correlated; the cross-section correlation coefficient between the market value of the equity and the variance of the rate of growth of earnings of the sample firms, averaged over the 11 years, is −0.20 (significant at the 5% level). Plausibly, the correlation coefficient in the population (which is more diversified in terms of size) is even more negative.

Appendix: Specification of the exponential smoothing models and their application to the sample

The selection of the order and coefficient of the exponential smoothing model was based on the findings of Brooks and Buckmaster (1976) — hereafter referred to as B & B. The stratification was according to the normalized first difference, defined as

$$d_t = (A_t - A_{t-1})/\sigma_{t-1},$$

where A_t is the EPS in year t and σ_{t-1} is the standard deviation of A over the available history of the company from 1959. Table 8 presents the distribution of company-years in the sample according to d, the comparative distribution in the much larger sample used by B & B, and the order and coefficient of the best smoothing model using the minimization of the mean-absolute-error as the optimization criterion. The distribution of cases in our sample is essentially similar to that of B & B. However, our sample has a somewhat lower percentage of extreme observations. This might be due to the special care that was taken in verifying the correctness of apparent anomalous earnings changes in the data. This verification procedure was obviously infeasible in the large sample of B & B. Note that for almost 70% of the cases (company-years) in our sample, the martingale process is the best predictor.

Table 8

Distribution of company-years by the magnitude of normalized first differences of earnings and the corresponding best predictor.

	This sample		B & B study[a]			
					Best smoothing model[b]	
Normalized first difference	No. of cases	% of cases	No. of cases	% of cases	Order	Constant
$9 < \text{difference}$		1.3			1	0.90
$6 < \text{difference} \leq 9$		2.2			1	1.00
$4 < \text{difference} \leq 6$		6.2			1	1.00
$2 < \text{difference} \leq 4$		23.6			1	1.00
$1 < \text{difference} \leq 2$		20.5			1	1.00
$0 \leq \text{difference} \leq 1$		24.8			1	1.00
$-1 \leq \text{difference} < 0$		9.8			1	1.00
$-2 \leq \text{difference} < -1$		5.6			1	0.65
$-4 \leq \text{difference} < -2$		4.8			1	0.45
$-6 \leq \text{difference} < -4$		1.0			1	0.33
$-9 \leq \text{difference} < -6$		0.2			1	0.1
$\text{difference} > -9$					2	0.2

The smoothing models for the nth order is

$${}_nE(A_t)=\alpha A_{t-1}+(1-\alpha)_nE(A_{t-1}),$$

where ${}_nE(A_t)$ is the smoothing function of the nth order model at time t (see footnote 6).

References

Abdel-Khalik, Rashad and J. Espejo, 1978, Expectations data and the predictive value of interim reporting, Journal of Accounting Research, Spring, 1–13.

Albrecht, Steve N., Larry L. Lookbill and James C. McKeown, 1976, The time series properties of earnings, Journal of Accounting Research, Autumn, 226-244.

Ball, Ray and Philip Brown, 1968, An empirical evaluation of accounting index numbers, Journal of Accounting Research, Autumn, 159–178.

Ball, Ray and Ross L. Watts, 1979, Some additional evidence on survival biases, Journal of Finance, March, 197-206.

Barefield, R.M. and E. Comiskey, 1975, The accuracy of analysts' forecasts of earnings per share, Journal of Business Research, July, 241–252.

Barnea, Amir, Joshua Ronen and Simcha Sadan, 1976, Classificatory smoothing of income with extraordinary items, The Accounting Review, Jan., 110–122.

Beaver, William H. and R.E. Dukes, 1972, Interperiod tax allocation, earnings expectations and the behavior of security prices, The Accounting Review, April, 320–332.

Beaver, William H., Roger Clark and William F. Wright, 1979, The association between unsystematic security returns and the magnitude of earnings forecasts, Journal of Accounting Research, Autumn, 316–340.

Brooks, Leroy D. and Dale E. Buckmaster, 1976, Further evidence of the time series properties of accounting income, Journal of Finance, Dec., 1359-1373.

Brown, Lawrence, D. and Michael S. Rozeff, 1978, The superiority of analyst forecasts as measure of expectations: Evidence from earnings, Journal of Finance, March, 1–16.

Brown, Lawrence, D. and Michael S. Rozeff, 1979, Adaptive expectation, time series models and analyst forecasts revision, Journal of Accounting Research, Autumn, 341-351.

Brown, Lawrence, D., John S. Hughes, Michael S. Rozeff and James H. Vanderweide, 1980, Expectation data and the predictive value of interim reporting: A comment, Journal of Accounting Research, Spring, 278-288.

Brown, Philip and J. Kennelly, 1972, The information content of quarterly earnings: An extension and some further evidence, Journal of Business, July, 403–415.

Collins, Daniel W., 1975, SEC product line reporting and market efficiency, Journal of Financial Economics, June, 121–164.

Collins, William A. and William S. Hopewood, 1980, A multiple analysis of annual earnings forecasts generated from quarterly forecasts of financial analysts and univariate time series models, Journal of Accounting Research, Autumn, 390-406.

Critchfield, Timothy, Thomas Dyckman and Josef Lakonishok, 1978, An evaluation of security analysts' forecasts, The Accounting Review, July, 651–668.

Elton, Edwin J., Martin J. Gruber and Mustafa Gultekin, 1981, Professional expectations: Accuracy and diagnosis of errors, Working paper, June (Graduate School of Business Administration, New York University, New York).

Foster, George, 1977, Quarterly accounting data: Time series properties and predictive ability results, The Accounting Review, Jan., 1–21.

Foster, George, 1978, Asset pricing models: Further tests, Journal of Financial and Quantitative Analysis, March, 39–53.

Freund, John E., 1962, Mathematical statistics (Prentice-Hall, Englewood Cliffs, NJ).

Givoly, Dan and Josef Lakonishok, 1979, The information content of financial analysts' forecasts of earnings, Journal of Accounting and Economics, Winter, 1–21.

Givoly, Dan and Dan Palmon, 1981, Timeliness of annual reports — Some empirical evidence, The Accounting Review, forthcoming.

Gonedes, N., 1973, Properties of accounting numbers: Models and tests, Journal of Accounting Research, Autumn, 212–237.

Jaffe, Jeffrey F., 1976, Insiders and market efficiency, Journal of Finance, Sept., 1141–1148.

Mandelker, Gershon, 1974, The effect of mergers on the value of the firm, Journal of Financial Economics, Dec., 303–336.

McDonald, C., 1973, An empirical examination of the reliability of published predictions of future earnings, The Accounting Review, July, 502–510.

Mincer, Jacob, ed., 1969, Economic forecasts and expectations (National Bureau of Economic Research, New York).

Patell, James M., 1976, Corporate forecasts of earnings per share and stock price behavior: Empirical tests, Journal of Accounting Research, Autumn, 246–276.

Patell, James M., 1979, The API and the design of experiments, Journal of Accounting Research, Autumn, 528–549.

Reinganum, Marc R., 1981, Misspecification of capital asset pricing empirical anomalies based on earnings' yields and market values, Journal of Financial Economics 9, 19–46.

Securities and Exchange Commission, 1977, Report of the advisory committee on corporate disclosure (S.E.C., Washington, DC).

Standard and Poor, 1967–, Earnings Forecaster, weekly (S & P, New York).

Theil, Henry, 1966, Applied economic forecasting (North-Holland, Amsterdam).

Watts, Ross L., 1978, Systematic abnormal returns after quarterly earnings announcements, Journal of Financial Economics, June–Sept., 127–150.

Watts, Ross L. and Richard W. Leftwich, 1977, The time series of annual accounting earnings, Journal of Accounting Research, Autumn, 253-271.

R. Charles Moyer, Robert E. Chatfield, Gary D. Kelly
"The Accuracy of Long-Term Earnings Forecasts
in the Electric Utility Industry"
***International Journal of Forecasting* Vol. 1**
1985

international journal of forecasting

INTERNATIONAL INSTITUTE OF FORECASTERS

NORTH-HOLLAND

International Journal of Forecasting 1 (1985) 195-196
North-Holland

CONTRIBUTORS

Jobs in the 1980's and Beyond 197

Janet L. NORWOOD

Janet Norwood is the Commissioner of Labor Statistics in the U.S. Department of Labor since 1979. She has a PhD. from Tufts University and in addition to her government experience she has written extensively for a variety of statistical and labor review journals.

Forecasting Automobile Insurance Paid Claim Costs Using Econometric and ARIMA Models 203

J. David CUMMINS and Gary L. GRIEPENTROG

J. David Cummins is the Harry J. Loman Professor of Insurance at the Wharton School, University of Pennsylvania. He is also Director of the Center for Research on Risk and Insurance and Associate Director of the S.S. Huebner Foundation for Insurance Education. Dr. Cummins currently serves as Vice President of the American Risk and Insurance Association and Associate Editor of the *Journal of Risk and Insurance*. His recent books include *Strategic Planning and Modeling in Property-Liability Insurance* and *Risk Classification in Life Insurance*.

Gary L. Griepentrog is Associate Professor of Business Administration at the University of South Carolina. He holds a PhD. degree in economics from Purdue University. His articles have appeared in numerous journals including the *Journal of Financial and Quantitative Analysis*, the *American Statistician*, and the *Southern Economic Journal*.

Comparative Analysis of Company Forecasts and Advanced Time Series Techniques Using Annual Electric Utility Energy Sales Data 217

William R. HUSS

William R. Huss is the Manager of Forecasting and Planning at Battelle Columbus Division where he has led a number of projects in planning and forecasting for the electric and gas utility industries. Mr. Huss is currently involved in a study for the Electric Power Research Institute which is developing a methodology for forecasting and simulating electrical energy, peak demand, and patterns of use by service area and establishment in the manufacturing sector. Mr. Huss also has been a key participant in transferring Battelle's SHAPES end use load forecasting package to the personal computer. In addition, he has led a project for a large gas utility to forecast the amount of fuel switching (from gas to oil) which will occur in the industrial sector given various future price scenarios. Mr. Huss holds a BA., summa cum laude in mathematics and physics from Gettysburg College as well as an M.A. in public administration and an MS. in industrial and systems engineering from The Ohio State University. He is completing a PhD. in forecasting and decision analysis from Ohio State.

The Accuracy of Long-Term Earnings Forecasts in the Electric Utility Industry 241

R. Charles MOYER, Robert E. CHATFIELD and Gary D. KELLEY

International Journal of Forecasting 1 (1985) 241–252
North-Holland

THE ACCURACY OF LONG-TERM EARNINGS FORECASTS IN THE ELECTRIC UTILITY INDUSTRY

R. Charles MOYER *
Texas Tech University, Lubbock, TX 79409, USA

Robert E. CHATFIELD *
Texas Tech University, Lubbock, TX 79409, USA

Gary D. KELLEY *
West Texas State University, Canyon, TX 79016, USA

This paper examines the accuracy of various methods of forecasting long-term earnings growth for firms in the electric utility industry. In addition to a number of extrapolative techniques, *Value Line* analyst forecasts are also evaluated. *Value Line* analyst forecasts for a five-year time horizon are found to be superior to many of the extrapolative models. Among the extrapolative models examined, implied growth and historical book value per share growth rate models performed best. These results provide strong support for using *Value Line* growth forecasts in cost of capital estimates for electric utilities in the context of utility rate cases. *Value Line* forecast errors could be explained by changes in dividend payout ratios, the firm's regulatory environment and bond rating changes.

Keywords: Earnings forecasting, Utility forecasting, Analysts' forecasts, Electric utilities.

1. Introduction

A central issue in most public utility rate cases is the determination of the cost of equity capital for the utility. In the regulatory process the return required by investors is considered a legitimate cost of doing business that is appropriately charged to customers. Other things being equal, the lower the rate of return which a utility is permitted to earn from its customers, the higher the level of customer welfare. However, if the utility does not have the opportunity to earn investor-required rates of return on capital, investment in plant and equipment will lag and the demand for service at the established price will be greater than the utility can supply. Accordingly, it is important to permit a utility to earn a fair return on its invested capital in order to assure that adequate levels of service will be provided.

Two landmark judicial decisions have provided the general framework within which this analysis must be done. The Supreme Court concluded in the Bluefield Water Works case [Bluefield Water Works (1923)] that the 'return must be reasonably sufficient to ... support its credit and enable it to raise the money necessary for the proper discharge of its public duties.' Recognition must be given to the returns currently earned 'on investments in other business undertakings which are attended by

* The authors wish to acknowledge the assistance of Phil Sisneros and Jesse Reyes for their fine data collection and computer analysis work. We also appreciate the helpful comments of Editor Scott Armstrong, Professor Mike Rozeff, Associate Editor Lawrence D. Brown and two anonymous reviewers. Any errors are the sole responsibility of the authors.

corresponding risks and uncertainties ...'. In the Hope Natural Gas case [Federal Power Commission (1944)] the Supreme Court stated that the return must also enable a firm to 'maintain its credit and attract capital'.

These judicial guidelines provide a general framework for implementing the determination of the cost of equity capital in utility rate cases. Neither the Hope nor the Bluefield decisions provides guidance about what specific method(s) should be used to establish the cost of equity. In the Hope case, the Court stated that 'under the statutory standard of 'just and reasonable' it is the result reached not the method employed which is controlling' [Federal Power Commission (1944, p. 603)].

In contrast, the rich academic literature in this area has emphasized the appropriateness of various methods employed to determine the cost of equity capital [Brigham and Gordon (1968), Elton and Gruber (1971), Gordon (1974), Gordon and Gould (1978), Litzenberger, Ramaswamy and Sosin (1980), Myers (1972) and Robichek, Higgins and Kinsman (1973)]. In practice, three models have dominated recent utility rate cases. These are the capital asset pricing model, the comparable earnings model, and the constant-growth form of the dividend valuation model (often called the DCF or discounted cash flow methodology).

This paper focuses on the DCF model as it is commonly applied in utility rate cases. Specifically, we examine the long-term accuracy of a number of forecasting techniques which are used to estimate the growth rate component in the DCF cost of equity model. [1] Based on a rational expectations view of the formation of investor expectations, [2] we find support for the use of *Value Line* analyst forecasts, [3] implied growth techniques, and historical book value growth rate models. However, *Value Line* forecast accuracy deteriorates significantly if the forecast is evaluated over a three or four year time horizon rather than the maximum five year horizon reported by *Value Line*.

Section 2 of the paper develops the DCF model as it is normally applied in rate cases. Section 3 describes the data used, and Section 4 discusses the various forecasting techniques tested. In Section 5 the statistical tests used in the analysis are discussed; Section 6 presents the results of the tests. Section 7 reports the results of tests conducted to explain the errors in *Value Line* analyst forecasts. Section 8 offers conclusions and implications.

2. The DCF model

The DCF model of valuation is based on the proposition that the value of a share of stock is equal to the present value of all expected future dividends, discounted at the shareholders' required rate of return. Expert witnesses in utility rate cases commonly rely on a constant growth form of the basic dividend valuation model, such as $k_e = D_1/P_0 + g$, as the basis for their cost of equity recommendations. [4] Expert witnesses do so because it is thought that many utility firms meet or nearly meet the requirements necessary to use the constant growth DCF model. Whether the constant growth DCF

[1] There is an extensive literature, including Brown and Rozeff (1978), Cragg and Malkiel (1968), Elton and Gruber (1972), Johnson and Schmitt (1974) and Ruland (1980) that considers the accuracy of short-term forecasting models. With the exception of a recent paper by Rozeff (1983), there has been very little analysis of the accuracy of long-term earnings forecasts.

[2] We use the term 'rational expectations' in the same sense as Sargent (1972, p. 74), and Brown and Rozeff (1978, p. 1). We use the term, basically, to mean that rational investors' expectations are the same as the best availabe forecasts.

[3] *Value Line* is a well-known, widely available, investment advisory service which is published quarterly and includes, among other things, five year earnings forecasts for the over 1700 firms followed by the service.

[4] Twenty-four witnesses who were authorities on the cost of capital testified before the Federal Energy Regulatory Commission in eleven separate rate cases between 1980 and 1982. An analysis of their testimony showed that all used $k_e = D_1/P_0 + g$ as the basis of their DCF analysis, where k_e is the cost of equity capital, D_1 is dividends expected over the next period, P_0 is the current market price of the firm's stock and g is the long-term perpetual growth rate in dividends.

model or the non-constant growth model is employed, long-term (three to five year) earnings and dividend growth forecasts are essential inputs.

The application of this model invariably results in considerable controversy among expert witnesses regarding the appropriate method by which to estimate the growth (g) component. Theoretically, this growth component is the growth rate expected by investors at the margin. Since expectations cannot be directly observed, experts focus on a wide range of alternative techniques as a proxy for g. According to the rational expectations hypothesis [Sargent (1972)], the best forecasting method should be used to estimate g. In practice, proxies for g have included historical earnings and dividend growth rates, historical book value growth rates, implied growth rates (the product of the retention ratio times the return on book equity), and analysts' forecasts such as *Value Line*.

This paper examines the long-term accuracy of different methods of forecasting earnings growth of electric utility corporations and compares the results with *Value Line* forecasts of future earnings growth. On an ex-post basis the different methods are evaluated to determine the most accurate, long-range (three to five year) forecast. [5]

3. The data

The sample consists of the ninety-eight electric utilities that *Value Line* followed between 1971 and 1976 and the ninety-three electric utilities followed by *Value Line* between 1977 and 1982. Per share data have been adjusted for stock splits and dividends. Generally, *Value Line* reports on each firm four times a year. The *Value Line* data come from its second quarterly report of each year since this is the first *Value Line* report which generally includes actual data for the previous year. For example, *Value Line* earnings forecasts for 1976 are those reported in its second quarterly report in 1972.

All data, both actual earnings and forecasts of earnings, have been converted to compound annual growth rates. Hence, all comparisons of forecast accuracy are based on annual growth rates. Two five-year forecast horizons are used in the analysis: 1971–1976 and 1977–1982. *Value Line* makes its earnings per share forecasts for a three-year range, e.g., the forecast made in 1972 (which is conditional on actual 1971 data) is for the 1974–1976 time period. Thus, forecasted *Value Line* growth rates can be computed assuming a three, four, or five-year horizon. We considered each possible *Value Line* horizon in the paper, i.e., earnings forecasting accuracy is evaluated for the 1971–1974, 1971–1975 and the 1971–1976 time periods, as well as the 1977–1980, 1977–1981, and the 1977–1982 time periods.

These time periods are especially important for the electric utility industry because of the unsettled conditions prevailing in that industry through the 1970s. These conditions include the effects of rapidly escalating fuel costs, the need to convert large amounts of capacity from natural gas and oil to coal and nuclear power, and the impact of high inflation and rapidly rising capital costs.

4. Forecasting methods

The forecasting methods tested have been selected for analysis because of their use in prior studies and because of the extent to which they are commonly used in utility rate cases. These methods are:

X2. *Value Line* 3, 4, and 5-year earnings forecast.
X3. The 5-year historical compound dividend per share growth rate; for example, the 1971–1976 forecast horizon uses the actual annual compound growth rate from 1966–1971.

[5] The three to five year horizon was chosen since this is the longest forecast horizon available from *Value Line* analysts.

X4. The 5-year historical compound earnings per share growth rate.
X5. The 5-year historical compound book value per share growth rate.
X6. The 10-year historical compound dividend per share growth rate.
X7. The 10-year historical compound earnings per share growth rate.
X8. The 10-year historical compound book value per share growth rate.
X9. The 5-year average implied earnings growth rate, i.e., the 5-year historical average return on equity times the 5-year historical average retention rate.
X10. The 10-year average implied earnings growth rate.
X11. The current implied earnings growth rate (e.g., the implied growth rate for the 1971–1976 forecasting horizon is equal to the return on equity in 1971 times the 1971 retention rate).
X12. Brigham–Shome method of smoothing to compute the implied earnings growth rate [Brigham and Shome (1981)]; for example, the implied growth rate for the 1971–1976 forecasting horizon is equal to smoothed ROE times smoothed retention rate and the smoothed ROE is computed as

$$0.1ROE_{t-4} + 0.2ROE_{t-3} + 0.3ROE_{t-2} + 0.4ROE_{t-1} = ROE \text{ forecast}.$$

A similar computation is done for the retention rate forecast.
X13. The growth rate computed from the following trend line in book value per share (BPS) over a five year period

$$\ln BPS = a + bt.$$

X14. Same as X13 except for the use of 10 years of historical data.
X15. The growth rate computed from a trend line in dividends per share over a 5-year period.
X16. Same as X15 except for the use of 10 years of historical data.
X17. The growth rate computed from a trend line in earnings per share over a 5-year period.
X18. Same as X17 except for the use of 10 years of historical data.

X1 is defined as the actual 3, 4 or 5-year compound annual growth rate in earnings per share, e.g., the growth rate for the 1971 to 1976 time horizon is the actual compound annual growth computed using 1971 earnings per share as the start point and 1976 earnings per share as the end point. Similar computations are made for each horizon.

5. Statistical tests

First we examined the directional relationship between individual forecasts and actual earnings per share (EPS) growth rates. Kendall rank order correlations were calculated between the forecasted growth rates for each of the forecasting methods and the actual earnings growth rates. Next, similar to Rozeff (1983), the average deviation (average forecast growth minus average actual growth), mean absolute error ($MABE$) and root mean square error ($RMSE$) were calculated for each forecasting method. The $MABE$ is the sample average of the absolute value of the forecast error calculated for each forecast method on the entire sample of firms. The $RMSE$ is the square root of the sample average of the squared forecast error. As such, $RMSE$ gives more weight to large forecast errors than does $MABE$.

A method similar to that used by Brown and Rozeff (1978) was employed to test for significant differences in the accuracy of each forecasting model and of *Value Line*. The measure of forecast accuracy used was the absolute value of the difference between forecasted growth in *EPS* for each of n forecast methods (for each time horizon) over i firms (g_{in}) and actual growth in *EPS* over the same horizon (a_i), or $|g_{in} - a_i|$. The forecast errors were then compared across firms.

We used the Friedman test [Friedman (1937)] to test for the relative accuracy of all forecasting methods. The test criterion was the magnitude of forecast error. In practice the distribution of the Friedman test statistic is usually approximated by the chi-square distribution as in Brown and Rozeff (1978), but recent studies by Iman and Davenport (1980) show that the F-distribution approximation is superior to the chi-square approximation. Hence, the F-distribution approximation to the Friedman test is employed to test the null hypothesis that all seventeen forecasts are equally accurate. If the null hypothesis is rejected, we may conclude that at least one forecasting method is superior to at least one other.

The next step in evaluating the relative accuracy of the forecasting methods was to compare forecast accuracy across firms using pairwise comparisons between forecasts. These comparisons test the accuracy of a method's forecasts against each of the other methods' forecasts using a least significant difference test statistic developed by Conover (1980, p. 300). The Wilcoxian signed ranks test can also be used for these pairwise comparisons as in Brown and Rozeff (1978), but this least significant difference test is more powerful [Conover (1980)]. The null hypothesis tested is that one method's forecasts are as accurate as another method's forecasts.

6. Empirical results

Exhibit 1 reports the Kendall rank order correlations between each of the forecasting methods and the actual earnings per share growth for the two five-year forecast horizons. In both five-year periods,

Exhibit 1
Kendall rank order correlations between actual 5-year annual earnings growth rates and earnings forecasts.

Method	Period 1 (1971–1976)	Period 2 (1977–1982)
X2	0.214 [a]	0.269 [a]
X3	−0.153 [b]	−0.118 [c]
X4	−0.093	−0.058
X5	0.013	0.151 [b]
X6	0.021	0.105
X7	−0.020	0.084
X8	0.013	0.033
X9	−0.137 [b]	0.078
X10	−0.091	0.042
X11	−0.209 [a]	−0.164 [b]
X12	−0.149 [b]	0.024
X13	−0.010	0.112
X14	0.006	0.077
X15	0.020	0.193 [a]
X16	0.007	0.109
X17	−0.132 [c]	−0.108
X18	−0.085	0.065

[a] Significant at 1% or better.
[b] Significant at 5%.
[c] Significant at 10%.

Exhibit 2
Summary of error statistics 1971–1976.

Method	Average deviation (forecast–actual)	MABE	RMSE
X2	0.021	0.036	0.044
X3	−0.013	0.047	0.066
X4	0.013	0.042	0.053
X5	0.006	0.038	0.057
X6	0.016	0.039	0.048
X7	0.003	0.037	0.046
X8	0.013	0.039	0.050
X9	−0.002	0.036	0.046
X10	0.000	0.035	0.045
X11	−0.007	0.040	0.056
X12	−0.004	0.037	0.049
X13	0.007	0.038	0.046
X14	0.009	0.036	0.045
X15	0.000	0.038	0.050
X16	0.015	0.039	0.047
X17	−0.017	0.050	0.070
X18	0.007	0.040	0.050

Value Line forecasts (X2) are positively and significantly correlated with actual earnings growth.

In period 1, no other forecasting method is both significant and positively correlated with actual earnings growth. In period 2, methods X5 (five-year compound book value per share growth) and X15 (five-year trend line growth in dividends per share) also have statistically significant positive correlations.

Exhibit 1 provides strong cross-sectional evidence of the superiority of *Value Line* forecasts in capturing movement in the *direction* of earnings growth rates. Thus, *Value Line* forecasts higher growth for firms which later show higher growth, and lower growth for firms which later show lower growth. During the highly unstable periods included in the forecast horizons, only *Value Line* forecasts consistently reflected the direction of movement in actual earnings growth rates for the electric utility industry.

Exhibit 1 does not, however, show any indication of the accuracy of *Value Line* relative to alternative forecasting techniques. From a cost of capital perspective, accuracy in forecasting is of greatest importance. Exhibits 2 and 3 report the average deviation, mean absolute error and root mean square error for the two five-year forecast horizons.

The *Value Line* average deviation is the largest in period 1 at 2.1%, but the lowest in period 2 at 1%. In both periods it is positive, indicating that *Value Line* forecasts tend to be on the high side. Hence, it appears that in the long-term (five years) *Value Line* is relatively successful in forecasting the *direction* of future earnings movements, but there is a tendency to overestimate the size of this earnings growth. In order to verify this initial conclusion we next look at two other measures of overall forecasting accuracy – the *MABE* and *RMSE*.

Value Line has a relatively low *MABE* in period 1. Only X10 (ten-year average implied growth of *EPS*) is lower; X9 (five-year average implied growth) and X14 (ten-year trend line growth in book value) are equivalent. In period 2 *Value Line* has the lowest *MABE*. *Value Line* appears even better when accuracy is evaluated using *RMSE*. In both periods *Value Line* has the lowest *RMSE*.

Thus, in addition to forecasting successfully the direction of movement, *Value Line* is relatively accurate as a predictor of the future growth rate itself. Its forecasts tend to be on the high side but

Exhibit 3
Summary of error statistics 1977–1982.

Method	Average deviation (forecast–actual)	*MABE*	*RMSE*
X2	0.010	0.039	0.059
X3	−0.030	0.067	0.094
X4	−0.019	0.053	0.075
X5	−0.013	0.044	0.063
X6	−0.013	0.044	0.063
X7	−0.024	0.051	0.070
X8	−0.011	0.045	0.065
X9	−0.016	0.046	0.067
X10	−0.013	0.045	0.065
X11	−0.015	0.052	0.074
X12	−0.017	0.048	0.070
X13	−0.027	0.052	0.070
X14	−0.014	0.045	0.065
X15	−0.012	0.045	0.068
X16	−0.016	0.046	0.065
X17	−0.015	0.065	0.093
X18	−0.020	0.049	0.071

when compared to the sixteen mechanical forecasting methods, it is among the most accurate.

Finally, we consider two statistical tests of relative accuracy – the Friedman test and the least significant difference test. Exhibits 4 and 5 report the results from these two tests for periods 1 and 2 respectively. The Friedman test rejects the null hypothesis at the 1% level for both periods. Thus, the alternative hypothesis that at least one forecasting method is more accurate than at least one other forecasting method may be accepted.

The least significant difference test of the multiple pairwise comparisons is performed at a 5% significance level. The results indicate that *Value Line* is dominated only by X10 (ten-year average implied growth) in period 1 and is not dominated by any forecasting method in period 2.

Several of the forecasting methods performed exceedingly well in the multiple pairwise comparisons. X5, X8 (five and ten-year compound book value per share growth), X9, X10 (five and ten-year average implied growth), X14 (ten-year trend line growth in book value), and X15 (five-year trend line growth in dividends) are not dominated by any other forecasting method in either period.

In summary, *Value Line* performs very well relative to the 16 extrapolative forecasting methods in the five-year forecast horizons. It is relatively successful at forecasting the direction of future earnings growth. Also, the *MABE*, *RMSE*, and multiple pairwise comparisons indicate that *Value Line* is relatively accurate in predicting the actual future growth rate.

Value Line forecasts are made for a three to five-year forecast horizon. The preceding results have focused on the five-year horizon. Identical statistical tests were performed for two three-year horizons (1971–1974 and 1977–1980) and two four-year horizons (1971–1975 and 1977–1981). Because *Value Line* forecasts per share earnings for a three to five-year horizon, the calculated growth rate will be greater the shorter the horizon. Since the *Value Line* forecasts tended to overestimate the actual growth rate for five-year horizons, one would expect the same dollar earnings forecast for a three or four-year horizon to perform less well.

The correlation results for three and four-year horizons are similar to those for five years. *Value Line* forecasts are positively and significantly correlated with actual earnings growth in both periods for both the three and four-year horizons. In addition to *Value Line*, only X5 and X10 are significant

Exhibit 4
Multiple pairwise comparisons period 1 (1971–1976). [a]

	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	Times superior	Times inferior
X2			+						–									1	1
X3								–	–		–		–	–				0	5
X4	–			–		–	–	–	–	–	–		–	–				0	10
X5			+													+		2	0
X6								–	–									0	2
X7			+													+		2	0
X8			+													+		2	0
X9		+	+		+							+			+	+	+	7	0
X10	+	+	+		+							+			+	+	+	8	0
X11			+													+		2	0
X12		+	+													+		3	0
X13								–	–									0	2
X14		+	+													+		3	0
X15		+	+													+		3	0
X16								–	–									0	2
X17				–		–	–	–	–	–	–		–	–				0	9
X18								–	–									0	2

[a] Friedman test. F-value is 2.63, significant at 1% level. A plus sign (negative sign) in the table indicates the forecast method represented by the row is superior (inferior) to the forecast method represented by the column at a significance level of 5%.

Exhibit 5
Multiple pairwise comparisons period 2 (1977–1982). [a]

	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	Times superior	Times inferior
X2		+	+			+				+		+				+		6	0
X3	–			–	–	–	–	–	–	–	–	–	–	–	–		–	0	14
X4	–			–	–		–	–	–		–		–	–	–		–	0	11
X5		+	+			+				+		+				+	+	7	0
X6		+	+			+				+		+				+	+	7	0
X7	–	+		–	–		–	–	–		–		–	–	–			1	10
X8		+	+			+				+		+				+		6	0
X9		+	+			+				+		+				+		6	0
X10		+	+			+				+		+				+		6	0
X11	–	+		–	–		–	–	–				–	–	–			1	9
X12		+	+			+						+		–		+		5	1
X13	–	+		–	–		–	–	–		–		–	–	–			1	10
X14		+	+			+				+		+				+		6	0
X15		+	+			+				+	+	+				+	+	8	0
X16		+	+			+				+		+				+		6	0
X17	–			–	–		–	–	–		–		–	–	–		–	0	11
X18		+	+	–	–									–		+		3	3

[a] Friedman test. F-value is 8.24, significant at 1% level. A plus sign (negative sign) in the table indicates the forecast method represented by the row is superior (inferior) to the forecast method represented by the column at a significance level of 5%.

and positively correlated. Moreover, this phenomenon persists only in period 2 for three and four-year horizons.

The average deviation, *MABE*, and *RMSE* show *Value Line*'s forecast to decline appreciably in relative accuracy. With the exception of the *RMSE* in period 2 of the three and four-year horizons, *Value Line* is outperformed in these measures of relative accuracy by all or most of the sixteen forecasting methods.

The multiple pairwise comparisons for the four-year horizon still show *Value Line* to be relatively accurate. It is less accurate than only one method in both periods. However, for the three-year horizon, it is less accurate than all the other methods in period 1 and less accurate than 14 of 16 methods in period 2.[6]

These results indicate that, whether it is intentional or not, *Value Line* tends to forecast most accurately to the five-year end of their three to five-year forecast horizon. In forecasting earnings for a five-year horizon, *Value Line* is very successful relative to the sixteen extrapolative forecasting methods examined in this study.

7. Error analysis of value line forecasts

The results reported in section 6 indicate that *Value Line* earnings growth rate forecasts for a five-year horizon are significantly, positively correlated with actual earnings growth rates. In addition, *Value Line* forecasts have mean absolute errors and root mean square errors which are among the lowest when compared with the sixteen extrapolative models. The multiple pairwise comparison tests reported in exhibits 4 and 5 indicate that *Value Line* forecasts are less accurate than only one other forecast method in the 1971–1976 period, and are not less accurate than any other method during the 1977–1982 period.

In this section we perform a micro-analysis of errors in order to discover causes for over and under-estimates of forecasted earnings growth rates made by *Value Line*. This analysis can help users of *Value Line* earnings forecasts to identify instances where *Value Line* forecasts are likely to be least reliable.

We have examined a number of firm-specific/regulatory environment variables which might be expected to influence the accuracy of *Value Line* forecasts. These variables are

(1) *Regulatory environment. Value Line* rates the regulatory environment faced by each firm as either above average, average, or below average. It is possible that regulatory environments that are perceived to be more (less) favorable cause the analysts to over-(under-)estimate actual earnings growth potential for the firm. Two dummy variables are used to represent regulatory environment at the end of each forecast horizon ($D_1 = 1$ if above average, 0 otherwise; $D_2 = 1$ if average, 0 otherwise; below average is the excluded class).

(2) *Percent of electric revenues from residential customers* (measured at the end of each forecast horizon). Residential electric revenue is less subject to cyclical fluctuations than commercial and industrial electric revenue. Hence, firms with a high proportion of residential demand might be expected to have more stable and easily forecasted earnings.

(3) *Percent of revenues from electric sales* (measured at the end of each forecast horizon). Some firms in the sample had a significant portion of total revenues attributable to natural gas distribution services and/or other diversified business efforts. During the 1971–1982 time period, natural gas demand was highly volatile because of shortages and large price increases. Hence, firms that

[6] Complete statistical results for the three and four-year horizons are available on request from the authors.

concentrated on providing electric service might also be expected to have more stable and easily forecasted earnings.

(4) *Percent of generation from oil and gas capacity* (measured at the end of each forecast horizon). Oil and gas prices increased dramatically during the time periods examined. and not all firms had the benefit of perfectly effective fuel adjustment clauses. Hence, it is hypothesized that those firms with a greater proportion of oil and gas generating capacity were faced with more volatile and less easily forecasted earnings during this period.

(5) *Nuclear construction.* Firms with a significant nuclear construction program [defined with a dummy variable (D_3) as a firm having a greater than 10% ownership interest in a nuclear plant under construction at the end of each forecast horizon] were expected to have more volatile and less easily forecasted earnings than non-nuclear firms. This is particularly true during the 1977–1982 period when, following the accident at Three Mile Island, the Nuclear Regulatory Agency ordered plant shutdowns. At that time, also, cancelled projects began to affect adversely the earnings of electric utilities.

(6) *Percentage change in dividend payout ratio* (defined as the 1976 payout ratio minus the 1971 payout ratio for the first period and the 1982 payout ratio minus the 1977 payout ratio for the second period). An increase in the payout ratio reduces funds for reinvestment in the firm and is hypothesized to be directly related to overestimates of earnings made by *Value Line.*

(7) *Percentage change in net plant* (measured as the percentage increase (decrease) in net plant over the period). The hypothesized direction of the effect of this variable is indeterminant since a rapid growth in net plant might be associated with growth in demand and future earnings. Alternatively, firms with large construction programs during the 1970s and 1980s have been under heavy financing and regulatory pressures that have negatively influenced earnings.

(8) *Change in bond ratings* (measured from the beginning to the end of each period by two dummy variables: $D_4 = 1$ if downgraded by Moody's, 0 otherwise; $D_5 = 1$ if upgraded by Moody's, 0 otherwise; firms with no rating change are the excluded set). When a firm is upgraded (downgraded), this indicates an improvement (decline) in its financial profile. Hence, upgradings (downgradings) might be associated with underestimates (overestimates) of future earnings.

(9) *Coefficient of variation of earnings per share* (measured over the ten years prior to the start of each forecast horizon). Highly volatile earnings are expected to be positively related to *Value Line* earnings forecasting errors.

For each forecasting horizon (1971–1976 and 1977–1982), two regressions were run using the above independent variables and (1) positive forecasting errors (*Value Line* minus actual) and (2) negative forecasting errors as the dependent variables.

During the 1971–1976 period, the factors identified above explained 24% (adjusted) of the variation in the positive *Value Line* errors and 13% (adjusted) of the variation in negative *Value Line* errors. The only factor significant at the 5% or better level was the percentage change in the payout ratio. Increases in a firm's payout ratio were significantly associated with overestimates of earnings (positive errors) made by *Value Line* analysts. This result is consistent with the support found for the use of implied growth techniques for forecasting future earnings. No factors were found to be statistically significant in explaining negative *Value Line* forecast errors during the 1971–1976 period.

During the 1977–1982 horizon, the percentage change in the payout ratio again was associated significantly with positive *Value Line* errors. In addition, there was a significant, positive relationship between bond downgradings and positive *Value Line* errors. Negative *Value Line* errors were significantly associated with bond upgradings. There was also evidence that *Value Line* significantly underestimated future earnings growth for firms with a high coefficient of variation of earnings.

In sum, this evidence suggests the *Value Line* earnings forecasts adequately consider each of the factors identified above except the impact of changes in a firm's dividend payout ratio, the effects of bond rating changes, and, to a lesser extent, the volatility of past earnings. Consequently, users of *Value Line* data should be aware of potential biases in *Value Line* earnings forecasts for firms likely to change significantly their dividend payout policy, for firms likely to have a bond downgrading or upgrading over the forrecast horizon, and for firms with historically volatile earnings. Unfortunately, forecasting changes in dividend payout ratios and bond ratings is itself a difficult matter. It can be noted, however, that although the explanatory variables examined were not generally significantly correlated with each other, there were significantly positive (+0.287 and +0.317) correlations between downgradings and nuclear construction during the 1971–1976 and 1977–1982 period respectively) and significantly negative correlations (−0.212 and −0.170) between upgradings and nuclear construction. This suggests that *Value Line* earnings forecasts were less reliable for firms with significant nuclear construction programs. Additional support for this fact can be inferred by observing that during the 1977–1982 time period, 62% (32 of 52) of the firms whose earnings were overestimated by *Value Line*, were involved with nuclear construction while only 37% (14 of 38) of the firms where *Value Line* underestimated earnings were involved with nuclear construction.

8. Summary

type whole para

Value Line performed very well in forecasting earnings per share in the 1971–1976 and 1977–1982 time horizons relative to extrapolative forecasting methods. It was clearly superior in forecasting the direction of future earnings growth and provided forecasts that were among the best when evaluated using various tests of accuracy. Among the extrapolative models, implied growth and historical book value growth rate models performed best.

The results are from two specific past time periods, but *Value Line* performed consistently well in both periods. The evidence supports the use of five-year *Value Line* earnings forecasts as an estimate of future growth rates in future cost of capital rate cases. *Value Line* forecasts based on three and four-year time horizons appear to have a significant upward bias.

The results of the micro-analysis of *Value Line* forecast errors might assist users to detect biases in the *Value Line* forecasts. In this study *Value Line* forecasts overestimated future earnings when firms increased their payout ratios or if a firm's bonds were downgraded. They underestimated when a firm's bonds were upgraded or if a firm had very volatile earnings prior to the beginning of the forecast horizon. As is true with all empirical studies, the results may pertain only to the industry and time-periods studied. Additional work is needed to ascertain whether the findings will prove applicable to other industries, time-periods, and analyses.

References

Bluefield Water Works and Investment Company v. Public Service Commission of the State of West Virginia, 262 US 679, 1923.

Brigham, Eugene F. and M.J. Gordon, 1968, Leverage, dividend policy and the cost of capital, Journal of Finance 23, 85–105.

Brigham, Eugene F. and Delip K. Shome, 1981, The risk premium approach to estimating the cost of common equity capital, Research paper (University of Florida, Gainesville, FL).

Brown, Lawrence D. and Michael S. Rozeff, 1978, The superiority of analyst forecasts as measures of expectations: Evidence from earnings, Journal of Finance 33, 1–16.

Conover, W.J., 1980, Practical non-parametric statistics, 2nd ed. (Wiley, New York).

Cragg, J.G. and Burton G. Malkiel, 1968, The consensus and accuracy of some predictions of the growth of corporate earnings, Journal of Finance 23, 67–84.

Elton, Edwin J. and Martin J. Gruber, 1972, Earnings estimates and the accuracy of expectational data, Management Science 18, B-409–B-424.

Elton, Edwin J. and Martin J. Gruber, 1971, Valuation and the cost of capital for regulated industries, Journal of Finance 26, 661–670.

Federal Power Commission v. Hope Natural Gas Company, 320 U.S. 591, 1944.

Friedman, M., 1937, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, Journal of the American Statistical Association 32, 675–701.

Gordon, Myron J., 1974, The cost of capital to a public utility (MSU Public Utility Studies, East Lansing, MI).

Gordon, Myron J. and L.I. Gould, 1978, The cost of equity capital: A reconsideration, Journal of Finance 33, 849–861.

Iman, R.L. and J.M. Davenport, 1980, Approximations of the critical region of the Friedman statistic, Communications in Statistics A9, 571–595.

Johnson, Timothy E. and Thomas G. Schmitt, 1974, Effectiveness of earnings per share forecasts, Financial Management 3, 64–72.

Litzenberger, Robert, Krishna Ramaswamy and Howard Sosin, 1980, On the CAPM approach to the estimation of a public utility's cost of equity capital, Journal of Finance 35, 369–383.

Miller, Merton and Franco Modigliani, 1966, Some estimates of the cost of capital to the electric utility industry, American Economic Review 56, 333–391.

Myers, S.C., 1972, The application of finance theory to public utility rate cases, Bell Journal of Economics and Management Science 3, 58–97.

Robichek, A.A., R.C. Higgins and M. Kinsman, 1973, The effect of leverage on the cost of equity of electric utility firms, Journal of Finance 28, 353–367.

Rozeff, Michael S., 1983, Predicting long-term earnings growth: Comparisons of expected return models, submartingales and Value Line analysts, Journal of Forecasting 2, 425–435.

Ruland, William, 1980, On the choice of simple extrapolative model forecasts of annual earnings, Financial Management 9, 30–37.

Sargent, Thomas J., 1972, Rational expectations and the term structure of interest rates, Journal of Money, Credit and Banking 4, 74–97.

International Jou
North-Holland

MODELLING
SOFT DRINK

Katarina JUS
Swedish School of

A dynamic regress
seasonal model wi
ARMA process wi
and long-run adapt
consumers as real
these trends in th
response paramete

Keywords: App
income distribut
Comparative me

1. Introduction

The aim of t
short- and long
based on an ad
unobserved co
upon a similar
components, w
ity is present.

In this paper
the per capita
grows. This ty
for which there
high level. Wi
function of a c
not appropriat
predictions wh
applied to mo
In section 2
The specificati
introduced by
(1981). The m
adaptation to
effects is mode

0169-2070/85 $3

Robert S. Harris
"Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return"
Financial Management
Spring 1986

Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return

Robert S. Harris

Robert S. Harris is a member of the faculty of the University of North Carolina at Chapel Hill. He is also an Associate Editor of Financial Management.

I. Introduction

Shareholder required rates of return play key roles in establishing economic criteria for resource allocation in many corporate and regulatory decisions. Theory dictates that such returns should be forward-looking return requirements that take into account the risk of the specific equity investment.

Estimation of such returns, however, presents numerous and difficult problems. Although theory clearly calls for a forward-looking required return, investigators, lacking a superior alternative, often resort to averages of historical realizations. One primary example is the determination of equity required return as a "least risk" rate plus a risk premium where an equity risk premium is calculated as an average of past differences between equity returns and returns on debt instruments. The historical studies of Ibbotson *et al.* [9] have been used frequently to implement this approach.[1] Use of such historical risk premia assumes that past realizations are a good surrogate for future expectations and that risk premia are roughly constant over time. Additionally, the choice of a time period over which to average data under such a procedure is essentially arbitrary. Carleton and Lakonishok [3] demonstrate empirically some of the problems with such historical premia when they are disaggregated for different time periods or groups of firms.

Recently Brigham, Shome, and Vinson [2] surveyed work on developing *ex ante* equity risk premia with particular emphasis on regulated utilities. They presented their own risk premia estimates, which make use of financial analysts' forecasts as surrogates for investor expectations.

The current paper follows an approach similar to Brigham *et al.* and derives equity required returns and risk premia using publicly available expectational

Thanks go to Ed Bachmann, Rich Harjes, and Hamid Mehran for computational assistance and to Bill Carleton, Pete Crawford, and Steve Osborn for many discussions. I gratefully acknowledge financial support from the UNC Business Foundation and the Pogue Foundation and thank Bell Atlantic for supplying data for this project. Finally, I thank colleagues at UNC for their helpful comments.

[1] Many leading texts in financial management use such historical risk premia to estimate a market return. See for example, Brealey and Myers [1]. Often a market risk premium is adjusted for the observed relative risk of a stock.

data. The estimation makes use of dividend growth models but incorporates expected rather than historical growth rates. A consensus forecast of financial analysts is used as a proxy for investor expectations. While Brigham *et al.* focus on utility securities, this paper also provides estimates of risk premia for a broad market index. Equity risk premia for both the market and for utilities are shown to vary over time with changes in the perceived riskiness of corporate activity relative to U.S. government bonds. In addition, the estimated risk premia at any given time are shown to vary across groups of stocks. The paper also provides results using the dispersion of analysts' forecasts as an *ex ante* proxy for equity risk.

Section II discusses related literature on financial analysts' forecasts (FAF) and the estimation of required returns using such forecasts. In Section III models and data are discussed. Following a comparison of the results to those of earlier studies (including historical risk premia), the estimates are subjected to economic tests of both their time-series and their cross-sectional characteristics in Section V. Finally, conclusions are offered.

II. Background and Literature Review

In finance, it is often convenient to use the notion of a shareholder's required rate of return. Such a rate (k) is the minimum level of expected return necessary to compensate the investor for bearing risks and receiving dollars in the future rather than in the present. In general, k will depend on returns available on alternative investments (*e.g.*, bonds or other equities) and the riskiness of the stock. To isolate the effects of risk it is often useful (both theoretically and empirically) to work in terms of a risk premium (rp), defined as

$$rp = k - i, \tag{1}$$

where i = required return for a zero risk investment. Theoretically, i is a risk free rate, though empirically its proxy (*e.g.*, yield to maturity on a government bond) is only a "least risk" alternative that is itself subject to risk.[2] While models such as the capital asset pricing model offer explicit methods for varying risk premia across securities, they provide little practical advice on establishing some benchmark market risk premium. Other models, such as the dividend growth model (hereafter referred to as the discounted cash flow, or DCF, model), can be used to provide direct estimates of k, and hence implied values of rp, but are silent on how rp ought to vary across firms. In this paper DCF models are used to establish risk premia both for the market and for utility stocks. Since the DCF analysis uses a consensus measure of FAF of earnings as a proxy for investor expectations, a brief review of research on FAF is appropriate.

A. Literature on FAF

Much of the burgeoning literature on properties of FAF is surveyed by Givoly and Lakonishok [8]. Of primary importance for this work is the relationship between FAF and investor expectations that determine stock prices. Such forecast data are readily available. That they are used by investors is evidenced by the commercial viability of services that provide such forecasts and by the results of studies of investors' behavior (Touche, Ross and Company [16], Stanley, Lewellen and Schlarbaum [15]). Moreover, a growing body of knowledge shows that analysts' earnings forecasts are indeed reflected in stock prices. Such studies typically employ a consensus measure of FAF calculated as a simple average[3] of forecasts by individual analysts. Elton, Gruber, and Gultekin [5] show that stock prices react more to changes in analysts' forecasts of earnings than they do to changes in earnings themselves, suggesting the usefulness of FAF as a surrogate for market expectations. In an extensive NBER study using analysts' earnings forecasts, Cragg and Malkiel [4, p. 165] conclude "the expectations formed by Wall Street professionals get quickly and thoroughly impounded into the prices of securities. Implicitly, we have found that the evaluations of companies that analysts make are the sorts of ones on which market valuation is based." Updating Cragg and Malkiel's work, Vander Weide and Carleton [17] recently compare consensus FAF of earnings growth to 41 different historical growth measures.[4] They con-

[2]In this development the effects of tax codes and inflation on required returns are ignored.

[3]Mayshar [14] discusses the problems of explaining equilibrium prices of securities when there is divergence of opinion among investors. One issue is whether it is the expectation of the marginal investor or the average investor that determines security prices. Mayshar shows that, in general given divergence of opinion and trading costs, not all investors trade in all assets and that equilibrium prices and the identity of investors trading in each asset are jointly determined. In this sense, equilibrium prices can be considered as "determined simultaneously by the average and marginal investors."

[4]Both Cragg and Malkiel [4] and Vander Weide and Carleton [17] show that an average measure of analysts' forecasts of growth in earnings is powerful in explaining cross-sectional variation in price earnings ratios of stocks.

clude that "there is overwhelming evidence that the consensus analysts' forecast of future growth is superior to historically-oriented growth measures in predicting the firm's stock price . . . consistent with the hypothesis that investors use analysts' forecasts, rather than historically-oriented growth calculations, in making stock buy and sell decisions." [17, p. 15].

B. Use of FAF to Estimate Equity Required Returns

Given the demonstrated relationship of FAF to equity prices and the direct theoretical appeal of expectational data, it is no surprise that FAF have been used in conjunction with DCF models to estimate equity return requirements. Typically such approaches have estimated an *ex ante* risk premium (rp) calculated as the difference between required return and a least risk rate as shown in Equation (1).

Malkiel [13] estimated such risk premia for the Dow Jones Industrial Index using a nonconstant growth version of the DCF model. Initial years of growth were based on Value Line's five-year earnings growth forecasts with subsequent growth approaching a long-run real national growth rate of 4%. More recently, Brigham, Vinson, and Shome [2] used a two stage DCF growth model to estimate *ex ante* risk premia for electric utilities and the Dow Jones Industrial Index. For the period 1966–1984, they report annual risk premia for both Dow Jones Industrial and Electric Indices using Value Line's forecasts. Beginning in 1980 they report monthly risk premia for electric utilities with the source of FAF varying over time; starting with Value Line, adding Merrill Lynch and Salomon Brothers in 1981 and finally, in mid-1983, adding IBES data. IBES (Institutional Broker's Estimate System) is a collection of analysts' forecasts and is discussed in the next section. The resultant risk premia vary over time. In addition, Brigham *et al.* present evidence that their estimated risk premia vary cross-sectionally with a stock's risk (as proxied by bond rating) and over time with the level of interest rates. FAF also have been used in conjunction with DCF models by a number of expert witnesses in rate of return determination for regulated utilities. Recently, the Federal Communications Commission [6] tentatively endorsed the use of consensus FAF in DCF determinations of required return on equity.[5]

This paper adds to earlier work in a number of important respects. First, while Malkiel and Brigham *et al.* focus on electric utilities or the Dow Jones Industrial Index, this paper estimates risk premia for a broadly defined market index — the Standard and Poor's 500. Thus, the results are directly comparable to historical "market" risk premia typically estimated on a similar sample of stocks. Second, the study uses a large sample of FAF (beginning in 1982 when the necessary data first became available). This provides the ability to use a consensus measure of expectations as would be suggested by financial theory. Third, the results show that the derived risk premia change over time and that these changes are related to proxies for risk, which would be expected to be associated with equity risk premia. Although such changes have been noted by earlier studies (*e.g.*, Brigham *et al.*), there is little work explaining the patterns of change. Finally, the paper shows the usefulness of the dispersion of FAF as a proxy for risk. Such a measure is a direct expectational measure of risk and does not rely on assumptions of risk stability over time as do most operational methods of deriving risk surrogates.

III. Models and Data

A. Model for Estimation

The DCF model states that the current market price is the present value of expected future cash flows from ownership. The simplest and most commonly used version estimates shareholders' required rate of return, k, as the sum of dividend yield and expected growth in dividends, or

$$k = (D_1/P_0) + g, \qquad (2)$$

where D_1 = dividend per share expected to be received at time one, P_0 = current price per share (time 0), and g = expected growth rate in dividends per share. The limitations of this model are well known, and it is straightforward to derive expressions for k based on more general specifications of the DCF model.[6] The primary difficulty in using the DCF model is obtaining an estimate of g, since it should reflect market expecta-

[5] In response to the FCC's *Notice of Proposed Rulemaking* [6] to determine authorized rates of return, AT&T used an approach driven by FAF growth estimates from IBES. Also see, for example, W.T. Carleton, *Testimony before the Vermont Public Service Board*, Docket No. 4865 (January 1984) and R.S. Harris, *Testimony filed with the Delaware Public Service Commission*, Docket 84–33 (November 1984). In its *Supplemental Notice* [6], the FCC tentatively endorsed substantial reliance on FAF for use in DCF determination of cost of equity.

[6] As stated, Equation (2) requires expectations of either an infinite horizon of dividend growth at rate g or a finite horizon of dividend growth at rate g and special assumptions about the price of the stock at the end of that horizon. Essentially, the assumption must ensure that the stock price grows at a compound rate of g over the finite horizon

tions of future performance. Without a ready source for measuring such expectations, application of the DCF model is fraught with difficulties even if the simple version shown in Equation (2) fits the equity investment in question. This paper uses published FAF of long-run growth in earnings as a proxy for g.

B. Data

Many analysts publish forecasts of corporate earnings. Such forecasts are widely disseminated and are the subject of considerable interest both to investors and researchers (see Givoly and Lakonishok [8]). In recent years, this interest has led to a viable market for services that collect and disseminate such FAF. FAF for this research come from IBES (Institutional Broker's Estimate System), which is a product of Lynch, Jones, and Ryan, a major brokerage firm. Data in IBES represent a compilation of earnings per share (EPS) estimates of about 2000 individual analysts from 100 brokerage firms on over 2000 corporations. IBES data are provided to clients in a number of forms, including on-line data bases provided by vendors. The client base, which currently numbers more than 300, includes most large institutional investors such as pension funds, banks, and insurance companies. Representative of industry practice, IBES contains estimates of (i) EPS for the upcoming fiscal year, (ii) EPS for the subsequent year, and (iii) a projected five-year growth rate in EPS. Each item is available at monthly intervals.

IBES collection procedures are designed to obtain timely forecasts made on a consistent basis. IBES requests "normalized" five-year growth rates from analysts. Such normalization is designed to remove short-term distortions that might stem from using an unusually high or low earnings year as a base. These growth and other earnings forecasts are updated when analysts formally change their stated predictions. IBES does, however, verify prior forecasts monthly to make sure that analysts still hold to them. Despite these procedures, there remain potential difficulties in using IBES data to the extent that some analysts fail to normalize growth projections or fail to continually review and revise their earnings estimates. To control for some of these potential difficulties, this analysis uses averages of analysts' forecasts for a wide range of companies over an extended number of months.

In this research, the mean value of individual analyst's forecasts of five-year growth rate in EPS will be used as a proxy for g in the DCF model.[7] The five-year horizon is the longest horizon over which such forecasts are available from IBES and often is the longest horizon used by analysts. One could make alternate assumptions about growth after five years and use a more general version of a DCF model, but unfortunately, there is no source for obtaining market estimates of this expected growth. As a result, the current analysis applies the five-year growth rate as a proxy for g in Equation (2). Given no objective basis for predicting a change in growth (see Footnote 6), this avoids the introduction of *ad hoc* assumptions about future growth. Importantly, however, the approach is applied to portfolios of stocks rather than to individual securities, since future growth patterns may be expected to have drastic changes for some specific securities. Stock prices were obtained from Chase Econometrics and dividend and other firm-specific information from COMPUSTAT. Interest rates (both government and corporate) were gathered from Federal Reserve Bulletins and from Moody's Bond Record. Exhibit 1 describes key variables used in the study. Data collected cover all dividend paying stocks in the Standard and Poor's 500 stock (SP500) index plus approximately

Exhibit 1. Variable Definitions

k	= equity required rate of return
P_0	= average daily price per share*
D_1	= expected dividend per share measured as current indicated annual dividend from COMPUSTAT multiplied by $(1+g)$†
g	= average financial analysts' forecasts of five-year growth rate in earnings per share (from IBES)
σ_g	= cross-sectional standard deviation of analysts' forecasts of growth in earnings per share (from IBES)
N_g	= number of analysts' forecasts of g (from IBES)
i_{20}	= yield to maturity on 20-year U.S. government obligations. Source: Federal Reserve Bulletin, constant maturity series
i_c	= yield to maturity on long-term corporate bonds: Moody's average
i_u	= yield to maturity on long-term public utility bonds: Moody's average
rp	= equity risk premium calculated as $rp = k - i_{20}$

*In results reported P_0 is the average daily price for a stock from the beginning of the month up to and including the date of publication of monthly IBES data (typically half a month). Almost identical results were found using the average price for the entire month.

†See Footnote 8 at the end of the paper for a discussion of the $(1+g)$ adjustment.

[7]While the model calls for expected growth in dividends, no source of data on such projections is readily available. In addition, in the long run, dividend growth is sustainable only via growth in earnings. As long as payout ratios are not expected to change, the two growth rates will be the same. Vander Weide and Carleton [17] also use the IBES growth rate in earnings per share.

Exhibit 2. Required Rates of Return and Risk Premia

		SP500		SPUT	
	Bond Yield*	Required† Return	Risk‡ Premium	Required† Return	Risk‡ Premium
1982					
Quarter 1	14.27	20.81	6.54	18.83	4.56
Quarter 2	13.74	20.68	6.94	18.51	4.77
Quarter 3	12.94	20.23	7.29	18.55	5.61
Quarter 4	10.72	18.58	7.86	17.20	6.48
Average	12.92	20.08	7.16	18.28	5.36
1983					
Quarter 1	10.87	18.07	7.20	16.71	5.84
Quarter 2	10.80	17.76	6.96	16.52	5.72
Quarter 3	11.79	17.90	6.11	16.39	4.60
Quarter 4	11.90	17.81	5.91	16.00	4.10
Average	11.34	17.88	6.54	16.41	5.07
1984					
Quarter 1	12.09	17.22	5.13	16.48	4.39
Quarter 2	13.21	17.42	4.21	16.99	3.78
Quarter 3	12.83	17.34	4.51	16.62	3.79
Quarter 4	11.78	17.05	5.27	15.18	4.04
Average	12.48	17.26	4.78	16.48	4.00
Average 1982–1984	12.25	18.41	6.16	17.06	4.81

*i_{20} = Yield on U.S. Treasury obligation, 20 year constant maturity.
†Monthly required return (k) calculated as value weighted average. Quarterly values are simple averages of monthly figures.
‡Risk premium calculated as $k - i_{20}$.

[8]The construction of D_1 is controversial since dividends are paid quarterly and may be expected to change during the year; whereas, Equation (2), as is typical, is being applied to annual data. Both the quarterly payment of dividends (due to investors' reinvestment income before year's end, see Linke, and Zumwalt [11]) and any growth during the year require an upward adjustment of the current annual rate of dividends to construct D_1. If quarterly dividends grew at a constant rate, both factors could be accommodated straightforwardly by applying Equation (2) to quarterly data (with a quarterly growth rate) and then annualizing the estimated quarterly required return. Unfortunately, with lumpy changes in dividends, the precise nature of the adjustment depends, on both an individual company's pattern of growth during the calendar year and an individual company's required return (and hence reinvestment income in that risk class).

In this work, D_1 is calculated as $D_0(1 + g)$. The full g adjustment is a crude approximation to adjust for both growth and reinvestment income. For example, if one expected dividends to have been raised, on average, six months ago, a "½ g" adjustment would allow for growth, the remaining "½ g" would be justified on the basis of reinvestment income. Any precise accounting for both reinvestment income and growth would require tracking each company's dividend change history and making explicit judgments about the quarter of the next change. Since no organized "market" forecasts of such a detailed nature exist, such a procedure is not possible. To get a feel for the magnitudes involved, the average dividend yield (D_1/P_0) and growth (market value weighted 1982–1984) for the SP500 were 5.8% and 12.5%. Comparable figures for the SP utility index were 10.4% and 6.7%. As a result, a "full g" adjustment on average increases the required return by 60–70 basis points (relative to no g adjustment) for both indices.

[9]Brigham, Shome, and Vinson [2] also use this interest rate to create equity risk premia. The results were robust to changes in weighting. For the SP500, equal weighting (rather than value weighting) increased the 1982–1984 risk premium by two basis points while for the SPUT equal weighting resulted in a 21 basis point increase. As a further test, the SP500 stocks were ranked on g and the upper and lower deciles deleted. The resulting risk premium (1982–84 average) was 5.94%. A similar procedure used to rank dividend yield produced an SP500 risk premium of 6.18%.

Exhibit 3. Results of Related Studies: Historical Returns and Estimated Risk Premia

	Geometric	Arithmetic
A. Historical Return Realizations (1926–1980)*		
Common Stocks	9.4%	11.7%
Long-Term Government Bonds	3.0%	3.1%
U.S. Treasury Bills	2.8%	2.8%

	Dow Jones Industrials		Dow Jones Electrics	
	Average	Range	Average	Range
B. DCF risk premia using one analyst†				
1966–1970	5.45	4.97–6.81	3.91	3.46–4.13
1971–1975	5.51	4.95–6.92	5.95	4.52–8.72
1976–1980	6.23	5.09–6.88	5.82	5.55–6.21
1981	5.38		5.62	
1982	5.30		3.70	
1983	5.87		5.64	
1984	3.75		4.06	
Average 1982–1984	4.97		4.47	

C. DCF risk premia using three analysts‡	
1981	3.73
1982	4.52
1983	5.17
1984 (through June)	5.01

*Ibbotson, Sinquefield, and Siegel [9].
†Analyst is Value Line. Data are annual estimates using two-stage growth DCF model. Source: Brigham, Shome, and Vinson [2].
‡Analysts are Value Line, Merrill Lynch and Salomon Brothers. Data are averages of monthly values from Brigham, Shome, and Vinson [2].

ly, their work does not include a broad market index directly comparable to the SP500. Rather, they use the Dow Jones Industrial Index based on 30 large industrial concerns. Though the SPUT includes a broader set of utilities than the electrics covered by Brigham *et al.*, their average risk premium estimates are also in the 4 to 5% range for the early 1980s.

While the estimates in Exhibit 2 are quite plausible, the question still remains as to whether they satisfy economic criteria one would expect of risk premia. In the following section, the estimated risk premia are subjected to a series of tests to see if they vary both cross-sectionally and over time with changes in risk. The tests are ultimately joint tests of the estimates as useful risk premia, the measured proxies for risk and the validity of the economic hypothesis. Nonetheless, if the tests using the risk premia have results conforming to theoretical expectation, the comfort level in using them is increased accordingly.

Exhibit 4. Risk Premia by Moody's Bond Ratings*

Electric Utilities: SIC's 4911 and 4931	Aaa	Aa	A	Baa
Risk Premia				
Risk Premium (Expectational g)	3.60	4.33	4.81	4.90
Risk Premium (Historical g†)	6.10	3.28	3.09	5.24
Financial Data				
Debt Ratio‡	0.46	0.48	0.50	0.51
Beta§	0.58	0.61	0.62	0.61
Variability¶				
Operating Cash Flow	0.009	0.016	0.022	0.059
Equity Cash Flow	0.006	0.013	0.019	0.024
Standard Deviation** of Analysts' Forecasts	1.00	1.26	.33	1.79

*Moody's ratings as of January 1984 from *Moody's Bond Record*, February 1984. The number of companies by rating is Aaa (2), Aa (22), A (32), Baa (22). Risk premia are averages of monthly values, January 1982–September 1983.
†Historical Growth is past five-year earnings growth, based on 20 quarters of past data. Source: IBES.
‡Debt Ratio = Long-Term Debt ÷ Total Capital, average 1978–1982 from COMPUSTAT.
§Beta from *Value Line*, January 29, 1982.
¶Measure of variability around trend growth: variance of residuals of regressions on quarterly COMPUSTAT data (1978–1982). Regressions are log of variable regressed on time and seasonal dummies.
**This is the average value of the standard deviation around the mean long-term growth forecast. Such standard deviations are reported for each company in each month. Note it is *not* the cross-sectional standard deviation of growth rates among companies.

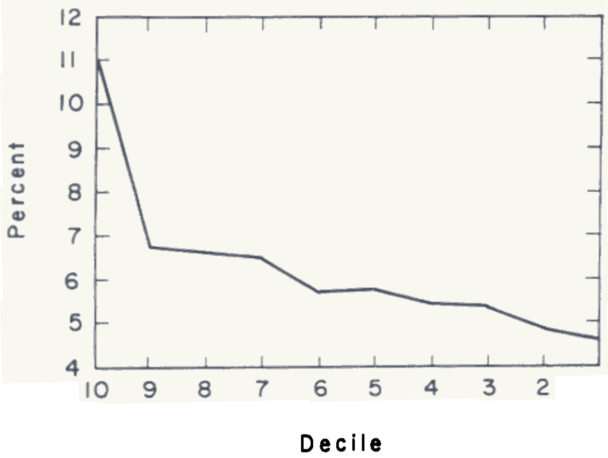
V. Characteristics of Risk Premia

A. Cross-Sectional Tests

Brigham *et al.* show that risk premia (IBES estimates for first half of 1984) for electric utilities are lower the higher the bond rating of the company, confirming the expected tradeoff between risk and return. A similar experiment for electrics, using the current data stretching back to January 1982, confirmed this relationship for a longer time period. Exhibit 4 reports selected results of that analysis. As a contrast, Exhibit 4 also shows the results of using historical growth rates (rather than FAF) in a DCF model. Risk premia derived from historical growth are actually higher for companies with very safe debt, suggesting the clear inferiority of historical to expectational growth rates. With the exception of beta, which is roughly constant across groups, other measures of risk noted in Exhibit 4 confirm the risk differentials associated with bond rating groups.

A further test of the cross-sectional variation in risk premia was performed by dividing the universe of

Exhibit 5. Equity Risk Premia: Deciles Based on Standard Deviation of Financial Analysts Forecasts* (Companies with at least three analysts)



*Risk premia were calculated as equally weighted averages for each decile (10 = highest dispersion) for each of three months: January 1982, December 1982, and September 1983 (approximately 50 companies per decile). These premia were then averaged across deciles. A similar downward pattern was evident in each month.

stocks (industrial plus utility) according to the dispersion of analysts' forecasts, σ_g. This cross-sectional measure of analysts' disagreement should be positively related to the uncertainty of future growth prospects and hence to the riskiness of equity investment. Elsewhere, Malkiel [12] has discussed the rationale and usefulness of such dispersion as an *ex ante* measure of risk. Malkiel argues that σ_g may be a proxy for systematic risk and shows that it bears a closer empirical relationship to expected return than does beta or other risk measures. Most of Malkiel's work is, however, based on data from the 1960s. Exhibit 5 reports risk premia by decile based on σ_g for companies having at least three analysts' forecasts. The three months were chosen as representative. The results show a consistent positive relationship between risk premia and dispersion of analysts' forecasts.

The results in Exhibits 4 and 5 show that the estimated risk premia conform to theoretical relationships between risk and required return that are expected when investors are risk averse. This strengthens the case for using such risk premia, and provides encouragement for further study of their structure.[10]

B. Time Series Tests

A potential benefit of using *ex ante* risk premia is the estimation of changes in risk premia over time. Brigham *et al.* [2] note such changes for utility stocks and relate them to changes in interest rates. They conclude that prior to 1980 utility risk premia increased with the level of interest rates, but that this pattern reversed thereafter, resulting in an inverse correlation between risk premia and interest rates. They explain this turnaround as the outcome of changes in bond markets and adaptation of utilities and their regulators to an inflationary environment. Brigham *et al.* do not, however, analyze changing risk premia for stocks in general. Furthermore, they do not provide direct empirical proxies for changes in equity risks that would explain changes in equity risk premia over time.[11]

C. Changes in Risk Premia

One would expect changes in measured equity risk premia to be related to changes in perceived riskiness. First, with changes in the economy and financial markets, equity investments may be perceived to change in risk. Second, since government bonds are risky investments themselves, their perceived riskiness may change. For example, the large increase in interest rate volatility in the last decade has undoubtedly made fixed income investments more risky holdings than they were in a world of relatively stable rates. Measured equity risk premia (relative to government bonds) could thus be reduced due to increases in perceived riskiness of bonds, even if equities displayed no shifts in risk.

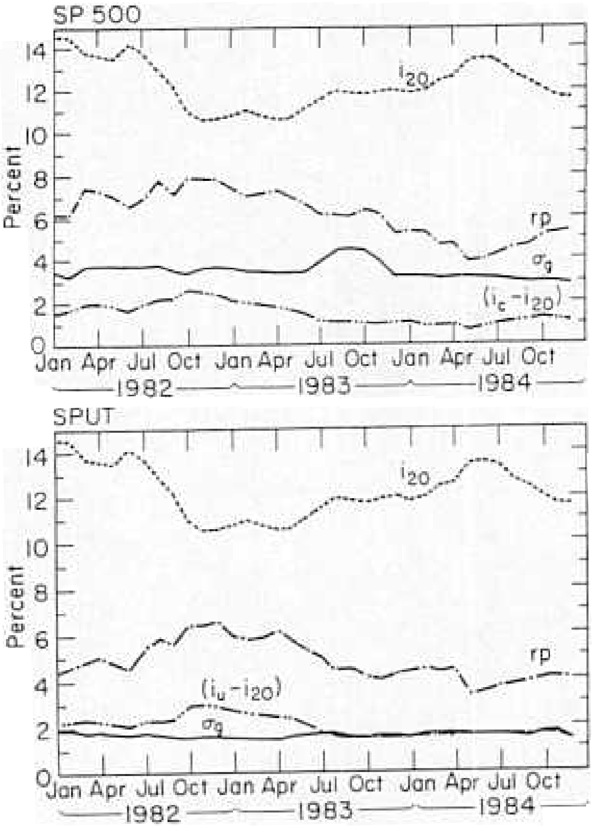
One measure of risk, the standard deviation of FAF, σ_g, was shown previously to be related to cross-sectional differences in risk premia. To test its usefulness as a time series measure of risk, the average value of σ_g was calculated each month for the SP500 index and the SPUT index. The results are graphed in Exhibit 6.[12]

[10]Such *ex ante* required returns offer a useful alternative to *ex post* data typically used in tests of asset pricing models. See Friend, Westerfield, and Granito [7] for a test of the CAPM using survey data rather than *ex post* holding period returns.

[11]In addition, Brigham *et al.* do not report on their treatment of serial correlation in reported regression results, making it more difficult to interpret their findings. As an example, monthly data are used for the 1980–1984 period in a time series regression of a risk premium on the level of interest rates. Similar regressions using data in this paper (1982–1984 monthly data) showed significant positive autocorrelation with Durbin Watson Statistics well below 1.0.

[12]The average values of σ_g are the market value weighted averages of the σ_g for individual stocks. If one looked at a direct estimate of g made by individual analysts for the index, one would expect to find a lower amount of dispersion because some of the differences on individual securities would cancel out. Such data are not available. One would suspect, however, that the calculated average would move up and down in tandem with this unobservable measure of dispersion.

Exhibit 6. Equity Risk Premia, Interest Rates and Risk



Another possible time series proxy for equity risk is the set of yield spreads between corporate and government bonds. As the perceived riskiness of corporate activity increases, the difference between yields on corporate bonds and government bonds should increase. One would expect the sources of increased riskiness to corporate bonds to also increase risks to shareholders.[13] Exhibit 6 graphs two series of yield spreads. The first is the difference between the yield on Moody's corporate average series and the yield on 20-year U.S. Treasury obligations. This series includes debt of both industrial and utility companies and thus would be appropriate as a risk proxy for a broad market index such as the SP500. The second is the spread between the yields on Moody's public utility series and 20-year U.S. Treasury bonds. This series should reflect relative risks of utility stocks as proxied by SPUT.[14]

Exhibit 7 reports results of analyzing the relationship between risk premia, interest rates, and proxies for risk for both the SP500 and SPUT. All regressions are corrected for serial correlation.[15] For stocks in general, Panel A shows that risk premia are negatively related to the level of interest rates — as proxied by i_{20}. Such a negative relationship may result from increases in the perceived riskiness of investment in government debt at high levels of interest rates. A direct measure of uncertainty about investments in government bonds would be necessary to test this hypothesis directly.

The results also show the significant positive relationship between the two proxies for risk and the estimated risk premia. For example, regression 4 of Panel A shows that the equity premium on the SP500 increases with the dispersion of FAF (σ_g) and the yield spread between corporate and government bonds ($i_c - i_{20}$). Evidently, these two risk measures capture somewhat different dimensions of risk, both of which appear important in explaining risk premia on stocks in general. The simple correlation coefficient between the two risk measures is 0.19 and is insignificantly different from zero. The addition of the yield spread risk proxy also dramatically lowers the magnitude of the coefficient on government bond yields, as can be seen by comparing Equations 1 and 3 of Panel A. Apparently, a large part of the effect of changes in government bond rates on equity risk premia may be explained through the narrowing of the yield spread between corporate and government bonds. This suggests that such increases in government yields may often be associated with a reduction in the *difference* in risk between investment in government bonds and in corporate activity.

Panel B shows that utility risk premia are also inversely related to the level of interest rates as was found by Brigham *et al.* [2]. Unlike the results for stocks in general, however, changes in the dispersion of FAF over time are not significantly related to changes in these utility risk premia. This may be be-

[13]Of course, counterexamples could be constructed but one would expect an overall positive correlation across companies. Additionally, the cross-sectional relationship between bond ratings and equity risk premia reported earlier in the paper supports the link between corporate debt risks and risks on equity.

[14]Note that these two series reflect both changes in the ratings of corporate bonds as well as yield spreads for a given bond rating. The two series proved better in explaining equity risk premia than use of two comparable series for AA-rated debt.

[15]Ordinary least squares regressions showed severe positive autocorrelation in many cases with Durbin Watson Statistics typically below one. Estimation used the Prais-Winsten method. See Johnston [10], pp. 321–325.

Exhibit 7. Changes in Equity Risk Premia Over Time — Entries are Coefficient (t-value)

Regression	Intercept	i_{20}	σ_g	$i_c - i_{20}$	R^2
A. SP500: Dependent Variable is Equity Risk Premium*					
1.	0.140	−0.632			
	(8.15)†	(−4.95)†			
2.	0.118	−0.660	0.754		
	(7.10)†	(−5.93)†	(3.32)†		
3.	0.069	−0.235		1.448	
	(3.44)†	(−1.76)		(4.18)†	
4.	0.030	−0.177	0.855	1.645	0.79
	(2.17)†	(−2.07)†	(4.68)†	(7.63)†	

Regression	Intercept	i_{20}	σ_g	$i_u - i_{20}$	R^2
B. SPUT: Dependent Variable is Equity Risk Premium*					
1.	0.110	−0.510			0.37
	(7.35)†	(−4.41)†			
2.	0.101	−0.543	0.805		0.41
	(6.28)†	(−4.68)†	(1.42)		
3.	0.051	−0.259		1.432	0.80
	(5.54)†	(−4.05)†		(8.87)†	
4.	0.049	−0.287	0.387	1.391	0.80
	(5.15)†	(−3.87)†	(0.75)	(8.14)†	

*All variables are defined in Exhibit 1 and graphed in Exhibit 6. Regressions were estimated for the 36 month period January 1982–December 1984 and were corrected for serial correlation using the Prais-Winsten method. For purposes of this regression variables are expressed in decimal form, *e.g.*, 14% = 0.14.

†Significantly different from zero at 0.05 level using two-tailed test.

cause of lower variability over time in the dispersion of FAF for utility stocks as compared to equities in general. The yield spread between utility and government bonds is significantly positively related to utility equity risk premia. And, as in the case of stocks in general, introduction of this spread substantially reduces the independent effect of interest rate levels on equity risk premia.

Given the short time series (36 months), tests for the stability of the relationships found in Exhibit 7 present difficulties. As a check, the relationships were reestimated dividing the data into two 18-month periods. For stocks in general (SP500), coefficients on σ_g and $(i_c - i_{20})$ were positive in all regressions and significantly so, except in the case of $(i_c - i_{20})$ for the second 18-month period. The coefficient of i_{20} was significantly negative in both periods. This confirms the general findings for the SP500 in Panel A of Exhibit 7. For utility stocks, results for the subperiods also matched the entire period results. The coefficients of $(i_u - i_{20})$ were significantly positive in both subperiods while those of σ_g were insignificantly different from zero. The level of interest rates (i_{20}) had a significant negative effect in both subperiods.

In summary, the estimated risk premia change over time and the patterns of such change are directly related to changes in proxies for the risks of equity investments. Risk premia for both stocks in general and utilities are inversely related to the level of government interest rates but positively related to the bond yield spreads which proxy for the incremental risk of investing in equities rather than government bonds. For stocks in general, risk premia also increase over time with increases in the general level of disagreement about future corporate performance.

VI. Conclusions

Notions of shareholder required rates of return and risk premia are based in theory on investors' expectations about the future. Research has demonstrated the usefulness of financial analysts' forecasts for such expectations. When such forecasts are used to derive equity risk premia, the results are quite encouraging. In addition to meeting the theoretical requirement of using expectational data, the procedure produces estimates of reasonable magnitude that behave as economic theory would predict. Both over time and across stocks, the risk premia vary directly with the perceived riskiness of equity investment.

The approach offers a straightforward and powerful aid in establishing required rates of return either for corporate investment decisions or in the regulatory arena. Since data are readily available on a wide range of equities, an investigator can analyze various proxy groups (*e.g.*, portfolios of utility stocks) appropriate for a particular decision. An additional advantage of the estimated risk premia is that they allow analysis of changes in equity return requirements over time. Tracking such changes is important for managers facing changing economic climates.

References

1. R. Brealey and S. Myers, *Principles of Corporate Finance* 2nd edition, New York, McGraw-Hill, 1984.
2. E. Brigham, D. Shome, and S. Vinson, "The Risk Premium Approach to Measuring a Utility's Cost of Equity," *Financial Management* (Spring 1985), pp. 33–45.
3. W. T. Carleton and J. Lakonishok, "Risk and Return on Equity: The Use and Misuse of Historical Estimates," *Financial Analysts Journal* (January/February 1985).
4. J. Cragg and B. G. Malkiel, *Expectations and the Structure of Share Prices*, National Bureau of Economic Research, Chicago, University of Chicago Press, 1982.
5. E. J. Elton, M. J. Gruber, and M. Gultekin, "Expectations and Share Prices," *Management Science* (September 1981).
6. Federal Communications Commission, *Notice of Proposed Rulemaking*, FCC Docket No. 84–800 (August 1984), and *Supplemental Notice* (August 1985).
7. I. Friend, R. Westerfield, and M. Granito, "New Evidence on the Capital Asset Pricing Model," *Journal of Finance* (June 1978), pp. 903–917.
8. D. Givoly and J. Lakonishok, "Earnings Expectation and Properties of Earnings Forecasts — A Review and Analysis of the Research," *Journal of Accounting Literature*, Vol. 3 (Spring 1984), pp. 85–107.
9. R. Ibbotson, R. A. Sinquefield, and L. B. Siegel, "Historical Returns on Principal Types of Investments," University of Chicago, CRSP Working Paper #71, February 1982.
10. J. Johnston, *Econometric Methods*, 3rd edition, New York, McGraw-Hill, 1984.
11. C. Linke and J. Zumwalt, "Estimation Biases in Discounted Cash Flow Analyses of Equity Capital Cost in Rate Regulation," *Financial Management* (Autumn 1984), pp. 15–21.
12. B. Malkiel, "Risk and Return: A New Look," in B. Friedman, ed., *The Changing Role of Debt and Equity in Financing U.S. Capital Formation*, National Bureau of Economic Research, Chicago, University of Chicago Press, 1982.
13. ——— "The Capital Formation Problem in the United States," *Journal of Finance* (May 1979), pp. 291–306.
14. J. Mayshar, "On Divergence of Opinion and Imperfections in Capital Markets," *American Economic Review* (March 1983), pp. 114–128.
15. K. Stanley, W. Lewellen, and G. Schlarbaum, "Further Evidence on the Value of Professional Investment Research," *Journal of Financial Research* (Spring 1981), pp. 1–9.
16. Touche, Ross and Company, *Proxy Disclosures and Stockholder Attitude Survey*, Washington, D.C., National Association of Corporate Directors, May 1982.
17. J. Vander Weide and W.T. Carleton, "Investor Growth Expectations and Stock Prices," Working paper, December 1985, Duke University and University of Arizona.

James H. Vander Weide and William T. Carleton
"Investor Growth Expectations: Analysts vs. History"
The Journal of Portfolio Management
Spring 1988

Investor growth expectations: Analysts vs. history

Analysts' growth forecasts dominate past trends in predicting stock prices.

James H. Vander Weide and Willard T. Carleton

For the purposes of implementing the Discounted Cash Flow (DCF) cost of equity model, th[redacted] analyst must know which growth estimate is embodied [redacted] price [redacted] and M[redacted] suggests that the stock valuation process embodies analysts' forecasts rather than historic[redacted] ten-year historical growth in dividends per share or the five-year growth in book value per share. The Cragg and Malkiel study is based on data for the 1960s, however, a deca[redacted]ly more stable than the recent [redacted]

As the issue of which growth rate to use in implementing the DCF model is so important to applications of the model, we decided to investigate whether the Cragg and Malkiel conclusions continue t[redacted]iods. This paper describes the results of our study.

STATISTICAL MODEL

The DCF model suggests that the firm's stock price is equal to the present value of the stream of dividends that investors expect to receive from owning the firm's shares. Under the assumption that investors expect dividends to grow at a constant rate, g, in perpetuity, the stock price is given by the following simple expression:

$$P_S = \frac{D(1+g)}{k-g} \tag{1}$$

where:

P_S = current price per share of the firm's stock;

D = current annual dividend per share;

g = expected constant dividend growth rate; and

k = required return on the firm's stock.

Dividing both sides of Equation (1) by the firm's current earnings, E, we obtain:

$$\frac{P_S}{E} = \frac{D}{E} \cdot \frac{(1+g)}{k-g} \tag{2}$$

Thus, the firm's price/earnings (P/E) ratio is a nonlinear function of the firm's dividend payout ratio (D/E), the expected growth in dividends (g), and the required rate of return.

To investigate what growth expectation is embodied in the firm's current stock price, it is more convenient to work with a linear approximation to Equation (2). Thus, we will assume that:

$$P/E = a_0(D/E) + a_1 g + a_2 k. \tag{3}$$

(Cragg and Malkiel found this assumption to be reasonable throughout their investigation.)

Furthermore, we will assume that the required

JAMES H. VANDER WEIDE is Research Professor at the Fuqua School of Business at Duke University in Durham (NC 27706). WILLARD T. CARLETON is Karl Eller Professor of Finance at the University of Arizona in Tucson (AZ 85721).
Financial support for this project was provided by BellSouth and Pacific Telesis. The authors wish to thank Paul Blalock at BellSouth, Mohan Gyani at Pacific Telesis, Bill Keck at Southern Bell, and John Carlson, their programmer, for help with this project.

rate of return, k, in Equation (3) depends on the values of the risk variables B, Cov, Rsq, and Sa, where B is the firm's Value Line beta; Cov is the firm's pretax interest coverage ratio; Rsq is a measure of the stability of the firm's five-year historical EPS; and Sa is the standard deviation of the consensus analysts' five-year EPS growth forecast for the firm. Finally, as the linear form of the P/E equation is only an approximation to the true P/E equation, and B, Cov, Rsq, and Sa are only proxies for k, we will add an error term, e, that represents the degree of approximation to the true relationship.

With these assumptions, the final form of our P/E equation is as follows:

$$P/E = a_0(D/E) + a_1 g + a_2 B + a_3 Cov + a_4 Rsq + a_5 Sa + e. \quad (4)$$

The purpose of our study is to use more recent data to determine which of the popular approaches for estimating future growth in the Discounted Cash Flow model is embodied in the market price of the firm's shares.

We estimated Equation (4) to determine which estimate of future growth, g, when combined with the payout ratio, D/E, and risk variables B, Cov, Rsq, and Sa, provides the best predictor of the firm's P/E ratio. To paraphrase Cragg and Malkiel, we would expect that growth estimates found in the best-fitting equation more closely approximate the expectation used by investors than those found in poorer-fitting equations.

DESCRIPTION OF DATA

Our data sets include both historically based measures of future growth and the consensus analysts' forecasts of five-year earnings growth supplied by the Institutional Brokers Estimate System of Lynch, Jones & Ryan (IBES). The data also include the firm's dividend payout ratio and various measures of the firm's risk. We include the latter items in the regression, along with earnings growth, to account for other variables that may affect the firm's stock price.

The data include:

Earnings Per Share. Because our goal is to determine which earnings variable is embodied in the firm's market price, we need to define this variable with care. Financial analysts who study a firm's financial results in detail generally prefer to "normalize" the firm's reported earnings for the effect of extraordinary items, such as write-offs of discontinued operations, or mergers and acquisitions. They also attempt, to the extent possible, to state earnings for different firms using a common set of accounting conventions.

We have defined "earnings" as the consensus analyst estimate (as reported by IBES) of the firm's earnings for the forthcoming year.[1] This definition approximates the normalized earnings that investors most likely have in mind when they make stock purchase and sell decisions. It implicitly incorporates the analysts' adjustments for differences in accounting treatment among firms and the effects of the business cycle on each firm's results of operations. Although we thought at first that this earnings estimate might be highly correlated with the analysts' five-year earnings growth forecasts, that was not the case. Thus, we avoided a potential spurious correlation problem.

Price/Earnings Ratio. Corresponding to our definition of "earnings," the price/earnings ratio (P/E) is calculated as the closing stock price for the year divided by the consensus analyst earnings forecast for the forthcoming fiscal year.

Dividends. Dividends per share represent the common dividends declared per share during the calendar year, after adjustment for all stock splits and stock dividends). The firm's dividend payout ratio is then defined as common dividends per share divided by the consensus analyst estimate of the earnings per share for the forthcoming calendar year (D/E). Although this definition has the deficiency that it is obviously biased downward — it divides this year's dividend by next year's earnings — it has the advantage that it implicitly uses a "normalized" figure for earnings. We believe that this advantage outweighs the deficiency, especially when one considers the flaws of the apparent alternatives. Furthermore, we have verified that the results are insensitive to reasonable alternative definitions (see footnote 1).

Growth. In comparing historically based and consensus analysts' forecasts, we calculated forty-one different historical growth measures. These included the following: 1) the past growth rate in EPS as determined by a log-linear least squares regression for the latest year,[2] two years, three years, . . ., and ten years; 2) the past growth rate in DPS for the latest year, two years, three years, . . ., and ten years; 3) the past growth rate in book value per share (computed as the ratio of common equity to the outstanding common equity shares) for the latest year, two years, three years, . . ., and ten years; 4) the past growth rate in cash flow per share (computed as the ratio of pretax income, depreciation, and deferred taxes to the outstanding common equity shares) for the latest year, two years, three years, . . ., and ten years; and 5) plowback growth (computed as the firm's retention ratio for the current year times the firm's latest annual return on common equity).

We also used the five-year forecast of earnings per share growth compiled by IBES and reported in mid-January of each year. This number represents the consensus (i.e., mean) forecast produced by analysts from the research departments of leading Wall Street and regional brokerage firms over the preceding three months. IBES selects the contributing brokers "because of the superior quality of their research, professional reputation, and client demand" (IBES *Monthly Summary Book*).

Risk Variables. Although many risk factors could potentially affect the firm's stock price, most of these factors are highly correlated with one another. As shown above in Equation (4), we decided to restrict our attention to four risk measures that have intuitive appeal and are followed by many financial analysts: 1) B, the firm's beta as published by Value Line; 2) Cov, the firm's pretax interest coverage ratio (obtained from Standard & Poor's Compustat); 3) Rsq, the stability of the firm's five-year historical EPS (measured by the R^2 from a log-linear least squares regression); and 4) Sa, the standard deviation of the consensus analysts' five-year EPS growth forecast (mean forecast) as computed by IBES.

After careful analysis of the data used in our study, we felt that we could obtain more meaningful results by imposing six restrictions on the companies included in our study:

1. Because of the need to calculate ten-year historical growth rates, and because we studied three different time periods, 1981, 1982, and 1983, our study requires data for the thirteen-year period 1971-1983. We included only companies with at least a thirteen-year operating history in our study.
2. As our historical growth rate calculations were based on log-linear regressions, and the logarithm of a negative number is not defined, we excluded all companies that experienced negative EPS during any of the years 1971-1983.
3. For similar reasons, we also eliminated companies that did not pay a dividend during any one of the years 1971-1983.
4. To insure comparability of time periods covered by each consensus earnings figure in the P/E ratios, we eliminated all companies that did not have a December 31 fiscal year-end.
5. To eliminate distortions caused by highly unusual events that distort current earnings but not expected future earnings, and thus the firm's price/earnings ratio, we eliminated any firm with a price/earnings ratio greater than 50.
6. As the evaluation of analysts' forecasts is a major part of this study, we eliminated all firms that IBES did not follow.

Our final sample consisted of approximately sixty-five utility firms.[3]

RESULTS

To keep the number of calculations in our study to a reasonable level, we performed the study in two stages. In Stage 1, all forty-one historically oriented approaches for estimating future growth were correlated with each firm's P/E ratio. In Stage 2, the historical growth rate with the highest correlation to the P/E ratio was compared to the consensus analyst growth rate in the multiple regression model described by Equation (4) above. We performed our regressions for each of three recent time periods, because we felt the results of our study might vary over time.

First-Stage Correlation Study

Table 1 gives the results of our first-stage correlation study for each group of companies in each of the years 1981, 1982, and 1983. The values in this table measure the correlation between the historically oriented growth rates for the various time periods and the firm's end-of-year P/E ratio.

The four variables for which historical growth rates were calculated are shown in the left-hand column: EPS indicates historical earnings per share growth, DPS indicates historical dividend per share growth, BVPS indicates historical book value per share growth, and CFPS indicates historical cash flow per share growth. The term "plowback" refers to the product of the firm's retention ratio in the currennt year and its return on book equity for that year. In all, we calculated forty-one historically oriented growth rates for each group of firms in each study period.

The goal of the first-stage correlation analysis was to determine which historically oriented growth rate is most highly correlated with each group's year-end P/E ratio. Eight-year growth in CFPS has the highest correlation with P/E in 1981 and 1982, and ten-year growth in CFPS has the highest correlation with year-end P/E in 1983. In all cases, the plowback estimate of future growth performed poorly, indicating that — contrary to generally held views — plowback is not a factor in investor expectations of future growth.

Second-Stage Regression Study

In the second stage of our regression study, we ran the regression in Equation (4) using two different measures of future growth, g: 1) the best historically oriented growth rate (g_h) from the first-stage correlation study, and 2) the consensus analysts' forecast (g_a) of five-year EPS growth. The regression results, which are shown in Table 2, support at least

TABLE 1

Correlation Coefficients of All Historically Based Growth Estimates by Group and by Year with P/E

Historical Growth Rate Period in Years

Current Year	1	2	3	4	5		7	8	9	10
1981										
EPS	−0.02	0.07	0.03	0.01	0.03	0.12	0.08	0.09	0.09	0.09
DPS	0.05	0.18	0.14	0.15	0.14	0.15	0.19	0.23	0.23	0.23
BVPS	0.01	0.11	0.13	0.13	0.16	0.18	0.15	0.15	0.15	0.15
CFPS	−0.05	0.04	0.13	0.22	0.28	0.31	0.30	0.31	−0.57	−0.54
Plowback	0.19									
1982										
EPS	−0.10	−0.13	−0.06	−0.02	−0.02	−0.01	−0.03	−0.03	0.00	0.00
DPS	−0.19	−0.10	0.03	0.05	0.07	0.08	0.09	0.11	0.13	0.13
BVPS	0.07	0.08	0.11	0.11	0.09	0.10	0.11	0.11	0.09	0.09
CFPS	−0.02	−0.08	0.00	0.10	0.16	0.19	0.23	0.25	0.24	0.07
Plowback	0.04									
EPS	−0.06	−0.25	−0.25	−0.24	−0.16	−0.11	−0.05	0.00	0.02	0.02
DPS	0.03	−0.10	−0.03	0.08	0.15	0.21	0.21	0.21	0.22	0.24
BVPS	0.03	0.10	0.04	0.09	0.15	0.16	0.19	0.21	0.22	0.21
CFPS	−0.08	0.01	0.02	0.08	0.20	0.29	0.35	0.38	0.40	0.42
Plowback	−0.08									

two general conclusions regarding the pricing of equity securities.

[redacted]ce. In every case, the R^2 in the regression containing the consensus analysts' forecast is higher than the R^2 in the regression containing the historical growth measure. The regression coefficients in the equation containing the consensus analysts' forecast also are considerably more significant than they are in the alternative regression. These results are consistent with those found by Cragg and Malkiel for data covering the period 1961-1968. Our results also are consistent with the hypothesis that investors use analysts' forecasts, rather than historically oriented growth calculations, in making stock buy–and–sell decisions.

TABLE 2

Regression Results
Model I

Part A: *Historical*

$$P/E = a_0 + a_1 D/E + a_2 g_h + a_3 B + a_4 Cov + a_5 Rsq + a_6 Sa$$

Year	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	R^2	F Ratio
1981	−6.42*	10.31*	7.67*	3.24	0.54*	1.42*	57.43	0.83	46.49
	(5.50)	(14.79)	(2.20)	(2.86)	(2.50)	(2.85)	(4.07)		
1982	−2.90*	9.32*	8.49*	2.85	0.45*	−0.42	3.63	0.86	65.53
	(2.75)	(18.52)	(4.18)	(2.83)	(2.60)	(0.05)	(0.26)		
	−5.96*	10.20*	19.78*	4.85	0.44*	0.33	32.49	0.82	45.26
	(3.70)	(12.20)	(4.83)	(2.95)	(1.89)	(0.50)	(1.29)		

Part B: *Analysis*

$$P/E = a_0 + a_1 D/E + a_2 g_a + a_3 B + a_4 Cov + a_5 Rsq + a_6 Sa$$

Year	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{a}_6$	R^2	F Ratio
1981	−4.97*	10.62*	54.85*	−0.61	0.33*	0.63*	4.34	0.91	103.10
	(6.23)	(21.57)	(8.56)	(0.68)	(2.28)	(1.74)	(0.37)		
1982	−2.16*	9.47*	50.71*	−1.07	0.36*	−0.31	119.05*	0.90	97.62
	(2.59)	(22.46)	(9.31)	(1.14)	(2.53)	(1.09)	(1.60)		
1983	−8.47*	11.96*	79.05*	2.16	0.56*	0.20	−34.43	0.87	69.81
	(7.07)	(16.48)	(7.84)	(1.55)	(3.08)	(0.38)	(1.44)		

Notes:
* Coefficient is significant at the 5% level (using a one-tailed test) and has the correct sign. T-statistic in parentheses.

Second, there is some evidence that investors tend to view risk in traditional terms. The interest coverage variable is statistically significant in all but one of our samples, and the stability of the operating income variable is statistically significant in six of the twelve samples we studied. On the other hand, the beta is never statistically significant, and the standard deviation of the analysts' five-year growth forecasts is statistically significant in only two of our twelve samples. This evidence is far from conclusive, however, because, as we demonstrate later, a significant degree of cross-correlation among our four risk variables makes any general inference about risk extremely hazardous.

Possible Misspecification of Risk

The stock valuation theory says nothing about which risk variables are most important to investors. Therefore, we need to consider the possibility that the risk variables of our study are only proxies for the "true" risk variables used by investors. The inclusion of proxy variables may increase the variance of the parameters of most concern, which in this case are the coefficients of the growth variables.[4]

To allow for the possibility that the use of risk proxies has caused us to draw incorrect conclusions concerning the relative importance of analysts' growth forecasts and historical growth extrapolations, we have also estimated Equation (4) with the risk variables excluded. The results of these regressions are shown in Table 3.

Again there is overwhelming evidence that the consensus analysts' growth forecast is superior to the historically oriented growth measures in predicting the firm's stock price. The R^2 and t-statistics are higher in every case.

TABLE 3

Regression Results
Model II

Part A: *Historical*

$P/E = a_0 + a_1 D/E + a_2 g_h$

Year	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	R^2	F Ratio
1981	−1.05	9.59	21.20	0.73	82.95
	(1.61)	(12.13)	(7.05)		
1982	0.54	8.92	12.18	0.83	167.97
	(1.38)	(17.73)	(6.95)		
1983	−0.75	8.92	12.18	0.77	107.82
	(1.13)	(12.38)	(7.94)		

Part B: *Analysis*

$P/E + a_0 + a_1 D/E + a_2 g_a$

Year	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	R^2	F Ratio
1981	3.96	10.07	60.53	0.90	274.16
	(8.31)	(8.31)	(20.91)	(15.79)	
1982	−1.75	9.19	44.92	0.88	246.36
	(4.00)	(4.00)	(21.35)	(11.06)	
1983	−4.97	10.95	82.02	0.83	168.28
	(6.93)	(6.93)	(15.93)	(11.02)	

Notes:
* Coefficient is significant at the 5% level (using a one-tailed test) and has the correct sign. T-statistic in parentheses.

CONCLUSION

The relationship between growth expectations and share prices is important in several major areas of finance. The data base of analysts' growth forecasts collected by Lynch, Jones & Ryan provides a unique opportunity to test the hypothesis that investors rely more heavily on analysts' growth forecasts than on historical growth extrapolations in making security buy–and–sell decisions. With the help of this data base, our studies affirm the superiority of analysts' forecasts over simple historical growth extrapolations in the stock price formation process. Indirectly, this finding lends support to the use of valuation models whose input includes expected growth rates.

We also tried several other definitions of "earnings," including the firm's most recent primary earnings per share prior to any extraordinary items or discontinued operations. As our results were insensitive to reasonable alternative definitions of "earnings" we report only the results for the IBES consensus.

[2] For the latest year, we actually employed a point-to-point growth calculation because there were only two available observations.

[3] We use the word "approximately," because the set of available firms varied each year. In any case, the number varied only from zero to three firms on either side of the figures cited here.

[4] See Maddala (1977).

REFERENCES

Bower, R. S., and D. H. Bower. "Risk and the Valuation of Common Stock." *Journal of Political Economy*, May-June 1969, pp. 349-362.

Cragg, J. G., and Malkiel, B. G. "The Consensus and Accuracy of Some Predictions of the Growth of Corporate Earnings." *Journal of Finance*, March 1968, pp. 67-84.

Cragg, J. G., and Malkiel, B. G. *Expectations and the Structure of Share Prices*. Chicago: University of Chicago Press, 1982.

Elton, E. J., M. J. Gruber, and Mustava N. Gultekin. "Expectations and Share Prices." *Management Science*, September 1981, pp. 975-987.

Federal Communications Commission. *Notice of Proposed Rulemaking*. CC Docket No. 84-800, August 13, 1984.

IBES *Monthly Summary Book*. New York: Lynch, Jones & Ryan, various issues.

Maddala, G. E. *Econometrics*. New York: McGraw-Hill Book Company, 1977.

Malkiel, B. G. "The Valuation of Public Utility Equities." *Bell Journal of Economics and Management Science*, Spring 1970, pp. 143-160.

Peterson, D., and P. Peterson. "The Effect of Changing Expectations upon Stock Returns." *Journal of Financial and Quantitative Analysis*, September 1982, pp. 799-813.

Theil, H. *Principles of Econometrics*. New York: John Wiley & Sons, 1971.

David Gordon,Myron Gordon and Lawrence Gould
“Choice Among Methods of Estimating Share Yield”
The Journal of Portfolio Management
Spring 1989

Choice among methods of estimating share yield

The search for the growth component in the discounted cash flow model.

David A. Gordon, Myron J. Gordon, and Lawrence I. Gould

The yield at which a share of stock is selling, also called its expected return or required return, is an important statistic in finance. Firms use it in choosing among investment opportunities and financing alternatives, and investors use it in making portfolio decisions. Nevertheless, the yield at which a share is selling is a difficult quantity to measure, which has limited its use in the practice of finance. This paper develops and tests a basis for choice among alternative methods of estimating a share's yield.

A share's yield, like a bond's yield, is the discount rate that equates its expected future payments with its current price. A bond's yield is easy to measure under the common practice of ignoring default risk, as the future payments are then known with certainty. The future payments on a share, however, are dividends and market price, and these payments are uncertain.

The common practice is to represent these future dividend payments with estimates of two numbers: One is the coming dividend, and the other is a growth rate. The latter can be an estimate of the long-run growth rate in the dividend or of the growth rate in price over the coming period. In the latter case, the estimate is called the expected holding-period return (EHPR); in the former case, it is called the discounted cash flow yield (DCFY).[1] In either case, the estimate of a share's yield reduces to the sum of its dividend yield and a future growth rate, with the latter inferred in some way from historical data.

There is a wide variety of acceptable methods for using historical data to estimate future growth. This variation in method is illustrated in the testimony of expert witnesses before public utility commissions on the fair return for a public utility. In these cases, the estimates and the methods used are a matter of public record. Some idea of the various methods can be found in Morin (1984) and Kolbe, Read, and Hall (1984). The performance of alternative estimating methods has been examined in Gordon (1974), Kolbe, Read, and Hall (1984), Brigham, Shome, and Vinson (1985), and Harris (1986).

We have derived our basis for comparing the accuracy of alternative methods for estimating the DCFY on a share from the generally accepted propositions that yield should vary according to risk, and that beta is the best estimate of risk. Hence, the DCFY should vary among shares with beta, and, between two methods for estimating growth, the superior method is the one for which the variation in yield among shares is explained better by the variation in beta among the shares.

First we present simple, plausible, and objective measurement rules for implementing four popular and/or attractive methods for estimating the DCFY. We then describe how sample statistics may be used to judge the accuracy of each method. We also describe how the CAPM model has been used to estimate share yield and explain why we do not compare it with the various DCFY methods. The following section carries out the comparison with samples of utility and industrial shares, and the last section pre-

DAVID A. GORDON is in charge of transaction finance at Scotia McLeod, a subsidiary of the Bank of Nova Scotia in Toronto. MYRON J. GORDON is Professor of Finance at the Faculty of Management at the University of Toronto (Ontario M5S 1V4). LAWRENCE I. GOULD is Professor and Head of Accounting and Finance at the University of Manitoba in Winnipeg (Manitoba R3T 2N2).

sents the conclusions that may be drawn from the findings.

ALTERNATIVE MEASUREMENT RULES FOR A SHARE'S YIELD

Under the DCF method or model for estimating the expected return on a stock, the yield for the jth stock is:

$$DCFY_{jt} = DYD_{jt} + GR_{jt} \tag{1}$$

where

$DCFY_{jt}$ = DCF yield on the jth stock at time t,

DYD_{jt} = dividend yield on the jth stock at time t, and

GR_{jt} = long-run growth rate in the dividend on the jth stock that investors expect at time

In what follows, we omit the time and firm subscripts on the variables when they are not required. Also, DCFY will refer to the unknown true yield on a share.

The difficult problem in arriving at the DCFY is estimation of the long-run growth rate that investors expect. Four estimates of that quantity are:

EGR = rate of growth in earnings per share over a prior time period, usually the last five years;

DGR = rate of growth in dividend per share over a prior time period, usually the last five years;

FRG = consensus among security analyst forecasts of the growth rate in earnings, over the next five years; and

BRG = an average over the prior five years of the product of the retention rate b and rate of return on common equity r on a stock.

The estimate of share yield that incorporates each of these estimates of growth is denoted KEGR, KDGR, KFRG, and KBRG, respectively.

A case can be made for each of the four methods for estimating growth. KEGR, KDGR, and KBRG have been widely used in public utility testimony and in research on stock valuation models. The rationale for KEGR is the belief that the past growth rate in earnings is the best predictor of future growth in earnings and dividends. The rationale for KDGR is that the future growth rate in dividends is the statistic we want to estimate, and the past dividend record is free of the noise in past earnings.[2] The rationale for KBRG is that all variables will grow at this rate if the firm earns r and retains b. Furthermore, as Gordon and Gould (1980) show, KEGR and KDGR will be biased one direction or another if r and b have changed over the last five years. As for KFRG, security analysts are professionals employed to forecast future performance; their forecasts are widely accepted by investors. The IBES collection of forecast growth rates of security analysts compiled by Lynch, Jones, and Ryan has increased the popularity of this estimate.

As stated earlier, we may also take the yield on a share as the sum of the dividend yield and the expected rate of growth in price over the coming period. This estimate of a share's yield is widely used in testing the CAPM, with the average HPR over the prior five years commonly used in such empirical work. On the other hand, this estimate of a share's yield varies so widely among firms and over time as to be patently in error as an estimate of share yield.[3]

BASIS OF COMPARISON

To compare the accuracy of the four estimates of the DCFY stated above, we regress the data under each estimate on beta for a sample of shares. If KEGR is the estimate,

$$KEGR_j = \alpha_0 + \alpha_1 BETA_j + \epsilon_j. \tag{2}$$

The rationale for this expression lies in the risk premium theory of share yield, where the share yield is equal to the interest rate plus a risk premium that varies with the share's relative risk. Hence, if BETA is an error-free index of relative risk, α_0 is equal to the interest rate, and α_1 is the risk premium on the market portfolio or standard share.[4]

The higher the correlation between KEGR and BETA, assuming that α_1 is positive, the greater the confidence we may have in KEGR as an estimate of DCFY. We cannot rely solely on the correlation, though, in selecting among the methods for estimating DCFY. Errors in KEGR as a basis for estimating the DCFY on the jth share have random and systematic components. The former is ϵ_j, and its average value can be taken as the root mean square error of the regression (MSE). The larger the root MSE of the regression, the less attractive KEGR is as an estimate of share yield, because the error makes the problem of choice between $KEGR_j$ and $KEGR_j - \epsilon_j$ more acute. (That problem will be discussed shortly.)

The systematic error is the difference between the unknown true yield on the jth share, $DCFY_j$, and the value predicted by Equation (2). There is no obvious measure of the systematic error, as we do not know $DCFY_j$, but sample values of α_0 may provide information on its average value. The difference between α_0 and the interest rate is an indicator of systematic error, because the difference is zero under the risk premium theory. Error in the measurement of BETA biases α_0 upward, but, with the same BETA for each share used in all four regressions, differences in α_0 are indicators of systematic error.[5]

In addition to regression statistics, the sample mean and standard deviation of KEGR is a source of information on its accuracy as a method for the estimation of DCFY. If the mean departs radically from the long-term bond rate, or if the standard deviation indicates an unreasonable range of variation among shares, the accuracy of the method is open to question. Also, the sample mean may be a source of information on the systematic error for a method of estimation. Hence, sample values for the mean, standard deviation, correlation, root MSE, and constant term all contribute to a judgment on a method's accuracy for estimating the DCFY on a share. Unfortunately, there is no simple criterion for choice among the alternatives.

Once a conclusion is reached on the most accurate method for estimating DCFY — say, KEGR — we then have the problem of choice between $KEGR_j$ and $KEGR_j - \epsilon_j$ for the jth share. If the random error in $KEGR_j$ is due to error in its measurement for the jth share, we simply use the value predicted by Equation (2), which is $KEGR_j - \epsilon_j$. On the other hand, KEGR and DCFY may vary among shares with other (omitted) variables as well as BETA, in which case ϵ_j is also due to the omitted variables, and $KEGR_j$ may be the better estimate of DCFY. Unfortunately, we have no basis for choice among these two hypotheses, and the smaller the root MSE the less troublesome the problem of choice between them.

A more favorable tax treatment of capital gains over dividends should make investors prefer capital gains to dividends. As Brennan (1973) has shown, the yield investors require on a share would then vary with the excess of its dividend yield over the interest rate. To recognize this, Equation (2) becomes

$$KEGR_j = \alpha_0 + \alpha_1 BETA_j + \alpha_2 DMI_j + \epsilon_j, \qquad (3)$$

with DMI_j the excess of the dividend yield over the interest rate for the jth firm. Although the tax effect should make α_2 positive, its information in DMI on share risk would tend to make α_2 negative. That is, dividend yield varies inversely with expected growth, and we would find α_2 negative insofar as growth is risky. To the extent that these two influences of the dividend yield offset each other, α_2 will tend toward zero.

The CAPM theory of how expected return varies among shares has been proposed as an alternative to the DCF model for measuring yield. Its value for the jth stock is

$$EHPR_j = INTR + BETA_j[EHPR_m - INTR], \qquad (4)$$

where:

$EHPR_j$ = expected holding-period return on the jth share,

INTR = one-period risk-free interest rate,

$EHPR_m$ = expected holding-period return on the market portfolio.

There is an important difference between this CAPM model of share yield and the DCF model represented by Equation (1). The latter is merely an instrument for measuring share yield: There is nothing in the DCF model that explains the variation in yield among shares. The CAPM, on the other hand, is a theory on why and how yield varies among shares, but one must go outside of the theory to estimate the variables on the right-hand side of Equation (4). Given rules for estimating the variables, EHPR and BETA, empirical work then provides a joint test of the theory and the estimating rules, such as we are carrying out here.[6]

The CAPM nonetheless has been used to estimate share yield in testimony before regulatory commissions by assigning numbers to each of the quantities on the right-hand side of Equation (4). For INTR, a long-term bond yield is sometimes used instead of a one-period rate. BETA is estimated by conventional methods.

The big problem is the expected return on the market portfolio. Here the practice has been to use the average realized risk premium over a period of about fifty years as the estimate of $EHPR_m$ − INTR in Equation (4). Although the implicit assumption is that the risk premium is a constant over time, we would expect the premium to change from one period to the next for various reasons, among them changes in the interest rate, the risk premium on the market portfolio, and the relative taxation of interest and share income. Hence, this estimate of share yield is more or less in error at any particular time, but we have no way of estimating this error and comparing the method with the others.

COMPARATIVE PERFORMANCE

We carried out our empirical work with a sample of 75 large electric and gas utility firms and a sample of 244 firms that includes 169 industrial firms drawn from the S&P 400. We obtained share yield under the four methods for estimating it as of the start of the year for the years 1984, 1985, and 1986.

For the explanatory variables, BETA for each share on each date was obtained by regressing the monthly HPRs for the share on the monthly HPRs for the S&P 500 over the prior five years. DMI for a share is its dividend yield less the interest rate on the one-month Treasury bill at the start of each year. EGR and DGR are the growth rates in earnings and in dividends per share, respectively, over the prior five years as reported on the Value Line Tape. BRG is a weighted average of the retention growth rates over the prior five years,[7] and FRG is the average of forecast growth rates in earnings over the next five years reported by IBES. The corresponding estimates of share yield were obtained by adding the dividend yield at the start of each year to the estimate of growth.

Table 1 presents the statistics that we obtained with KBRG and KFRG as the estimates of DCFY for the sample of utility shares and of all shares. The means of KBRG for the utility shares seems reasonable, with the interest rate on ten-year government bonds the standard of comparison, the latter being 11.67%, 10.43%, and 9.19% at the start of 1984, 1985, and 1986, respectively.[8] The standard deviations for KBRG are small enough to make its range of variation well within the bounds of reason. The lower means for all shares reveal that the means for industrial shares are below the means for utility shares.[9] This casts doubt on the accuracy of KBRG as a basis for estimating the DCFY on industrial shares, because industrials are riskier than utility shares.

The beta model explains none of the variation in KBRG among utility shares, but the two-factor model is a substantial improvement. The DMI coefficient, α_2, is positive and significant in every year, meaning that the unfavorable tax effect of a high dividend yield dominates the favorable risk effect. The coefficient on BETA is positive and significant in two of the three years. The only disturbing feature of the data is the sharp fall in R^2 and the corresponding rise in the root MSE relative to the standard deviation of KBRG as we go from 1984 to 1986.

The KBRG statistics for all shares are substantially inferior to the utility share statistics. This forces the unhappy conclusion that, for industrial shares, BETA is a poor measure of risk, or KBRG is a poor measure of DCFY, or both.

The KFRG statistics for the utility sample are superior to the KBRG statistics. The means are reasonable under the two criteria of being above the interest rate and moving with it. The range of variation of KFRG suggested by its standard deviations seems reasonable. The statistics for the beta model are a slight improvement on the corresponding statistics for KBRG. Furthermore, the two-factor model does a good job of explaining the variation in KFRG among

TABLE 1

Sample and Regression Statistics for KBRG and KFRG, Utility Shares and All Shares, 1984, 1985, and 1986

	KBRG			KFRG		
	1984	1985	1986	1984	1985	1986
UTILITY SHARES (75)						
Mean	14.64	14.38	12.93	15.64	14.56	12.93
Standard Deviation	2.51	1.87	1.80	2.26	1.43	1.42
Beta Model a_0	14.26	13.96	13.05	15.14	13.48	12.74
a_1	1.44	1.21	−0.28	1.25	3.09	0.42
t-statistic	(0.97)	(1.12)	(0.19)	(0.93)	(4.14)	(0.37)
Root MSE	2.52	1.87	1.81	2.26	1.29	1.43
R^2	0.013	0.017	0.001	0.012	0.190	0.002
Two-Factor Model a_0	12.45	12.75	12.42	13.30	12.46	11.97
a_1	3.43	2.11	0.11	3.28	3.85	0.89
t-statistic	(3.13)	(2.19)	(0.08)	(3.83)	(6.33)	(0.88)
a_2	0.68	0.45	0.34	0.66	0.58	0.41
t-statistic	(8.22)	(4.88)	(2.81)	(10.73)	(6.52)	(4.65)
Root MSE	1.82	1.63	1.73	1.41	1.05	1.26
R^2	0.491	0.262	0.100	0.620	0.491	0.232
ALL SHARES (244)						
Mean	12.98	13.19	11.86	16.17	15.87	14.31
Standard Deviation	3.86	3.21	3.52	2.60	2.32	2.30
Beta Model a_0	15.00	14.71	13.90	15.56	14.50	12.57
a_1	−2.47	−1.91	−2.40	0.74	1.72	2.05
t-statistic	(4.22)	(4.15)	(4.25)	(1.83)	(5.29)	(5.70)
Root MSE	3.73	3.10	3.40	2.59	2.20	2.16
R^2	0.069	0.066	0.069	0.014	0.104	0.118
Two-Factor Model a_0	14.34	14.42	13.95	15.40	14.61	12.75
a_1	0.09	−1.18	−2.51	1.37	1.44	1.61
t-statistic	(0.13)	(2.04)	(3.45)	(2.69)	(3.52)	(3.49)
a_2	0.48	0.17	−0.02	0.12	−0.06	−0.10
t-statistic	(6.04)	(2.09)	(0.24)	(2.01)	(1.12)	(1.53)
Root MSE	3.49	3.08	3.41	2.57	2.20	2.16
R^2	0.191	0.083	0.070	0.030	0.108	0.127

utility shares. The R^2s are higher here than for KBRG in every year. Finally, α_1 is positive and significant in every year, and α_2 is not significant only in 1986.

The implicit means of KFRG for the industrial shares seem high but not beyond reason. On the other hand, the regression statistics for the all-shares sample are not good, which leads to the same unhappy conclusion for industrial shares as we reached for KBRG.

Table 2 presents the statistics that we obtained using KEGR and KDGR as estimates of the DCFY on the shares in our samples. Comparison of the regression statistics with those in Table 1 reveals that KEGR and KDGR, particularly the former, fall short by a wide margin of the performance of KBRG and KFRG as estimates of the DCFY on a share.

CONCLUSION

We have compared the accuracy of four methods for estimating the growth component of the discounted cash flow yield on a share: past growth rate in earnings (KEGR), past growth rate in dividends (KDGR), past retention growth rate (KBRG), and forecasts of growth by security analysts (KFRG). Criteria for the comparison were the reasonableness of sample means and standard deviations and the success of beta and dividend yield in explaining the variation in DCF yield among shares. For our sample of utility shares, KFRG performed well, with KBRG, KDGR, and KEGR following in that order, and with KEGR a distant fourth. If we had used past growth in price, it would have been an even more distant fifth. Nevertheless, none of the four estimates of growth performed well under the criteria for a sample that included industrial shares.

Before closing, we have three observations to make. First, the superior performance by KFRG should come as no surprise. All four estimates of growth rely upon past data, but in the case of KFRG a larger body of past data is used, filtered through a group of security analysts who adjust for abnormalities that are not considered relevant for future growth. We assume this is done by any analyst who develops retention growth estimates of yield for a firm. If we had done this for all seventy-five firms in our utility sample, it is likely that the correlations

TABLE 2

Sample and Regression Statistics for KEGR and KDGR, Utility Shares and All Shares, 1984, 1985, and 1986

	KEGR			KDGR		
	1984	1985	1986	1984	1985	1986
			UTILITY SHARES (75)			
Mean	16.16	0.32	14.91	16.49	15.76	14.13
Standard Deviation	3.31	3.47	4.66	3.12	2.41	2.21
Beta Model α_0	15.45	16.18	0.51	15.75	14.53	12.30
α_1	1.75	0.40	−7.87	1.85	3.53	3.99
t-statistic	(0.89)	(0.20)	(2.16)	(0.99)	(2.64)	(2.32)
Root MSE	3.32	3.49	4.55	3.12	2.32	2.15
R^2	0.010	0.001	0.060	0.013	0.067	0.069
Two-Factor Model α_0	14.20	15.83	18.76	14.10	13.56	12.64
α_1	3.13	0.66	−8.03	3.65	4.25	3.78
t-statistic	(1.66)	(0.32)	(2.18)	(2.23)	(3.25)	(2.20)
α_2	0.47	0.13	−0.13	0.61	0.35	−0.18
t-statistic	(5.32)	(0.66)	(0.42)	(5.02)	(2.86)	(1.21)
Root MSE	3.11	3.50	4.58	2.70	2.21	2.14
R^2	0.142	0.007	0.063	0.269	0.180	0.087
			ALL SHARES (244)			
Mean	11.14	9.42	7.88	15.08	13.63	11.35
Standard Deviation	10.67	11.67	11.45	6.08	6.30	6.71
Beta Model α_0	15.96	18.28	19.53	15.15	0.04	15.39
α_1	−5.90	−11.16	−13.70	−0.09	−1.78	−4.74
t-statistic	(3.62)	(7.07)	(8.10)	(0.09)	(1.92)	(4.41)
Root MSE	10.41	10.65	10.18	6.09	6.27	6.47
R^2	0.051	0.171	0.213	0.000	0.015	0.074
Two-Factor Model α_0	14.84	18.01	19.91	14.31	14.11	14.79
α_1	−1.56	−10.49	−14.62	3.17	0.63	−3.25
t-statistic	(0.77)	(5.27)	(6.72)	(2.73)	(0.55)	(2.36)
α_2	0.31	0.15	−0.21	0.61	0.55	0.34
t-statistic	(3.51)	(0.55)	(0.67)	(4.57)	(3.47)	(1.72)
Root MSE	10.18	10.67	10.19	5.86	6.13	5.45
R^2	0.097	0.172	0.215	0.080	0.062	0.085

would have been as good or better than those obtained with the analyst forecasts of growth.

Second, we examined shares and not portfolios, because our objective is to estimate the DCFY for shares and not for portfolios. As common practice in testing the CAPM has been to execute tests on portfolios instead of shares, we classified our population of shares into ten portfolios on the basis of their beta values. Regression statistics were substantially unchanged, except that correlations increased dramatically.

Finally, we must acknowledge that we have no basis for estimating the expected HPR or DCF yield for industrial shares with any confidence. Theories on financial decision-making in industrial corporations that rely on that statistic have a weak empirical foundation.

[1] The EHPR is a one-period return, while the DCFY is a yield to maturity measure. The two may differ in actuality because of measurement problems, but they also may differ in theory. That is, they may differ in the same way that interest rates on bonds of different maturities may differ. See Gordon and Gould (1984a). This source of difference between EHPR and DCFY will be ignored here.

A widely accepted hypothesis is that dividends contain information on earnings, because management sets the dividend to pay out a stable fraction of normal or permanent earnings.

[3] Over a five-year period, there may even be a negative rate of growth in price for a large number of firms. Furthermore, this negative growth rate may be larger in absolute value than the dividend yield, which leads to the conclusion that investors are holding such shares to earn a negative return. The frequency of negative rates of growth in price is reduced as the prior time period used in its calculation increases in length. As that takes place, however, the estimate of the expected return for a firm approaches a constant or a constant plus the dividend yield. The expected return on a share is one statistic for which it is an error to assume that expectations are on average realized.

[4] Equation (2) is similar to the CAPM according to Sharpe, Lintner, and Mossin. They arrived at this expression under very rigorous assumptions. The heuristic risk premium model is adequate for our purposes.

[5] It may be thought that Theil's (1966) decomposition of the difference between the actual and predicted values of a variable can be used here, but in fact that decomposition applies to a different problem. It assumes that the observed (actual) past values of a variable are free of error, and it decomposes the error in a model that is employed to explain the past values. The purpose of Theil's decomposition is to cast light on the possible error in using the model to predict future values of the dependent variable. Our problem is to determine which set of observed values is closest to the true values, with the risk premium theory of share yield and BETA as the source of information on the true values. Theil's method would be appropriate for decomposing the difference between the actual and predicted values of the *realized* holding-period return on a share. The actual values here can be observed without error.

[6] There is an enormous volume of empirical work devoted to discovering whether the theory is true, but this empirical work does not provide useful estimates of the EHPR on a share. To test the truth of Equation (4), the practice has been to regress EHPR on BETA for a sample of firms with the average realized HPR over the prior five or so years used as an estimate of the EHPR. Because of the large error in the realized HPR over a prior time period, as noted earlier, neither the actual values of the dependent variable nor the values predicted by the model are usable as estimates of share yield. See Fama and MacBeth (1973) and Friend, Westerfield, and Granito (1978).

[7] BRC for a year is earnings less dividend divided by the end-of-year book value. The estimate of the expected value as of the start of 1986 is 0.3BRC85 + 0.25BRC84 + 0.20BRC83 + 0.15BRC83 + 0.10BRC82. If any value of BRC was negative, it was set equal to zero.

[8] We expect the yields on shares to be above the risk-free interest rate, but with a high enough interest rate the more favorable tax treatment of shares can reduce the yield below the interest rate. Interest rates were not that high in these years. See Gordon and Gould (1984b).

[9] The statistics reported for all shares and for utility shares were also obtained for industrial shares. All methods of estimation performed so poorly for industrial shares, however, as to suggest no confidence can be placed in any of them. To save space, we do not present statistics for the industrial shares. Whatever we want to know about them can be deduced by comparing the data for all shares and utility shares.

REFERENCES

Brennan, M.J. "Taxes, Market Valuation and Corporate Financial Policy." *National Tax Journal*, 23 (1973), pp. 417-427.

Brigham, E., D. Shome, and S. Vinson. "The Risk Premium Approach to Measuring a Utility's Cost of Equity." *Financial Management*, Spring 1985, pp. 33-45.

Fama, E., and J.D. MacBeth. "Risk, Return and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (May 1973), pp. 607-636.

Friend, I., R. Westerfield, and M. Granito. "New Evidence on the Capital Asset Pricing Model." *Journal of Finance*, 33 (June 1978), pp. 903-917.

Gordon, M.J. *The Cost of Capital to a Public Utility*. East Lansing, Michigan: Michigan State University, 1974.

Gordon, M.J., and L.I. Gould. "Comparison of the DCF and HPR Measures of the Yield on Common Shares." *Financial Management*, Winter 1984a, pp. 40-47.

——. "The Nominal Yield and Risk Premium on the TSE-300, 1956-1982." *Canadian Journal of Administrative Sciences*, 1 (1984b), pp. 50-60.

——. Testimony Before the Federal Communications Commission in the Matter of American Telephone and Telegraph Company. FCC Docket No. 79-63, April 1980.

Harris, R.S. "Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return." *Financial Management*, Spring 1986, pp. 58-67.

Kolbe, A.L., J.A. Read, and G.R. Hall. *The Cost of Capital: Estimating the Rate of Return for Public Utilities*. Cambridge, MA: MIT Press, 1984.

Morin, R.A. *Utilities' Cost of Capital*. Arlington, VA: Public Utilities Reports, Inc., 1984.

Theil, H. *Applied Economic Forecasting*. Chicago: North Holland, 1966.