

FINANCIAL ANALYST FORECAST SUPERIORITY:  
IS THERE AN INFORMATION INTERPRETATION?

Lawrence D. Brown\*

Gordon D. Richardson\*\*

Steven J. Schwager\*\*\*

\*University at Buffalo

\*\*University of British Columbia

\*\*\*Cornell University

April 1986

The comments of David Chen, Peter Easton, Chi-Ying Han, Calvin Lai, Rex Thompson, Robert Verrecchia, Gregory Waymire, G. Peter Wilson, Mark Zmijewski, and a reviewer are appreciated. We also appreciate the comments of participants at the Cornell University, University of Oregon, University of Rochester, University of Toronto, University of Washington, and Washington University workshops. We are grateful for the editing and word processing support of Nancy Schell and Cheryl Tubisz and the research assistance of Benjamin Lee. Any remaining errors are the sole responsibility of the authors.

## 1. Introduction

Much of the empirical evidence involving the accuracy of financial analysts' forecasts (hereafter FAF) of upcoming firm earnings has addressed FAF accuracy relative to the predictions of univariate time-series (hereafter TS) forecasting models. While earlier studies have produced mixed results (see Abdel-khalik and Thompson (1977-1978)), recent empirical evidence is consistent with FAF superiority (see Brown, Griffin, Hagerman, and Zmijewski (1985)).

Brown et al. (1985) identified two potential sources of FAF superiority: (1) better utilization of information existing at the forecast initiation date for the time-series models, a contemporaneous advantage, and (2) acquisition and use of information after the forecast initiation date for the time-series models, a timing advantage. Brown et al. provided evidence that FAF superiority is attributable to both types of informational advantages, but they did not provide any evidence regarding the characteristics of the information set underlying FAF.

Brown, Foster, and Noreen (1984) reported that little has been done in modeling determinants of FAF error magnitude.<sup>1</sup> Moreover, no work has been done to date either in modeling FAF superiority or in relating FAF superiority to attributes of the firm's information environment.

This study models the determinants of FAF superiority in the context of the firm's information environment. FAF and TS model predictions are both adversely affected by the current stochastic disturbance term or random shock in earnings. The source of FAF superiority can be understood intuitively as the analyst's knowledge

of the nature of the current disturbance term, something TS models cannot have. More formally, we present a Bayesian model in which FAF superiority is positively related to (1) the dimensionality of the analyst's information set, and negatively related to both (2) the variance of information observations and (3) the covariance between information observations.

Empirically, surrogates for the theoretical factors are used in order to explain FAF superiority for two samples of firms: individual analysts (Value Line) and consensus analysts (IBES). FAF superiority is defined as  $\sigma^2_{RW}/\sigma^2_{FAF}$ , the ratio of the variance of a random walk (RW) error to the variance of FAF error. The empirical proxies for the three informational variables discussed above are: (1) firm size, (2) the prior dispersion of FAF forecasts, and (3) the number of lines of business. The test results are generally consistent with an information interpretation of FAF superiority.

The existence of differential firm information environments is important to accountants. Atiase (1985) demonstrated that the price reaction to earnings is negatively related to firm size, a crude proxy for the extent of prior information availability. Our study explores finer partitions of the firm's information environment through the use of multiple empirical proxies.

Our finding that FAF superiority is related to firm size is consistent with Atiase's firm-size related differential information hypothesis. Our finding that FAF superiority is also related to the prior dispersion of analyst forecasts and the number of lines of business suggests that multiple empirical proxies specify the firm's information environment more correctly than does a single empirical proxy (i.e., size).

The remainder of the paper is organized as follows: Section 2 describes the model formulation and implications. Section 3 describes surrogate selection and hypotheses. Section 4 describes the sample and data. Methodological considerations are discussed in Section 5. Section 6 presents empirical results. A summary and conclusion constitute Section 7.

## 2. Model Formulation and Implications

The objective of the model is to examine the relationship between FAF superiority and the firm's information environment. The model provides theoretical determinants of FAF superiority, conditional on the model's assumptions. The model is limited in the sense that it does not incorporate such factors as: (1) the supply and demand of forecasts and forecast accuracy, (2) macro factors that affect forecast accuracy, (3) the interaction between macro and micro factors that affect forecast accuracy, and (4) multiperiod intricacies. Thus, while the model does provide a basis to examine the effects of the information environment, it is not a complete model.

We make the following assumptions in developing the Bayesian model: (1) the information available to the analyst is exogenous; (2) the analyst uses all available information to generate a forecast; (3) the firm has existed for  $T-1$  prior periods; (4) the analyst is interested in forecasting firm earnings for period  $T$ , denoted by  $R_T$ ; (5) earnings behave as a random walk process:  $R_T = R_{T-1} + a_T$ , where the stochastic error terms  $a_T$  are a sequence of identically distributed, uncorrelated random variables with common mean 0 and variance  $\sigma^2_0$ ; (6) prior to observing the  $n$  interim information signals  $(Y_1, \dots, Y_n) = Y$ , the analyst's prior predictive distribution for  $R_T$  is

normal with mean  $E[R_T | R_{T-1} = r_{T-1}] = r_{T-1}$  and variance of  $\sigma^2_0$ ; (7) after observing  $Y$ , the analyst's point estimate of  $R_T$  is the posterior mean,  $R_T = E[R_T | Y]$ ; (8) the  $n$  interim information signals,  $Y_1, \dots, Y_n$ , have a multivariate normal distribution and a common mean  $r_T$ , which is the draw of  $R_T$  determined by nature, but not yet known to the analyst; and (9) the  $n$  interim information signals have a common variance,  $\sigma^2_y$ , and a common correlation,  $\rho$ , between each pair of signals.<sup>2</sup>

Given these assumptions, it can be shown that:

$$(1) \quad \sigma^2_{RW} = \sigma^2_0 = 1/\tau_0$$

where  $\sigma^2_{RW}$  denotes the variance of the random walk model forecast error, and  $\tau_0$  is defined as  $1/\sigma^2_0$ ;

$$(2) \quad \sigma^2_{FAF} = 1/(\tau_0 + \tau_1),$$

where  $\sigma^2_{FAF}$  denotes the variance of financial analysts' forecast error, and

$$\tau_1 = n/[1 + (n-1)\rho]\sigma^2_y$$

is a function of the three information variables ( $n$ ,  $\sigma^2_y$ ,  $\rho$ ) in assumption (9)<sup>3</sup>; and

$$(3) \quad \text{FAFSUP} = \frac{\sigma^2_{RW}}{\sigma^2_{FAF}} = \frac{\tau_0 + \tau_1}{\tau_0}.$$

Equation (3) expresses FAF superiority (FAFSUP) as the ratio of RW error variance to FAF error variance. FAFSUP increases towards  $\infty$  (decreases towards 1) as the analyst's improvement in forecast precision over the random walk model increases (decreases), where precision is the reciprocal of variance.

FAFSUP is positively related to the dimensionality of the FAF information set ( $n$ ), and negatively related to both the variance of

the interim information observations ( $\sigma_y^2$ ) and the covariance between information observations ( $\rho$ ).

If the analyst does not observe information (i.e.,  $n = 0$  in equation (2)), then  $\tau_1 = 0$  and the variance of an analyst's forecast error simplifies from  $1/(\tau_0 + \tau_1)$  to  $1/\tau_0 = \sigma_0^2$ , the variance of stochastic disturbance terms for the firm's earnings. However, as the amount of interim information available to the analyst increases (i.e., as  $n$  increases), the variance of the analyst's forecast error ( $\sigma_{FAF}^2$ ) declines monotonically from  $\sigma_0^2$  (this follows from inspection of the expression  $1/(\tau_0 + \tau_1)$ ), explaining one potential source of analyst superiority over a random walk model.<sup>4</sup>

Expressed more informally, the variance of stochastic disturbance terms adversely affects the forecasting performance of both the random walk model and the analyst. One source of analyst superiority over the random walk model is that more interim information is available to the analyst, enabling FAF to learn about the nature of the current disturbance or "random shock" (due to litigation, award of a new contract, and so on). In this way, FAF error variance is reduced below that of RW error variance.

While  $n$ ,  $\sigma_y^2$ , and  $\rho$  reflect different aspects of the firm's information environment, a given value of  $\tau_1$  can be obtained from numerous combinations of these variables. Also, as discussed below in Section 3, our empirical surrogates do not bear a one-to-one correspondence to their theoretical counterparts. Thus, while the model does provide a basis for examining the effects of the firm's information environment, it cannot be used to determine unambiguously which specific attributes of the information environment are relevant.

### 3. Surrogate Selection and Hypotheses

As the theoretical model is derived for a single analyst, the empirical tests should, ideally, pertain to the forecasts of a single analyst. Forecasts published in The Value Line Investment Survey (Value Line) are made by one or two analysts, and may be subject to review by other analysts. As such, Value Line forecasts are not strictly the forecasts of a single analyst, but they are the closest publicly available source of individual analyst forecasts and we use them for this purpose. We selected a second sample to determine whether the model is empirically valid for consensus FAF. For this purpose, we utilized the Institutional Brokers Estimate System (IBES) tape provided by Lynch, Jones & Ryan. IBES forecasts are a consensus of as many as 60 analysts.

Each set of forecasts utilizes predictions made over three horizons. Value Line forecasts are predictions made one, two, and three quarters in advance; IBES forecasts are final (for the year), one, and two years ahead. The Value Line forecasts are obtained in the appropriate issue of The Value Line Investment Survey; IBES final forecasts are obtained in the last month of the firm's fiscal year; IBES one- and two-year-ahead forecasts are made in the sixth month of the firm's fiscal year.

Ideally, we would like to observe the information set of a given analyst in order to obtain the covariance matrix of all information signals underlying the FAF. The dimensionality of the matrix would provide the number of observed information sources ( $n$ ), while the diagonal and off-diagonal elements would provide the variance ( $\sigma_y^2$ ) of and covariance ( $\rho$ ) between information observations, respectively.

However, such a matrix is unobservable, necessitating surrogate selection.

Following Atiase (1985), we select firm size (SIZE), defined as the market value of common stock and the book value of long-term debt in the fiscal year preceding the year of forecast (source: Compustat Industrial tape) as a surrogate for the dimensionality ( $n$ ) of the analyst's information set. We expect FAF superiority to be positively related to firm size.<sup>5</sup>

We use prior dispersion of forecasts across various analysts following the firm as a proxy for the variance ( $\sigma_y^2$ ) of information observations. Cukierman and Givoly (1982) demonstrated theoretically that the divergence of forecasts increases as the variance of information observations increases. In the present context, this suggests that the dispersion among analysts' forecasts reflects their ex ante uncertainty about upcoming earnings. We use the dispersion in IBES forecasts (coefficient of dispersion) as a proxy for  $\sigma_y^2$ , and expect FAF superiority to be negatively related to the coefficient of dispersion.<sup>6</sup>

For the purpose of the annual IBES sample, dispersion (DISP) is observed prior to observing IBES final, one-, and two-year-ahead forecasts for a fiscal year.<sup>7</sup> For the purpose of the quarterly Value Line sample, it would be ideal to observe FAF dispersion regarding upcoming quarterly earnings. However, IBES dispersion is available for upcoming fiscal years but not quarters. Accordingly, DISP was estimated in the month prior to the month in which the Value Line forecast was generated, and pertained to consensus analyst forecasts for the fiscal year that Value Line was forecasting. Dispersion is measured as the standard deviation across analysts of upcoming annual

forecasts for firm  $j$ , deflated by the absolute value of the mean forecast across analysts. We expect FAF superiority to be negatively related to DISP.

Unlike  $n$  and  $\sigma_y^2$ , there does not exist a literature to guide us in surrogate selection for the covariance ( $\rho$ ) between information observations. As a proxy for the covariance measure, we selected the number of lines of business reported by the firm in accordance with SFAS No. 14. We reason as follows. Suppose that the  $n$  information sources available to a particular analyst relate to  $m$  lines of business. If  $m$  is small, the  $n$  information signals will be highly dependent, and in the extreme may be considered to be repeated measures of the same information signal. On the other hand, if  $m$  is large and the  $n$  information signals are randomly assigned to  $m$ , the  $n$  information signals will tend to represent more diverse types of information. Thus, we expect the covariance among information signals to be a decreasing function of  $m$ . The number of lines of business (LOB) is defined as the number of distinct lines of business disclosed in the fiscal year preceding the year of forecast (source: The Value Line Investment Survey).<sup>8</sup> We expect FAF superiority to be positively related to the LOB variable.

Brown, Griffin, Hagerman, and Zmijewski (1985) showed that FAF superiority is positively related to FAF timing advantage. The FAF timing advantage for tests involving Value Line is controlled for by adding a timing variable (TIME) for tests, measured as the number of days lapsed between the end of the previous fiscal quarter and the date of the Value Line forecast. The FAF timing advantage for tests involving IBES is controlled for by requiring all IBES firms to have

their forecasts appear on the IBES tape the same number of days after their fiscal year end.<sup>9</sup> We expect FAF superiority to be negatively related to TIME.

Imhoff and Pare' (1982) have shown that FAF superiority is negatively related to forecast horizon. In an effort to control for horizon (HOR), we use dummy variables for one- and two-quarter-ahead forecasts for Value Line, and final and one-year-ahead forecasts for IBES. The effect of Value Line's three-quarter-ahead forecasts and IBES's two-year-ahead forecasts are included in the intercept. We expect FAF superiority to be negatively related to HOR.

#### 4. Sample and Data

The firms in the Value Line sample are a subset of the 233 firms used by Brown, Griffin, Hagerman, and Zmijewski (1985), hereafter BGHZ. Value Line forecasts included in this study are made one, two and three quarters in advance of the reported earnings numbers. Of the 233 BGHZ firms, 65 were eliminated from our sample because they were not covered by either the IBES tape (July 1984) or the Compustat Industrial tape (1983), the sources of data for FAF dispersion and firm size, respectively. The 168 Value Line firms are covered for the twelve fiscal quarters from first quarter 1977 to fourth quarter 1979.

The firms in the IBES sample are all firms with continuously available data on the IBES tape (i.e., every month) for the 1977-82 period, and that are covered by the Compustat Industrial tape (1983). A total of 673 firms met these criteria. Of these firms, 139 were also part of the aforementioned 168 firm Value Line sample. As we want to compare the Value Line and IBES sample results for the 168 firms, we added 29 firms included in the Value Line sample, that did

not have continuously available IBES forecast data.<sup>10</sup> Thus, all 168 Value Line firms were included in the 702 firm IBES sample (673+29). For each IBES sample firm, we collected three sets of forecasts for each year, 1977-82: the final forecast for the fiscal year made in the last month of the fiscal year; and forecasts of the current and following fiscal year made six months through the (fiscal) year.

### 5. Methodological Considerations

We now examine empirically the relationship between FAF superiority and its theoretical determinants. This requires both the specification of a model relating FAF superiority to its theoretical determinants and the selection of empirical surrogates for these quantities. The model to be used will have the form:

$$(4) \text{ FAFSUP} = f(\overset{+}{\text{SIZE}}, \overset{-}{\text{DISP}}, \overset{+}{\text{LOB}}) + e,$$

where SIZE is the firm's total capitalization, DISP is prior dispersion of forecasts across analysts, LOB is the firm's number of lines of business, and the superscript signs indicate the a priori direction of the predictor variables. As discussed in Section 4, these are empirical surrogates for  $n$ ,  $\sigma_y^2$ , and  $\rho$ , respectively, the theoretical determinants of FAFSUP.

A surrogate for FAFSUP will be defined for each firm  $j$  at each time period  $t$ . Take (FAF error)<sup>2</sup> and (RW error)<sup>2</sup> as proxies for  $\sigma_{\text{FAF}}^2$  and  $\sigma_{\text{RW}}^2$  respectively, where FAF error is the difference between the analyst's forecast (from either Value Line or IBES) and the actual value of earnings (EPS) for firm  $j$  at time  $t$ , and RW error is the difference between EPS at time  $t$  and EPS at time  $t-1$  for firm  $j$ . Motivated by equation (3), we define the empirical surrogate FAFSUPS for FAFSUP as the ratio of squared errors:

$$(5) \text{ FAFSUPS}_{jt} = (\text{RW error}_{jt} / \text{FAF error}_{jt})^2.$$

We define two additional surrogates that are transformations of FAFSUPS, the ratio of absolute errors and the logarithm of FAFSUPS:

$$(6) \text{ FAFSUPS}_{jt}^{1/2} = |\text{RW Error}_{jt} / \text{FAF Error}_{jt}|$$

$$\text{LOG}_{10}(\text{FAFSUPS}_{jt}) = \text{LOG}_{10}[(\text{RW Error}_{jt} / \text{FAF Error}_{jt})^2].$$

Where necessary, a small constant was added to the value of FAFSUPS before taking the logarithm.

The surrogate for FAFSUP was calculated from Value Line and IBES data that do not overlap the data used to calculate SIZE, DISP, and LOB. We substitute these empirical surrogates into equation (4) and use regression to analyze the resulting data. If we obtain significant effects attributable to SIZE, DISP and LOB, these will constitute evidence supporting both the model specification and the surrogate selection.

In order to combine time-series and cross-sectional data in one overall linear regression, we adopt a model with constant slope coefficients and an intercept that varies with respect to forecast horizon, industry membership, and calendar year (quarter) as explained by Judge, Hill, Griffith, and Lee (1980, p. 338). The approach involves fitting the following model for  $\text{LOG}_{10}(\text{FAFSUPS})$ :

$$(7) Y_{jt} = \beta_0 + \beta_1 \text{LSIZE}_{jt} + \beta_2 \text{LDISP}_{jt} + \beta_3 \text{LOB}_{jt} + \beta_4 \text{LTIME}_{jt}$$

$$+ \beta_5 \text{HOR}_{1jt} + \beta_6 \text{HOR}_{2jt} + \sum_{i=1}^8 \gamma_i I_{ijt} + \sum_{k=1}^6 \delta_k T_{kjt} + \epsilon_{jt}$$

where  $j = \text{firm } j, j=1, \dots, 168$  (Value Line, IBES);  $j=1, \dots, 702$  (IBES).

t = quarter t, t=1,...,12 (Value Line); year t, t=1,...,6 (IBES).

$Y_{jt}$  =  $\text{LOG}_{10}(\text{FAFSUPS}_{jt})$  for firm j in quarter t (Value Line); year t (IBES).

$\text{LSIZE}_{jt}$  =  $\text{LOG}_{10}(\text{SIZE}_{jt})$ .

$\text{LDISP}_{jt}$  =  $\text{LOG}_{10}(\text{DISP}_{jt})$ .

$\text{LTIME}_{jt}$  =  $\text{LOG}_{10}(\text{TIME}_{jt})$ . (Value Line sample only.)

$\text{HOR}_{1jt}$  = 1 or 0, a horizon dummy that is 1 if forecast is one-quarter-ahead (Value Line) or final forecast (IBES).

$\text{HOR}_{2jt}$  = 1 or 0, a horizon dummy that is 1 if forecast is two-quarters-ahead (Value Line) or one-year-ahead (IBES). The effects of Value Line's three-quarter-ahead forecasts and IBES's two-year-ahead forecasts are included in the intercept.

$I_{ijt}$  = 1 or 0, an industry dummy for firm j in quarter t representing industry category i, i=1,...,8. The effect of industry category 9 is included in the overall intercept.

$T_{kjt}$  = 1 or 0, a time dummy for firm j in quarter (year) t, representing quarter (year) k, k = quarter 1,...,11 (Value Line); k, k = year 1,...,5 (IBES). The effect of quarter 12 (year 6) is included in the overall intercept.

$e_{jt}$  = model disturbance term, assumed to be serially independent, independent of the predictor variables, and distributed normally.<sup>11</sup>

The logarithmic transformation of the dependent and predictor variables was chosen after considering several alternatives; it resulted in a modest improvement in model fit, compared to results (not reported) using the untransformed variables. Moreover, standard

tests of normality of model error terms could not reject normality after taking log transformations. In its raw form, FAFSUPS exhibits substantial positive skewness. (For a discussion of transformations, see Box and Tidwell (1962).) Where necessary, a small constant was added to the original variable before taking the log transformation.

The inclusion of industry membership and year (quarter) dummies attempts to control for intercept differences due to industry and calendar effects, factors that our intuition suggests are related to FAF superiority. If the frequency or magnitude of random shocks affecting firm earnings is not uniform across industry or calendar time, and if such shocks do not affect FAF error and RW error equally, then FAF superiority could depend on industry and year.<sup>12</sup>

## 6. Empirical Results

### 6.1 Data Description<sup>13</sup>

Table 1 contains a profile of FAF superiority for the 168 firm quarterly (Value Line) sample across the predictor variables: SIZE, DISP, LOB, and TIME. Except for LOB, strong patterns are evident, with relationships being nearly monotonic for all variables. More specifically, looking down the columns (i.e., holding forecast horizon constant), the relationship is strictly monotonic for SIZE and DISP for all horizons, and for TIME for the one- and two-quarter-ahead horizons. Looking across the rows (i.e., letting forecast horizon increase), the relationship is strictly monotonic for all the predictor variables. The profile suggests that FAF superiority is related to the predictors in the hypothesized direction: SIZE (+), DISP (-), TIME (+), and HOR (-). No clear pattern of association is evident for LOB. Table 2 contains similar results for the same 168

firms and the 702 firms using annual (IBES) data. Once again, except for LOB, the relationships are nearly monotonic for all variables, and FAF's superiority is related to the predictors in the hypothesized direction.<sup>14</sup>

## 6.2 Multivariate Tests of Association

Table 3 contains the results of pooled time-series cross-sectional linear regressions relating FAF superiority to the predictor variables for the 168 firm quarterly (Panel A), 168 firm annual (Panel B), and 702 firm annual samples (Panel C), respectively. Residual plots were examined and a linear fit appeared to be adequate.

The results confirm the patterns suggested by the profiles in tables 1 and 2. Firm size (LSIZE) has the predicted positive coefficient and is significant at the 5 percent level in all three regressions. Prior analyst dispersion (LDISP) has the predicted negative coefficient and is significant at the 5 percent level in all three regressions. Number of lines of business (LOB) has the predicted positive coefficient in one case (Panel B), but the opposite sign in another case (Panel A). The LOB variable is not significant in either case. The timing variable (LTIME) has the expected positive sign and is significant (see Panel A). Finally, the two forecast horizon dummy variables have the expected sign.<sup>15</sup> These results are consistent with the univariate tests and are generally consistent with an information interpretation of FAF superiority.

## 7. Summary and Implications

Past research focusing upon FAF superiority has been primarily evaluative, comparing analysts' forecasts to those of other (primarily

time-series) models. Little work has been done to date that theoretically or empirically models the determinants of FAF superiority. This study developed a Bayesian model that relates FAF superiority to three attributes of the information set underlying FAF (with expected sign of association indicated): (1) the dimensionality of the information set (+), (2) the variance of information observations (-), and (3) the covariance among information observations (-).

Empirically, surrogates for the theoretical factors were chosen. The ratio of RW error variance to FAF error variance (FAFSUPS) was our surrogate for FAF superiority, the dependent variable. Firm size (SIZE), prior dispersion of FAF forecasts (DISP), and the number of lines of business (LOB) proxy for the three predictor variables.<sup>16</sup> Significant effects were found for SIZE and DISP, but not for LOB. We attribute the poor findings for LOB to its inability to serve as an adequate proxy for  $\rho$ , rather than to  $\rho$  being an unimportant attribute of the analyst's information set. (See footnote 8.) Thus, we conclude that an information interpretation underlies the association between FAF superiority and our predictor variables.

Our findings have implications for studies that require measures of earnings expectations as well as for studies that incorporate differential information environments into capital markets research. If a researcher requires an earnings predictor for the purpose of generating an accurate forecast, and if the choice is limited to either an analyst's forecast or the forecast of a random walk model, the researcher should consider the firm's information environment when making the choice. More specifically, the random walk model generally will be a poorer choice vis a vis the analyst's forecast when

the firm is large and the prior dispersion of analysts' forecasts is small. On the other hand, the researcher is less likely to commit a type two error (fail to reject the null hypothesis of equal forecast accuracy of the analyst and the random walk model when it should be rejected) if the firm is small and the prior dispersion of analysts' forecasts is large.

Relying upon the work of Ohlson (1979), several recent studies have incorporated differential information environments into capital markets research. These studies have shown that the capital market reaction to the firm's reported earnings depends upon one particular aspect of the firm's information environment, the number of available information signals (e.g., Atiase (1985); Kross and Schroeder (1985); Shores (1985)). For example, Atiase (1985) explores the hypothesis of firm-size related differential availability. He suggests that firm size is a crude proxy for the firm's information environment. We agree. Our theoretical model demonstrates that the number of available information sources is only one of three aspects of the firm's information environment. Our model suggests that the individual quality (or "variance") of information sources and the extent of overlap (or "covariance") between information sources must also be considered in addition to the number ("dimensionality") of information sources. Empirically, we explore finer partitions of the firm's information environment, through the use of multiple empirical proxies.

Our evidence suggests that, after including firm size in the regression model, prior dispersion of FAF forecasts is a significant determinant of FAF superiority. Thus, we show both theoretically and

empirically that multivariate approaches to explaining FAF superiority in an information context are more appropriate than univariate approaches.

Studies that have incorporated differential information environments into capital markets research have investigated the capital market reaction to the release of the report. As examples, Grant (1980) observed a larger capital market reaction to reports of OTC firms than to reports of NYSE firms; McNichols and Manegold (1983) found the capital market reaction to annual reports to be smaller when they are preceded by interim reports than when interim reports were unavailable; Atiase (1985) showed that the magnitude of capital market reaction is negatively related to firm size.

The relationship between the release of an earnings report and the capital market reaction to the report has two facets. The first facet is the extent of the "unexpected earnings" revealed through the report; the second is the mapping of "unexpected earnings" to capital market reaction. Our evidence suggests that there exists an information interpretation to the first facet, and as such our findings are consistent with those of the above studies. More specifically, a sufficient condition for the findings of Grant (1980), McNichols and Manegold (1983), and Atiase (1985) is that it is more difficult to predict the earnings of OTC than NYSE firms, of annual earnings preceded by interim reports than annual earnings not preceded by interim reports, and earnings of small firms vis a vis those of large firms, respectively.

In our model, the firm's information environment affects the amount of "surprise" (unexpected earnings) revealed by the report. Conditional upon the sign and magnitude of the

"surprise", its mapping to share price may be independent of the firm's information environment. Whether or not the mapping of "surprise" to capital market reaction depends upon the firm's information environment is beyond the scope of this study and is a subject for further research.

## FOOTNOTES

1. Albrecht, Johnson, Lookabill, and Watson (1977) suggested (but did not test) a number of factors (e.g., year, forecast horizon, industry, firm size, number of lines of business, and variance of firm earnings) that potentially explain FAF errors. Their list of factors employed intuition and the results of prior empirical studies rather than formal modeling techniques.
2. "Interim information" is defined as all information available to analysts that is not contained in the sequence of past earnings numbers. Thus, forecasts based solely on past earnings have an  $n$  equal to 0. The model's assumptions imply that the expressions "information set underlying FAF" and "firm's information environment" are equivalent, since the analyst uses all available information regarding the firm. Assumptions 5 and 6 pertain to annual data. For quarterly data, the seasonal random walk model  $R_T = R_{T-4} + a_T$  will be used (assumption 5), and the mean of the prior predictive distribution for  $R_T$  becomes  $E[R_T | R_{T-4} = r_{T-4}] = r_{T-4}$  (assumption 6).
3. The model's implications are consistent with Holthausen and Verrecchia (1981), who demonstrated that the variance of price reaction to earnings is a negative function of the precision of interim information, and with Ohlson (1979), who demonstrated that the variance of price reaction to announced earnings is greater in a coarser interim disclosure environment than in a finer information environment.
4. While the model explicitly addresses FAF superiority relative to a random walk model, the results are easily generalized to earnings processes for which more complicated univariate time-

series models apply. For example, for an autoregressive time-series model of order 1,  $R_T = \phi R_{T-1} + a_T$  where the  $a_T$ 's are identically distributed, uncorrelated random variables with mean 0 and variance  $\sigma_0^2$ , the variance of a k-step-ahead TS forecast error is  $\sigma_0^2 \left( \frac{1-\phi^{2k}}{1-\phi^2} \right)$ . See Abraham and Ledolter (1983, p. 241).

The reciprocal of this expression would become  $\tau_0$  in the analysis of Section 2. The expression for  $\sigma_0^2$  changes with other TS specifications, but the essential point remains: FAF superiority equals  $\sigma_{TS}^2 / \sigma_{FAF}^2$ , the ratio of univariate time series (TS) model error variance to FAF error variance. We selected RW because it is the simplest of the univariate time-series models.

5. Kross and Schroeder (1985) found that the number of inches in the Wall Street Journal provided incremental information content beyond that contained in firm size when attempting to explain cross-sectional differences in the stock market's reaction to quarterly earnings announcements. When variable selection in the model (7) of Section 5 was performed with stepwise inclusion and elimination of variables, firm size entered the model but the number of millimetres in The Wall Street Journal Index did not. When the number of WSJ millimetres was included in the initial model, the firm size entered the model and the number of WSJ millimetres was deleted in a subsequent step. Thus, we did not find the number of millimetres in The Wall Street Journal Index to be a significant determinant of FAF superiority.

6. One potential problem with the IBES dispersion measure is the possibility that the forecasts used to calculate dispersion may not be contemporaneous, since the dates of forecasts vary (see O'Brien (1985)). The problem is not severe if the age distribution of forecast is random cross-sectionally. Unfortunately, we do not know the age distribution of the IBES forecasts, so we are unable to ascertain the severity of this problem. A second potential problem with the IBES dispersion measure is that it relates to annual rather than quarterly forecasts. Thus, it is likely to be a more suitable proxy for IBES than for Value Line data. However, no measure of the coefficient of dispersion of analysts' quarterly forecasts is available. As the Section 6 findings suggest that the dispersion in IBES forecasts appears to be a suitable proxy for the ex ante uncertainty about upcoming quarterly earnings, the use of an annual measure of dispersion for quarterly data does not appear to be a critical problem.
7. In order to ensure that prior dispersion was measured, dispersion was observed prior to observing IBES final, one-, and two-year-ahead forecasts for a fiscal year. For example, consider fiscal 1978 for a December year-end firm. Forecasts for fiscal 1978 were observed in December 1977, and the dispersion of those forecasts (DISP78) was calculated. In June 1978 (mid-year), forecasts for fiscal 1978 were observed and a one-year-ahead forecast error was calculated for fiscal 1978. In June 1978, forecasts were also observed for fiscal 1979 (since IBES reports forecasts for one and two fiscal years ahead) and a two-year-ahead forecast error was calculated for fiscal 1979. In December

1978, final forecasts for fiscal 1978 were observed and the final forecast error for fiscal 1978 was calculated. Associations were tested for between DISP78 and one-year-ahead forecast error for fiscal 1978, final forecast error for fiscal 1978, and two-year-ahead forecast error for fiscal 1979. This procedure ensures that prior forecast dispersion is related to errors in subsequently generated forecasts.

8. There does not exist a uniform definition of a line of business. Each firm defines LOB in its own way. For example, one oil company may list oil and gas as a separate LOB from refining and processing; another oil company may treat them as a single LOB. In order to conduct cross-sectional tests, we attempted to define lines of business in a uniform manner. Whenever possible, we adopted the LOB coding scheme contained in the 1984 version of the Value Line Data Base.
9. It is not strictly true that all IBES firms have the same timing advantage. IBES forecasts are a consensus of individual forecasts and each forecast is made on a different date (see O'Brien (1985)). Thus, our procedure for adjusting for FAF timing advantage for IBES is only an approximation.
10. Adding firms without continuously available data on the IBES tape involved mechanically adjusting each firm's data. As this was a very time-intensive task, we did not attempt to add any other IBES firms with missing forecast data.
11. The reported t-statistics in Section 6 are derived under the assumption that cross-sectional dependencies do not exist, an assumption that is clearly violated for FAFSUPS. We computed t-

statistics corrected for estimates of cross-sectional correlation in OLS residuals. The significance of reported results, at the .05 level, remained intact. The correction procedure involved calculating the sample covariance matrix,  $S$ , using residuals from the first-pass OLS regression. Because we have fewer time-series observations than firms,  $S$  is singular and thus cannot be inverted. This prevents us from using GLS procedures to estimate and draw inferences about model parameters. However, because  $S$  converges to  $\Sigma$ , the full cross-sectional covariance matrix of regression disturbances, asymptotic inferences can be drawn about the significance of OLS regression coefficients. Rex Thompson suggested and programmed the modification for cross-sectional dependencies.

12. The empirical model adopted [or (7)] fits the data significantly better than the model with a constant intercept over industries or years (quarters). This supports the intuitive belief that a model with constant intercept omits terms that should be present. Such omissions lead to biased parameter estimates and inaccurate tests (Seber, 1977, Section 6.1). For the sake of parsimony in table presentation, the industry and calendar time effects are not reported. Pooling of the data assumes homogeneity of the distribution of FAFSUPS across industries, time periods, and forecast horizons. We tested for and could not reject homogeneity.
13. For each of the three samples (168 quarterly, 168 annual, 702 annual), the maximum possible number of FAF observations is 2016 (168 x 12), 672 (168 x 4), 4212 (702 x 6), respectively. Missing values come about because of missing FAF errors and

- missing DISP observations (mostly cases where DISP is undefined because a single analyst follows the firm). Due to missing FAF observations, the number of non-missing FAF errors declines depending on the forecast horizon. For the annual samples, 1975 FAF observations were unavailable on the IBES tape; accordingly, the two-year-ahead FAF error for fiscal 1977 is always missing.
14. The correlations between predictor variables (while sometimes statistically significant) are not severe. For example, for the one-quarter-ahead horizon, the pairwise Pearson correlation coefficients between LSIZE and LDISP, LSIZE and LOB, and LDISP and LOB are -0.25, -.12, and -0.10, respectively.
  15. The horizon dummies have a positive sign because of our procedure of reflecting the effect of the longest horizon (i.e., three-quarters-ahead for Value Line; two-years-ahead for IBES) in the intercept term. Thus, the dummy variables reflect the effect of the shorter horizons and are expected to have positive coefficients.
  16. The a priori signs of the SIZE, DISP, and LOB variables are positive, negative, and positive, respectively. The LOB variable is of the opposite sign to its theoretical counterpart because we expect LOB to be negatively related to  $\rho$ .

## REFERENCES

- Abdel-khalik, A. R. and R. B. Thompson. "Research on Earnings Forecasts: The State of the Art." The Accounting Journal (1977-78): 180-209.
- Abraham, B. and J. Ledolter. Statistical Methods for Forecasting. New York: Wiley, 1983.
- Albrecht, W. S., D. Johnson, L. L. Lookabill, and D. J. H. Watson. "A Comparison of the Accuracy of Corporate and Security Analysts' Forecasts of Earnings: A Comment." The Accounting Review (July 1977): 736-40.
- Atiase, R. K. "Predisclosure Information, Firm Capitalization, and Security Price Behavior Around Earnings Announcements." Journal of Accounting Research (Spring 1985): 21-36.
- Box, G. E. P. and P. W. Tidwell. "Transformation of the Independent Variables." Technometrics (November 1962): 531-50.
- Brown, L., P. Griffin, R. Hagerman, and M. Zmijewski. "The Existence and Potential Sources of Analyst Forecast Superiority." Working Paper, State University of New York at Buffalo, 1985.
- Brown, P., G. Foster, and E. Noreen. Security Analyst Multi-Year Earnings Forecasts and the Capital Market. Sarasota: American Accounting Association, 1984.

Cukierman, A. and D. Givoly. "Heterogenous Earnings Expectations and Earnings Uncertainty - Theory and Evidence." Working Paper, Tel Aviv University, 1982.

Holthausen, R. W. and R. W. Verrecchia. "The Change in Price Resulting from a Sequence of Information Releases." Working paper, University of Chicago, 1981.

Imhoff Jr., E. A. and P. V. Pare'. "Analysis and Comparison of Earnings Forecast Agents." Journal of Accounting Research (Autumn 1982): 429-39.

Judge, G., R. Hill, W. Griffiths, and T. Lee. The Theory and Practice of Econometrics. New York, Wiley, 1980.

Kross, W. and D. A. Schroeder. "Firm Prominence and the Differential Information Content of Quarterly Earnings Announcements." Working Paper, Purdue University, 1985.

McNichols, M. and J. G. Manegold. "The Effect of the Information Environment on the Relationship Between Financial Disclosure and Security Price Variability." Journal of Accounting and Economics (April 1983): 49-74.

O'Brien, P. "An Empirical Analysis of Forecasts of Earnings per Share." Unpublished Dissertation, University of Chicago, 1985.

Ohlson, J. "On Financial Disclosure and the Behavior of Security Prices." Journal of Accounting and Economics (December 1979): 211-32.

Seber, G. A. F. Linear Regression Analysis. New York: Wiley, 1977.

Shores, D. "Differential Information and Security Returns on the Over-the-Counter Market." Working Paper, University of Chicago, 1985.

TABLE 1

Profile of Value Line FAF Superiority<sup>a</sup>  
by Categories of Predictor Variables  
(Pooled Time-Series Cross-Sectional)

186 Firms, Quarterly Data (1977-1979)

Predictor Variables	Grouping	Cell Median (Number of Observations) <sup>b</sup>		
		1 Quarter Ahead	2 Quarter Ahead	3 Quarter Ahead
SIZE	Under 100 million	1.29 (113)	1.07 (113)	1.00 (113)
	100-500 million	2.00 (727)	1.47 (726)	1.28 (710)
	501 million-1 billion	2.04 (466)	1.57 (465)	1.36 (430)
	Over 1 billion	2.29 (595)	1.79 (593)	1.50 (587)
	TOTAL	2.00(1901)	1.56(1897)	1.33(1840)
DISP	QUINTILE 1	2.90 (347)	2.01 (339)	1.88 (317)
	2	2.00 (347)	1.61 (340)	1.58 (317)
	3	2.00 (347)	1.59 (340)	1.33 (317)
	4	1.94 (348)	1.40 (340)	1.20 (317)
	5	1.83 (348)	1.30 (340)	1.07 (317)
	TOTAL	2.05(1737)	1.57(1699)	1.35(1585)
LOB	1	2.05 (773)	1.53 (770)	1.33 (728)
	2	1.90 (473)	1.57 (469)	1.33 (460)
	3	2.11 (294)	1.48 (294)	1.28 (294)
	4	2.16 (306)	1.68 (308)	1.45 (303)
	5 or more	2.07 (55)	1.78 (56)	1.47 (55)
	TOTAL	2.00(1901)	1.56(1897)	1.33(1840)
TIME	0-30 days	1.63 (118)	1.33 (125)	1.33 (126)
	31-60 days	2.00 (697)	1.49 (697)	1.34 (686)
	61-90 days	2.00 (496)	1.64 (479)	1.34 (427)
	over 90 days	2.29 (590)	1.67 (596)	1.33 (601)
	TOTAL	2.00(1901)	1.56(1897)	1.33(1840)

$${}^a \text{FAFSUPS}_{jt}^{1/2} = |\text{Seasonal RW Error}_{jt} / \text{FAF Error}_{jt}|$$

<sup>b</sup>The number of cases is less than the total possible number of cases (168 firms x 12 quarters = 2016) due to missing values.

TABLE 2

Profile of IBES Consensus FAF Superiority<sup>a</sup>  
 by Categories of Predictor Variables  
 (Pooled Time-Series Cross-Sectional)

Predictor Variables	Grouping	Cell Median (Number of Observations) <sup>b</sup>		
		Final Forecast	1 Year Ahead	2 Year Ahead
A. 168 Firms, Annual Data (1977-1979)				
SIZE	Under 100 million	1.72 (23)	.87 (20)	.65 (9)
	100-500 million	4.44 (169)	1.55 (164)	1.41 (78)
	501 million-1 billion	5.15 (127)	2.04 (126)	2.16 (77)
	Over 1 billion	6.99 (173)	2.63 (173)	2.25 (115)
	TOTAL	4.85 (492)	1.86 (483)	1.83 (279)
DISP	QUINTILE 1	9.01 (87)	3.72 (87)	3.50 (51)
	2	5.99 (87)	2.62 (87)	1.95 (52)
	3	3.91 (88)	1.65 (88)	1.55 (52)
	4	4.78 (88)	1.61 (88)	1.54 (52)
	5	4.29 (88)	1.45 (88)	1.30 (52)
	TOTAL	5.00 (438)	1.88 (438)	1.86 (259)
LOB	1	4.90 (195)	1.78 (187)	1.76 (105)
	2	5.24 (126)	2.01 (126)	2.16 (78)
	3	3.66 (73)	1.71 (73)	1.36 (40)
	4	5.52 (84)	1.89 (84)	2.21 (49)
	5 or more	13.89 (14)	3.86 (13)	2.59 (7)
	TOTAL	4.85 (492)	1.86 (483)	1.83 (279)
B. 702 Firms, Annual Data (1977-1982)				
SIZE	Under 100 million	4.00 (654)	1.60 (651)	1.27 (345)
	100-500 million	4.50(1753)	1.82(1748)	1.57(1201)
	500-1000 million	4.07 (941)	1.87 (940)	1.50 (735)
	Over 1 billion	5.15 (824)	2.14 (824)	1.69 (673)
	TOTAL	4.50(4172)	1.85(4163)	1.55(2954)
DISP	QUINTILE 1	6.29 (762)	2.80 (762)	2.69 (565)
	2	4.63 (762)	2.16 (762)	1.83 (565)
	3	4.22 (762)	1.81 (762)	1.44 (565)
	4	4.00 (763)	1.52 (763)	1.23 (565)
	5	4.00 (763)	1.54 (763)	1.08 (566)
	TOTAL	4.50(3812)	1.86(3812)	1.53(2826)

$${}^a_{\text{FAFSUPSP}}_{jt}^{1/2} = |\text{Annual RW Error}_{jt} / \text{FAF Error}_{jt}|$$

<sup>b</sup>The number of cases is less than the total possible number of cases (168 firms x 3 years = 504; 702 firms x 6 years = 4212) due to missing values.

TABLE 3

Results of Regressions of FAF Superiority Measures<sup>a</sup>  
(Pooled Time-Series Cross-Sectional)

	PREDICTOR VARIABLES					MODEL STATISTICS			
	LSIZE <sub>jt</sub>	LDISP <sub>jt</sub>	LOB <sub>jt</sub>	LTIME <sub>jt</sub>	HOR <sub>1jt</sub>	HOR <sub>2jt</sub>	N	Adj.R <sup>2</sup>	F
A. Value Line Quarterly Sample (168 Firms)									
Betas	.150	-.295	-.021	.284	.301	.132	5021	.04	17.37
T-Stat	(5.23)*	(-6.53)*	(-1.79)	(3.46)*	(9.03)*	(3.94)*			
B. IBES Annual Sample (168 Firms)									
Betas	.440	-.396	.030		.661	.082	1135	.17	24.96
T-Stat	(7.31)*	(-4.09)*	(1.33)		(9.06)*	(1.12)			
C. IBES Annual Sample (702 Firms)									
Betas	.084	-.256			.741	.209	10450	.13	109.14
T-Stat	(4.42)*	(-8.20)*			(30.40)*	(8.57)*			

\*Significant at  $\alpha = .05$ .

<sup>a</sup>For the Value Line quarterly sample, the dependent variable is  $\text{LOG}_{10}(\text{FAFSUPS})$ , measured as  $\text{LOG}_{10} [(\text{Seasonal RW Error}_{jt} / \text{FAF Error}_{jt})^2]$ . For the IBES sample, the dependent variable is  $\text{LOG}_{10}(\text{FAFSUPS})$ , measured as  $\text{LOG}_{10} [(\text{Annual RW Error}_{jt} / \text{FAF Error}_{jt})^2]$ . Prior to taking the log transformation, a small positive constant was added due to zero FAF errors.