THE ACCOUNTING REVIEW
Vol. 78, No. 3
2003
pp. 679–706

# Sophistication-Related Differences in Investors' Models of the Relative Accuracy of Analysts' Forecast Revisions

*Sarah E. Bonner*
*University of Southern California*

*Beverly R. Walther*
*Northwestern University*

*Susan M. Young*
*Emory University*

**ABSTRACT:** The accuracy of sell-side analysts' forecast revisions is related to a number of factors, including characteristics of the analyst and the age of the forecast. In this study we examine whether there are differences in how sophisticated and unsophisticated investors use these factors to predict the relative accuracy of forecast revisions. We adapt the lens model methodological approach from the judgment and decision-making literature to investigate these differences in an archival setting. Our results suggest that sophisticated investors have greater knowledge overall about the relation of the factors to forecast accuracy. Further, our evidence is consistent with sophisticated investors relying more on the specific factors that provide the most benefits (relative to their costs) for predicting relative forecast accuracy.

**Keywords:** *analyst forecast revisions; market reaction; investor sophistication; lens model.*

**Data Availability:** *Data are available from public sources.*

## I. INTRODUCTION

Prior research has shown that the accuracy of sell-side analysts' forecast revisions is systematically associated with characteristics of the analyst and the age of the forecast (e.g., O'Brien 1988; Stickel 1992; Clement 1999; Jacob et al. 1999). In this study,

We thank Linda Bamber, Jan Barton, Michael Clement, Joel Demski, Artur Hugon, Hai Lu, Mark Nelson, Grace Pownall, Leonard Soffer, Geoff Sprinkle, Kristy Towry, Sydney S. Y. Young, two anonymous reviewers, participants at the Behavioral-Financial Research Conference at Emory University, and workshop participants at Michigan State University, Southern Methodist University, University of California, Irvine, University of Illinois at Chicago, and Washington University for helpful comments and discussions. We appreciate the financial support, respectively, of BDO Seidman, LLP and the Marshall School of Business, the Kellogg School of Management, and the Goizueta Business School. We are grateful for the research assistance of Artur Hugon and Wai-Ming Loke. Earnings forecasts are from Zacks Investment Research. Errors or omissions are our responsibility.

**Editor's note:** This paper was accepted by Terry Shevlin, Senior Editor.

*Submitted August 2001*
*Accepted January 2003*

we examine the extent to which investors consider these factors when reacting to forecast revisions. Specifically, we investigate whether there are differences between sophisticated and unsophisticated investors in how they use these factors to predict the relative accuracy of analysts' forecast revisions. We first examine whether sophisticated investors have greater knowledge overall about the factors associated with forecast accuracy. We then examine whether sophisticated investors are better able to identify factors that provide large benefits (relative to their costs) for predicting forecast accuracy, and therefore focus more on these factors.[1]

We address these questions by examining the quality of the models investors implicitly use to predict relative forecast accuracy. We infer sophisticated and unsophisticated investors' prediction models from their stock price reactions to analyst forecast revisions. We then compare these prediction models to a statistical model, which specifies relative forecast accuracy as a function of the factors shown by previous research to be associated with forecast accuracy. We base our inferences regarding differences between sophisticated and unsophisticated investors' prediction models on the "matching index" from Brunswik's lens model of human judgment. The matching index represents the correlation between the fitted values from investors' (implied) models of relative forecast accuracy and the fitted values from the statistical model of relative forecast accuracy. Our measure of investor sophistication is based on five firm characteristics: analyst following, percentage of institutional ownership, number of institutions holding shares, number of shares held by institutions, and dollar value of shares traded. We use a factor analysis of these characteristics to obtain a single measure of investor sophistication for each firm-year observation, and then compute matching indexes separately for sophisticated and unsophisticated investors.

Our first hypothesis is that sophisticated investors have greater knowledge overall about how to use the information in the set of factors that can be used to predict analysts' forecast accuracy. To test this hypothesis, we compare the matching indexes for sophisticated and unsophisticated investors. As predicted, the matching index for sophisticated investors is statistically higher than the matching index for unsophisticated investors.

Our second hypothesis is that sophisticated investors are more knowledgeable about the benefits of individual factors for predicting analysts' relative forecast accuracy. To test this hypothesis, we compare the matching indexes from a model that includes only the two most beneficial factors for predicting forecast accuracy (the age of the forecast and the analyst's prior accuracy) to those from more comprehensive models that also include less beneficial factors. If sophisticated investors ignore or underweight the information in these less beneficial factors more than unsophisticated investors, the decrease in their matching index will be statistically greater than the decrease in unsophisticated investors' matching index when these factors are added to the model. The results are consistent with our prediction, suggesting that sophisticated investors exhibit more "adaptive decision-making." Taken together, the findings suggest that sophisticated investors not only have more knowledge overall about the set of factors related to analysts' forecast accuracy, but they also have greater knowledge of the individual factors that are most beneficial to use.

Our findings contribute to the literature in the following ways. First, we provide information about how different types of investors use analysts' forecasts when predicting the earnings of a firm. Sell-side analysts' earnings forecasts are an important source of information for investors making predictions about the future earnings and profitability of a firm

[1] Although we use the terminology "sophisticated" and "unsophisticated" investors, we recognize that our "unsophisticated" investors are not necessarily untrained, but are simply less sophisticated than those we define as "sophisticated."

and investment decisions based on these predictions (SRI International 1987; Williams et al. 1996; Hodge 2001). While recent research has examined investors' reactions to analysts' forecast revisions conditional on factors that are related to forecast accuracy (Stickel 1992; Mikhail et al. 1997; Park and Stice 2000; Clement and Tse 2003; Gleason and Lee 2003), no previous study has examined sophistication-related differences in these reactions. Investigating sophistication-related differences in investors' reactions, specifically their implied prediction models for the accuracy of sell-side analysts' earnings forecasts, can provide insights into whether and how unsophisticated investors' decisions can be improved. Additionally, our factor analysis of various sophistication proxies may provide guidance to researchers interested in examining other sophistication-related differences in investor behavior.

Our study further contributes to the literature in that we illustrate how a methodological approach from behavioral research (the lens model) can be used to study investor behavior in an archival setting. Because this approach focuses on the extent to which the *output* from investors' implied prediction models matches that from the statistical model, it is appropriate in any setting in which investors are making predictions using factors that are correlated. The lens model accounts for the fact that investors can exploit the correlation among factors and, consequently, make good predictions even if their models contain inaccurate individual factor weightings. However, as with any technique that infers investors' prediction models from market reactions, our approach requires several assumptions. While additional analyses indicate that the assumptions likely hold in our setting, we are unable to rule out all alternative explanations for our findings.

This paper is organized as follows. The next section reviews the literature, followed by a description of our methodological approach in Section III. Section IV develops our hypotheses. Section V discusses the sample, variables, and models we use. Section VI presents the results of our analyses. Section VII contains robustness tests, including analyses to address the validity of the assumptions underlying our empirical tests. Section VIII concludes.

## II. LITERATURE REVIEW

Sell-side analysts' earnings forecasts and forecast revisions are an important source of information for investors making predictions about the future earnings and profitability of a firm and investment decisions based on these predictions (SRI International 1987; Williams et al. 1996; Hodge 2001). Prior research has documented that there is a significant stock market reaction associated with analyst forecast revisions (see, for example, Gonedes et al. 1976; Givoly and Lakonishok 1979; Brown et al. 1985). Overall, this literature indicates that sell-side analysts' forecast revisions provide information to investors.

More recent research has controlled for the sign and magnitude of the forecast revision, and documented that investors' reaction to these revisions varies with factors that are associated with the accuracy of the forecasts. Stickel (1992) finds that there is a stronger price reaction to upward, but not downward, forecast revisions issued by analysts named to the *Institutional Investor* All-American team (see also Gleason and Lee 2003). Park and Stice's (2000) and Chen et al.'s (2001) results indicate that investors react more strongly to forecast revisions issued by analysts with high prior forecasting accuracy. Finally, in an experimental study, Maines (1996) finds that investors consider the historical accuracy of analysts' forecasts when combining them to derive a consensus forecast. These findings suggest that investors appear to be attempting to predict the accuracy of analyst forecasts when reacting to them.[2]

In contrast to prior studies that examined investors' reaction to analyst forecast revisions conditional on only one factor associated with forecast accuracy, Clement and Tse (2003) examine investors' reaction conditional on multiple factors associated with forecast accuracy. They find that the association of investors' reaction with some factors is inconsistent with those factors' association with forecast accuracy. For example, their results indicate that investors react more strongly to forecasts issued earlier; however, earlier forecasts tend to be less accurate (e.g., O'Brien 1990). Consequently, their work suggests that investors do not properly incorporate all of the information in factors for forecast accuracy when reacting to forecast revisions.

Prior studies of investors' reaction to forecast revisions are somewhat limited in providing insights about the quality of investors' prediction of forecast accuracy for three reasons. First, some prior studies have documented that investors' reaction to forecast revisions is a function of factors associated with the accuracy of these revisions, but have not directly assessed the quality of investors' prediction model by examining, for example, whether investors' weightings on the factors correspond to the statistical association between the factors and forecast accuracy (e.g., Stickel 1992). Other studies have investigated the quality of investors' prediction model, but have used the Mishkin (1983) technique to examine the similarity of individual coefficients on factors in the model of investors' reaction to forecast revisions to the corresponding coefficients in the model of forecast accuracy (e.g., Clement and Tse 2003). By contrast, we use the lens model, described in detail in the next section, to assess the quality of investors' prediction model. This approach focuses on the similarity of *fitted values* from investors' prediction model of forecast accuracy to the *fitted values* obtained from a statistical model of forecast accuracy. The lens model implicitly allows for a setting in which investors exploit correlation among the factors by testing the accuracy of the *output* from investors' model rather than the accuracy of their weightings on individual factors.

Second, prior studies on the relation between investors' reaction to forecast revisions and the factors associated with forecast accuracy do not allow for heterogeneity in how investors incorporate the information in factors for forecast accuracy when predicting the accuracy of a revised forecast. However, recent studies have suggested that a firm's stock price reaction to information is associated with proxies for the sophistication of the firm's marginal investor. For example, Bartov et al. (2000) find that the magnitude of the post-earnings announcement drift decreases as the sophistication of the firm's investors (proxied by the percentage of stock held by institutions) increases (also see Hand 1990). We examine differences in the quality of investors' prediction models for forecast accuracy between subsamples formed on the basis of the sophistication of the marginal investor. Finally, there are multiple facets of the "quality" of prediction models. In addition to examining the quality of models in terms of the accuracy of the model's overall predictions for forecast accuracy, we examine whether sophisticated investors better consider the costs and benefits of using individual factors in constructing their models than do unsophisticated investors. In the next two sections, we describe the methodology we use to investigate these issues and our hypotheses concerning these differences in model quality.

[2] Additionally, Brown and Mohammed (2001) find that investors can develop a profitable trading strategy based on an understanding of the accuracy of analysts' forecasts.

## III. METHODOLOGY

To investigate differences between sophisticated and unsophisticated investors as to the quality of their models for predicting analysts' forecast accuracy, we use Brunswik's lens model. Brunswik (and later Hammond) developed the lens model to examine how well individuals assess the relation between factors and outcomes in the environment in making predictions about future outcomes (see Hammond and Stewart [2001] for a review of Brunswik's work). In the traditional application of the lens model, individuals are given a set of factors and are asked to make predictions of outcomes (e.g., analyst forecast accuracy) in a laboratory setting. Using these data, the researcher would calculate a statistical model and each individual decision-maker's model of analyst forecast accuracy, both based on the given set of factors. The correlation between the fitted values from the statistical model and the fitted values from an individual decision-maker's model is called the matching index (see Ashton [1982] for an excellent discussion of the lens model). The matching index measures the quality of the linear component of the decision-maker's model, or, put more simply, the accuracy of the factor signs and weights *only* to the extent they affect the accuracy of the model's output (e.g., predictions of analyst forecast accuracy).[3]

### Adaptation of the Lens Model

The lens model was developed to study individuals' predictions. By contrast, in our setting, we have archival market-level data, requiring us to infer investors' predictions from market prices. Many assumptions must hold in order for our application of this technique to yield valid inferences. For example, we must assume that all investors have equal access both to analyst forecast revisions and to the factors that can be used to predict their accuracy. In Section VII, we outline the necessary assumptions, provide empirical evidence on the appropriateness of these assumptions where possible, and assess the sensitivity of our results to these assumptions.

To illustrate how we adapt the lens model to our setting, assume for ease of exposition that there is only one variable that is informative for predicting the accuracy of the analyst's forecast revision, the analyst's prior accuracy relative to all other analysts who follow the firm (*PRMAPE*). Therefore, the equation for the relative accuracy of the analyst's forecast revision is:

$$PRMAPE_{i,j,q} = \alpha_0 + \alpha_1 PRMAPE_{i,j,q-1} + \varepsilon_{i,j,q} \quad (1)$$

where:

$PRMAPE_{i,j,q}$ = the absolute value of analyst $i$'s forecast error for firm $j$ and quarter $q$ minus the mean absolute forecast error for firm $j$ and quarter $q$, deflated by the mean absolute forecast error for firm $j$ and quarter $q$; and

$\varepsilon_{i,j,q}$ = error term with assumed zero mean and constant variance.

The fitted values from Equation (1) are the statistical model's predictions of the relative accuracy of the analyst's forecast revision ($PR\hat{M}APE^{STAT}_{i,j,t}$).

[3] Many studies calculate "achievement" rather than or in addition to the matching index. Achievement is the correlation of an individual's predictions and the actual outcomes, and is influenced by noise in the environment (how highly associated with the outcome the set of factors is), matching of factor signs and weights, and noise in the individual's use of his prediction policy. We focus solely on the matching component of overall achievement in order to examine directly the quality of investors' models.

Further assume that there is only one type of investor, and that investors consider the predicted relative accuracy of an analyst's earnings forecast revision when reacting to its release:[4]

$$CAR_{i,j,t} = \beta_0 + \beta_1 FCREV_{i,j,t} \cdot PRMAPE^{PRED}_{i,j,t} + \xi_{i,j,t} \tag{2}$$

where:

$CAR_{i,j,t}$ = three-day ($t-1$, $t+1$) size-adjusted return surrounding the issue of the revised quarterly earnings forecast by analyst $i$ for firm $j$ at time $t$;
$FCREV_{i,j,t}$ = the price-deflated revision in analyst $i$'s forecast for firm $j$ issued at time $t$;
$PRMAPE^{PRED}_{i,j,t}$ = investors' predicted relative accuracy of analyst $i$'s forecast for firm $j$ issued at time $t$; and
$\xi_{i,j,t}$ = error term with assumed zero mean and constant variance.

In the lens model framework, the matching index is the correlation between the fitted values from the statistical model and the fitted values from an individual decision-maker's model. Loosely speaking, this is equivalent to calculating the correlation between $PRMAPE^{PRED}_{i,j,t}$ in Equation (2) with the fitted values, $PR\hat{M}APE^{STAT}_{i,j,t}$, from Equation (1).

To infer investors' prediction of the relative accuracy of the forecast revision, we substitute Equation (1) into Equation (2):

$$CAR_{i,j,t} = \beta_0 + \beta_1 FCREV_{i,j,t} \cdot (\alpha'_0 + \alpha'_1 PRMAPE_{i,j,q-1}) + \xi_{i,j,t} \tag{3}$$

where $\alpha'_0$ and $\alpha'_1$ equal investors' weightings used to predict the relative accuracy of the analyst's forecast revision. Therefore, to infer investors' prediction of the relative accuracy of the forecast revision ($PR\hat{M}APE^{INV}_{i,j,t}$), we replace the estimated coefficients in Equation (1) with these estimated $\alpha'$'s from Equation (3). The matching index can be expressed as:

$$Matching\ Index = corr\ (PR\hat{M}APE^{STAT}_{i,j,t}, PR\hat{M}APE^{INV}_{i,j,t}). \tag{4}$$

The lens model has not been used previously in an empirical study of market-level prices, and differs from the Mishkin (1983) approach traditionally used in the literature.[5] We do not use the Mishkin approach because, as discussed in the next section, our hypotheses rely on the existence of different types of investors (sophisticated and unsophisticated) and predict that sophisticated investors' model is of higher quality than unsophisticated investors' model. One reason this result could occur is if sophisticated investors' factor weightings are closer to the weights from a statistical model of forecast accuracy than unsophisticated investors' weightings.[6] In the extreme, we could use the Mishkin (1983) approach to test whether sophisticated investors' $\alpha'$'s from Equation (3)

[4] Consistent with behavioral studies that use the lens model methodology, we assume investors use a linear decision model.

[5] The Mishkin (1983) approach tests whether the estimated coefficients from the statistical model (the $\alpha$'s from Equation (1)) equal the estimated coefficients from the investor model (e.g., the $\alpha'$'s from Equation (3)).

[6] Stated in the context of Equations (1) and (3), if sophisticated investors' $\alpha'$'s from Equation (3) are closer to the $\alpha'$'s from Equation (1) than are unsophisticated investors' $\alpha$'s from Equation (3) (i.e., $|\alpha'^{SOPH} - \alpha| < |\alpha'^{UNSOPH} - \alpha|$), then the fitted values derived from sophisticated investors' model would more closely match the fitted values derived from the statistical model.

equal the $\alpha$'s from Equation (1), and test separately whether unsophisticated investors' $\alpha''$s from Equation (3) equal the $\alpha$'s from Equation (1). However, if we reject the null hypothesis that $\alpha = \alpha'$ for both investor types, then we have not necessarily rejected (or failed to reject) the null hypothesis that the quality of sophisticated investors' model of the relative accuracy of an analyst's forecast revision is equal to the quality of unsophisticated investors' model. The lens model allows us to compare the models of sophisticated and unsophisticated investors without this difficulty.

Second, research has identified factors aside from the analyst's prior accuracy that are related to the relative accuracy of analysts' forecast revisions (e.g., Clement 1999; Jacob et al. 1999; Brown 2001). The environment in which analysts work is one that creates correlation among these factors. For example, firm-specific experience and general experience tend to be highly correlated; this occurs because analysts tend to cover firms for lengthy periods of time rather than periodically switching the firms they cover. Consequently, investors may weight correlated factors (Equation (3)) in a manner that differs from their weighting in the statistical model (Equation (1)) but still make predictions of the relative accuracy of analysts' forecast revisions that are similar to those that would be derived from the statistical model. For example, an investor who relies mostly on firm-specific experience when predicting forecast accuracy (and, thus, puts little weight on general experience) could make very similar predictions to those derived from a statistical model that relies mostly on general experience. As discussed in the previous section, the lens model methodology implicitly allows for the situation in which investors exploit the correlation among the factors in the environment by focusing on investors' overall predictions, rather than their weightings of individual factors. In essence, the lens model methodology accounts for differences in factor signs and weights between investors and the statistical model, but only to the extent they have an effect on the accuracy of investors' predictions. Further, to the extent we omit factors that are correlated with the factors we include, the fitted values from investors' models likely reflect the use of these correlated factors.

While the lens model and its matching index have many desirable properties for testing our hypotheses, as discussed above, many assumptions must hold in order for the application of a technique that seeks to infer investor behavior from market reaction data to yield valid inferences. In essence, there must be a link between individual investors' predictions of the accuracy of an analyst forecast based on observable factors and the market reaction to that forecast revision. Therefore, the conclusions we draw from our empirical results are subject to the assumptions necessary for our methodological approach, as outlined in Section VII.

## IV. HYPOTHESES

We use the lens model to investigate two differences in the quality of sophisticated and unsophisticated investors' (implied) models for predicting the accuracy of analysts' forecast revisions. Our first hypothesis is that sophisticated investors have greater knowledge overall than unsophisticated investors about how to use the information in the set of factors that predict relative forecast accuracy, resulting in a higher matching index for sophisticated investors than for unsophisticated investors. Sophisticated investors may have greater knowledge because they have more experience with analyst forecasts through forecasting earnings on a full-time basis (Yunker and Krehbiel 1988; Potter 1992) and through following a larger number of stocks (Barber and Odean 2000); more experienced individuals have more knowledge about factor signs and weights (e.g., Slovic 1969; Einhorn 1974; Johnson et al. 1981; Johnson 1988; Bonner 1990). Further, sophisticated investors may be better able to learn the appropriate signs and weights for factors associated with forecast accuracy because they have superior abilities (Beaver 1998), and people with superior abilities learn better from experience (Ackerman et al. 1989; Bonner and Walker 1994). In addition, sophisticated investors may have access to resources that assist them in learning about analysts' forecasts (such as analyst rating services) that unsophisticated investors do not. Finally, sophisticated investors may have greater incentives to learn about the factors associated with analysts' forecast accuracy since their decisions based on analysts' forecasts likely have greater financial consequences, assuming they manage larger portfolios (Cready 1988; Hassel and Norman 1992; Brous and Kini 1994; El-Gazzar 1998). Therefore, our first hypothesis, stated in alternative form, is:

**H1:** The matching index for sophisticated investors' model of relative forecast accuracy will be greater than the matching index for unsophisticated investors' model of relative forecast accuracy.

Our second hypothesis concerns differences in the degree to which sophisticated and unsophisticated investors' models reflect cost-benefit trade-offs related to the individual factors associated with forecast accuracy. Psychology research suggests people will employ "adaptive decision making" by choosing factor-weighting strategies based on their expected benefits and costs (Bettman et al. 1990; Payne et al. 1993). The key potential benefit from correctly incorporating information in a given factor is making superior predictions of the accuracy of forecasts; the costs include the cognitive effort of acquiring and processing the information needed to use the factor in a prediction model. The net benefit of factors investors use to predict analyst forecast accuracy can vary substantially. For example, the age of the forecast is very beneficial for predicting its accuracy (e.g., O'Brien 1988; Brown 2001), and it is not costly to obtain and use—the investor simply needs to know the date on which the forecast is issued and the time period to which it pertains. By contrast, the explanatory power of the number of firms followed by the analyst for forecast accuracy is not as great, and the costs of acquiring and using this information are greater—the investor must gather information on all firms the analyst follows, regardless of whether the investor follows all the firms himself.

We expect that sophisticated investors' model will better reflect the net benefit of using individual factors because they have greater knowledge of the net benefit of the factors that predict forecast accuracy. The psychology literature suggests that the sophistication (experience) of the decision maker may create differences in factor-weighting strategies because experience affects the accuracy of the benefit expectations. Payne et al. (1993) and Fennema and Kleinmuntz (1995) note that people must have extensive experience making and receiving feedback about their judgments in order to learn to anticipate the extent to which focusing on various factors results in better predictions. These studies also have shown few sophistication-related differences in expectations about the costs of using factors for predictions. In our setting, this finding implies that sophisticated and unsophisticated investors may be equally knowledgeable about the costs of acquiring and processing various factors, but that there will be sophistication-related differences in investors' knowledge of the expected benefits. These differences would lead to the investor model better distinguishing between high net benefit factors and low net benefit factors when the marginal investor is more likely to be sophisticated.

To illustrate how we derive our second hypothesis, consider an extreme case in which sophisticated investors only focus on the two most beneficial factors (e.g., forecast age and prior accuracy) to predict relative forecast accuracy, and ignore all other factors. When the statistical model includes only forecast age and prior accuracy, sophisticated investors' matching index should be very high. When another (less beneficial and more costly) factor (e.g., number of firms followed) is added to the statistical model, we predict that sophisticated investors' matching index will decrease because they ignore or underweight the information in the number of firms followed. Therefore, to examine whether sophisticated investors' model better reflects cost-benefit trade-offs related to the factors associated with forecast accuracy, we examine whether sophisticated investors' matching index decreases more than unsophisticated investors' matching index as factors with increasingly lower net benefit are added to the statistical model.[7] Our second hypothesis, stated in alternative form, is:

**H2:** The matching index for sophisticated investors' model of relative forecast accuracy will decrease more as factors lower in net benefit are added to a statistical model that includes only high-benefit factors than will the matching index for unsophisticated investors' model of relative forecast accuracy.

The next section discusses the data we use to test our hypotheses.

## V. DATA AND DESCRIPTIVE STATISTICS

### Sample Selection

Our initial sample contains 208,240 revisions in quarterly earnings forecasts from the Zacks Investment Research database for the period 1981–1999. These observations represent forecast revisions issued closest in time to the quarterly earnings announcement date by individual analysts with sufficient data to calculate our dependent and independent variables (described below).[8] We eliminate 5,625 revisions that are less than five calendar days before the quarterly earnings announcement. We impose this minimum forecast horizon so that we can attribute the market reaction to the forecast revision and not the release of actual earnings. We also require that the revised forecast be issued after the beginning of the fiscal quarter to which it pertains to ensure that our results are not due to the inclusion of "stale" forecasts in our sample. This requirement eliminates 47,432 observations.[9] We control for time-period and firm effects in our empirical tests by mean-adjusting each variable; this procedure eliminates the need to include any firm characteristics or time-period variables in the regression model (see Clement 1999). Because we mean-adjust each variable, we further require a minimum of two analysts for each firm and quarter. This requirement eliminates 25,744 observations. We eliminate 22,863 observations with missing values for our proxies for investor sophistication (described below). Finally, to minimize

[7] We predict that the benefits and costs of incorporating the information in various factors when predicting analysts' forecast accuracy vary across *factors*. However, the benefits and costs of incorporating information from various factors also may vary between sophisticated and unsophisticated *investors*. It is reasonable to expect that sophisticated investors may have greater benefits from making good predictions about the accuracy of analysts' forecasts (Cready 1988; Hassel and Norman 1992; Brous and Kini 1994; El-Gazzar 1998) and/or lower costs of acquiring and processing factors to make predictions about forecast accuracy (Wilson 1975; Lev 1988). However, this would bias against finding results consistent with our hypothesis. If, for example, the net benefit of using the number of firms followed is greater for sophisticated investors than for unsophisticated investors, then we would expect a *smaller* decrease (or larger increase) in sophisticated investors' matching index compared to unsophisticated investors' matching index when this factor is added to the statistical model.

[8] We eliminate forecast revisions that are associated with 345 analyst codes on Zacks that correspond either to an unidentified individual (e.g., Value Line) or a brokerage house. Our results are unchanged if we retain all quarterly forecast revisions issued by individual analysts in our analysis instead of only the one issued closest in time to the quarterly earnings announcement date.

[9] Our results are unchanged if we retain forecasts issued prior to the beginning of the fiscal quarter in our analyses.

the effect of outliers on our results, we eliminate 5,338 observations in the extreme 1 percent of the distribution of our two dependent variables, forecast accuracy and the market reaction to the forecast revision (see Clement 1999).[10] Our final sample includes 101,238 observations, representing 3,290 analysts and 1,757 firms from 1981 to 1999.

### Variable Definitions and Descriptive Statistics

Table 1 provides descriptive statistics for our variables. The mean (median) absolute percentage forecast error, defined as the absolute value of actual quarterly earnings per share minus the quarterly earnings forecast, deflated by price ten trading days before the release of the forecast, is 0.63 percent (0.17 percent). Consistent with prior findings that analysts revise their forecasts downward throughout the period (Richardson et al. 2001), the forecast revisions tend to be negative. The mean forecast revision (*FCREV*), defined as the revised forecast minus the prior forecast, deflated by price ten trading days before the release of the revised forecast, is –0.0016. The median forecast revision, however, is 0.0000. Consistent with the negative mean forecast revision, the market reaction at the forecast

**TABLE 1**
**Descriptive Statistics for Quarterly Forecast Revisions (n = 101,238)**

| Variable[a] | Mean | Median | First Quartile | Third Quartile | Std. Dev. |
|---|---|---|---|---|---|
| *MAPE* | 0.0063 | 0.0017 | 0.0006 | 0.0058 | 0.0130 |
| *FCREV* | −0.0016 | 0.0000 | −0.0021 | 0.0000 | 0.0059 |
| *CAR* | −0.0021 | −0.0017 | −0.0234 | 0.0209 | 0.0497 |
| *FCAGE* | 49.7878 | 44.0000 | 22.0000 | 78.0000 | 31.0522 |
| *FIRMEXP* | 13.8311 | 10.0000 | 6.0000 | 19.0000 | 11.1747 |
| *GENEXP* | 327.3060 | 244.0000 | 113.0000 | 452.0000 | 311.0560 |
| *TURNOVER* | 0.1235 | 0.0000 | 0.0000 | 0.0000 | 0.3290 |
| *FCFREQ* | 4.6819 | 4.0000 | 3.0000 | 6.0000 | 2.5195 |
| *NOFIRM* | 18.0610 | 16.0000 | 12.0000 | 21.0000 | 10.3890 |
| *NOIND* | 5.2896 | 5.0000 | 3.0000 | 7.0000 | 3.2588 |
| *BROKSIZE* | 41.1839 | 35.0000 | 18.0000 | 58.0000 | 28.5140 |
| *IIAWARD* | 0.2161 | 0.0000 | 0.0000 | 0.0000 | 0.4116 |

[a] Variable definitions: The Mean Absolute Percentage Error (*MAPE*) is the absolute value of actual quarterly earnings per share minus the quarterly earnings forecast, deflated by price ten trading days before the release of the forecast. *FCREV* is the revised forecast minus the prior forecast, deflated by price ten trading days before the release of the revised forecast. *CAR* is the three-day size-adjusted return centered on the forecast release date. *FCAGE* is the number of calendar days between the firm's earnings announcement date and the forecast release date. *FIRMEXP* is the number of prior quarters for which the analyst has issued an earnings forecast for the firm. *GENEXP* is the number of prior firm quarters the analyst has issued a quarterly forecast for any firm on the Zacks database. *TURNOVER* is an indicator variable that equals 1 if the analyst changed brokerage houses during the year, and 0 otherwise. *FCFREQ* is the number of quarterly earnings forecast issued by the analyst for the firm and quarter. *NOFIRM* is the number of firms for which the analyst issues quarterly earnings announcement forecasts for the quarter. *NOIND* is the number of two-digit SIC codes containing firms for which the analyst issues quarterly earnings announcements. *BROKSIZE* is the number of analysts issuing quarterly earnings forecasts at the brokerage house at which the analyst is employed. *IIAWARD* is an indicator variable that equals 1 if the analyst was named to the *Institutional Investor* All-American team in the previous year.

[10] Our results are not sensitive to using alternative outlier detection techniques, such as eliminating observations on the basis of studentized residuals or Cook's D in the regression estimations.

revision date also tends to be negative. The mean (median) three-day cumulative size-adjusted return centered on the forecast release date, defined as the firm's compounded raw return minus the compounded return on the size decile portfolio to which the firm belongs at the beginning of the calendar year, is −0.21 percent (−0.17 percent).[11]

Table 1 also provides descriptive statistics on the factors associated with relative forecast accuracy. We base our selection of the factors on prior research; these factors include prior accuracy, forecast age, firm experience, general experience, forecast frequency, number of firms followed, number of industries followed, and brokerage house size (e.g., Clement 1999; Jacob et al. 1999). In addition to the factors considered in prior studies, we also include indicator variables for analyst turnover and award status. We include turnover because Mikhail et al. (1999) find that turnover is negatively related to relative forecast accuracy. We include award status because Stickel (1992) documents that analysts named to the *Institutional Investor* All-American team issue more accurate forecasts.

The forecast revisions are issued an average of 49.79 calendar days prior to the quarterly earnings announcement date; the median forecast age (*FCAGE*) is 44 calendar days. The sample analysts tend to have followed the firm for about three years; the mean (median) number of prior quarters for which the analyst has issued a quarterly earnings forecast for the firm, *FIRMEXP*, is 13.83 (10.0). The mean (median) level of general experience (*GENEXP*), defined as the number of prior firm-quarters the analyst has issued a quarterly forecast for any firm on the Zacks database, is 327.31 (244.0).[12]

*TURNOVER* is an indicator variable that equals 1 if the analyst changed brokerage houses during the year, 0 otherwise. The mean value of *TURNOVER* indicates that 12.35 percent of our observations correspond to forecast revisions issued by analysts who changed employers during the year. Examining the frequency of turnover of our sample analysts, we find that approximately 33 percent of our sample analysts change brokerage houses at least once during our sample period.

The sample analysts tend to be active forecasters; the mean (median) number of quarterly forecasts issued by the analyst for a firm per quarter, *FCFREQ*, is 4.68 (4.0). The sample analysts issue quarterly earnings forecasts for an average of 18.06 firms during the quarter (*NOFIRM*) and follow an average of 5.29 different two-digit SIC codes (*NOIND*). *BROKSIZE* measures the size of the brokerage house, defined as the number of analysts issuing earnings forecasts at the brokerage house during the year. The average brokerage house represented in the sample has 41.18 analysts issuing earnings forecasts; the median brokerage house size is 35 analysts. The mean value of *IIAWARD* indicates that 21.61 percent of our sample observations correspond to forecast revisions issued by analysts who were named to the *Institutional Investor* All-American team in the prior year. Examining the frequency of award listing of our sample analysts, we find that approximately 17 percent of our sample analysts were named to the *Institutional Investor* All-American team at least once during our sample period.

## VI. EMPIRICAL RESULTS

We first provide the results from using the lens model methodology on the entire sample, including estimation results for the factors associated with relative forecast accuracy

---

[11] All empirical results hold if we define CAR as the firm's cumulative return less the cumulative return on the value- or equal-weighted market index.

[12] These measures of experience are calculated from the Zacks database, which begins only in 1980. As a result, the experience measures may be understated, especially for analyst-firm observations in the early years of our sample. All conclusions, however, hold if we estimate our empirical models using observations from 1993 to 1999 or from 1997 to 1999.

and investors' reaction to forecast revisions conditional on these factors. This analysis ensures that the findings in prior research on the factors related to forecast accuracy hold in our sample. Then, we present the results from testing our hypotheses.

**Results: Overall Sample**

In order to calculate the matching index, we first estimate the following expanded version of Equation (1), which represents the statistical model of relative forecast accuracy:

$$\begin{aligned} PRMAPE_{i,j,q} = {} & \alpha_0 + \alpha_1 RFCAGE_{i,j,q} + \alpha_2 PRMAPE_{i,j,q-1} + \alpha_3 RFIRMEXP_{i,j,q} \\ & + \alpha_4 RGENEXP_{i,j,q} + \alpha_5 RTURNOVER_{i,j,q} + \alpha_6 RFCFREQ_{i,j,q} \\ & + \alpha_7 RNOFIRM_{i,j,q} + \alpha_8 RNOIND_{i,j,q} + \alpha_9 RBROKSIZE_{i,j,q} \\ & + \alpha_{10} RIIAWARD_{i,j,q} + \varepsilon_{i,j,q} \end{aligned} \tag{5}$$

where:

$PRMAPE_{i,j,q}$ = the absolute value of analyst $i$'s forecast error for firm $j$ and quarter $q$ minus the mean absolute forecast error for firm $j$ and quarter $q$, deflated by the mean absolute forecast error for firm $j$ and quarter $q$;

$RFCAGE_{i,j,q}$ = the number of calendar days between the issuance of analyst $i$'s forecast for firm $j$ and quarter $q$ minus the mean forecast age for firm $j$ and quarter $q$;

$RFIRMEXP_{i,j,q}$ = the firm experience of analyst $i$ for firm $j$ at quarter $q$ minus the mean firm experience for firm $j$ and quarter $q$;

$RGENEXP_{i,j,q}$ = the general experience of analyst $i$ at quarter $q$ minus the mean general experience for firm $j$ and quarter $q$;

$RTURNOVER_{i,j,q}$ = an indicator variable that equals 1 if analyst $i$ changed brokerage houses during the year in which quarter $q$ falls (0 otherwise) minus the mean of the indicator variable for firm $j$ and quarter $q$;

$RFCFREQ_{i,j,q}$ = the number of forecasts issued by analyst $i$ for firm $j$ and quarter $q$ minus the mean forecast frequency for firm $j$ and quarter $q$;

$RNOFIRM_{i,j,q}$ = the number of firms followed by analyst $i$ at quarter $q$ minus the mean number of firms followed for firm $j$ and quarter $q$;

$RNOIND_{i,j,q}$ = the number of industries followed by analyst $i$ at quarter $q$ minus the mean number of industries followed for firm $j$ and quarter $q$;

$RBROKSIZE_{i,j,q}$ = the size of the brokerage house at which analyst $i$ is employed at quarter $q$ minus the mean brokerage house size for firm $j$ and quarter $q$;

$RIIAWARD_{i,j,q}$ = an indicator variable that equals 1 if analyst $i$ was named to the *Institutional Investor* All-American list during the year prior to that in which quarter $q$ falls (0 otherwise) minus the mean of the indicator variable for firm $j$ and quarter $q$; and

$\varepsilon_{i,j,q}$ = error term with assumed zero mean and constant variance.

The fitted values from the statistical model, $\widehat{PRMAPE}^{STAT}_{i,j,t}$, are those from Equation (5).

To obtain fitted values from investors' (implied) prediction model of relative forecast accuracy, we separately estimate the following expanded version of Equation (3):

$$\begin{aligned} CAR_{i,j,t} = \beta_0 &+ \alpha_0' FCREV_{i,j,t} + \alpha_1' FCREV_{i,j,t} * RFCAGE_{i,j,q} \\ &+ \alpha_2' FCREV_{i,j,t} * PRMAPE_{i,j,q-1} + \alpha_3' FCREV_{i,j,t} * RFIRMEXP_{i,j,q} \\ &+ \alpha_4' FCREV_{i,j,t} * RGENEXP_{i,j,q} + \alpha_5' FCREV_{i,j,t} * RTURNOVER_{i,j,q} \\ &+ \alpha_6' FCREV_{i,j,t} * RFCFREQ_{i,j,q} + \alpha_7' FCREV_{i,j,t} * RNOFIRM_{i,j,q} \\ &+ \alpha_8' FCREV_{i,j,t} * RNOIND_{i,j,q} + \alpha_9' FCREV_{i,j,t} * RBROKSIZE_{i,j,q} \\ &+ \alpha_{10}' FCREV_{i,j,t} * RIIAWARD_{i,j,q} + \zeta_{i,j,q} \end{aligned} \tag{6}$$

where:

$CAR_{i,j,t}$ = three-day ($t-1$, $t+1$) size-adjusted return surrounding the issue of the revised quarterly earnings forecast by analyst $i$ for firm $j$ at time $t$;
$FCREV_{i,j,t}$ = the price-deflated revision in analyst $i$'s forecast for firm $j$ issued at time $t$; and
$\zeta_{i,j,t}$ = error term with assumed zero mean and constant variance.

Investors' fitted values, $PR\hat{M}APE_{i,j,t}^{INV}$, are obtained from replacing the estimated coefficients in Equation (5) with the estimated coefficients from Equation (6).[13] As described in Equation (4), the matching index is the correlation between the fitted values from the statistical model (Equation (5)) and the fitted values from investors' implied prediction model (Equation (6)).

The results from estimating Equation (5) on the entire sample are provided in the first two columns of Table 2. Consistent with prior research (e.g., O'Brien 1988), the relative age of the forecast, *RFCAGE*, is positively associated with *PRMAPE*. Since higher levels of *PRMAPE* represent less accurate forecasts, the positive coefficient indicates that forecasts issued earlier tend to be less accurate.

Although relative forecast age is the dominant explanatory variable for relative forecast accuracy, consistent with Brown and Mohammed (2001), the other independent variables do add significant explanatory power beyond that contained in relative forecast age. The estimated coefficient on relative prior accuracy is positive and statistically significant, indicating that forecasts issued by analysts who were less accurate in the past are less accurate in the current period.[14] This finding supports Brown (2001) who documents that prior accuracy is an important determinant of current accuracy.[15]

The results for the two experience variables, *RFIRMEXP* and *RGENEXP*, suggest that forecasts issued by more experienced analysts are more accurate. Inconsistent with Mikhail et al.'s (1999) result that analysts who change brokerage houses tend to be less accurate,

[13] In Equation (5), higher levels of *PRMAPE* indicate less accurate forecasts. In Equation (6), we expect that investors react more strongly to forecasts that are predicted to be more accurate. Therefore, we multiply the estimated coefficients in Equation (6) by $-1$ before substituting them into Equation (5).

[14] Replacing $PRMAPE_{q-1}$ by a variable representing the analyst's historical accuracy for the firm (measured over the prior four or eight quarters) yields virtually identical results.

[15] Despite the importance of prior accuracy and forecast age, the other independent variables add significant explanatory power beyond that contained in relative forecast age and relative prior accuracy. The F-statistic testing the null hypothesis that all independent variables except *RFCAGE* and *PRMAPE* equal zero is significantly different from zero ($F = 20.14$, $p < 0.01$, not tabulated).

**TABLE 2**
**Factors Associated with Relative Forecast Accuracy and the Market Reaction to Analyst Forecast Revisions: Entire Sample (n = 101,238)[a]**

| Equation (5): Forecast Accuracy | | Equation (6): Market Reaction | |
|---|---|---|---|
| Intercept | 0.0000<br>(0.00) | Intercept | −0.0008**<br>(−4.82) |
| | | *FCREV* | 0.8491**<br>(31.75) |
| *RFCAGE* | 0.0038**<br>(46.71) | **RFCAGE* | 0.0004<br>(0.38) |
| $PRMAPE_{q-1}$ | 0.0399**<br>(13.53) | *$PRMAPE_{q-1}$ | −0.1645**<br>(−4.39) |
| *RFIRMEXP* | −0.0009**<br>(−2.83) | **RFIRMEXP* | 0.0081*<br>(1.72) |
| *RGENEXP* | −0.0000**<br>(−3.28) | **RGENEXP* | −0.0003<br>(−1.35) |
| *RTURNOVER* | 0.0036<br>(0.51) | **RTURNOVER* | −0.0600<br>(−0.60) |
| *RFCFREQ* | −0.0020*<br>(−2.05) | **RFCFREQ* | 0.0758**<br>(5.80) |
| *RNOFIRM* | 0.0013**<br>(3.48) | **RNOFIRM* | −0.0053<br>(−0.98) |
| *RNOIND* | 0.0042**<br>(4.05) | **RNOIND* | −0.0269*<br>(−1.85) |
| *RBROKSIZE* | 0.0005**<br>(5.36) | **RBROKSIZE* | 0.0011<br>(0.80) |
| *RIIAWARD* | −0.0209**<br>(−3.18) | **RIIAWARD* | −0.0845<br>(−1.01) |
| Adjusted $R^2$ | 2.54% | Adjusted $R^2$ | 1.00% |
| F-statistic[b] | 39.45** | F-statistic[b] | 7.86** |
| | Matching Index = 0.2739 | | |

*, ** $p < 0.10$ and $p < 0.01$, respectively, two-tailed.

[a] This table provides the estimated coefficients, t-statistics (in parentheses below the estimated coefficient), and adjusted $R^2$s from estimating Equations (5) and (6) on the sample of 101,238 quarterly forecast revisions. In Equation (5), the dependent variable is *PRMAPE*, defined as the analyst's absolute forecast error for the firm and quarter, minus the mean absolute forecast error of other analysts following the firm for that quarter, deflated by the mean absolute forecast error. In Equation (6), the dependent variable is *CAR*, defined as the three-day size-adjusted return centered on the forecast release date. The prefix "R" for the independent variables indicates that the mean value of the variable for the firm and quarter is subtracted from the observation's value. Variables are defined in Table 1.

[b] For Equation (5), the F-statistic tests the null hypothesis that all independent variables except *RFCAGE* equal zero. For Equation (6), the F-statistic tests the null hypothesis that all interaction terms equal zero.

the estimated coefficient on *RTURNOVER* is insignificant in the full sample estimations. The estimated coefficient on *RFCFREQ* is significantly negative, indicating that analysts who revise their forecasts more frequently are more accurate. The estimated coefficients on *RNOFIRM* and *RNOIND* are positive and significant, indicating that forecasts issued by analysts who follow more firms or more industries tend to be less accurate. In contrast to prior work, we find that the estimated coefficient on *RBROKSIZE* is significantly positive, indicating that forecasts issued by analysts who are employed by larger brokerage houses tend to be less accurate.[16] Finally, the significantly negative coefficient on *RIIAWARD* is consistent with award-winning analysts issuing more accurate forecasts.[17] Overall, these results indicate that the relations between forecast accuracy and these factors documented in prior research (e.g., Clement 1999; Jacob et al. 1999) generally hold in our sample.

Columns three and four in Table 2 provide the results for estimating Equation (6) on the entire sample. The market reaction surrounding the release of the revised forecast is, as expected, positively associated with the signed magnitude of the forecast revision. The interaction terms of *FCREV* with the determinants of forecast accuracy provide incremental explanatory power for the market reaction beyond that contained in *FCREV*; the null hypothesis that all interaction terms are zero is rejected at two-tailed $p < 0.01$.

The results on whether investors, on average, react to factors associated with relative forecast accuracy in a manner consistent with how these factors are related to accuracy are mixed. For example, investors react less strongly to forecasts issued by analysts who were less accurate in the prior quarter, consistent with the relation of this factor to accuracy. Investors also react more strongly to forecasts issued by analysts who have more firm-specific forecasting experience, who are frequent forecasters, and who follow relatively fewer industries. However, they tend not to react differentially to forecasts of different age, inconsistent with this variable's clear importance in predicting accuracy. Further, the results indicate that investors, on average, do not react differentially to forecast revisions issued by analysts who change employers, by analysts who follow more firms, or by analysts with *Institutional Investor* All-American status.

Based on these estimations, the matching index for the entire sample is 0.2739, which is significantly less than 1 at $p < 0.01$ based on the Fisher Z-statistic (Zar 1999). These findings are consistent with those of Clement and Tse (2003), who reject the null hypothesis that the estimated coefficients on the factors they consider in the statistical model equal the coefficients on their factors in the market reaction model using the Mishkin (1983) technique. In the next section, we present the results of our main investigation into differences in the quality of sophisticated and unsophisticated investors' models for predicting forecast accuracy.

## Results: Hypothesis Tests for the Effects of Investor Sophistication

In this section, we first test H1 by examining whether the fitted values from investors' implied prediction model of the relative accuracy of forecast revisions more closely match the fitted values from the statistical model when the marginal investor is more likely to be sophisticated. Second, to test H2, we compare the decrease in the matching index from the model that contains the two high-benefit factors to more comprehensive models for sophisticated investors to the same decrease for unsophisticated investors.

### *Sophistication Characteristics*

Table 3 provides descriptive statistics on the investor sophistication characteristics we examine. We use firm characteristics to proxy for the sophistication of the marginal investor,

[16] This finding, however, appears to be concentrated in observations from a few years. When Equation (5) is estimated by year, the coefficient on *RBROKSIZE* is significantly positive in only four of the 19 years.

[17] The results are qualitatively similar if we replace *RIIAWARD* with *RWSJAWARD*, an indicator variable that equals 1 if the analyst was named to the *Wall Street Journal* All-Star Analyst list during the year prior to that in which quarter $q$ falls minus the mean of the indicator variable for firm $j$ and quarter $q$. We tabulate the results for *RIIAWARD* since data on the membership of the *Institutional Investor* All-American team are available for our entire sample period, whereas data to calculate *RWSJAWARD* are only available beginning in 1993.

**TABLE 3**
**Descriptive Statistics on Investor Sophistication Characteristics**

**Panel A: Descriptive Statistics**

| Variable[a] | Mean | Median | First Quartile | Third Quartile | Std. Dev. |
|---|---|---|---|---|---|
| *ANFOLL* | 13.24 | 12.00 | 8.00 | 18.00 | 7.49 |
| *INST%* | 52.59 | 55.06 | 38.00 | 67.31 | 21.61 |
| *INST#* | 191.82 | 124.00 | 53.00 | 272.00 | 186.06 |
| *INST_SH* | 59.76 | 24.81 | 7.48 | 66.98 | 92.30 |
| *TRADE_DV* | 19,993.53 | 17,062.50 | 11,250.00 | 26,025.00 | 11,630.67 |

**Panel B: Pearson and Spearman Rank Correlations[b]**

| | *ANFOLL* | *INST%* | *INST#* | *INST_SH* | *TRADE_DV* |
|---|---|---|---|---|---|
| *ANFOLL* | | 0.2104 | 0.5495 | 0.4927 | 0.3105 |
| | | (<0.001) | (<0.001) | (<0.001) | (<0.001) |
| *INST%* | 0.2533 | | 0.2241 | 0.1565 | 0.3362 |
| | (<0.001) | | (<0.001) | (<0.001) | (<0.001) |
| *INST#* | 0.6196 | 0.3880 | | 0.8667 | 0.2675 |
| | (<0.001) | (<0.001) | | (<0.001) | (<0.001) |
| *INST_SH* | 0.6516 | 0.4277 | 0.9517 | | 0.1161 |
| | (<0.001) | (<0.001) | (<0.001) | | (<0.001) |
| *TRADE_DV* | 0.3222 | 0.3773 | 0.3679 | 0.3147 | |
| | (<0.001) | (<0.001) | (<0.001) | (<0.001) | |

[a] Variable definitions: The number of analysts following the firm (*ANFOLL*), the percentage of the firm's common shares held by institutions (*INST%*), the number of institutions holding shares in the firm (*INST#*), the number of shares held by institutions in millions (*INST_SH*), and the median dollar value of shares traded (*TRADE_DV*).

[b] Pearson (Spearman) correlations are above (below) the diagonal. Two-tailed probability values are in parentheses below the correlations. The sample size for all variables except *TRADE_DV* is 101,238; the sample size for *TRADE_DV* is 59,689.

and focus on those firm characteristics identified in prior empirical studies that are associated with the efficiency of a firm's stock price: analyst following, percentage of stock held by institutions, number of institutions owning shares, number of shares held by institutions, and the dollar value of shares traded (see Hand 1990; Walther 1997; Bartov et al. 2000; Brown and Han 2000; Ayers and Freeman 2001; Bhattacharya 2001; Jiambalvo et al. 2002; Piotroski and Roulstone 2002).[18] Analyst following (*ANFOLL*) is the number of analysts on Zacks issuing quarterly earnings forecasts for the firm in the prior year. The percentage of stock held by institutions (*INST%*) is the number of shares held by institutions from Spectrum divided by the common shares outstanding from Compustat as of the beginning of the year. Data on the number of institutions holding shares (*INST#*) and the number of shares held by institutions (*INST_SH*) for the firm as of the beginning of the year are obtained from Spectrum. We measure trade size (*TRADE_DV*) as the median dollar value of shares traded over the prior year. As shown in Panel A, our sample firms are followed by an average of 13.24 analysts. On average, approximately 192 institutions hold

[18] Prior research also has used firm size as a proxy for investor sophistication (e.g., Walther 1997). We do not use firm size in our study because this variable also may reflect differences in information environments between firms we classify as having sophisticated and unsophisticated investors. Instead, we control for the effects of firm size by defining each variable on a relative basis, which controls for firm and time-period effects.

59.76 million, or 53 percent, of the sample firms' outstanding common shares. The average of the dollar value of shares traded for our sample firms is $19,994.

Panel B of Table 3 provides the Pearson (Spearman rank) correlations among the sophistication proxies above (below) the diagonal. As expected, all investor sophistication characteristics are positively correlated, suggesting that these firm variables reflect a similar underlying construct. Therefore, we perform principal factor analysis to identify common factor(s) and to classify observations into groups based on these firm characteristics. This procedure reduces the dimensionality of the data by identifying linear combinations of the individual firm characteristics that explain the shared variance among the characteristics (see Bushee 1998). Since data on the dollar value of shares traded from the Trade and Quote Database are only available beginning in 1993, we perform two principal factor analyses, one using the four analyst following and institutional ownership variables over 1981–1999, and the other using all five variables over 1993–1999.

Panel A of Table 4 presents the results of the factor analysis, which produced one common factor.[19] Regardless of whether dollar value of shares traded is included in the analysis, the one common factor identified explains approximately all of the variability

**TABLE 4**
**Factor Analysis of Investor Sophistication Characteristics[a]**

**Panel A: Factor Analysis**

| Variable | 1981–1999 | 1993–1999 |
|---|---|---|
| *ANFOLL* | 0.5803 | 0.6096 |
| *INST%* | 0.2390 | 0.2720 |
| *INST#* | 0.9172 | 0.9230 |
| *INST_SH* | 0.8845 | 0.8706 |
| *TRADE_DV* | | 0.3260 |
| Variability Explained | 105.74% | 93.34% |

**Panel B: Descriptive Statistics Based on Factor Scores**

| | 1981–1999 | | 1993–1999 | |
|---|---|---|---|---|
| Variable | Low Soph. | High Soph. | Low Soph. | High Soph. |
| Mean *ANFOLL* | 7.62 | 19.03** | 7.51 | 19.83** |
| Mean *INST%* | 37.25 | 59.79** | 36.48 | 62.49** |
| Mean *INST#* | 40.49 | 405.81** | 39.72 | 400.03** |
| Mean *INST_SH* | 5.66 | 145.44** | 5.77 | 155.07** |
| Mean *TRADE_DV* | 14,450.40 | 24,463.45** | 13,313.81 | 26,002.46** |

** Statistically different from the corresponding value for the low sophistication group using a t-test for means at two-tailed $p < 0.01$.

[a] Panel A provides the pattern weights and variability explained from principal factor analysis of the investor sophistication characteristics. The first factor analysis excludes the median dollar value of shares traded (*TRADE_DV*), and is based on 101,238 observations from 1981–1999. The second factor analysis includes *TRADE_DV*, and is based on 59,689 observations from 1993–1999. Panel B provides the mean values of the investor sophistication characteristics for subsamples based on the standardized factor scores. Standardized factor scores in the bottom (top) 33 percent are defined to indicate low (high) investor sophistication. Variables are defined in Table 3.

[19] We used three criteria to determine the appropriate number of factors to retain: scree plots, minimum eigenvalues, and the proportion of variability explained. All criteria indicated that only one factor should be kept.

among the characteristics. Based on this analysis, we calculate standardized factor scores and rank our sample observations on the basis of these scores. We then divide the sample observations into three equal-sized groups. We classify observations in the top (bottom) group as high (low) investor sophistication. All results discussed below hold if we place the sample observations into two or four groups, or if we separately classify observations into groups based on the five firm characteristics, rather than the standardized scoring coefficients from the factor analysis. As expected, the mean and median analyst following, institutional ownership, and dollar value of shares traded is statistically higher for the high sophistication group than for the low sophistication group (Panel B, median results not tabulated). Therefore, our assumption that the marginal investor of firms with higher values of the standardized factor scores is more likely to be sophisticated is consistent with prior research that analyst following, institutional ownership, and dollar value of shares traded are positively correlated with investor sophistication.

***Tests of Hypothesis 1***

To test whether sophisticated investors have greater knowledge overall about how to use the information in factors associated with forecast accuracy, we separately estimate Equations (5) and (6) for the sophisticated and unsophisticated investor groups. We then use these estimations to calculate the matching index for each group. We estimate the forecast accuracy regression (Equation (5)) separately for each group to allow for the possibility that the factors related to forecast accuracy vary cross-sectionally with investor sophistication. We estimate the market reaction regression (Equation (6)) separately for each group to allow investors' implied models of relative forecast accuracy to vary cross-sectionally with investor sophistication.[20] We do not tabulate the results for estimating Equations (5) and (6) separately for these subsamples. Rather, we focus on the matching indexes.

Table 5 presents the matching indexes to test H1. We find that the matching index for the sophisticated group is more than two times that for the unsophisticated group. This finding holds regardless of whether the dollar value of shares traded is included in the factor analysis. For example, in the 1993–1999 analysis, the matching index is 0.3943 for sophisticated investors, compared to 0.1166 for unsophisticated investors. The matching index for the unsophisticated group is significantly lower than the corresponding matching index for the sophisticated group at $p < 0.01$ based on the Fisher Z-statistic to test for differences in correlations (Zar 1999, 387), providing strong support for H1. Untabulated results indicate that unsophisticated investors' weights on some factors, most notably *RFCAGE*, are opposite in sign to the weight in the statistical model. Clement and Tse (2003) find that the coefficient on forecast age becomes insignificant once they control for prior uncertainty, and interpret the incorrect sign on *RFCAGE* as suggesting that investors prefer earlier forecasts due to greater uncertainty. In contrast, we find that sophisticated (unsophisticated) investors' weight on *RFCAGE* is of the same (opposite) sign as that in the statistical model, even after controlling for prior uncertainty (see Section VII). Therefore, an alternative explanation for Clement and Tse's (2003) finding is that unsophisticated investors lack knowledge of the information in *RFCAGE* for forecast accuracy. In sum, we

[20] We replicated our tests using year-by-year regressions to better reflect the information available to investors. In this analysis, we estimated the accuracy regression using data from year $t-1$, and the market regression using data from year $t$. We then assessed how well investors' implied predictions of accuracy in year $t$ corresponded to the historical relation between forecast accuracy and the factors. All inferences are unchanged. Further, the results are unchanged if we compare each group of investors' models to the accuracy models estimated on the overall sample.

**TABLE 5**
**Matching Indexes Conditional on Investor Sophistication[a]**

| | Factor Score Based on *ANFOLL*, *INST%*, *INST#*, and *INST_SH* (1981–1999) | Factor Score Based on *ANFOLL*, *INST%*, *INST#*, *INST_SH*, and *TRADE_DV* (1993–1999) |
|---|---|---|
| Low Sophistication | 0.1714 | 0.1166 |
| High Sophistication | 0.3614** | 0.3943** |

** Statistically greater than the corresponding correlation for the bottom group at two-tailed $p < 0.01$ using the Fisher Z test.

[a] The matching index is the Pearson correlation between the fitted values for forecast accuracy from Equation (5) and the investors' implied predictions from Equation (6). In calculating the matching indexes, Equations (5) and (6) are estimated separately for each subsample formed on the basis of the standardized score from the principal factor analysis of the investor sophistication characteristics (see Table 4). The matching indexes are provided for the low and high sophistication subsamples, defined as observations with scores in the bottom 33 percent and top 33 percent, respectively.

conclude that sophisticated investors have greater knowledge overall than unsophisticated investors about the appropriate signs and weights for the set of factors that can be used to predict forecast accuracy.

### *Tests of Hypothesis 2*

Table 6 presents evidence related to H2, which proposes that sophisticated investors have greater knowledge of the net benefit of individual factors that can be used to predict relative forecast accuracy than do unsophisticated investors. To test this hypothesis, we first assess the benefit of each of the factors in Equation (5) for explaining relative forecast accuracy using stepwise model selection techniques.[21] Using the findings from this analysis, we order the factors based on their contribution to the explanatory power of Equation (5) for relative forecast accuracy. We then estimate variations of Equations (5) and (6). The stepwise model selection analysis indicates that relative forecast age has the greatest partial $R^2$, followed by relative prior accuracy. These findings are consistent with Brown (2001), who documents that this two-factor model performs as well (in terms of predictive ability and explanatory power) as a more complex model based on multiple factors. Therefore, we use this two-factor model to represent the set of high-benefit factors. Subsequent variations of the model progressively add factors that have decreasing benefits for explaining relative forecast accuracy. If sophisticated investors better understand the benefits of factors, then there will be a larger decrease in their matching index (compared to unsophisticated investors) when moving from the model that includes the two high-benefit factors for forecast accuracy (forecast age and prior accuracy) to models that also include lower benefit factors.

Table 6 presents the matching indexes for the overall sample and for subsamples based on investor sophistication for the various models. The conclusion based on the matching indexes for all the models estimated on the overall sample (Panel A) is consistent with findings discussed in the previous section. That is, regardless of the independent variables included in the model, the matching index is less than 1. This result suggests that investors,

[21] Ideally, this analysis would be based on the benefits of each factor net of any acquisition and processing costs. An assessment of the costs relative to the benefits would be highly subjective. Consequently, we limit this analysis to an ordering of the factors based on an objective estimate of their key benefit, that of predicting relative forecast accuracy. As a result, our analysis implicitly assumes that the ordering based on this benefit is identical to that based on net benefit.

TABLE 6

Matching Indexes For Simple and Complex Models: Overall Sample and Conditional on Investor Sophistication[a]

| | Variables Included in Model | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *RFCAGE* (Model 1) | + *RPMAPE* (Model 2) | + *RFIRMEXP* (Model 3) | + *RNOIND* (Model 4) | + *RBROKSIZE* (Model 5) | + *RIIAWARD* (Model 6) | + *RFCFREQ* (Model 7) | + *RNOFIRM* (Model 8) | + *RGENEXP* (Model 9) | + *RTURNOVER* (Model 10) |
| **Panel A: Overall Sample** | | | | | | | | | | |
| | 1.0000 | 0.3463 | 0.3797 | 0.3875 | 0.3842 | 0.3549 | 0.2938 | 0.2887 | 0.2751 | 0.2739 |
| **Panel B: Measure of Investor Sophistication: Factor Score Based on ANFOLL, INST%, INST#, and INST_SH (1981–1999)** | | | | | | | | | | |
| Low Soph. | −1.0000 | 0.1345 | 0.1951 | 0.2152 | 0.2038 | 0.1736 | 0.2041 | 0.2132 | 0.1730 | 0.1714 |
| | | | *0.0606* | *0.0807* | *0.0693* | *0.0391* | *0.0696* | *0.0787* | *0.0385* | *0.0369* |
| High Soph. | 1.0000** | 0.9916** | 0.9736** | 0.6235** | 0.6261** | 0.5406** | 0.4271** | 0.3798** | 0.3613** | 0.3614** |
| | | | ***−0.0180*** | ***−0.3681*** | ***−0.3655*** | ***−0.4510*** | ***−0.5645*** | ***−0.6118*** | ***−0.6303*** | ***−0.6302*** |
| **Panel C: Measure of Investor Sophistication: Factor Score Based on ANFOLL, INST%, INST#, INST_SH, and TRADE_DV (1993–1999)** | | | | | | | | | | |
| Low Soph. | −1.0000 | −0.0311 | 0.0545 | 0.0868 | 0.0616 | 0.0632 | 0.0986 | 0.1130 | 0.1006 | 0.1166 |
| | | | *0.0856* | *0.1179* | *0.0927* | *0.0943* | *0.1297* | *0.1441* | *0.1317* | *0.1477* |
| High Soph. | 1.0000** | 0.9928** | 0.9853** | 0.8468** | 0.8457** | 0.6930** | 0.4306** | 0.4090** | 0.4030** | 0.3943** |
| | | | ***−0.0075*** | ***−0.1460*** | ***−0.1471*** | ***−0.2998*** | ***−0.5622*** | ***−0.5838*** | ***−0.5898*** | ***−0.5985*** |

** Statistically greater than the corresponding correlation for the low sophistication subsample at two-tailed $p < 0.01$ using the Fisher Z test.

[a] This table provides the matching indexes for various models for the overall sample (Panel A) and subsamples formed on the basis of the standardized score from the principal factor analysis of the investor sophistication characteristics (see Table 4). Within Panels B and C, the indexes are provided for the low and high sophistication subsamples, defined as observations with scores in the bottom 33 percent and top 33 percent, respectively. The difference relative to Model 2 is provided in italics under the matching index. The difference for the high sophistication subsample is in bold when it is statistically smaller than the corresponding difference for the low sophistication subsample at two-tailed $p < 0.01$ using a Tukey-type test for differences in correlations.

on average, use some (but not all) of the information in these factors when predicting relative forecast accuracy.

Panels B and C of Table 6 present the results conditional on investor sophistication. Consistent with the findings provided in Table 5 to test H1, the matching index for the unsophisticated subsample is significantly lower than for the sophisticated subsample using a Fisher Z-statistic, regardless of the independent variables included. This finding holds regardless of whether the dollar value of shares traded is included in the factor analysis. Consistent with the untabulated results discussed above, the matching index for Model 1 in Panels B and C is $-1$ for unsophisticated investors, indicating that unsophisticated investors' weight on *RFCAGE* is of opposite sign to that in the statistical model.[22] Overall, our findings supporting H1 hold regardless of the factors included.

Table 6 also provides the change in the matching index (in italics) from Model 2 to Models 3 through 10, respectively, for each subsample. To test H2, we compare the change in the matching index for the sophisticated investor subsample to the change for the unsophisticated investor subsample as factors with decreasing benefits are progressively added as independent variables. In both Panels B and C, we find a consistently larger decrease in the matching index for the subsample in which the marginal investor is likely to be sophisticated. The decrease in the matching index is statistically greater for sophisticated investors than for unsophisticated investors using a Tukey-type test for differences in correlations (Zar 1999, 393). These results are consistent with H2, and suggest that sophisticated investors better understand the benefits of the individual factors that can be used to predict forecast accuracy.

The results in Table 6 indicate that the fitted values derived from sophisticated investors' model correspond less with the fitted values derived from the statistical model when low-benefit factors are included. While consistent with H2, we cannot determine based on these results whether sophisticated investors underweight low-benefit factors more than unsophisticated investors. To investigate this specific form of adaptive decision making, which underlies our prediction in H2, we must compare the coefficients on individual factors in the statistical model to the implied coefficients on those factors from sophisticated and unsophisticated investors' models. We define underweighting to occur when the implied weight sophisticated investors place on a factor is either insignificantly different from zero or less than the weight on that factor in the statistical model.[23] This analysis indicates that sophisticated investors underweight the following low-benefit factors: *RFIRMEXP*, *RGENEXP*, *RTURNOVER*, *RFCFREQ*, *RNOFIRM*, *RBROKSIZE*, and *RIIAWARD*. By contrast, unsophisticated investors overweight *RFIRMEXP*, *RGENEXP*, and *RFCFREQ*. These findings provide additional support for H2, and suggest that sophisticated investors' model better reflects the cost-benefit trade-offs for the factors associated with forecast accuracy than does that of unsophisticated investors.

## VII. ROBUSTNESS CHECKS

In this section, we discuss the assumptions underlying our empirical tests and conclusions based on those tests. Where empirically feasible, we provide evidence on the validity of each of these assumptions.

---

[22] This finding holds if we include a control for prior uncertainty in the model (see Section VII).

[23] Regression diagnostics indicate no evidence of the statistical problems of multicollinearity in either the statistical or the investor model, suggesting that the individual coefficient estimates upon which we make these inferences are not biased.

**Assumption 1: Sophisticated and Unsophisticated Investors Attempt to Predict Forecast Revision Accuracy**

We assume that both sophisticated and unsophisticated investors attempt to predict the accuracy of analysts' forecast revisions, and, in particular, that sophisticated and unsophisticated investors have equivalent incentives for predicting the accuracy of analysts' forecasts. We know of no way to measure directly investors' incentives for predicting accuracy. However, we would expect that, if sophisticated investors had greater incentives for predicting forecast accuracy (e.g., greater expected trading profits), then the correlation of *CAR* and *ex post* accuracy would be greater for sophisticated investors. Therefore, to investigate the validity of this assumption, we calculate the correlation of *CAR* and *ex post* accuracy of forecast revisions separately for our sophisticated and unsophisticated investor subsamples. We are unable to reject the null hypothesis that the correlation is equal across these two subsamples (two-tailed, $p > 0.10$). Additionally, if sophisticated investors face greater incentives to predict forecast accuracy, then we would not find in our tests of H2 that they ignore or underweight less beneficial cues more than do unsophisticated investors. These findings are consistent with both sophisticated and unsophisticated investors attempting to predict the accuracy of analysts' forecast revisions. However, we are unable to completely rule out differences in the incentives for predicting forecast revision accuracy across sophisticated and unsophisticated investors.

**Assumption 2: Sophisticated and Unsophisticated Investors Predict Forecast Revision Accuracy Using the Same Set of Observable Factors**

Our tests also assume that both sophisticated and unsophisticated investors predict forecast revision accuracy using the set of observable factors that we include in our models. In other words, we must assume that both groups of investors have access to information about all these factors. Further, we must assume that these factors predict forecast accuracy equally well in the sophisticated and unsophisticated subsamples. The first part of this assumption is supported by the similarity of the explanatory power of the market reaction regression (Equation (6)) across the two subsamples.[24] To investigate the second part of this assumption, we examine whether the factors we consider explain relative forecast accuracy equally well in the sophisticated and unsophisticated subsamples. We find that the explanatory power (both in-sample and out-of-sample) of the accuracy regression (Equation (5)) is similar across the two subsamples.[25] Overall, these results suggest that our results are not due to differences between sophisticated and unsophisticated investors as to the factors they use to predict forecast accuracy or to the predictive ability of observable factors.

**Assumption 3: Sophisticated and Unsophisticated Investors Obtain Forecast Revisions on a Timely Basis**

Our tests assume that unsophisticated and sophisticated investors have equally timely access to forecast revisions. It is possible that the higher matching index for sophisticated investors reflects greater or more timely access to analyst forecast revisions. To investigate

[24] As an alternative test of the assumption of equal access to information on the factors that can be used to predict forecast revision accuracy, we classify *FIRMEXP*, *GENEXP*, *FCFREQ*, *NOIND*, and *BROKSIZE* as "not equally accessible," and eliminate them from our models. All of our results for H1 and H2 hold when these "inaccessible" factors are excluded from the analyses. We thank the referee for this suggestion.

[25] We assess the out-of-sample predictive ability of these factors by estimating the accuracy regression using data from year $t-1$, and predicting relative forecast accuracy in year $t$ based on the estimated coefficients from year $t-1$. The absolute (squared) difference between the actual relative accuracy in year $t$ and the predicted relative accuracy based on the year $t$–1 regression for the unsophisticated subsample was either lower than or equal to the absolute (squared) difference for the sophisticated subsample.

this possibility, we perform three tests. First, we investigate whether the explanatory power of *FCREV* for *CAR* is higher in the sophisticated investor subsample. We regress *CAR* on *FCREV* separately for the sophisticated and unsophisticated subsamples and find no evidence that this is the case. Further, in a regression of *CAR* on an intercept, an indicator variable for sophistication (*SOPH* = 0 if unsophisticated, 1 if sophisticated), *FCREV*, and *FCREV*SOPH*, we find that the estimated coefficient on *FCREV*SOPH* is insignificant, suggesting that sophisticated investors do not react more strongly to analyst forecast revisions.[26]

Second, we replicate our analyses in Tables 5 and 6 using five- and seven-day returns accumulation windows (($-3$, $+3$) or ($-1$, $+3$), where $t = 0$ is the date the revised forecast is issued) to ensure that we fully capture the reaction of all investors. We extend the window to three days after the forecast release date because prior work suggests that virtually all of abnormal trading responses for all investor types occur within three days after announcements (Cready 1988; Lee 1992; Bhattacharya 2001). We extend the window to three days before the forecast release date in case sophisticated investors have access to this information in advance of the formal release date.[27] Third, we replicate our primary analyses for H1 and H2 using observations only from 1997 forward. We select this time period because many investor advice sites and analysts' forecasts became available on the Internet starting in 1997 (Kurtz 2000). All results continue to hold using either the expanded returns accumulation window and/or the more recent time period. Overall, these analyses do not suggest that sophisticated investors have greater or timelier access to analyst forecast revisions.

### Assumption 4: Sophisticated and Unsophisticated Investors Update Their Beliefs Based on Forecast Revisions and Predicted Accuracy

Our tests assume that investors use the revised analyst forecast as well as their prediction regarding its accuracy in revising their beliefs about future earnings. Previous research has shown that sophisticated and unsophisticated investors may differ in the extent to which they form earnings expectations on the basis of analysts' forecasts versus seasonal random walk forecasts (Walther 1997; Bhattacharya 2001). Consequently, to ensure that our use of the analyst's prior forecast as the expectation does not drive our results, we replicate our analyses for H1 and H2 using either the seasonal random walk forecast or the consensus analyst forecast as the expectation.[28] These analyses suggest that our assumption that both sophisticated and unsophisticated investors use the analyst forecast revision to update their beliefs does not confound our results.

[26] We do not find any evidence that differences in information environments related to firm size are confounded with differences in investor sophistication and, thus, create these findings. The explanatory power of *FCREV* for *CAR* remains similar even after we partition the sample based on firm size into three groups and re-estimate that regression. The coefficient on *FCREV*SOPH* remains insignificant when we include firm size (*SIZE*) and *FCREV*SIZE* in the regression.

[27] All results for H1 and H2 also hold using a ($-7$, $+7$) window.

[28] We measure the outstanding consensus forecast as the mean of all forecasts issued after the beginning of the quarter but before the revised forecast is released. We require a minimum of two forecasts to construct the consensus. Our results hold if we define the outstanding consensus forecast as the median of all forecasts issued during this window.

**Assumption 5: Sophisticated and Unsophisticated Investors Face a Similar Level of Information Uncertainty**

Our tests implicitly assume that sophisticated and unsophisticated investors face a similar level of information uncertainty about upcoming earnings. To investigate this assumption, we compare the mean squared forecast error (MSFE) across the sophisticated and unsophisticated subsamples. Following Clement and Tse (2003), we calculate the average of the squared price-deflated forecast error for all prior quarterly earnings forecasts outstanding at the date of each forecast revision (see also Barron et al. 1998). While the median MSFE is 0.0000 in both subsamples, the mean is significantly higher in the unsophisticated sample (0.0004) than in the sophisticated subsample (0.0002, $p < 0.01$). This finding provides some evidence that unsophisticated investors face a higher level of information uncertainty prior to the forecast revision date. To ensure that our results are not sensitive to our assumption, we replicate our analyses for H1 and H2 after including the mean-adjusted MSFE in our models. All results hold, suggesting that differences in information uncertainty between sophisticated and unsophisticated investors do not cause our findings.

**Assumption 6: Sophisticated and Unsophisticated Investors' Stock-Holding Decisions are Based on Revised Earnings Expectations**

We assume that individual investors' decisions about whether to buy or sell a stock are based on their revised expectations of future earnings. However, these decisions also could be affected by their risk preferences, the expected benefits (net of costs) of changing holdings, and the incentives they face. For example, sophisticated investors may face greater incentives to justify their investment decisions than do unsophisticated investors. Therefore, sophisticated investors may react more strongly, for example, to forecast revisions issued by analysts named to the *Institutional Investor* All-American team or by analysts employed at larger brokerage houses, not because these factors are associated with the accuracy of the revised forecast, but because these factors provide an easy justification to outside parties or superiors. To examine whether sophisticated investors react more strongly than unsophisticated investors to forecast revisions that would be easier to justify, we regress *CAR* on an intercept, an indicator variable for sophistication (*SOPH* = 0 if unsophisticated, 1 if sophisticated), *FCREV*, *FCREV*SOPH*, and *FCREV*SOPH*RIIAWARD* (or *FCREV*SOPH*RBROKSIZE*). The estimated coefficient on *FCREV*SOPH*RIIAWARD* (*FCREV*SOPH*RBROKSIZE*) is not significant, indicating that sophisticated investors do not react more strongly to forecast revisions issued by analysts who were members of the *Institutional Investor* All-American team (who work for larger brokerage houses). While these results provide some evidence that differences in accountability or incentives are not driving the higher matching index for sophisticated investors, we are unable to rule out that differences in risk preferences, the expected benefits of changing holdings, or other incentives could be an alternative explanation for our results.

Overall, our additional analyses suggest that the conclusions we draw regarding differences between sophisticated and unsophisticated investors' models for predicting the relative accuracy of analysts' forecast revisions are not confounded by the assumptions underlying our empirical tests.

## VIII. CONCLUSIONS

In this study we examine the extent to which sophisticated and unsophisticated investors' reactions to sell-side analysts' forecast revisions appear consistent with factors shown to be associated with forecast accuracy. We infer investors' model for predicting the relative accuracy of forecast revisions from their reactions to those revisions, and investigate two differences in the quality of sophisticated and unsophisticated investors' models. We base our inferences regarding differences in the quality of prediction models on the "matching index" from Brunswik's lens model of human judgment.

As predicted, we find that the matching index for the subsample in which the marginal investor is likely to be sophisticated is statistically higher than the matching index for the subsample in which the marginal investor is likely to be unsophisticated. We also find that the decrease in the matching index for a model that includes the two most beneficial factors to those for more comprehensive models is greater for the sophisticated investor subsample than for the unsophisticated investor subsample. Taken together, our results suggest that sophisticated investors not only have more knowledge overall about the factors related to analysts' forecast accuracy, but also that they are aware of the individual factors that are the most beneficial to use.

The findings of this study contribute to the literature in several ways. First, we investigate how different types of investors use an important source of accounting information provided by a key information intermediary—the sell-side analyst's quarterly earnings forecast revision. Consequently, we add to both the growing literature on behavioral differences between sophisticated and unsophisticated investors (e.g., Lee et al. 1999; Shapira and Venezia 2001) and the literature on investors' reactions to analysts' forecasts. Further, our factor analysis of various sophistication proxies may be useful to future research examining investor sophistication issues.

Additionally, we introduce the lens model to the archival accounting literature. The lens model methodology might be a fruitful approach for researchers to consider when investigating the quality of investors' predictions about future outcomes in situations in which the information signals are correlated, e.g., factors related to corporate bankruptcy. Further, we provide evidence on the validity of the assumptions underlying our application of the lens model. These analyses should be useful to future researchers who seek to make inferences about individual investor behavior using market-level data.

Our study has both benefits and limitations related to the archival market-level approach we employ. The most salient benefit is that we demonstrate differences between investors we classify as sophisticated and unsophisticated in a market setting, with its attendant incentives and disciplinary mechanisms. Further, we demonstrate that both groups of investors appear to be predicting the accuracy of analysts' forecasts in this setting. Although we are able to provide some empirical evidence on the validity of most of the assumptions necessary for inferring investors' prediction models from market reactions, we cannot examine all of them. In particular, we cannot investigate sophistication-related differences in risk preferences and other such factors. As a result, we are unable to rule out all alternative explanations for our findings. Future experimental research could attempt to replicate our findings in a setting that would control for differences between groups of investors. Finally, experimental work that probes the processes that underlie investors' development of models to predict forecast accuracy ultimately would be necessary to provide the basis for specific suggestions for investor education programs (Bonner 1999).

## REFERENCES

Ackerman, P., R. Sternberg, and R. Glaser. 1989. *Learning and Individual Differences: Advances in Theory and Research*. New York, NY: W.H. Freeman and Co.

Ashton, R. 1982. *Human Information Processing in Accounting*. Sarasota, FL: American Accounting Association.

Ayers, B., and R. Freeman. 2001. Evidence that price leads of earnings increase with analyst following and institutional ownership. Working paper, University of Georgia.

Barber, B., and T. Odean. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* (April): 773–806.

Barron, O., O. Kim, S. Lim, and D. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* (October): 421–433.

Bartov, E., S. Radhakrishnan, and I. Krinsky. 2000. Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review* (January): 43–63.

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution.* 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Bettman, J., E. Johnson, and J. Payne. 1990. A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes* (February): 111–139.

Bhattacharya, N. 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *The Accounting Review* (April): 221–244.

Bonner, S. 1990. Experience effects in auditing: The role of task-specific knowledge. *The Accounting Review* (January): 72–92.

———, and P. Walker. 1994. The effects of instruction and experience on the acquisition of auditing knowledge. *The Accounting Review* (January): 157–178.

———. 1999. Judgment and decision-making research in accounting. *Accounting Horizons* (December): 385–398.

Brous, P., and O. Kini. 1994. The valuation effects of equity issues and the level of institutional ownership: Evidence from analysts' earnings forecasts. *Financial Management* (Spring): 33–46.

Brown, L., and J. Han. 2000. Do stock prices fully reflect the implications of current earnings for future earnings for AR1 firms? *Journal of Accounting Research* (Spring): 149–164.

———. 2001. How important is past analyst earnings forecast accuracy? *Financial Analysts Journal* (November/December): 44–49.

———, and E. Mohammad. 2001. Profiting from predicting individual analyst earnings forecast accuracy. Working paper, Georgia State University.

Brown, P., G. Foster, and E. Noreen. 1985. *Security Analyst Multi-Year Earnings Forecasts and the Capital Markets.* Sarasota, FL: American Accounting Association.

Bushee, B. 1998. The influence of institutional investors on myopic R&D investment behavior. *The Accounting Review* (July): 305–333.

Chen, Q., J. Francis, and W. Jiang. 2001. Investor learning about analyst predictive ability. Working paper, Duke University.

Clement, M. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* (June): 285–303.

———, and S. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* (January): 227–249.

Cready, W. 1988. Information value and investor wealth: The case of earnings announcements. *Journal of Accounting Research* (Spring): 1–27.

Einhorn, H. 1974. Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology* (October): 562–571.

El-Gazzar, S. 1998. Predisclosure information and institutional ownership: A cross-sectional examination of market revaluations during earnings announcement periods. *The Accounting Review* (January): 119–129.

Fennema, M., and D. Kleinmuntz. 1995. Anticipations of effort and accuracy in multiattribute choice. *Organizational Behavior and Human Decision Processes* (July): 21–32.

Givoly, D., and J. Lakonishok. 1979. The information content of financial analysts' earnings forecasts. *Journal of Accounting and Economics* (March): 165–185.

Gleason, C., and C. Lee. 2003. Analyst forecast revisions and market price discovery. *The Accounting Review* (January): 193–225.

Gonedes, N., N. Dopuch, and S. Penman. 1976. Disclosure rules, information production, and capital market equilibrium: The case of forecast disclosure rules. *Journal of Accounting Research* (Spring): 89–137.

Hammond, K., and T. Stewart. 2001. *The Essential Brunswik: Beginnings, Extensions, Applications.* New York, NY: Oxford University Press.

Hand, J. 1990. A test of the extended functional fixation hypothesis. *The Accounting Review* (October): 764–780.

Hassel, C., and M. Norman. 1992. Financial characteristics of neglected and institutionally held stocks. *Journal of Accounting, Auditing, and Finance* (April): 313–330.

Hodge, F. 2001. Investor habits. Speech given at the Annual Meeting of the Puget Sound Chapter of the National Association of Investment Clubs, Women's University Club, Seattle, WA, June 14.

Jacob, J., T. Lys, and M. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* (November): 51–82.

Jiambalvo, J., S. Rajgopal, and M. Venkatachalam. 2002. Institutional ownership and the extent to which stock prices reflect future earnings. *Contemporary Accounting Research* (Spring): 117–145.

Johnson, E. 1988. Expertise and decision under uncertainty: Performance and process. In *The Nature of Expertise*, edited by M. Chi, R. Glaser, and M. Farr. Hillsdale, NJ: Lawrence Erlbaum.

Johnson, P., A. Duran, F. Hassebrock, J. Moller, M. Prietula, P. Feltovich, and D. Swanson. 1981. Expertise and error in diagnostic reasoning. *Cognitive Science* (September): 235–283.

Kurtz, H. 2000. *The Fortune Tellers: Inside Wall Street's Game of Money, Media, and Manipulation.* New York, NY: The Free Press.

Lee, C. 1992. Earnings news and small traders: An intraday analysis. *Journal of Accounting and Economics* (June/September): 265–302.

Lee, Y., J. Lin, and Y. Liu. 1999. Trading patterns of big versus small players in an emerging market: An empirical analysis. *Journal of Banking and Finance* (May): 701–725.

Lev, B. 1988. Toward a theory of equitable and efficient accounting policy. *The Accounting Review* (January): 1–22.

Maines, L. 1996. An experimental examination of subjective forecast combination. *International Journal of Forecasting* (June): 223–233.

Mikhail, M., B. Walther, and R. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* (Supplement): 131–157.

———, ———, and ———. 1999. Does forecast accuracy matter to security analysts? *The Accounting Review* (April): 185–200.

Mishkin, F. 1983. *A Rational Expectations Approach to Macroeconomics: Testing Policy Ineffectiveness and Efficient-Markets Models.* Chicago, IL: University of Chicago Press.

O'Brien, P. 1988. Analysts' forecasts as earnings expectations. *Journal of Accounting and Economics* (January): 53–83.

———. 1990. Forecast accuracy of individual analysts in nine industries. *Journal of Accounting Research* (Autumn): 286–304.

Park, C., and E. Stice. 2000. Analyst forecasting ability and the stock price reaction to forecast revisions. *Review of Accounting Studies* (September): 259–272.

Payne, J., J. Bettman, and E. Johnson. 1993. *The Adaptive Decision Maker.* New York, NY: Cambridge University Press.

Piotroski, J., and D. Roulstone. 2002. Informed traders and the incorporation of industry and firm-specific information into stock prices. Working paper, University of Chicago.

Potter, G. 1992. Accounting earnings announcements, institutional investor concentration, and common stock returns. *Journal of Accounting Research* (Spring): 146–155.

Richardson, S., S. Teoh, and P. Wysocki. 2001. The walkdown to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. Working paper, University of Michigan.

Shapira, Z., and I. Venezia. 2001. Patterns of behavior of professionally managed and independent investors. *Journal of Banking and Finance* (August): 1573–1587.

Slovic, P. 1969. Analyzing the expert judge: A descriptive study of a stockbroker's decision processes. *Journal of Applied Psychology* (August): 255–263.

SRI International. 1987. *Investor Information Needs and the Annual Report*. Morristown, NJ: Financial Executives Research Foundation.

Stickel, S. 1992. Reputation and performance among security analysts. *Journal of Finance* (December): 1811–1836.

Walther, B. 1997. Investor sophistication and market earnings expectations. *Journal of Accounting Research* (Autumn): 157–179.

Williams, P., G. Moyes, and K. Park. 1996. Factors affecting earnings forecast revisions for the buy-side and sell-side analyst. *Accounting Horizons* (September): 112–121.

Wilson, R. 1975. Informational economies of scale. *Bell Journal of Economics* (Spring): 184–195.

Yunker, J., and T. Krehbiel. 1988. Investment analysis by the individual investor. *Quarterly Review of Economics and Business* (Winter): 90–101.

Zar, J. 1999. *Biostatistical Analysis*. 4th edition. Upper Saddle River, NJ: Prentice Hall.