The Role of Forecast Patterns in Conveying Analysts’ Predictive Ability

Qi Chen*
(Duke University)

Jennifer Francis
(Duke University)

Wei Jiang
(Columbia University)

We investigate the ability of forecast patterns to convey information about an analyst's predictive ability. We establish an equilibrium strategy where the analyst issues a forecast only if the realization of his private signal exceeds a threshold. In equilibrium, higher-ability analysts choose higher thresholds than lower-ability analysts, and investors interpret all forecasts correctly. A testable implication of this equilibrium is that because they choose higher thresholds, higher-ability analysts issue fewer and bolder forecasts than their lower-ability peers. Empirical evidence is consistent with both predictions.

Draft: October 2005

* Corresponding author: Fuqua School of Business, Duke University, Durham, NC 27708. gc2@duke.edu

We have benefited from helpful comments provided by Moshe Bareket, Robert Bushman, John Graham, Thomas Hemmer, Michael Johannes, Chandra Kanodia, Ernst Maug, Michael Mikhail, Katherine Schipper, Michael Smith, Phillip Stocken, Mohan Venkatachalam, Jerry Zimmerman, and from workshop participants at the College of William and Mary, Columbia, Duke/UNC Fall Camp, INSEAD, Humboldt University at Berlin, University of Minnesota, Eastern Finance Association Annual Meeting, and American Accounting Association Annual Meeting. We thank Xu Li for research assistance. We received financial support from the Fuqua School of Business at Duke University and Columbia Business School at Columbia University. All errors remain ours.
The Role of Forecast Patterns in Conveying Analysts’ Predictive Ability

1. Introduction

An important element for financial analysts’ success is their predictive ability. Because analysts’ predictive ability is unobservable to investors, it is also important that some mechanism exist for investors to correctly gauge analysts’ ability, so that investors may correctly appraise the information conveyed in analysts’ reports. Several mechanisms have been proposed and studied in the literature. For example, Mikhail, Walther and Willis (1999), Park and Stice (2000), and Chen, Francis and Jiang (2005) suggest that analysts’ track records can be used to convey information about their ability. Other characteristics of analysts’ forecast behaviors, such as their forecasts’ deviation from consensus (a measure of boldness) as well as forecast timing have also been proposed as observable factors that investors might use to learn about analysts’ ability. This paper contributes to this literature by studying the ability of forecast patterns to convey an analyst’s ability.

Our analysis proceeds in two stages. First, we identify the conditions under which an equilibrium exists such that an analyst can convey his ability to the market by adopting a forecast pattern. The particular forecast pattern we examine is a threshold reporting strategy, under which the analyst issues a forecast only when the realized value of his private signal exceeds his chosen threshold. We show that a separating equilibrium exists in which higher-ability analysts adopt higher thresholds than their lower-ability peers, and investors correctly interpret all forecasts. The intuition behind this result is the following. Given the analyst's objective of maximizing a weighted average of his measured forecast accuracy and his expected market influence,\(^1\) a higher threshold is costly because it implies both lower expected accuracy (because only extreme signals are reported) and lower expected market influence (because it reduces the likelihood of issuing forecasts). Because both costs are lower for higher-ability

---

\(^1\) Support for this characterization of the analyst’s objective function is provided by Mikhail, Walther and Willis (1999) and Hong and Kubik (2003). These studies show that analysts’ compensation and reputation depend on their perceived ability (which is increasing in market influence (Gleason and Lee (2003)) and measured forecast accuracy. As detailed in section 2, a forecast’s expected market impact is an increasing function of both the forecast’s information content and investors’ beliefs about the analyst’s ability.
analysts (and therefore, can be born more easily by them), higher-ability analysts can distinguish themselves from lower-ability analysts by choosing higher thresholds.

While our model demonstrates how a threshold strategy may be used to convey ability, its empirical validity (like that of any theoretical model) must be assessed by how well its predictions fit the data. This approach is similar to tests of other theoretical models of analysts’ forecasting strategies. For example, empirical tests of an accuracy strategy (Mikhail, Walther and Willis, 1997; Park and Stice, 2000) presume that analysts convey their ability through their forecast accuracy, and then examine whether more able analysts (as proxied by some characteristic) have more accurate forecasts. As another example, the herding to consensus strategy argues that analysts use this strategy to convey their ability, resulting in predictions concerning boldness (Trueman, 1994; Erhrbeck and Waldmann, 1996); tests of this prediction (such as by Clement and Tse, 2005) examine whether more able analysts are characterized by bolder forecasts. We follow this same approach in the second stage of our analysis where we test two empirical implications from the threshold strategy: 1) by adopting higher thresholds, higher-ability analysts will issue fewer forecasts than lower-ability analysts; and 2) higher-ability analysts issue “bolder” forecasts (i.e., forecasts that are further from the consensus) than their lower-ability peers. Comparative statics for our model also generate predictions about how other factors affect forecast frequency and forecast boldness, holding constant analyst ability. These corollary hypotheses predict that analysts issue fewer forecasts and their forecasts are bolder when the firm's earnings are less predictable, when investors' priors about future earnings are less precise, and when the penalty for forecast errors is smaller.

Empirical results from a sample of about 65,000 quarterly forecasts issued during 1985-2000 are consistent with all of the model’s predictions. Controlling for other factors affecting forecast frequency and forecast boldness (e.g., the number of firms that the analyst follows, proxies for the size of the analyst's employer, as well as firm- and year effects), we find that higher-ability analysts issue fewer forecasts per firm than lower-ability analysts and their forecasts are bolder (both results are significant at the 1% level). We also find that forecast frequency (forecast boldness) is positively (negatively) related both to the predictability of the firm's earnings (as captured by smaller levels of dispersion in analysts'
forecasts) and to the precision of investors' prior beliefs about the firm (as proxied by the extent of analyst coverage and the percentage of institutional holdings). Forecast frequency (forecast boldness) is positively (negatively) related to the penalty for issuing inaccurate forecasts, as captured by the experience of the analyst, with more junior analysts bearing higher costs of inaccuracy than more established analysts. While our result concerning the relation between analyst ability and forecast boldness is not new to the literature, our finding concerning forecast frequency is distinct; moreover, our model demonstrates a common mechanism (the threshold reporting strategy) producing both the frequency and boldness results.

Our paper makes several contributions to the literature. First, we build on McNichols and O’Brien’s (1997) analysis of selection as an explanation for certain forecast patterns by providing a formal characterization of the analyst's endogenous choice of threshold. While our interest in forecast frequency and boldness (rather than forecast bias as in McNichols and O'Brien) produces a different threshold strategy than that assumed by these authors, our model shares many similarities with theirs. In particular, both studies posit a rational, utility-maximizing threshold reporting strategy whereby analysts truthfully reveal their private information when they forecast, but do not necessarily issue forecasts for all private signals received. Second, we believe the combination of our analytical and empirical results, together with results found in McNichols and O’Brien, supports the view that analysts use a threshold strategy. In particular, the threshold strategy is the only strategy which predicts both a positive relation

---

2 Hong, Kubik and Solomon (2000) and Clement and Tse (2005) find a positive relation between analysts’ experience and forecast boldness. The only study which we are aware of examining forecast frequency is Stickel (1992) who shows, in a univariate comparison, that higher ability analysts (as proxied by All-American status) issue more forecasts than do lower ability analysts (i.e., non-All-American analysts). As we discuss in section 3, we believe our finding of lower forecast frequency for higher ability analysts is explained by differences in the design of our tests versus Stickel’s tests.

3 McNichols and O’Brien argue that forecast optimism results from the analyst’s choice to not report their private signal when that signal is sufficiently unfavorable relative to some (unmodeled) threshold. They conjecture that the analyst chooses the threshold by weighing the potential effects of reporting bad news (such as harming relations with client management) against the effects of not issuing a forecast at all (such as lost brokerage commissions). Our characterization of the analyst’s objective function as a weighted average of measured forecast accuracy and market influence is broadly consistent with this specification, adjusted for the fact that we focus on forecast accuracy whereas McNichols and O’Brien model forecast bias.
between ability and boldness and a negative relation between ability and frequency, and our empirical evidence strongly supports both predictions.

Third, while our empirical results are specific to security analysts, the assumptions of our model are not. Consequently, the predictions of our model are generalizable to other forecaster's behavior where the forecaster cares about his forecast accuracy and the impact of his forecasts on some audience. For example, our results contribute to the management forecasting literature which puts forward several hypotheses for why managers infrequently issue earnings forecasts (e.g., proprietary costs (Verrecchia, 1983); uncertainty about information endowment (Dye, 1985; Jung and Kwon, 1987); real effects (Kanodia and Lee, 1998). Our results suggest the additional possibility that, by setting high thresholds (which result in infrequent reporting of their private signals), managers convey their higher-ability to the market. Our empirical tests focus on the forecasting behaviors of sell-side analysts because of the large cross-sectional variation within this group (due both to the number of analysts and to the dispersion in their reporting behaviors); hence, analysts offer a powerful setting to test the predictions of our model, relative to settings characterized by fewer forecasters and less variation in forecasting behavior.

Finally, our findings concerning the relation between ability and forecast patterns have several implications for research that uses analysts' forecasts, either as proxies for the market expectation of earnings or because they are of interest in their own right. Our finding that higher-ability analysts forecast less frequently than lower-ability analysts suggests that measures of consensus earnings expectations based on simple averages of all outstanding forecasts for a given period of time (e.g., the prior 30 days) may be systematically skewed toward the less accurate forecasts of lower-ability analysts – because lower-ability analysts forecast more frequently than their higher-ability peers. Our findings also indicate that the sample mean squared error (MSE) of an analyst's forecasts (a common measure of accuracy) and the dispersion of MSE are systematically understated; as a consequence, studies which focus on MSE alone may have low power and may fail to detect differences in forecast accuracy when they exist. Lastly, our findings suggest that differential use of threshold strategies by analysts – for example, using them early in their careers and discarding them later in favor of another strategy – may
give the appearance of learning-by-doing when no true improvement in forecast accuracy exists. The opposite is also true: the adoption of a threshold strategy by an analyst may make it hard to detect learning-by-doing when in fact true improvement exists.

The rest of the paper is organized as follows. Section 2 presents the model setup and its solution. Section 3 develops testable hypotheses based on the model’s predictions, and reports the results of our tests of these hypotheses. Section 4 concludes.

2. Model

Our model demonstrates the existence of a threshold strategy equilibrium in which higher-ability analysts adopt higher thresholds than lower-ability analysts. We lay out the structure of our model (the "forecasting game") in section 2.1. In section 2.2, we summarize the key results proving the existence of the equilibrium and demonstrating comparative static implications. Section 2.3 describes two implications of our model for prior research studying analyst forecast accuracy, and section 2.4 discusses our model in the context of other strategies that convey predictive ability.

2.1 The forecasting game

In our model, analyst \(i\) issues earnings forecasts for firm \(j\). Let \(z\) denote the earnings of firm \(j\)'s that are being forecasted. (We omit all \(i\) and \(j\) subscripts in this section for notational ease.) Both investors and the analyst have a common prior belief that \(z\) is normally distributed with mean zero and variance \(\sigma^2_z = 1/p_z\). We use \(x\) to denote investors' private signal about \(z\), \(x = z + \varepsilon_x\), where

\[
\varepsilon_x \sim N\left(0, 1/p_x\right)
\]

and is independent of \(z\). We use \(y\) to denote the analyst's private signal about \(z\), with

\[
y = z + \varepsilon_y, \quad \varepsilon_y \sim N\left(0, 1/p_y\right)
\]

and is independent of \(z\) and \(\varepsilon_x\).\(^4\)

We assume that investors do not observe the precision of the noise term in the analyst’s signal (i.e., \(p_y\)). Instead, investors observe the analyst’s forecast pattern and form rational beliefs about \(p_y\)

\(^4\) The assumption that \(\varepsilon_x\) and \(\varepsilon_y\) are independent is not crucial and is made to render the algebra tractable. \(x\) and \(y\) are still correlated via their common association with \(z\); as long as they are not perfectly correlated, we can always orthogonalize them.
based on that pattern. The specific strategy we consider is a threshold forecast strategy, where the analyst discloses his signal $y$ if and only if $|y| \geq b$ where $b \geq 0$. (We discuss other types of strategies in section 2.4.) For each threshold choice $b$, investors associate a precision level of $p_y(b)$. That is, when they receive a forecast $y$ from an analyst with a threshold choice of $b$, investors believe that $y = z + \varepsilon_y$, where $\varepsilon_y \sim N(0, 1/p_y(b))$. The focus of the remaining theoretical analysis is to identify a set of conditions under which: (i) the analyst’s strategy and investors’ beliefs are mutually consistent and (ii) a threshold strategy arises in a fully separating equilibrium.

We assume that both analysts and investors know that $z$ will be realized and observed at the end of the game. Before that, the sequence of events is as follows.

At time $t=0$, the analyst chooses his threshold to maximize his objective function. We assume that the analyst maximizes the weighted sum of the measured accuracy of his forecasts and the total market impact ($MI$) of his forecasts, conditional on his threshold choice. We use the mean squared error ($MSE$) of the analyst’s forecasts to measure the analyst’s forecast accuracy and use the squared revision in investors’ expectation about $z$ upon receiving these forecasts to measure market impact. (We detail the objective function shortly.)

Analysts care about market impact because it demonstrates (to their employers, for example) that investors value their forecasts, and because they benefit from the trading commissions associated with the market impact of investors’ revisions in expectations. To see the latter, note that since $z$ is a determinant of the stock price, changes in investors' expectations about $z$ cause changes in stock price. The larger the revisions of $z$, the more trading investors require to optimally rebalance their portfolios (Hayes (1998)). This characterization of the analyst’s objective function is consistent with McNichols and O’Brien’s conjecture about factors influencing analyst’s coverage selection. It is also

\footnote{Anecdotal evidence supporting the view that analysts benefit from forecasts that create trading activity is provided by a recent article in \textit{Business Week} which reports that "the more trading commissions [independent] analysts generate for their firms' broker/dealer arms, the more bacon they bring home." (\textit{Business Week}, September 8, 2003)}
consistent with evidence that both accuracy and market impact affect analysts’ current market values and future careers (Stickel (1992); Mikhail, Walther and Willis (1999); Hong and Kubik (2003)).

At time \( t=1 \), investors observe their signal \( x \) and update their expectation about \( z \) to

\[
\hat{z}_1 = E[z \mid x] = \frac{p_x}{p_x + p_z} x.
\]

This updating captures the intuitive notion that investors place greater weight on more precise signals.

At some time after \( t=1 \) (we call this \( t=2 \)), the analyst observes his private signal \( y \). We assume that investors do not observe \( y \), nor do they know whether and when the analyst has observed \( y \). The analyst issues a forecast if and only if \( |y| \geq b \). We assume that when the analyst discloses his signal, he discloses truthfully; hence, if the analyst forecasts, he forecasts \( y \). Upon observing \( y \), investors update their expectation about \( z \). Given investors' belief that the analyst has precision of \( \overline{p}_y(b) \), they will revise their expectation about \( z \) from \( z_1 \) to \( z_2 \) where \( z_2 \) is given by:

\[
\hat{z}_2 = E(z \mid x, y, \overline{p}(b)) = (1 - \overline{p}) \hat{z}_1 + \overline{p} y.
\]

where \( \overline{p}(b) = \frac{\overline{p}_y(b)}{(\overline{p}_y(b) + p_x + p_z)} \).

The market impact of the analyst's forecast is given by the squared revision in investors' expectations,

\[
(\hat{z}_2 - \hat{z}_1)^2 = \overline{p}^2 (y - \hat{z}_1)^2.
\]

That is, for given a forecast with news content of \( y - \hat{z}_1 \), investors revise their expectation more, or the forecast's market impact is larger, if investors believe that the analyst

---

6 Analysts' incentives to achieve both forecast accuracy and high trading commissions is also evident from a comment by Manny Korman, director of research at Dean Witter in 1991: "[w]e expect an analyst to do two things - be right, and produce transactions." He was further quoted as saying: "If you make superb recommendations but you’ve done little business, you’ve done the firm little good. ... If your performance is bad and you generate a lot of business, you’ve done a terrible job. ... If your performance is good, and you generate a lot of business, you’ve done a terrific job." ("Analysts devote more time to selling as firms keep scorecard on performance," Wall Street Journal, October 29, 1991, p. C1).

7 We use \( t=2 \) for notational ease to indicate any point in time [not a specific point in time] between the time when investors observe their private signal and the earnings announcement date.
has more precise signals (that is, if the analyst is of higher ability). For this reason, we use
\[ p = \frac{p_y}{p_y + p_x + p_z}, \]
the relative precision of the analyst’s signal, to denote the analyst's predictive ability, with higher values of \( p \) denoting higher ability. Further, for an analyst who has chosen a threshold of \( b \), \( \overline{p}(b) \) is investors' belief, or perception, of his ability.

Intuitively, a forecast’s market impact, \( \overline{p}^2 \left( y - \hat{z}_1 \right)^2 \), depends on both the forecast’s information content (i.e., how much \( y \) differs from the market prior of \( \hat{z}_1 \)) and how much faith the market has in the analyst, as represented by investors’ belief \( \overline{p}(b) \). Rational investors form their beliefs about the analyst’s predictive ability based on all signals about the analyst’s true ability, including the analyst’s past accuracy (if past forecast errors are available) and other observable aspects of forecasting behavior if past accuracy is not a sufficient statistic for true ability. Using forecast patterns to form beliefs about predictive ability may be especially useful for analysts early in their careers who do not have track records. We note, however, that the predictions of our model also imply a pattern of bold forecast behavior for superior analysts, which is consistent with anecdotal evidence that superior analysts are willing to make bold forecasts (see, for example, Blodget’s (2004) characterization of “great” analysts, or Yen’s (2004) discussion of the profitability of StarMine’s Bold Estimates).

We are now ready to formally define a fully separating equilibrium for this forecasting game:

**Definition:** In the forecasting game above, a fully separating equilibrium is a pair of functions: \( b(p) \) and \( \overline{p}(b) \) such that the following conditions hold: (i) Given investors’ perception \( \overline{p}(b) \), there exists a mapping from \( p \to b : b = b(p) \) such that \( b(p) = \arg \max U\left(b, p, \overline{p}(b)\right) \), where \( U\left(b, p, \overline{p}(b)\right) \) is the analyst’s objective function, defined as

\(^8\) Support for the view that forecasts issued by higher-ability analysts have greater market influence is provided by Stickel (1992), Park and Stice (2000), Mikhail, Walther and Willis (2004), and Chen, Francis and Jiang (2005).
\[ U(p, b, \bar{p}(b)) = -\lambda MSE + \Pr(|y| \geq b) MI \]
\[ = -\lambda E[(y - z)^2 | |y| \geq b] + \Pr(|y| \geq b) E\left[\bar{p}(b)^2 (y - \hat{z})^2 | |y| \geq b\right] \]

(ii) \( b(p) \) is strictly monotonic in \( p \) given \( \bar{p} \), and (iii) investors form belief about the analyst's ability based on the observed threshold \( b \), and this belief is correct in equilibrium, i.e., \( \bar{p} = \bar{p}(b(p)) = p, \forall p \).

Notice that the first term in the analyst's objective function (2), \(-\lambda MSE\), is proportional to the expected forecast accuracy of his forecasts. The second term, \( \Pr(|y| \geq b) MI \), is the expected market impact of those forecasts. \( \lambda \) is the penalty on forecast errors relative to the benefit from market impact.

We make two remarks on the objective function. First, the specification of the market impact term implies that the analyst benefits from market impact regardless of whether \( y \) is above or below \( \hat{z} \). It is straightforward to accommodate asymmetric gains by assuming a different \( \lambda \) for the case of \( y \geq b \) than for \( y < b \).9 Second, the market impact term is adjusted by the probability that the analyst issues a forecast, \( \Pr(y \geq b) \), because the analyst cannot enjoy any market impact unless he issues a forecast. It is possible that the absence of a forecast will also cause investors to revise their expectations about \( z \) (conditional on knowing \( y \leq b \)), if investors know, with certainty, that the analyst has observed \( y \). Because we assume that investors do not know whether or when the analyst observes his private signal(s), our model rules out investor revisions conditional on the absence of a forecast. Intuitively, this assumption precludes the analyst from claiming that any market reaction is a response to his silence. This objective function also implies that the \( MSE \) term should not be adjusted by the probability of issuing a forecast. With such an adjustment, the \( MSE \) term would measure the average forecast accuracy of both forecasts that are issued and forecasts about which the analyst keeps silent, assigning zero forecast errors (incorrectly) to the latter.

---

9 Asymmetric gains may lead to biased earnings forecasts made by individual analysts. However, this does not imply that the consensus, the average of all analysts' forecasts, is necessarily biased. For a consensus measure similar to that used here, Chen and Jiang (2005) show that the consensus is less biased and more accurate than a forecast derived from a mechanical forecasting model (which arguably is not subject to the institutional incentives that are typically thought to induce bias, e.g., cultivating management relations, investment banking business, etc.).
The equilibrium definition above follows the standard definition of a fully separating equilibrium as in Riley (1979). For completeness, we make the following standard assumptions for regularity conditions (Riley, 1979).

A1). The analyst's perceived ability can take a continuum of types, \( p \in [\underline{p}, \overline{p}] \), and

A2). The cumulative distribution function of the analyst's ability \( p \) is continuously differentiable of all orders and strictly increasing on \( [\underline{p}, \overline{p}] \).

We make two observations about the model before presenting the solution to the forecasting game. First, although our model is couched in terms of the analyst observing a single signal, it can be easily extended to a setting where the analyst receives a series of exogenous signals. Importantly, it is not necessary for our model that there be an earnings realization following each forecast made by the analyst; our model accommodates the commonly-observed scenario where an analyst issues multiple forecasts for the same firm-quarter prior to the announcement of that firm-quarter earnings. With multiple signals, analysts evaluate whether each signal exceeds the threshold and forecasts only those signals exceeding the threshold. Assuming the analyst’s signals are drawn independently from the same conditional distribution, the analyst’s expected number of forecasts is the product of the number of signals received and the probability that a signal exceeds the threshold. With the number of signals exogenously given, the analyst can increase the number of forecasts only by decreasing his threshold. One might also consider a setting where another dimension of the analyst’s ability is his ability to generate signals (i.e., better analysts generate more signals). While this extension is beyond the scope of our model, we note that an assumption that better analysts generate more signals would work against our empirical findings. Specifically, our model predicts that higher ability analysts issue fewer forecasts than lower ability analysts. If higher ability analysts generate more signals, we would expect a smaller difference in the number of forecasts issued by high and low ability analysts since the larger number of generated signals would offset the greater tendency of high ability analysts to issue only the most newsworthy signals.
Second, for ease of presentation, we describe a setting with one (representative) analyst following the firm; the intuition and the results of the model also apply to settings with multiple analysts. As long as there is uncertainty about each analyst's ability, all analysts can use their threshold choices to reveal their type. Because analysts choose thresholds before they receive any signals, our model does not consider forecast strategies where analysts condition their forecasts on other analysts' prior forecasts; this does not mean, however, that the analyst's threshold choice is independent of other analysts. As we describe in the next section, each analyst's optimal threshold is a function of \( p_x \), the precision of investors' information available before receiving the analyst's forecast. Since \( p_x \) is likely affected by other analysts' forecasts for the firms, each analyst will consider this when choosing his threshold.

2.2 Intuition and basic results

In this section, we solve the forecasting game described in section 2.1. We begin by establishing some basic results to illustrate the intuition; these results also show that the single crossing property is satisfied in our model. Because all variables are assumed to be normally distributed, we focus on the case where the analyst reports whenever \( y \geq b \); symmetry implies that all arguments apply to \( y \leq b \). The analyst's objective function becomes

\[
\max_b U(p, b, \bar{p}(b)) = -\lambda E[(y - z)^2 | y \geq b] + \Pr(y \geq b) E\left[\bar{p} \left(\frac{y - \hat{y}}{\hat{y}}\right)^2 | y \geq b\right]
\]  

(3)

Further, the first term, \( E[(y - z)^2 | y \geq b] \), can be simplified to

\[
MSE(p, b) = E[y^2 | y \geq b] = \frac{1}{p_y^2} \left(1 + \rho_{x,y}^2 \alpha H(\alpha)\right)
\]

(4)

where \( \rho_{x,y}^2 = \frac{p_x}{p_x + p_y} \), \( \alpha = \frac{b}{\sqrt{\text{var}(y)}} \) and \( H(\alpha) = \frac{\Phi(\alpha)}{1 - \Phi(\alpha)} \) is the hazard function of the standard normal distribution.\(^{10}\)

---

\(^{10}\) Expression (4) follows from the standard result that for any two random variables \( x \) and \( y \) that follow a bivariate normal distribution, \( E[x^2 | y \geq b] = \text{var}(x) \left(1 + \rho_{x,y}^2 \alpha H(\alpha)\right) \).
One of the main intuitions for the threshold reporting strategy is immediate from the following lemma. (All proofs are in the appendix.)

**Lemma 1:** For all parameter values: (1a) \( \frac{\partial \text{MSE}(p, b)}{\partial p} < 0 \); (1b) \( \frac{\partial \text{MSE}(p, b)}{\partial b} > 0 \); and (1c) \( \frac{\partial \text{MSE}(p, b)}{\partial p \partial b} < 0 \).

Lemma (1a) states the intuitive result that, for any given threshold, higher-ability analysts' forecasts are more accurate (i.e., have smaller MSE's). Lemma (1b) shows that an increase in the threshold, \( b \), results in lower accuracy for all analysts, and Lemma (1c) shows that this decline in accuracy is less for higher-ability analysts. These comparative static results are due to the varying information content of extreme signals for higher-ability versus lower-ability analysts. Specifically, when a signal deviates significantly from the common prior, it is more likely to contain valuable information if the underlying distribution for the signal has a high precision level; it is more likely to contain noise if the precision level is low. If lower-ability analysts maintain a high threshold, they constrain themselves to disclosing only extreme and (more likely) noisy signals, whereas a higher-ability analyst with the same high threshold discloses extreme and (more likely) informative signals.

Higher-ability analysts also enjoy a relative advantage in market influence. To see this, note that the market impact of the analyst's forecasts is proportional to \( y - \hat{z}_i \), which follows a normal distribution with mean zero and variance \( \frac{1}{p(p_x + p_z)} \). Factoring the \( p \) term out of the expectation sign and using the standard results for jointly distributed normal variables, we can express the market impact term in (3) as:

\[
\Pr(y \geq b) E[p^2 (y - \hat{z})^2 \mid y \geq b] = \Pr(y \geq b) \frac{\bar{p}^2}{p(p_x + p_z)} \left(1 + \rho_{y-z,y}^2 \alpha H(\alpha)\right) = p \Delta(p, b) \tag{5}
\]

where \( \Delta(p, b) \equiv \frac{1}{p(p_x + p_z)} \left[1 - \Phi(\alpha) + \rho_{y-z,y}^2 \alpha \phi(\alpha)\right] \) and \( \rho_{y-z,y}^2 = \frac{1}{p} (1 - \rho_{y,z}^2) \rho_{y,y}^2 \). That is, the market impact of the analyst’s forecasts is proportional to \( \Delta(p, b) \) and is strictly increasing in the analyst's
perceived precision $p$. Since the analyst takes the market’s perception of his ability as given, the effect of his threshold choice on the objective function (via the market impact term) is captured by the effect of the threshold on $\Delta(p,b)$. This effect is formalized in Lemma 2:

**Lemma 2:** For all parameter values, (2a) $\frac{\partial \Delta(p,b)}{\partial b} < 0$. Further, (2b), for any $p_1$ and $p_2$ there exists a $p^0$ such that for all $p > p^0$, $\frac{\partial \Delta(p,b)}{\partial p} > 0$.

Lemma (2a) states that for a given $p$, the analyst’s expected market impact is decreasing in the threshold $b$. That is, (2a) formalizes the intuition that analysts forego opportunities to influence the market when they choose not to issue forecasts. Lemma (2b) demonstrates that, for a given threshold, the marginal cost of not issuing a forecast (that is, the marginal cost of raising the threshold) is higher for lower-ability analysts than for higher-ability analysts. The intuition here is that for a given choice of $b$, the marginal cost of increasing from $b$ to $b' > b$, is higher for lower-ability analysts. This is because the distribution of lower-ability analysts’ signals has higher variance than does the distribution of higher-ability analysts’ signals. As a result, for lower-ability analysts, moving from $b$ to $b'$ eliminates more signals from being forecast than would a similar movement by higher-ability analysts. Hence, lower-ability analysts bear a larger marginal cost of moving to a higher threshold than do higher-ability analysts.

The condition in Lemma (2b) that $p > p^0$ is needed to rule out zero-ability analysts. For analysts with no forecasting ability, the equilibrium is trivial: the analyst’s forecasts are pure noise and investors pay no attention to them. As a consequence, the analyst’s threshold choice has no bearing on his market impact. Continuity of the underlying functions implies that this trivial equilibrium will prevail for analysts with ability close enough to zero (that is, $\frac{\partial \Delta(b,p)}{\partial b} > 0$ holds only for values of $p$ that is strictly bounded away from zero, or, when $p > p^0$). Thus, only when analyst ability is above a lower bound can the separating equilibrium arise, as demonstrated by Lemma (2b). Although $\frac{\partial \Delta(b,p)}{\partial b} > 0$ does not hold
globally for general parameterizations, we find that it holds for a reasonable range of \( p \) values (the details are provided in the appendix, see condition (25) in the Appendix). For example, if \( p_z = 1 \) and \( p_x = 2 \), the positive cross-partial property holds for all \( p_y \geq 1 \). Further, if \( p_z \to 0 \) (i.e., the unconditional information gets noisier relative to conditional information), the range of permissible \( p_y \) approaches the full set of \( p_y \in [0, \infty) \). The intuition for this last result is that what matters for our model predictions is the precision of analysts’ signals relative to \( p_z \). Thus, as \( p_z \to 0 \), even a close-to-zero \( p_y \) would convey some information.

Together, Lemmas 1 and 2 imply that higher-ability analysts can signal their ability by maintaining higher thresholds while achieving the same market impact and without appearing less accurate than lower-ability analysts. These results imply a separating equilibrium where analysts sort themselves based on forecast thresholds. The thresholds reveal analyst type before the realization of earnings (\( z \)), and the market interprets analysts’ abilities correctly. We formally prove this intuition in Proposition 1 below.

**Proposition 1:** Assume A1 and A2. Further, assume \( p > p^0 \) where \( p \) is the lower bound of analyst’s ability type and \( p^0 \) is as specified in Lemma 2. In the forecasting game described above there exists a fully separating equilibrium in which the analyst releases his private signal if and only if the signal he receives is greater than some threshold \( b \). Further, higher-ability analysts maintain higher reporting thresholds than lower-ability analysts.

The following corollary provides a complete characterization of the fully separating equilibrium solution.

**Corollary 1:** Assume A1, A2 and \( p > p^0 \). The analyst’s threshold strategy \( b \) as a function of \( p \) is the solution to the first-order differential equation

\[
\frac{db}{dp} = \frac{2 p \Delta(b)}{\lambda MSE(b) - p^2 \Delta'(b)}
\]

with the initial condition \( b(p) = 0 \), where \( MSE \) and \( \Delta(b) \) are given by (4) and (5). Further, \( \frac{db}{dp} > 0 \), and \( b > 0 \) with probability one, and this equilibrium is unique.
Corollary 2 summarizes the comparative statics of the optimal threshold $b^*$ with respect to the underlying parameters. Since these comparative static results point to cross-analyst and cross-firm factors affecting threshold levels, we add subscripts for the analyst ($i$) and for the firm ($j$).

**Corollary 2**: Under the threshold reporting strategy, the optimal threshold $b^*_{i,j}$ for analyst $i$ issuing forecasts for firm $j$ is decreasing in the precision of the market’s prior about $z$, $p_{zj}$, the precision of investors’ private information, $p_{zj}$, and the relative penalty for forecast errors ($\lambda_{ij}$). That is:

$$(2a) \frac{\partial b^*_{i,j}}{\partial p_{zj}} < 0, \forall i; \quad (2b) \frac{\partial b^*_{i,j}}{\partial p_{zj}} < 0, \forall i; \quad \text{and } (2c) \frac{\partial b^*_{i,j}}{\partial \lambda_{i,j}} < 0.$$

The intuition for the negative relation between $b^*_{i,j}$ and $p_{zj}$ given by Corollary (2a) is that a high $p_{zj}$ makes all analysts’ signals relatively less precise (because $p_{zj} = \frac{p_z}{p_y + p_z + p_{zj}}$ is decreasing in $p_{zj}$).

Therefore, an increase in $p_{zj}$ has the same effect on market impact as a decrease in $p_y$ (i.e., a decrease in analysts’ precision) for all analysts. The intuition for Corollary (2b) is that the more accurate is investors’ signal, $p_{zj}$, the less weight investors place on the analyst’s forecast; in this scenario, maintaining a high threshold does not help the analyst in increasing his market impact, but does hurt his measured forecast accuracy. Lastly, Corollary (2c) formalizes the intuition that high thresholds hurt measured accuracy and are more costly when $\lambda$ is high.

2.3 Implications of our model for prior research

Our threshold equilibrium results have at least two implications for prior studies of analyst forecast accuracy. The first is that the sample mean squared (or absolute) error of an analyst's forecasts, a common measure of accuracy, is not an unbiased estimator of the analyst's true forecasting precision. To see why, note that without adjusting for the analyst's optimal threshold choice, the sample $MSE$ is a conditional $MSE$ of the following form: $E[(y - z)^2 | y \geq b] = \frac{1}{p_y} + \frac{p_z}{p_y(p_z + p_y)} + \alpha H(\alpha)$, where
\[ \alpha = \frac{b}{\sqrt{\text{var}(y)}} \]. Because the second term in the expression is always positive whenever \( b > 0 \), the inverse of sample \( \text{MSE} \) (which is often used as a proxy for the analyst’s accuracy or ability) underestimates the analyst’s true forecasting ability \( (p_\gamma) \). The second implication is that the dispersion of analysts' forecast accuracy is also underestimated. This result obtains because measured forecast errors are increasing in the threshold at the same time that higher-ability analysts choose higher thresholds (i.e., \( \frac{\partial \text{MSE}}{\partial b} > 0 \) and \( \frac{db^*}{dp} > 0 \)). These countervailing forces make it difficult to detect significant differences in analyst forecast accuracy even when such differences exist. Figure 1 illustrates both points. The 45-degree line represents the true precision level and the dotted line shows the measured precision level (the inverse of \( \text{MSE} \)) that result from the analyst’s optimal threshold choice. The measured precision line, with its slope everywhere less than one, always lies below the true precision line, indicating that measured precision underestimates true precision. At the same time, the vertical distance between any two points on the measured precision line (a measure of the dispersion) is smaller than that on the true precision line, indicating that measured precision also underestimates true dispersion in ability.

These results have implications for two strands of the analyst forecasting literature. First are studies that seek to identify high-ability analysts by reference to their forecasting accuracy. This body of research provides mixed results about whether significant differences exist among analysts in terms of forecast accuracy.\(^{11}\) One possible reason for the mixed evidence is provided by our first implication: testing analysts' ability using \( \text{MSE} \) alone, as is done in most studies, has low power and may fail to detect differences when they exist. Our model implies that measures relating to analysts’ forecast patterns (such

\(^{11}\) Studies differ in terms of whether and how they define superior analysts. Using no ex ante (or ex post) classification of superior analysts, O’Brien (1990) concludes that there are no systematic differences in forecast accuracy across individual analysts. Using Institutional Investor All American rankings as his measure of superiority, Stickel (1992) finds that All-American analysts are more accurate than non-All American analysts. In contrast, using lead analysts as their definition of superior analysts, Shroff, Venkataraman and Xin (2003) find that lead analysts are less accurate than follower analysts.
as forecast frequency, forecast distance, or forecast impact) can increase the power of such tests, consistent with the analyses used by Cooper et al. (2001) to identify lead analysts.

Second are studies that examine the effect of experience on forecast accuracy. Here our model provides two insights that help reconcile empirical findings about whether analysts learn from experience. Mikhail, Walther and Willis (1997) and Clement (1999) find that analysts' forecast accuracy (as measured by the inverse sample $MSE$) improves as analysts make more forecasts, consistent with learning-by-doing; in contrast, Jacob, Lys and Neale (1999) argue that this effect disappears when one controls for the analyst’s aptitude and brokerage characteristics. The first insight is that because it is costly to use a threshold strategy, high-ability analysts may stop using the threshold strategy after having built their reputation early in their careers. Discarding the threshold strategy implies that the analyst’s measured forecast accuracy would improve (because they are no longer constrained to issue only extreme forecasts), without any improvement in true forecast ability. Consequently, a finding that forecast accuracy increases with experience may reflect true learning-by-doing effects or it may reflect a change in the forecast strategy of the analyst, such as elimination of a threshold strategy. The second insight is that when there is a learning-by-doing effect, the presence of a threshold strategy reduces the power of empirical tests based on forecast accuracy. This result occurs because as an analyst's precision level, $p_j$, increases with experience, his optimal threshold, $b^*$, increases. These opposing forces reduce the rate of improvement in the ex post measured $MSE$, making it difficult for the researcher to detect learning effects when they exist. Our model suggests that tests of learning-by-doing can be sharpened by controlling for information in analysts' forecast patterns. This implication provides a different motivation for Jacob et al.’s (1999) inclusion of forecast frequency as a control variable in their tests of learning-by-doing.

2.4 Other strategies to convey ability

Although our model shows that an analyst can adopt a threshold strategy to convey his predictive ability, it does not imply that this is the only strategy available to the analyst. Evidence in prior studies suggests at least four other strategies can be used to convey ability: a pure accuracy strategy (Mikhail,
Walther, and Willis, 1999; Park and Stice, 2000; Chen, Francis, and Jiang, 2005); a herding to consensus strategy (Trueman, 1994; Ehrbeck and Waldmann, 1996); a herding to prior own forecast strategy (Ehrbeck and Waldmann, 1996; Prendergast and Stole, 1996); and a timing strategy (Cooper, Day, and Lewis (2001); Lys and Sohn (1990); Shroff, Venkataraman and Xin (2003); Guttman (2005)). Although our model does not speak to which of these four strategies the analyst chooses, it does speak to the consequences of this choice. Specifically, our threshold strategy predicts a positive relation between ability and boldness and a negative relation between ability and frequency; in contrast, other strategies either do not predict both relations or have opposite predictions.

In particular, under a “pure accuracy strategy” analysts optimally use their private signals and issue forecasts that maximize their forecast accuracy. This strategy predicts no relation between ability and frequency, but does predict a positive (mechanical) relation between ability and boldness, because higher ability analysts (by definition) have more precise signals which (by definition) are expected to deviate more from consensus. Under a “herding to consensus strategy” lower ability analysts issue forecasts that are closer to the consensus. This strategy implies a similar prediction concerning boldness as our model (i.e., higher ability analysts issue bolder forecasts), but makes no prediction about frequency.12 Under a “herding to (analyst’s own) prior forecast strategy” (such as Prendergast and Stole, 1996) the equilibrium outcome is that higher ability analysts deviate more (less) from their previous forecast when inferences about their ability are diffuse (tight). As such, this model does not predict any direct relation between ability level and deviations from own prior forecasts; nor does it predict any relations between ability level and frequency. Finally, we consider the predictions from a “timing strategy” whereby higher ability analysts issue forecasts earlier than lower ability analysts (Cooper, Day, and Lewis, 2001). While this strategy has implications for the clustering of analysts’ forecasts in time, it has no implications for either frequency or boldness.

Clement and Tse (2005) find a positive relation between analyst ability and boldness (measured relative to consensus) which they attribute entirely to a herding to consensus strategy. Chen and Jiang (2005) suggest this interpretation may be premature insofar as they show that boldness is not a good measure for herding precisely because of the alternative explanation for the relation between ability and boldness predicted by a pure accuracy strategy.

12 Clement and Tse (2005) find a positive relation between analyst ability and boldness (measured relative to consensus) which they attribute entirely to a herding to consensus strategy. Chen and Jiang (2005) suggest this interpretation may be premature insofar as they show that boldness is not a good measure for herding precisely because of the alternative explanation for the relation between ability and boldness predicted by a pure accuracy strategy.
In summary, the threshold strategy is the only equilibrium strategy considered in the literature which implies both a positive relation between ability and boldness and a negative relation between ability and frequency. Our empirical tests (detailed in the next section) provide evidence about both predictions and, as such, are used to support inferences about whether our model of threshold strategy provides a reasonable description of analysts’ behaviors. An additional feature of our model is its identification of firm-specific and analyst-specific factors affecting boldness and frequency, holding ability constant: notably, dispersion in beliefs, institutional holdings, extent of analyst coverage, and experience of the analyst. As such, our model points to variables that should be included in empirical tests of predictions concerning the relation between ability and either boldness or frequency. At the same time, our model does not, of course, explain all aspects of analysts’ behaviors. For example, because we assume that analysts' incentives are captured by their preferences for forecast accuracy and market influence, our results may not apply to analysts with different incentives (such as incentives to placate management of the firms they cover). In addition, tests of our model provide evidence on the average effects (concerning the relations between ability, frequency and boldness) observed in a broad sample of analysts’ forecasts; our tests do not speak to whether subsets of analysts use one strategy versus another.

3. **Empirical Analyses**

3.1 **Hypothesis development**

As previously discussed, analysts’ threshold choices (i.e., their values of $b$) are not directly observable. However, the threshold strategy has implications for two observable analyst behaviors: the distance between the analyst’s forecast and the consensus (i.e., the “boldness” of his forecast) and the analyst’s forecast frequency. Holding all else constant, a higher threshold implies bolder and less frequent forecasts. Thus, we examine whether analysts’ forecast frequency per firm and forecast boldness are consistent with a threshold strategy. Prior research has studied the correlations between forecast boldness, analysts’ experience and analyst ability (e.g., Hong, Kubik, and Solomon, 2000; Chen and Jiang,
Our analysis adds to this work both by considering other cross-sectional determinants of forecast boldness (predicted by our model) and by examining the relation between ability and forecast frequency.

Proposition 1 implies that higher-ability analysts issue bolder forecasts and forecast less frequently than lower-ability analysts. Testing this prediction calls for a measure of the analyst’s (true) ability. The typical measure of analyst ability used in the forecasting literature is the analyst’s measured forecast accuracy for a given firm (see, for example, Jacob, Lys, and Neale, 1999). Such measured forecast accuracy is problematic for our tests because it is endogenous to the analyst’s threshold choice and because it underestimates analysts’ true ability (as discussed in section 2.3). Specifically, we model the analyst’s true ability as affecting his threshold, which in turn affects both forecast accuracy and market impact; hence, in our setting, forecast accuracy, market impact and threshold are distinct constructs. Our predictions focus on how true ability relates to threshold levels, consequently, we require a proxy for ability that is robust to different forecasting strategies. We use three such measures. The first is the incremental predictive ability of analyst \(i\)’s forecasts to predicting firm \(j\)’s earnings, beyond the predictive ability conveyed by the consensus forecast, \(Ability(R^2)_{i,j}\). \(Ability(R^2)_{i,j}\) is the difference between the \(R^2\) obtained from regressing actual earnings on the consensus forecast and the analyst’s forecast and the \(R^2\) from regressing actual earnings on the consensus forecast alone, i.e., \(Ability(R^2)_{i,j} = R^2_{i,j} - R^2_{i,j} \) (equation 6a) - \(R^2_{i,j} \) (equation 6b) where:

\[
z_{i,j} = \alpha_0 + \alpha_1 \text{Consensus}_{i,j} + \alpha_2 y_{i,j} + \epsilon_{i,j}
\]

\[ (6a) \]

---

13 Park and Stice (2000) find that investors’ perceptions of analyst ability, as measured by forecast accuracy, are firm-specific (i.e., an analyst who is good at predicting the earnings of firm A need not be good at predicting the earnings of firm B, even if the two firms are in the same industry). Given this evidence, we measure analyst ability for each analyst-firm pairing. In sensitivity tests (described in section 3.2), we repeat our analyses using analyst-level measures of ability (which aggregate our analyst-firm level measures for each analyst); results are qualitatively similar for these measures.

14 This measure corresponds to the reduction in investors’ uncertainty about firm earnings upon receiving the analyst’s forecast, scaled by the total uncertainty, i.e., \(\frac{Var(E(z|x)) - Var(E(z|x, y))}{Var(z)}\). Zitzewitz (2001) uses a similar measure, except that he does not scale by total uncertainty. Notice that both in Zitzewitz (2001) and here, this variable is a measure of a forecast’s incremental information content in predicting earnings; it is not a measure of boldness (which is defined as the distance between forecast and the consensus). Importantly, there is no mechanical relation between the boldness measure and the measures of analyst ability.
\[ z_{i,j} = \alpha_0 + \alpha_{\text{Consensus}_{i,j,t}} + \xi_{i,j}. \]  

(6b)

\[ z_{i,j} = \text{firm } j \text{’s actual (quarterly) earnings that analyst } i \text{ is forecasting; } y_{i,j,t} = \text{analyst } i \text{’s forecast for } z_{i,j} \text{ issued at time } t; \text{ Consensus}_{i,j,t} = \text{consensus forecast for } z_{i,j} \text{ at the time that analyst } i \text{ issues } y_{i,j,t}. \)

(Consensus\textsubscript\text{\(i,j,t\})\text{ is subscripted by by } i, j \text{ and } t \text{ because we re-calculate the consensus prevailing at each } y_{i,j,t}. \text{ Both regressions are estimated using all forecasts issued by analyst } i \text{ for a given firm } j. \)

Our second measure is the \(\text{Ability(Dir)}_{i,j}\) measure developed in Chen and Jiang (2005):

\[
\text{Ability(Dir)}_{i,j} = \frac{1}{N_{i,j}} \sum_{t=1}^{N_{i,j}} \text{sign}(\text{FE\textsubscript{\text{Cons}}}_{i,j,t} \cdot \text{Dev}_{i,j,t})
\]

(7)

where \text{sign(.)} \text{ is the sign function; } \text{FE\textsubscript{\text{Cons}}}_{i,j,t} = z_{i,j} - \text{Consensus}_{i,j,t} \text{ (the forecast error of the consensus); } \text{Dev}_{i,j,t} = y_{i,j,t} - \text{Consensus}_{i,j,t} \text{ (the deviation between the analyst’s forecast for firm } j \text{ issued at time } t \text{ and the consensus forecast for firm } j \text{ at the time the forecast is issued); and } N_{i,j} \text{ is the total number of tested forecasts (tested forecasts are those whose errors are observed in the sample period) by analyst } i \text{ for firm } j. \text{ Intuitively, } \text{Ability(Dir)}_{i,j} \text{ measures the frequency that analyst } i \text{’s forecasts move the new consensus (after incorporating his forecasts) in the direction of reported earnings. Note that when a forecast moves the consensus towards reported earnings, } \text{sign(Dev)} = \text{sign(\text{FE\textsubscript{\text{Con}}})}, \text{ yielding a positive value for the product of the two terms. Chen and Jiang (2005, Proposition 2) formally show that this measure is positively related to the analyst’s true ability but unrelated to the analyst’s forecast strategy conditional on his true ability under weak conditions. Consequently, any relation found between } \text{Ability(Dir)}_{i,j} \text{ and measures of the forecasting strategy (such as boldness and frequency) cannot be due to any mechanical relation between these variables.}

Our last measure of ability is whether the analyst is a member of the All-American Research team chosen by Institutional Investor. Each year, Institutional Investor polls about 2,000 money managers to evaluate analysts on the basis of four criteria: stock picking, earnings forecasts, written reports, and overall service. Analysts receiving the highest evaluations in each industry are selected to the All-
American Research Team and their names are published in the October issue of Institutional Investor. Stickel (1992) shows that, on average, All-American analysts have more accurate earnings forecasts than non-All-American analysts, but finds that the market impact to All-American analysts’ stock recommendations is mixed, with larger reactions found for upgrades and no differential reaction found for downgrades. One problem with using All Star status as a measure of analyst ability is that two of the evaluation criteria (written reports and overall service) are not well-linked to objective notions of ability, such as accuracy or market impact. With this caveat in mind, we use the dummy variable AllStar equal to one if the analyst is selected as an All-American Research team member from 1988-2000 by Institutional Investor, and zero otherwise.

Results (2a) and (2b) of Corollary 2 imply that, among analysts following different firms, analysts issue more forecasts for firms where the precision of the market’s priors about earnings \( p_{zj} \) is high, and where the precision of investors’ private information \( p_{xj} \) is high. Tests of these predictions require proxies for \( p_{zj} \) and \( p_{xj} \). We interpret \( p_{zj} \) as the predictability of firm \( j \)’s earnings; following prior research, we proxy for this construct using the dispersion in analysts’ forecasts for firm \( j \) in year \( T \),\(^{15}\) where greater dispersion indicates lower earnings predictability.\(^{16}\) We interpret \( p_{xj} \) as the amount or quality of information that investors have before receiving analyst \( i \)’s forecast; following prior studies, we measure this construct using institutional holdings and analyst coverage.\(^{17}\) The reasoning is that, on average, the market has more precise information about a firm when more of the firm’s shares are held by institutions and when analyst coverage is high.

\(^{15}\) We use upper-case \( T \) to index our sample year (described in section 3.2 to be \( T = 1985, \ldots, 2000 \)). Lower-case \( t \) is used to index the point in time when the analyst issues a forecast.

\(^{16}\) For example, Abarbanell and Lehavy (2003), Ashbaugh and Pincus (2001), Bowen, Davis, and Matsumoto (2002), Chaney, Hogan, and Jeter (1999), Clement, Frankel, and Miller (2003), and Heflin, Subramanyam, and Zhang (2003) use analysts’ forecast dispersion to proxy for firms’ earnings’ predictability. To address concerns that forecast dispersion may be influenced by the same threshold issues that we examine, we repeat our analyses using the standard deviation of the firm’s historical earnings as the proxy for earnings predictability; results (not reported) are unchanged.

\(^{17}\) For example, Lang, Lins, and Miller (2003), Elgers, Lo and Pfeiffer (2001) and Gleason and Lee (2003) use the number of analysts following to proxy for investors’ quality of information; whereas Balsam, Bartov, and Marquardt (2002), Bushee and Noe (2000), Bartov, Radhakrishnan, and Krinsky (2000), and Collins, Gong, and Hribar (2003) use institutional holdings to proxy for the same construct.
Result (2c) of Corollary 2 implies that analysts with relatively small penalties for forecast errors ($\lambda$) issue fewer forecasts. We use analyst $i$’s experience in forecasting firm $j$’s earnings as an inverse proxy for the penalty, or costliness, of the analyst’s forecast errors ($\lambda_{i,j}$). Support for this proxy comes from research on reputation concerns (e.g. Holmstrom (1999) and Chevalier and Ellison (1999)) which shows that inexperienced analysts are likely to suffer more (than experienced analysts) from making inaccurate forecasts for two reasons: they benefit more from establishing a pattern of accurate forecasts because their expected tenure in the profession is longer; and the market’s assessment of their true ability is more diffuse (due to the paucity of past forecasts on which to base an assessment of accuracy) and therefore, large forecast errors inflict greater damage on their reputation.

In summary, our model’s predictions lead to four testable hypotheses concerning forecast frequency and forecast boldness:

**H1a:** Higher ability analysts (as measured by larger values of $\text{Ability(Dir)}$ and $\text{Ability}(R^2)$ or $\text{AllStar} = 1$) issue fewer forecasts than lower-ability analysts in a given period.

**H1b:** Holding ability constant, analysts issue fewer forecasts for firms where there is more dispersion in outstanding forecasts, less institutional holdings and less analyst following, and when they [analysts] are more established in their careers.

**H2a:** Higher ability analysts (as measured by larger values of $\text{Ability(Dir)}$ and $\text{Ability}(R^2)$ or $\text{AllStar} = 1$) issue bolder forecasts (as measured by deviations from consensus) than lower-ability analysts.

**H2b:** Holding ability constant, analysts issue bolder forecasts for firms where there is more dispersion in outstanding forecasts, less institutional holdings and less analyst following, and when they [analysts] are more established in their careers.

### 3.2 Sample description, methodology, and results

We obtain analyst quarterly earnings forecasts from the Zacks Investment Research database, firms’ earnings and stock price data from Compustat and CRSP, and firms’ institutional holding data from Spectrum. We begin by identifying analysts whose first forecasts (for any firm) recorded by Zacks were made on or after January 1, 1985. This constraint ensures that the experience of the analyst can be measured precisely. Following the standard practice in the literature, we eliminate forecasts for firms
whose average share prices are less than $5 and average market capitalizations are less than $100 million (both in 2001 CPI-deflated dollars) to mitigate the influence of extreme outliers.

Our first dependent variable, \( Freq_{i,j,T} \) is the number of quarterly earnings forecasts (i.e., \( y_{i,j} \)) issued by analyst \( i \) for firm \( j \) in year \( T \) (\( T = 1985, \ldots, 2000 \)).\(^{18}\) We measure frequency for analyst-firm pairings based on Park and Stice’s (2000) finding that analyst forecasting ability is firm-specific. Because an analyst may initiate or terminate coverage of a firm at any time during the year, we exclude the first year and the last year an analyst appears in the sample to avoid truncation due to the analyst’s coverage decision. Our final sample contains 65,201 observations, covering 2,549 firms, 2,779 analysts, and 20,759 firm-analyst pairings.

Our second dependent variable, \( \text{Boldness}_{i,j,T} \), is calculated as follows:

\[
\text{Boldness}_{i,j,T} = \frac{\text{Average } |\text{Dev}/P_5| \text{ for analyst-firm pair } i,j \text{ in year } T}{\text{Average } |\text{Dev}/P_5| \text{ for all analyst-firm pairs in the same 2-digit SIC industry in year } T},
\]

where \( P_5 \) is the stock price five days before the forecast date, and \( \text{Dev} \) was defined previously.

For each forecast (\( y_{i,j} \)) in an analyst-firm pairing, we calculate the consensus earnings forecast for firm \( j \) at the time the forecast was issued (\( \text{Consensus}_{i,j} \)) as the weighted average of all outstanding forecasts. In combining forecasts to form the consensus, we use the linear forecast order weighting scheme described in Chen, Francis and Jiang (2005). Specifically, if there are \( n = 1, \ldots, N \) prevailing forecasts that are issued \( d_n \) days before the current forecast date, with \( d_n = d_{N-1} > \ldots > d_1 \), the linear forecast order weighting assigns weight \( w_n = \frac{N - n + 1}{\sum_{n=1}^{N} (N - n + 1)} \) to the \( n \)’th forecast. This weighting assigns larger weights to more recent forecasts. (Our results are not sensitive to alternative ways of calculating

---

\(^{18}\) If the analyst receives multiple signals in a given period, then the expected number of forecasts (i.e., his forecast frequency during the year), is \( Freq = \text{Num\_Signal} \times \text{Prob(signal exceeds threshold)} \), where \( \text{Num\_Signal} \) is the number of signals the analyst receives during the period. Our frequency tests use \( Freq \) to proxy for the threshold, controlling for the number of signals each analyst may receive by including in the regressions firm-dummies, variables controlling for firm-year specific amount of information (such as market size and market to book value), and variables controlling for analyst specific characteristics such as the number of firms an analyst cover and the size of his employing brokerage firm.
We use Consensus\_\textsubscript{i,j}, together with \(y_{i,j}\), to calculate Ability\(\textsuperscript{(Dir)}\)\_\textsubscript{i,j} as specified in equation (7), and to calculate Ability\(\textsuperscript{(R^2)}\)\_\textsubscript{i,j} estimated for each analyst-firm pairing, using all forecasts issued by the analyst over the sample period.

For each \(y_{i,j}\), we also calculate the standard deviation among all outstanding forecasts at time \(y_{i,j}\) is issued, scaled by the stock price five days before \(y_{i,j}\) was issued. We take the average of these standard deviations about firm \(j\) in year \(T\) to obtain our measure of dispersion (\(STD_{i,j,T}\)). We measure institutional holdings as the percent of firm \(j\)’s shares held by institutional investors in year \(T\) (\(%INST_{i,j,T}\)), and measure analyst coverage as the number of analysts issuing at least one forecast about firm \(j\) in year \(T\) (\(OTHER\_ANLST_{i,j,T}\)). Lastly, for each year \(T\), we measure analyst \(i\)’s experience with respect to forecasting firm \(j\)’s earnings as the number of years since he issued his first forecast for firm \(j\) (\(EXP_{i,j,T}\)).

Table 1 reports summary statistics on these variables. The mean (median) number of quarterly earnings forecasts issued by our sample analysts is 12.6 (10) per year with a standard deviation of 9.7. Our first proxy for analysts’ true predictive ability, \(Ability\(\textsuperscript{(R^2)}\)\_\textsubscript{i,j}\), has a mean value of 0.12, indicating that on average, analysts add an additional 12% explanatory power to the consensus in predicting earnings. Our second proxy for analyst’s true predictive ability, \(Ability\(\textsuperscript{(Dir)}\)\_\textsubscript{i,j}\), has a mean (median) value of 0.62 (0.63), indicating that on average, 62% of analysts’ forecasts move the consensus in the direction of realized earnings. Our third proxy for analysts’ true predictive ability, \(AllStar\)\_\textsubscript{i,j}, has a mean value of 0.35, indicating that 35% of our sample forecasts are made by All-Stars analysts. In unreported tests, we find that the average values of \(Ability\(\textsuperscript{(Dir)}\)\_\textsubscript{i,j}\) and \(Ability\(\textsuperscript{(R^2)}\)\_\textsubscript{i,j}\) are both higher than those for non-All-Star analysts (differences significant at the 1% level).

Table 2 reports Pearson and Spearman rank correlation coefficients among all test variables. We note that all three ability variables are positively correlated to each other, although the correlation between \(Ability\(\textsuperscript{(Dir)}\)\_\textsubscript{i,j}\) and \(Ability\(\textsuperscript{(R^2)}\)\_\textsubscript{i,j}\) (of 0.26-0.29) is larger than the correlation between either of these
variables and AllStar (of 0.030-0.06). More importantly, all of the ability measures are negatively related to Freq and positively related to Boldness. Table 2 also shows that Freq is positively correlated with OTHER ANLST, %INST, and BROKER, and that Boldness is negatively correlated with OTHER ANLST and %INST. Except for the positive correlations between EXP and Freq, these correlations are consistent with predictions given by Hypotheses 1 and 2.

To test our hypotheses, we estimate the following firm-year fixed effects equation:

\[
Freq_{i,j,T} \text{ or } Boldness_{i,j,T} = \alpha_j + \beta_T + \gamma_1 Ability_{i,j} + \gamma_2 EXP_{i,j,T} + \gamma_3 STD_{i,j,T} \\
+ \gamma_4 %INST_{j,T} + \gamma_5 OTHER\_ANLST_{j,T} + \gamma_6 BROKER_{i,T} + \gamma_7 NUM\_COVER_{i,T} \\
+ \gamma_8 \log(MV_{j,T}) + \gamma_9 MB_{j,T} + \xi_{i,j,T} \tag{8}
\]

where \( \alpha_j \) is the firm-specific intercept and \( \beta_T \) is the year-specific intercept. We include firm and year dummies to control for cross-firm and cross-year differences in the amount of information available about a given firm. Further, we include two control variables capturing the amount of firm-specific information about firm \( j \): the size of the firm \( j \), measured as the logarithm of firm \( j \)’s beginning of year \( T \) market equity value (Log(MV\(_j,T\))), and firm \( j \)’s market-to-book ratio, measured as the ratio of beginning of year market value of equity to beginning of year book value of equity (MB\(_j,T\)). If more information is available for large, high growth firms, we expect more frequent forecasts for these firms; we have no predictions as to how these variables affect boldness. We also include two variables to control for factors specific to individual analysts. The first is a measure of the size of the analyst’s employer (BROKER\(_i,T\), equal to the number of analysts employed by the brokerage in year \( T \)) to control both for the amount of resources available to the analyst (we expect larger brokerages have more resources) and for reporting behaviors specific to the brokerage (for example, if the brokerage requires analysts to issue at least one forecast every quarter). While we allow BROKER\(_i,T\) to affect Freq, we do not predict the direction of this relation.

---

19 The lower correlations found for AllStar are likely due to the fact that the criteria for All-American status include attributes other than those reflected in our model’s objective function (e.g., written reports and overall service). To the extent that an analyst is named as an All-American because of these other attributes, we would not necessarily expect AllStar to correlate highly with Ability(Dir)\(_{i,j}\) and Ability(R\(^2\))\(_{i,j}\) because the latter measures focus more on the analyst’s ability to predict earnings and to move the consensus in the correct direction.

20 For example, Bamber, Barron, and Stober (1997), Barth and Kasznik (1999), Collins, Rayburn, and Kothari (1987), and Rees and Elgers (1997), among others, use firms’ market value of equity and market to book ratios as proxies for the general amount of information available about the firms.
The second variable is the number of firms that an analyst covers in year $T$ (i.e., $NUM\_COVER_{i,T}$) to control for the fact that analysts covering multiple firms have, on average, less time per covered firm, and therefore are likely to issue fewer forecasts per firm.$^{21}$ (Although it is also possible that analysts covering multiple firms issue more forecasts because the covered firms are in the same or related industries and thus afford economies of scale.) It is unclear how $BROKER$ and $NUM\_COVER$ affect $Boldness$.

For $Freq$, Hypothesis 1a implies that $\hat{\gamma}_1 < 0$, and Hypothesis 1b implies that $\hat{\gamma}_2 < 0,\hat{\gamma}_3 < 0,\gamma_4 > 0$ and $\hat{\gamma}_5 > 0$. Table 3 presents the regression results for $Freq$ using $Ability(R^2)$ (column 1), $Ability(Dir)$ (column 2), and $AllStar$ (column 3) as the proxy for analyst ability. All $t$-statistics are based on heteroskedasticity adjusted robust standard error estimates. Turning first to tests of Hypothesis 1a, we find that the coefficients on all ability measures are significantly negative at the 1% level. These results indicate that higher-ability analysts issue fewer forecasts than lower-ability analysts. In terms of economic magnitude, the coefficient estimate of -2.08 on $AllStar$ means that, all else constant, All-Star analysts make two fewer quarterly earnings forecasts each year than do non All-Star analysts.$^{22}$

Tests of Hypothesis 1b focus on the coefficient estimates on $EXP$, $%INST$, $OTHER\_ANLYST$ and $STD$. In Table 3, results using $Ability(R^2)$ and $Ability(Dir)$ as proxies for analyst ability show the expected negative relations between $Freq$ and both $EXP$ and $STD$ (significant at less than the 1% level for

---

$^{21}$ Our model takes the analyst’s decision to follow the firm as given. As such, our empirical tests treat $NUM\_COVER$ and $OTHER\_ANLYST$ as exogenous variables. In reality, analysts have (at least some) discretion over how many and which firms they follow, suggesting that $NUM\_COVER$ and $OTHER\_ANLYST$ are likely to be endogenous. To address this endogeneity would require modeling the analyst’s coverage decision within the threshold reporting strategy. The coverage decision can be modeled as follows: we know that for each set of exogenous parameters (such as uncertainty about firms’ earnings), there will be an equilibrium threshold choice, which is the optimal choice for the analyst conditional on choosing to cover the firm. This threshold choice will determine the analyst’s equilibrium utility level. Assuming an exogenous cost of starting coverage, the analyst will initiate coverage if and only if the startup cost is lower than his expected utility. It is straightforward to show that the parameters that would make an analyst choose a high threshold in equilibrium will also lead to lower expected utility, and therefore, to fewer firms being followed.

$^{22}$ Using forecasts for annual earnings issued from 1981 to 1985, Stickel (1992) finds that All American Analysts issue more forecasts for annual earnings than non-All American analysts both before and during (but not after) their time as All American Analysts. Our finding that higher ability analysts issuer fewer forecasts conflicts with Stickel’s results. Although there are several differences between our study and Stickel’s (including time period, use of quarterly versus annual forecasts, and sample size), we believe the two main difference are that: 1) our tests control for other factors affecting frequency, such as time-specific effects, firm-specific effects (e.g., growth, firm size), and analyst-specific effects (e.g., tenure, size of employer); and 2) our tests compare the frequency of reporting of All-American Analysts over their lifetime with that of non-All American Analysts over their lifetimes.
and 5% level for STD), and the expected positive relations between Freq and both %INST and OTHER_ANLYST (significant at the 1% level). The negative associations are consistent with the prediction that more junior analysts issue more forecasts than do more established analysts and that analysts issue fewer forecasts for firms with less predictable earnings. The positive associations found for %INST and OTHER_ANLYST are consistent with the prediction that analysts issue more forecasts when investors have more precise prior information about earnings. Results based on AllStar as the measure of analyst ability are similar, with the exception that we find no relation between Freq and EXP.

Table 4 reports results testing Hypotheses 2a and 2b. When Boldness is the dependent variable, Hypothesis 2a implies that \( \gamma_1 > 0 \), and Hypothesis 2b implies that \( \gamma_1 > 0, \gamma_2 > 0, \gamma_4 < 0 \) and \( \gamma_5 < 0 \).

Columns (1) and (2) in Table 4 show that the coefficients for both Ability(R²) and Ability(Dir) are positive and significant at less than the 1% level. The coefficient for AllStar is positive, but not significant at less than 10% level for a two-sided test (it is significant at less than 10% level for a one-sided test). Overall, these results indicate that forecasts issued by higher-ability analysts, on average, deviate more from consensus than forecasts issued by lower-ability analysts. Table 4 also shows that the coefficients for EXP and STD are positive and significant at less than 1% level, consistent with Hypothesis 2b that more experienced analysts issue bolder forecasts and that analysts issue bolder forecasts for firms with less predictable earnings. The coefficient for OTHER_ANLST is significantly negative at less than 1%, consistent with the prediction that analysts issue bolder forecasts when investors have less precise prior information about earnings. The coefficient for %INST is positive, seemingly inconsistent with Hypothesis 2b. However, Table 2 shows that %INST and OTHER_ANLST are highly correlated. In results not shown, we re-estimate Table 4 without OTHER_ANLST and find that the coefficient for %INST is negative, although not significant at any conventional level. Further, in Table 5 where we use the Fama-MacBeth regressions (discussed next), the coefficient for %INST is significantly negative.

We examine whether our results are robust to estimating equation (8) by year (and excluding the firm- and year- fixed effects), and testing the across-year mean value of the coefficient estimates (Fama
and MacBeth (1973)). We use Healy et al.’s (1987) Z-score to assess the statistical significance of the annual t-statistics; this procedure alleviates the impact on standard error estimates from cross-sectional correlation in error terms. The Z-score is \[ Z = \sqrt{N-1} \frac{\bar{t}}{\sigma_t} \] where \( \bar{t} \) is the mean of the t-statistics from the yearly regressions, \( \sigma_t \) is the standard deviation of those annual t-statistics, and \( N \) is the number of sample years (\( N=16 \)). Results of the annual estimations are summarized in Table 5. We note that for \( \text{Freq} \), except for the coefficients for \( \text{STD} \) which are not reliably different from zero, all coefficient estimates retain their signs and statistical significance. For \( \text{Boldness} \), all coefficient estimates retain their signs and statistical significance.

We also examine whether our results extend to analyst-level ability, rather than analyst-firm level ability. For these tests, the analyst-specific ability measure is calculated as the average of each analyst-firm ability measure for all firms analyst \( i \) covers in our sample. The results from examining the analyst-specific measures (not reported) are qualitatively similar to those documented for the analyst-firm measure of ability. Specifically, we find that higher ability analysts (as measured by all three ability measures) issue fewer and bolder forecasts than do lower ability analysts; and we find that, holding analyst ability constant, analysts issue fewer and bolder forecasts, for firms with greater earnings uncertainty and less sophistication and when the analyst is less established.

In summary, consistent with Hypotheses 1a and 2a, we find that higher-ability analysts issue bolder forecasts and forecast less frequently than lower-ability analysts. We also find that, conditional on ability, analysts issue fewer and bolder forecasts for firms with high forecast dispersion, low analyst coverage and institutional holdings, and when they [analysts] are more experienced (consistent with Hypotheses 1b and 2b).

4. Conclusion

We establish a role for forecast patterns to credibly convey analysts’ forecasting abilities. We show that analysts can signal their ability by adopting a threshold strategy where they issue forecasts only
when their private signals exceed a threshold level, with higher-ability analysts adopting higher thresholds. All else equal, higher thresholds imply lower forecast frequency and greater forecast boldness. Given this relation, we report empirical evidence on how frequency and boldness (our proxies for threshold levels) vary with analyst ability. Using a large sample of analyst forecast data, we find strong evidence supporting the predicted relations: higher-ability analysts issue fewer and bolder forecasts than do lower ability analysts. We also find evidence consistent with other comparative static predictions of the model. Notably, we find that controlling for analyst ability, forecast frequency is higher and forecasts are closer to the consensus, when there is less uncertainty about a firm’s future earnings, when there is more precise information available to investors, and when the analyst is less experienced.

Taken together, we believe the combination of the theoretical model and the empirical evidence provide consistent and strong evidence suggesting analysts’ use of a threshold strategy in their forecast behavior. In particular, no other forecasting strategy considered by prior research yields predictions about both frequency and boldness, and both of these predictions are supported by our empirical tests. Our analysis is also distinctive in that we formally model and test other variables affecting the two main predicted relations. In contrast, other analytical studies do not explicitly model these other factors; and most empirical studies, while including some of these variables as controls in their tests, provide only qualitative arguments for their association with the test variable.
References


We use the following parameterization: \( p_x = 1, p_e = 2, \) and \( \lambda = 1. \) Measured precision lies everywhere below true precision because when analysts use a threshold strategy, their measured forecast accuracy underestimates the true precision of their private signals.
Table 1: Summary statistics

Sample analysts consist of those whose first forecasts recorded on Zacks Investment Research Database appeared on or after January 1, 1985. Number of observations in the sample is 65,201 analyst-firm-year. \( Freq_{i,j,T} \) is the number of forecasts for quarterly earnings issued by an analyst \( i \) for a firm \( j \) in year \( T \), excluding the first and last year the analyst appears in the dataset. \( Boldness_{i,j,T} \) is the year \( T \) average distance between analyst \( i \)’s forecast and consensus (scaled by price five days before forecast date), relative to the average distance of all analyst forecasts in the same SIC 2-digit industry (in percentage). \( NUM\_COVER_{i,T} \) is the number of firms covered by analyst \( i \) in year \( T \). \( Ability_{j} \) is the ability measure, where we use three proxies: \( Ability(R^2) = \text{difference between the } R^2 \text{ from regressing firm earnings on the consensus forecast and the } R^2 \text{ from regressing firm earnings on the consensus forecast and the analyst’s forecasts, using all forecasts the analyst has made for the firm in the sample period}; \) \( Ability(Dir) = \text{the frequency that an analyst’s forecasts move the new consensus (after incorporating his forecasts) in the direction of reported earnings}; \) \( AllStar_{j} \) is a dummy that equals to 1 if the analyst has been selected as an All-American Research Team from 1988-2000 by Institutional Investor, and 0 otherwise. \( EXP_{i,j,T} \) is the number of years since the analyst’s first forecast. \( STD_{i,j,T} \) is the standard deviation among all prevailing forecasts, scaled by stock price five days ago before the analyst’s forecast, and then averaged across all forecasts made by the analyst in a given year. \( %INST_{j,T} \) is the percentage of shares in firm \( j \) held by institutional investors in year \( T \). \( OTHER\_ANLST_{j,T} \) is the number of analysts following firm \( j \) in year \( T \). \( BROKER_{i,j,T} \) is the number of analysts employed in year \( t \) by analyst \( i \)’s employer in year \( T \). \( MV_{j,T} \) and \( MB_{j,T} \) are firm \( j \)’s market equity value and market to book equity ratio at the beginning of year \( T \), respectively.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std dev</th>
<th>1%</th>
<th>25%</th>
<th>Median</th>
<th>75%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq</td>
<td>12.59</td>
<td>9.68</td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>17</td>
<td>44</td>
</tr>
<tr>
<td>Boldness</td>
<td>90.49</td>
<td>127.85</td>
<td>4.52</td>
<td>30.36</td>
<td>57.30</td>
<td>106.11</td>
<td>544.3</td>
</tr>
<tr>
<td>NUM_COVER</td>
<td>12.61</td>
<td>8.16</td>
<td>2</td>
<td>8</td>
<td>11</td>
<td>15</td>
<td>45</td>
</tr>
<tr>
<td>Ability(R²)</td>
<td>0.12</td>
<td>0.16</td>
<td>0.00</td>
<td>0.015</td>
<td>0.061</td>
<td>0.163</td>
<td>0.761</td>
</tr>
<tr>
<td>Ability(Dir)</td>
<td>0.62</td>
<td>0.11</td>
<td>0.32</td>
<td>0.56</td>
<td>0.63</td>
<td>0.70</td>
<td>0.89</td>
</tr>
<tr>
<td>AllStar</td>
<td>0.35</td>
<td>0.48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>EXP</td>
<td>3.23</td>
<td>2.51</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>STD</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.001</td>
<td>0.002</td>
<td>0.005</td>
<td>0.027</td>
</tr>
<tr>
<td>%INST</td>
<td>45.68</td>
<td>23.53</td>
<td>0</td>
<td>31.147</td>
<td>49.380</td>
<td>63.438</td>
<td>88.067</td>
</tr>
<tr>
<td>OTHER_ANLST</td>
<td>11.09</td>
<td>6.38</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>14</td>
<td>31</td>
</tr>
<tr>
<td>BROKER</td>
<td>33.91</td>
<td>23.39</td>
<td>1</td>
<td>14</td>
<td>29</td>
<td>51</td>
<td>96</td>
</tr>
<tr>
<td>MV (in Smil)</td>
<td>5427.81</td>
<td>15947.44</td>
<td>55.754</td>
<td>423.833</td>
<td>1284.241</td>
<td>3825.464</td>
<td>74526.96</td>
</tr>
<tr>
<td>MB</td>
<td>3.20</td>
<td>4.14</td>
<td>0.606</td>
<td>1.560</td>
<td>2.315</td>
<td>3.604</td>
<td>15.447</td>
</tr>
</tbody>
</table>
Table 2: Correlations Between All Variables

See Table 1 for variable descriptions. The number of observation is 65,201. Pearson correlation coefficients are shown in the upper right panel and spearman correlation in the lower left panel. All are significant at less than the 1% level except those shown in italic font (which are significant at less than the 5% level) and those underlined (which are not significant at conventional level).

<table>
<thead>
<tr>
<th></th>
<th>Freq</th>
<th>Boldness</th>
<th>Ability (Dir)</th>
<th>Ability (R²)</th>
<th>AllStar</th>
<th>EXP</th>
<th>STD</th>
<th>%INST</th>
<th>OTHER_ANLST</th>
<th>LOG (MV)</th>
<th>MB</th>
<th>BROKER</th>
<th>NUM_COVER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq</td>
<td>-0.02</td>
<td>-0.06</td>
<td>-0.17</td>
<td>-0.07</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.10</td>
<td>0.16</td>
<td>0.07</td>
<td>0.09</td>
<td>0.12</td>
<td>-0.05</td>
<td></td>
</tr>
<tr>
<td>Boldness</td>
<td>0.04</td>
<td>0.11</td>
<td>0.05</td>
<td>0.02</td>
<td>-0.01</td>
<td>0.59</td>
<td>-0.05</td>
<td>-0.15</td>
<td>-0.09</td>
<td>-0.08</td>
<td>0.01</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Ability (Dir)</td>
<td>-0.06</td>
<td>0.12</td>
<td>0.26</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability (R²)</td>
<td>-0.16</td>
<td>0.05</td>
<td>0.29</td>
<td>0.03</td>
<td>-0.04</td>
<td>0.03</td>
<td>-0.05</td>
<td>-0.12</td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.01</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>AllStar</td>
<td>-0.07</td>
<td>0.04</td>
<td>0.06</td>
<td>0.04</td>
<td>0.12</td>
<td>0.03</td>
<td>0.03</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.02</td>
<td>0.35</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>EXP</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.15</td>
<td>-0.01</td>
<td>0.12</td>
<td>0.10</td>
<td>0.14</td>
<td>-0.01</td>
<td>0.07</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>STD</td>
<td>-0.02</td>
<td>0.60</td>
<td>0.13</td>
<td>0.03</td>
<td>0.05</td>
<td>0.00</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-0.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>%INST</td>
<td>0.11</td>
<td>-0.05</td>
<td>0.01</td>
<td>-0.05</td>
<td>0.03</td>
<td>0.13</td>
<td>0.01</td>
<td>0.21</td>
<td>0.03</td>
<td>-0.03</td>
<td>0.08</td>
<td>-0.03</td>
<td></td>
</tr>
<tr>
<td>OTHER_ANLST</td>
<td>0.14</td>
<td>-0.20</td>
<td>-0.12</td>
<td>-0.12</td>
<td>0.11</td>
<td>-0.03</td>
<td>0.25</td>
<td>0.40</td>
<td>0.17</td>
<td>0.08</td>
<td>0.08</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>LOG(MV)</td>
<td>0.07</td>
<td>-0.23</td>
<td>-0.08</td>
<td>0.10</td>
<td>0.21</td>
<td>-0.17</td>
<td>0.21</td>
<td>0.62</td>
<td>0.22</td>
<td>0.21</td>
<td>0.04</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td>MB</td>
<td>0.17</td>
<td>-0.31</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-0.05</td>
<td>-0.29</td>
<td>0.06</td>
<td>0.21</td>
<td>0.24</td>
<td>0.05</td>
<td>-0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BROKER</td>
<td>0.15</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.01</td>
<td>0.37</td>
<td>0.07</td>
<td>0.00</td>
<td>0.08</td>
<td>0.06</td>
<td>0.10</td>
<td>0.05</td>
<td>-0.00</td>
<td></td>
</tr>
<tr>
<td>NUM_COVER</td>
<td>-0.04</td>
<td>0.07</td>
<td>-0.01</td>
<td>0.04</td>
<td>0.01</td>
<td>0.13</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.13</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Determinants of Forecast Frequency

See Table 1 for variable descriptions. The dependent variable is $Freq$, the number of quarterly earnings forecasts an analyst issued for a firm in a given year. The estimated equation (and its variants) is:

$$ Freq_{i,t} = \alpha + \beta T + \gamma_1 Ability_{i,t} + \gamma_2 Ability_{i,t}^2 + \gamma_3 NUM\_COVER_{i,t} + \gamma_4 EXP_{i,t} + \gamma_5 STD_{i,t} + \gamma_6 %INST_{i,t} + \gamma_7 OTHER\_ANLST_{i,t} + \gamma_8 BROKER_{i,t} + \gamma_9 LOG(MV)_{i,t} + \gamma_9 MB_{i,t} + \epsilon_{i,t} $$

Firm and year-specific intercept estimates not reported. Number of observations is 65,201. T-statistics (in parentheses) are based on heteroscedasticity robust standard error estimates. Superscript a, b, and c indicate a two-tailed test of significance level of less than 1%, 5%, and 10%, respectively.

<table>
<thead>
<tr>
<th>Indep. Variable</th>
<th>Pred. Sign</th>
<th>(1) Ability($R^2$)</th>
<th>(2) Ability(DIR)</th>
<th>(3) AllStar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td>-</td>
<td>-8.30</td>
<td>-7.97</td>
<td>-2.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-40.17$^a$</td>
<td>-22.96$^a$</td>
<td>-21.44$^a$</td>
</tr>
<tr>
<td>EXP</td>
<td>-</td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.67$^a$</td>
<td>-2.22$^b$</td>
<td>0.01</td>
</tr>
<tr>
<td>STD</td>
<td>-</td>
<td>-25.00</td>
<td>-21.43</td>
<td>-26.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.21$^a$</td>
<td>-3.01$^a$</td>
<td>-3.33$^a$</td>
</tr>
<tr>
<td>%INST</td>
<td>+</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.80$^a$</td>
<td>6.13$^a$</td>
<td>5.59$^a$</td>
</tr>
<tr>
<td>OTHER_ANLST</td>
<td>+</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.52$^a$</td>
<td>6.70$^a$</td>
<td>6.57$^a$</td>
</tr>
<tr>
<td>BROKER</td>
<td>?</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24.51$^a$</td>
<td>24.55$^a$</td>
<td>29.88$^a$</td>
</tr>
<tr>
<td>NUM_COVER</td>
<td>-</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.12$^a$</td>
<td>7.85$^a$</td>
<td>8.01$^a$</td>
</tr>
<tr>
<td>LOG(MV)</td>
<td>+</td>
<td>1.50</td>
<td>1.70</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.95$^a$</td>
<td>20.28$^a$</td>
<td>18.19$^a$</td>
</tr>
<tr>
<td>MB</td>
<td>+</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.43$^c$</td>
<td>-1.72$^c$</td>
<td>-1.62</td>
</tr>
<tr>
<td>Within $R^2$</td>
<td></td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Adjust-$R^2$</td>
<td></td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Table 4: Determinants of Forecast Boldness

See Table 1 for variable descriptions. The dependent variable is *Boldness*, the average forecast distance for an analyst-firm, relative to the average distance of all analyst-firms in the 2-digit SIC industry in a given year. The estimated equation (and its variants) is:

\[
\begin{align*}
\text{Boldness}_{i,j,T} &= \alpha_j + \beta_T + \gamma_1 \text{Ability}_{i,j} + \gamma_4 \text{NUM \_COVER}_{i,T} + \gamma_5 \text{EXP}_{j,T} + \gamma_4 \text{STD}_{i,j,T} \\
&+ \gamma_5 \%\text{INST}_{j,T} + \gamma_7 \text{OTHER \_ANLST}_{j,i} + \gamma_7 \text{BROKER}_{i,j,T} + \gamma_9 \text{LOG(MV)}_{j,T} + \gamma_9 \text{MB}_{j,T} + \epsilon_{i,j,T}
\end{align*}
\]

Firm and year-specific intercept estimates not reported. Number of observations is 65,201. T-statistics (in parentheses) are based on heteroscedasticity robust standard error estimates. Superscript a, b, and c indicate a two-tailed test of significance level of less than 1%, 5%, and 10%, respectively.

<table>
<thead>
<tr>
<th>Indep. Variable</th>
<th>Pred. Sign</th>
<th>(1) Ability ((R^2))</th>
<th>(2) Ability ((Dir))</th>
<th>(3) AllStar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td>+</td>
<td>27.34</td>
<td>52.12</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.75(^a)</td>
<td>13.31(^a)</td>
<td>1.22</td>
</tr>
<tr>
<td>EXP</td>
<td>+</td>
<td>0.99</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.98(^a)</td>
<td>6.84(^a)</td>
<td>6.62(^a)</td>
</tr>
<tr>
<td>STD</td>
<td>+</td>
<td>7.75</td>
<td>7.72</td>
<td>7.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.34(^a)</td>
<td>10.37(^a)</td>
<td>10.33(^a)</td>
</tr>
<tr>
<td>%INST</td>
<td>-</td>
<td>0.16</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.22(^a)</td>
<td>4.10(^a)</td>
<td>4.16(^a)</td>
</tr>
<tr>
<td>OTHER_ANLST</td>
<td>-</td>
<td>-1.09</td>
<td>-1.11</td>
<td>-1.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-9.10(^a)</td>
<td>-9.27(^a)</td>
<td>-9.12(^a)</td>
</tr>
<tr>
<td>BROKER</td>
<td>?</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.74(^a)</td>
<td>5.60(^a)</td>
<td>4.89(^a)</td>
</tr>
<tr>
<td>NUM_COVER</td>
<td>?</td>
<td>0.07</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.44</td>
<td>1.60</td>
<td>1.44</td>
</tr>
<tr>
<td>LOG(MV)</td>
<td>?</td>
<td>-7.37</td>
<td>-8.45</td>
<td>-7.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.91(^a)</td>
<td>-5.55(^a)</td>
<td>-5.04(^a)</td>
</tr>
<tr>
<td>MB</td>
<td>?</td>
<td>-1.18</td>
<td>-1.16</td>
<td>-1.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-5.43(^a)</td>
<td>-5.42(^a)</td>
<td>-5.41(^a)</td>
</tr>
<tr>
<td>Within (R^2)</td>
<td></td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>Adjust-(R^2)</td>
<td></td>
<td>0.47</td>
<td>0.47</td>
<td>0.47</td>
</tr>
</tbody>
</table>
Table 5: Sensitivity Analysis

See Table 1 for variable descriptions. The following equation is estimated each year from 1986-2000:

\[
Freq_{i,j,t} = \alpha_j + \beta_j + \gamma_{j}Ability_{i,j} + \gamma_4NUM _{COVER}i_{j,t} + \gamma_5EXP_{i,j,t} + \gamma_6STD_{i,j,t} \\
+ \gamma_5%INST_{j,t} + \gamma_6OTHER _{ANLST}i_{j,t} + \gamma_7BROKER_{i,j,t} + \gamma_8LOG(MV)_{i,j,t} + \gamma_9MB_{i,j,t} + \epsilon_{i,j,t}
\]

\[
Boldness_{i,j,t} = \alpha_j + \beta_j + \gamma_{j}Ability_{i,j} + \gamma_4NUM _{COVER}i_{j,t} + \gamma_5EXP_{i,j,t} + \gamma_6STD_{i,j,t} \\
+ \gamma_5%INST_{j,t} + \gamma_6OTHER _{ANLST}i_{j,t} + \gamma_7BROKER_{i,j,t} + \gamma_8LOG(MV)_{i,j,t} + \gamma_9MB_{i,j,t} + \epsilon_{i,j,t}
\]

The Z-score is calculated as \(\frac{\sqrt{N-1}\bar{t}}{\text{stddev}(t)}\), where \(N=16\) is the number of yearly regressions, \(\bar{t}\) is the mean of t-statistics and \(\text{stddev}(t)\) is the standard deviation of the t-statistics. Superscript a, b, and c indicate a two-tailed test of significance level of less than 1%, 5%, and 10%, respectively.

<table>
<thead>
<tr>
<th>Indep. Variable</th>
<th>Pred. Sign</th>
<th>Pred.</th>
<th>Sign</th>
<th>Ability(R^2)</th>
<th>Ability(Dir)</th>
<th>AllStar</th>
<th>Pred. Sign</th>
<th>Ability(R^2)</th>
<th>Ability(Dir)</th>
<th>AllStar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability</td>
<td></td>
<td>(1)</td>
<td></td>
<td>-7.55</td>
<td>-3.99</td>
<td>-1.19</td>
<td>+</td>
<td>9.23</td>
<td>37.07</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2)</td>
<td></td>
<td>-10.52a</td>
<td>-5.46a</td>
<td>-2.58b</td>
<td></td>
<td>2.70a</td>
<td>8.04a</td>
<td>0.69</td>
</tr>
<tr>
<td>EXP</td>
<td></td>
<td>(3)</td>
<td></td>
<td>-0.04</td>
<td>-0.04</td>
<td>0.01</td>
<td>+</td>
<td>1.27</td>
<td>1.23</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-2.41b</td>
<td>-1.79c</td>
<td>-1.62</td>
<td></td>
<td>3.22a</td>
<td>3.13a</td>
<td>2.71a</td>
</tr>
<tr>
<td>STD</td>
<td></td>
<td></td>
<td></td>
<td>-0.94</td>
<td>2.16</td>
<td>-0.54</td>
<td>+</td>
<td>12.17</td>
<td>12.11</td>
<td>12.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.34</td>
<td>-1.30</td>
<td>-1.41</td>
<td></td>
<td>8.61a</td>
<td>8.66a</td>
<td>8.02a</td>
</tr>
<tr>
<td>%INST</td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>-</td>
<td>-0.10</td>
<td>-0.10</td>
<td>-0.10</td>
</tr>
<tr>
<td>OTHER_ANLST</td>
<td></td>
<td></td>
<td></td>
<td>4.83a</td>
<td>4.79a</td>
<td>5.76a</td>
<td>-</td>
<td>-2.90a</td>
<td>-2.93a</td>
<td>-2.84a</td>
</tr>
<tr>
<td>BROKER</td>
<td>?</td>
<td></td>
<td></td>
<td>0.19</td>
<td>0.22</td>
<td>0.22</td>
<td>-</td>
<td>-1.70</td>
<td>-1.71</td>
<td>-1.62</td>
</tr>
<tr>
<td>NUM_COVER</td>
<td>-</td>
<td></td>
<td></td>
<td>0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>?</td>
<td>0.14</td>
<td>0.13</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.82a</td>
<td>2.79a</td>
<td>4.68a</td>
<td></td>
<td>4.54a</td>
<td>4.20a</td>
<td>3.32a</td>
</tr>
<tr>
<td>LOG(MV)</td>
<td>+</td>
<td></td>
<td></td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.03</td>
<td>?</td>
<td>0.20</td>
<td>0.21</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1.51</td>
<td>-1.79c</td>
<td>-2.13b</td>
<td></td>
<td>2.88a</td>
<td>2.99a</td>
<td>3.13a</td>
</tr>
<tr>
<td>MB</td>
<td>+</td>
<td></td>
<td></td>
<td>-0.37</td>
<td>-0.39</td>
<td>-0.36</td>
<td>?</td>
<td>-5.37</td>
<td>-5.26</td>
<td>-5.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-4.60a</td>
<td>-4.75a</td>
<td>-3.96a</td>
<td></td>
<td>-5.58a</td>
<td>-5.48a</td>
<td>-5.48a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td>0.12</td>
<td>0.13</td>
<td>?</td>
<td>-0.65</td>
<td>-0.61</td>
<td>-0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.43a</td>
<td>5.45a</td>
<td>6.12a</td>
<td></td>
<td>-1.77c</td>
<td>-1.70c</td>
<td>-1.81c</td>
</tr>
</tbody>
</table>
A Appendix

Proof of Lemma 1. The following results are useful throughout the proof.

\[
\begin{pmatrix}
z \\
x \\
y
\end{pmatrix} \sim N\left[
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
\frac{1}{p_x} & \frac{1}{p_x} & \frac{1}{p_x} \\
\frac{1}{p_x} & \frac{1}{p_x} + \frac{1}{p_z} & \frac{1}{p_x} \\
\frac{1}{p_x} & \frac{1}{p_x} & \frac{1}{p_x} + \frac{1}{p_y}
\end{pmatrix}\right]
\] (9)

\[
\rho_{x,z} = \frac{p_x}{p_x + p_z}, \quad \rho_{z,y} = \frac{p_y}{p_y + p_z}, \quad \rho_{x,y} = \rho_{x,z} \rho_{z,y}
\]

Rewrite the expression for \(MSE\) (equation (4)) as

\[
MSE = \text{var}(\varepsilon_y)[1 + \rho_{\varepsilon_y \varepsilon_y}^2 H(\alpha)]
= \text{var}(\varepsilon_y) + \frac{\text{var}(\varepsilon_y)^2}{\text{var}(\varepsilon_y) + \text{var}(\varepsilon_y)} \alpha H(\alpha).
\] (10)

Using the fact

\[
H'(\alpha) = H^2 - \alpha H,
\] (11)

we have

\[\frac{\partial [\alpha H(\alpha)]}{\partial \alpha} = \alpha H' + H = (1 - \alpha^2) H + \alpha H^2.\] (12)

Further, \(\alpha H(\alpha)\) is an increasing and convex function of \(\alpha\), a fact we will use later.

Because \(\frac{\partial \alpha}{\partial b} = \frac{1}{\sqrt{\text{var}(\varepsilon_y) + \text{var}(\varepsilon_y)}}\), \(\frac{\partial MSE}{\partial \alpha} = \frac{\partial MSE}{\partial \alpha} \frac{\partial \alpha}{\partial b}\), substituting equation (12) we have

\[
\frac{\partial MSE}{\partial b} = \frac{\text{var}(\varepsilon_y)^2}{[\text{var}(\varepsilon_y) + \text{var}(\varepsilon_y)]^{3/2}} \left[H + \alpha H^2 - \alpha^2 H\right] > 0.
\]

Next we prove that \(\frac{\partial MSE}{\partial p_y} < 0\) and \(\frac{\partial^2 MSE}{\partial \alpha \partial p_y} < 0\). Without loss of generality, we normalize \(\text{var}(z) = 1\). We can fix \(b\) and let \(\alpha\) vary one-to-one with \(\text{var}(\varepsilon_y)\) or \(p_y\), i.e.,

\[
\alpha = \frac{b}{[\text{var}(\varepsilon_y) + 1]^{1/2}} \in [0, b],
\]

and we have

\[
\text{var}(\varepsilon_y) = \frac{b^2}{\alpha^2} - 1.
\]

For expositional ease, let \(w \equiv \text{var}(\varepsilon_y)\). Then \(MSE = w + \frac{w^2}{w+1} \alpha H(\alpha) = w + \frac{w^2}{(w+1)^{3/2}} b H(\alpha)\). To
The last step follows from $H' = H(H - \alpha)$. Gordon (1941) shows that

$$H \leq \frac{1 + \alpha^2}{\alpha} = \alpha + 1 < \alpha + \frac{4}{\alpha w} = \alpha + \frac{w + 4}{\alpha w}.$$ 

Therefore $[w + 4 - \alpha w(H - \alpha)] > 0$, so $\frac{\partial \text{MSE}}{\partial \alpha} = \frac{\partial \text{MSE}}{\partial \text{var}(\epsilon_y)} > 0$.

To prove $\frac{\partial^2 \text{MSE}}{\partial \alpha \partial \alpha} < 0$, let

$$g = \frac{\partial \text{MSE}}{\partial b} = \frac{\left(\frac{b^2}{\alpha^2} - 1\right)^2}{\frac{b^2}{\alpha^3}} [H + \alpha H^2 - \alpha^2 H]$$

$$= \left(\frac{b}{\alpha} - \frac{2\alpha}{b} + \frac{\alpha^3}{b^3}\right) [H + \alpha H^2 - \alpha^2 H]. \quad (14)$$

For any given $b, \alpha$ is increasing in $p_y$. $\frac{\partial g}{\partial \alpha} < 0$ would imply $\frac{\partial^2 \text{MSE}}{\partial \alpha \partial \alpha} < 0$. Expanding $\frac{\partial g}{\partial \alpha}$, we have

$$\frac{\partial g}{\partial \alpha} = \left(\frac{3\alpha^2 - \frac{b}{\alpha^2} - \frac{2}{b}}{\alpha^2} \right) [H + \alpha H^2 - \alpha^2 H]$$

$$+ \left(\frac{b}{\alpha} - \frac{2\alpha}{b} + \frac{\alpha^3}{b^3}\right) [H + \alpha H^2 - \alpha^2 H]' \quad (15)$$

$$= \frac{b}{\alpha} \left(\frac{\alpha^2}{b^2} - 1\right) \left(\frac{3\alpha^2}{b^2} + \frac{1}{\alpha}\right) [H + \alpha H^2 - \alpha^2 H] + \left(\frac{\alpha^2}{b^2} - 1\right) [H + \alpha H^2 - \alpha^2 H]'$$

Because $\frac{\alpha^2}{b^2} - 1 < 0$, it remains to be shown that

$$\psi = \left(\frac{3\alpha^2}{b^2} + \frac{1}{\alpha}\right) [H + \alpha H^2 - \alpha^2 H] - \left(\frac{1 - \frac{\alpha^2}{b^2}}{\alpha}\right) [H + \alpha H^2 - \alpha^2 H]' > 0 \quad (16)$$

for all $b$ and all $\alpha \in [0, b]$.

Rewrite $\psi$ as

$$\psi = \left(\frac{3\alpha^2}{b^2} + 1\right) \frac{\alpha}{A} - \left(1 - \frac{\alpha^2}{b^2}\right) \frac{\alpha}{C} [H + \alpha H^2 - \alpha^2 H]' - \left(\frac{1 - \frac{\alpha^2}{b^2}}{\alpha}\right) [H + \alpha H^2 - \alpha^2 H]' \quad (16)$$

41
Note that all terms $A$, $B$, $C$, and $D$ are positive and $A > 1 > C$. In addition, $B$ and $D$ only involve functions of standard normal distributions and do not concern any parameters. Finally, we verify that $B > D$ for all values of $\alpha$.

Let $\varphi(\alpha) = (\alpha H)'$. Then $B > D$ is equivalent to

$$\frac{\varphi(\alpha)}{\alpha} > \varphi'(\alpha).$$

(17)

Note that (17) is nothing but a comparison of the average and the marginal slope of the $\varphi(\alpha)$ function. By L’Hopital’s rule,

$$\lim_{\alpha \to \infty} \frac{\varphi(\alpha)}{\alpha} = \varphi'(\alpha).$$

Because $(\alpha H)' > 0$, $(\alpha H)'' > 0$ and $(\alpha H)''' > 0$, $\varphi(\alpha)$ is positive, increasing and convex on $[0, \infty)$, and has a positive intercept $\varphi(0)$. If we can show that $\lim_{\alpha \to \infty} \varphi'(\alpha)$ is finite, then for such a function the average slope must lie all the way above the marginal slope. It turns out to be true because

$$\varphi'(\alpha) = H' + (\alpha H)' \leq H' + \sup(H')$$

and $H'$ is bounded (because $\lim_{\alpha \to \infty} H'' = 0$). This proves (17), and hence (16) holds. ■

**Proof of Lemma 2.** For notational ease, we rewrite $\Delta(p, b) = \frac{1}{p(p_x + p_z)}L(p, b)$ where $L(p, b) = 1 - \Phi(\alpha) + \rho_{y-z_1,y}^2 \alpha \phi(\alpha) > 0$, then

$$\frac{\partial \Delta(p, b)}{\partial b} = \frac{1}{p(p_x + p_z)} \frac{\partial L(p, b)}{\partial \alpha} \frac{\partial \alpha}{\partial b}$$

$$= \frac{1}{p(p_x + p_z)} \left[ -\phi(\alpha) + \rho^2 \left[ \phi(\alpha) - \alpha^2 \phi(\alpha) \right] \right] \frac{\partial \alpha}{\partial b}$$

$$= \frac{-[\rho^2 \alpha^2 + (1 - \rho^2)]}{p(p_x + p_z) \sqrt{\text{var}(y)}} \frac{\phi(\alpha)}{\alpha} \frac{1}{\text{var}(y)} < 0,$$

(18)

where $\rho^2 = \rho_{y-z_1,y}^2 = \frac{1}{p} (1 - \rho_{x,z}^2) \rho_{z,y}^2 = \frac{p_x}{p_x + p_y} \frac{p_y}{p_y + p_z}$.

To show $\frac{\partial \Delta(p, b)}{\partial p} < 0$, we first obtain

$$\frac{\partial L}{\partial p} = -\phi(\alpha) \frac{\partial \alpha}{\partial p} + \rho^2 \frac{\partial \alpha \phi(\alpha)}{\partial p} \frac{\partial \alpha}{\partial \alpha} \frac{\partial \alpha}{\partial p}$$

$$= \rho^2 \frac{\partial \alpha \phi(\alpha)}{\partial p} \frac{\partial \alpha}{\partial \alpha} \frac{\partial \alpha}{\partial p}$$

$$< 0.$$
The first term is negative because $\frac{\partial^2}{\partial p^2} < 0$ as shown below.

$$\frac{\partial^2}{\partial p^2} = (1 - \rho^2_x)\left[\frac{1}{p_y} (1 - \rho^2_{y,z}) - \frac{p(1 - p)}{p_y} \frac{\rho^2_{y,z}}{p^2}\right]$$

$$= \frac{\rho^2}{p_y} (p - \rho^2_{y,z}) = \frac{\rho^2}{p_y} \left(\frac{p_y}{p_x + p_z + p_y} - \frac{p_y}{p_z + p_y}\right) < 0.$$  \hfill (19)

Hence

$$\frac{\partial}{\partial p} \left[\Delta(p, b) \right] \propto \frac{-1}{p^2} \frac{\partial}{\partial p} - \frac{1}{p} \frac{\partial L}{\partial p} \frac{\partial}{\partial p} < 0$$

We now derive $\frac{\partial^2\Delta(p, b)}{\partial b \partial p}$ based on equation (18). But first let us simplify the notations a little bit. Specifically, let

$$R = 1 - \rho^2 + \rho^2 \alpha^2 > 0,$$

and $D = \frac{\phi(\alpha)}{\sqrt{\text{var}(y)}}$. Then

$$\frac{\partial\Delta(p, b)}{\partial b} = \frac{-1}{(p_x + p_z)} D \frac{R}{p},$$

$$\frac{\partial^2\Delta(p, b)}{\partial b \partial p} \propto -\frac{\partial [D \frac{R}{p}]}{\partial p_y} = \left[\frac{\partial D \frac{R}{p}}{\partial p_y} + \frac{\partial (\frac{R}{p})}{\partial p_y} D\right].$$

Further expanding each of the terms in the bracket, we have

$$\frac{\partial D}{\partial p_y} = \phi(\alpha) \frac{1}{\sqrt{\text{var}(y)}} - \frac{1}{\sqrt{\text{var}(y)}} \frac{\partial}{\partial p_y} \phi(\alpha) \frac{\partial \alpha}{\partial p_y}$$

$$= \frac{\phi(\alpha)}{\sqrt{\text{var}(y)}} \left(1 - \rho^2_y\right) \frac{1}{2p_y} - \frac{\alpha \phi(\alpha)}{\sqrt{\text{var}(y)}} \left(1 - \rho^2_y\right) \frac{1}{2p_y}$$

$$= \frac{\phi(\alpha)}{\sqrt{\text{var}(y)}} \left(1 - \rho^2_{y,z}\right) \frac{1}{2p_y} (1 - \alpha^2) = \frac{1}{2p_y} (1 - \alpha^2),$$

and

$$\frac{\partial (\frac{R}{p})}{\partial p_y} = \frac{1}{p} \frac{\partial R}{\partial p_y} - \frac{1}{p^2} \frac{\partial p}{\partial p_y} R = \frac{1}{pp_y} \left(\frac{\partial R}{\partial p_y} p_y - R(1 - p)\right)$$

where

$$\frac{\partial R}{\partial p_y} = \frac{p^2}{p_y} \left(\alpha^2 - 1\right) \frac{\partial p^2}{\partial p_y} + \rho^2 \alpha \frac{\partial \alpha}{\partial p_y}$$

$$= \frac{p^2}{p_y} [\alpha^2 (\alpha - 1)(p - \rho^2_{y,z}) + \alpha^2 (1 - \rho^2_{y,z})].$$
Substituting in the expressions for \( \frac{\partial D}{\partial p_y} \) and \( \frac{\partial R}{\partial p_y} \), and taking out the common factor \( D \), we have

\[
\frac{\partial D R}{\partial p_y} + \frac{\partial R}{\partial p_y} D = D \frac{(1 - \rho_{yz}^2)}{2p_y} (1 - \alpha^2) R \frac{\rho_{p}}{p} + \frac{\partial R}{\partial p_y} \]

\[
= \frac{D}{2p_p} [(1 - \rho_{yz}^2)(1 - \alpha^2) R + 2(\frac{\partial R}{\partial p_y} p_y - R(1 - p))].
\]

(24)

Since \( \frac{D}{2p_p} > 0 \), \( \text{sign}(\frac{\partial^2 \Delta(p,b)}{\partial \alpha \partial p_y}) = -\text{sign}(F) \), we need to prove that there exists a \( p^0 \) such that \( F < 0 \) for all \( p > p^0 \) and all \( \alpha \geq 0 \). Since \( F \) is quadratic in \( \alpha^2 \), a sufficient condition for this to be true is: \( F|_{\alpha=0} \leq 0 \) and \( \frac{dF}{d\alpha}|_{\alpha=0} < 0 \). When \( \alpha = 0 \),

\[
F|_{\alpha=0} = (1 - \rho^2)(1 - \rho_{yz}^2) - 2 \frac{\rho^2}{1 - \rho^2} (p - \rho_{yz}^2) - 2(1 - p).
\]

Therefore \( F|_{\alpha=0} \leq 0 \) if and only if the term in the bracket is non-positive. With tedious but straightforward algebra and the fact that \( \frac{1 - \rho^2}{1 - \rho^2} (p - \rho_{yz}^2) = 1 - \rho_{yz}^2 \), we further simplify this term as below

\[
F|_{\alpha=0} \leq 0 \iff (1 - \rho_{yz}^2) + 2(p - \rho_{yz}^2) \leq 0
\]

\[
\iff p_y \geq \frac{p_x + p_z}{2p_x - p_z} \equiv p_1.
\]

(25)

For notational ease, let \( q = \alpha^2 - 1 \). To establish the condition for \( \frac{dF}{d\alpha}|_{\alpha=0} < 0 \), we can rewrite \( R = 1 + \rho^2 q \) and

\[
F = -(1 - \rho_{yz}^2)q + 2(1 - p)(\rho^2 q + 1) + 2\rho^2[(p - \rho_{yz}^2)q + (1 - \rho_{yz}^2)(1 + q)]
\]

\[
= -(1 - \rho_{yz}^2)q^2 - [(1 - \rho_{yz}^2) + 4\rho^2(\rho_{yz}^2 - p)]q + 2\rho(1 - \rho_{yz}^2)^2 - 2(1 - p)
\]

\[
\Rightarrow \frac{F}{(1 - \rho_{yz}^2)} = -q^2 - (1 + 4\rho_{xy}^2)q + 2\rho^2 - 2 \frac{1 - p}{1 - \rho_{yz}^2}.
\]

Note that the requirement \( \frac{dF}{d\alpha}|_{\alpha=0} \leq 0 \) is equivalent to \( \frac{dF}{dq}|_{q=-1} \leq 0 \),

\[
\frac{dF}{dq}|_{q=-1} = -(2q + 1 + 4\rho_{xy}^2)|_{q=-1} \leq 0
\]

\[
\iff \rho_{xy}^2 \geq \frac{1}{4} \iff p_y \geq \frac{p_x + p_z}{3p_x - p_z} \equiv p_2.
\]

The proof is completed by noticing \( p_1 > p_2 \) and by designating \( p^0 = p_1 \). ■

**Proof of Proposition 1.** Let \( \hat{p}(b) \) be the inferred expected precision of the analyst when he plays \( b \).
Suppose Proposition 1 does not hold. Then there exists \( p' > p \) but \( b' < b \). The analyst maximizes

\[
U(p, b) = -\lambda \text{MSE}(p, b) + \bar{p}^2 \Delta(p, b)
\]

In equilibrium, the market has the correct inference that \( \bar{p}(b') > \bar{p}(b) \). Given the partial derivative results established in Lemmas 1 and 2, when \( p > p^0 \), by revealed preference, we have

\[
U(p, b) + U(p', b') - U(p' , b' ) - U(p, b') > 0.
\] (26)

We will show that this is a contradiction.

\[
U(p, b) + U(p', b') - U(p' , b' ) - U(p, b') = \bar{p}(b)^2[(\Delta(p, b) - \Delta(p', b)) + \bar{p}(b')^2[(\Delta(p', b') - \Delta(p, b))] \\
- \lambda[MSE(p, b) - MSE(p', b)] - \lambda[MSE(p', b') - MSE(p, b')].
\]

However, the cross-partial signs imply that

\[
|\Delta(p, b) - \Delta(p', b)| < |\Delta(p', b') - \Delta(p, b)|, \text{ and}
\]

\[
|MSE(p, b) - MSE(p', b)| > |MSE(p', b') - MSE(p, b')|.
\]

Therefore, if \( \bar{p}(b') > \bar{p}(b) \), (26) can not be positive, a contradiction. ■

**Proof of Corollary 1.** Consider a simple example where there are only two possible precision levels for an analyst: \( p \) and \( \bar{p} \) with \( p < \bar{p} \). Because maintaining a non-zero threshold reduces both the market impact and the measured accuracy, the \( p \)-type, if unable to mimic the \( \bar{p} \)-type in equilibrium, will choose \( b = 0 \). The \( \bar{p} \)-type, on the other hand, needs to maintain \( \bar{b} \) high enough so that the \( p \)-type is just indifferent between mimicking and not mimicking. Accordingly, \( \bar{b} \) is the solution to the following equation:

\[
-\lambda E[(y - z)^2 | y \geq \bar{b}, p] + \bar{p}^2 \Pr(y \geq \bar{b}|p) E[(y - \bar{z}_1)^2 | y \geq \bar{b}, p] \\
= -\lambda E[(y - z)^2 | p] + \bar{p}^2 E((y - \bar{z}_1)^2 | p).
\] (27)

Equation (27) is the incentive compatible constraint for the lower type and can be expanded to continuous types between \( p \) and \( \bar{p} \). Specifically, let \( b = b(p) \) and \( b' = b + \Delta b \). Then the following
condition must hold when $\Delta b \to 0$:

$$-\lambda MSE(p, b') + \hat{p}(b')^2 \Delta(p, b') = -\lambda MSE(p, b) + \hat{p}(b)^2 \Delta(p, b).$$

Taking the derivative with respect to $b$ and rearranging terms, we have

$$\frac{d\hat{p}}{db} = \frac{\lambda MSE'(b) - \hat{p}(b)^2 \Delta'(b)}{2p(b) \Delta(b)},$$

where $\Delta'(b) = \frac{\partial \Delta}{\partial b}$ and $MSE'(b) = \frac{\partial MSE}{\partial b}$. The first-order differential equation in Corollary 1 is obtained by substituting the equilibrium condition $\hat{p}(b) = b$ in equation (28). $MSE'(b) > 0$ and $\Delta'(b) < 0$ (as shown in the appendix) imply that $\frac{db}{dp} > 0$, and continuity of analysts' type distribution implies $|b| > 0$ with probability one. Finally, by Mailath (1987), the initial condition $b(p) = 0$ and the single crossing property of the objective function (i.e., $\frac{\partial^2 U}{\partial b \partial p} > 0$) implies uniqueness.

**Proof of Corollary 2.** $\frac{\partial b^*}{\partial p_x} < 0$, $\frac{\partial b^*}{\partial p_x} < 0$, and $\frac{\partial b^*}{\partial \lambda} < 0$.

The first-order condition for the optimal $b^*$ is given by:

$$FOC = 2\hat{p} \frac{\partial \Delta}{\partial b} + \hat{p}^2 \frac{\partial \Delta}{\partial p} - \lambda \frac{\partial MSE}{\partial b}.$$

Accordingly $\frac{db^*}{dp_x} = \frac{\partial FOC}{\partial p_x}$ and is of the same sign as $\frac{\partial FOC}{\partial p_x}$ which is equal to

$$\frac{\partial FOC}{\partial p_x} = 2\hat{p} \frac{\partial \Delta}{\partial p_x} + \hat{p}^2 \frac{\partial^2 \Delta}{\partial \Delta \partial p_x}.$$

We have shown in Proposition 1 that in equilibrium $\frac{\partial^2 \Delta}{\partial b \partial p_x} > 0$, where $p = \frac{p_y}{p_x + p_y + p_z}$. Note that

$$\frac{\partial^2 \Delta}{\partial b \partial p_x} = \frac{\partial}{\partial b} \left( \frac{\partial \Delta}{\partial p_x} \right) = \frac{\partial}{\partial b} \left( \frac{\partial \Delta}{\partial p} \frac{\partial p}{\partial p_x} \right)$$

$$= \frac{\partial^2 \Delta}{\partial p \partial b} \frac{\partial p}{\partial p_x} + \frac{\partial \Delta}{\partial p} \frac{\partial^2 p}{\partial p_x \partial b} < 0.$$

Therefore $\frac{db^*}{dp_x} < 0$.

Similarly, $\frac{db^*}{dp_x}$ is of the same sign as

$$\frac{\partial FOC}{\partial p_x} = 2\hat{p} \frac{\partial \Delta}{\partial p_x} + \hat{p}^2 \frac{\partial^2 \Delta}{\partial \Delta \partial p_x} - \lambda \frac{\partial^2 MSE}{\partial b \partial p_x}.$$
Using the same method as in (29), we can shown that $\frac{\partial^2 \Delta}{\partial b \partial p} < 0$. Similarly, using the fact that $\frac{\partial^2 \text{MSE}}{\partial b \partial p} < 0$, we have

$$\frac{\partial^2 \text{MSE}}{\partial b \partial p_z} = \frac{\partial \left( \frac{\partial \text{MSE}}{\partial p_z} \right)}{\partial b} = \frac{\partial \left( \frac{\partial \text{MSE}}{\partial p} \frac{\partial p}{\partial p_z} \right)}{\partial b} = \frac{\partial^2 \text{MSE}}{\partial p \partial b} \frac{\partial p}{\partial p_z} + \frac{\partial \text{MSE}}{\partial p} \frac{\partial^2 p}{\partial p_z \partial b} = 0 > 0.$$ 

Therefore, $\frac{\partial \text{FOC}}{\partial p_z} < 0$, which implies $\frac{\partial^*}{\partial p_z} < 0$. ■