Financial Analysts and the False Consensus Effect*

Jared Williams†

March 20, 2013

Abstract

Social psychologists have documented a tendency for people to overestimate their similarity to others. I investigate whether financial analysts' forecast errors are consistent with this bias. I model the bias by assuming analysts overestimate the correlation of the private signals they receive about a firm's future earnings. My model predicts the likelihood of an analyst's revised forecast being too close to his earlier forecast is increasing in the number of other analysts issuing forecasts between the time of the analyst's initial forecast and the time of his revised forecast. I empirically confirm this prediction and consider several alternative explanations.

Keywords: **Information Aggregation, Analyst Underreaction, Forecast Errors, False Consensus Effect**

JEL Classification: D03, D8, G2

*Accepted by Haresh Sapra. I thank Laura Field, Dave Haushalter, Michelle Lowry, Annette Vissing-Jørgensen, and an anonymous referee for their helpful comments and suggestions. I have also benefited from conversations and feedback from Snehal Banerjee, Nick Barberis, Orie Barron, Sudipta Basu, Larry Brown, Fritz Burkhardt, Joey Engelberg, Grant Farnsworth, Paul Fischer, Mike Fishman, Paul Gao, Dan Givoly, Zhiguo He, Andrew Hertzberg, Steve Huddart, Peter Iliev, Jiro Kondo, Cami Kuhnen, Henock Louis, Mitchell Petersen, Jeff Pontiff, Paola Sapienza, Phil Shane, Joel Vanden, Raisa Velthuis, Svetla Vitanova, Beverly Walther, Adam Welker, and seminar participants at the 2011 Penn State Accounting Conference. All errors are mine.

†Department of Finance, Smeal College of Business, Penn State University. 344 Business Building, University Park, PA. E-mail: jmw52@psu.edu. Phone: (814) 865-2185. Fax: (814) 865-3362.

1 Introduction

Do financial analysts efficiently incorporate public information about other analysts' beliefs when issuing their forecasts? How do they update their beliefs based on other analysts' research reports? In addition to analyzing information from primary information sources such as SEC filings, communication with management, and macroeconomic conditions, financial analysts can also glean information about firms' future earnings by reading other analysts' research reports. Combining the information that analysts gather themselves (their own "private signal") with other analysts' research reports (other analysts' "private signals") is not a straightforward task because analysts cannot know with certainty what information the other analysts rely on when issuing their research reports. They face uncertainty over the quality of others' information as well as over the similarity of other analysts' information to their own information. In this paper, I focus on how analysts update their beliefs based on information about other analysts' beliefs. In particular, I focus on the problem that analysts do not know how *similar* their information is to others'.

More formally, consider a standard differential information model. Each analyst receives a private signal, which is equal to the firm's earnings plus an error term. There are two key parameters that should affect how an analyst updates his beliefs about the firm's future earnings after observing his signal and others' signals: the *precisions* of the analysts' signals and the *correlation* of the analysts' signal errors.

Drawing from the evidence in cognitive psychology that people tend to be overconfident, behavioral financial economists have modelled overconfidence (in the context of traders and financial markets) by assuming that people overestimate the precisions of their signals. See, e.g., Daniel, Hirshleifer, and Subrahmanyam (1998) and Odean (1998). Generally, for overconfidence to generate significant observable biases, there must be a small number of signals observed by the agents. Otherwise, if an overconfident agent observes sufficiently many signals from others, he recognizes that the aggregate informational content of others' signals is sufficiently high that he should essentially ignore his own private signal, even though he overestimates its precision. Unlike biases in perceived signal precisions, biases in perceived signal error correlations can cause significant biases even when many signals are observed by agents. Indeed, biases in signal error correlations cannot affect agents' behavior until they combine information from others with their own information. Because of this fundamental difference, I focus on analysts' beliefs about how correlated their signal errors are with others' signal errors.

Research in social psychology suggests that people overestimate their similarity to others. For example, if people are asked to estimate the percent of the population that prefers wheat bread to white bread, the average estimate given by people who themselves prefer wheat bread is 52.5%, whereas the average estimate given by people who themselves prefer white bread is only 37.4% (Ross, Greene, and House (1977)). This bias is referred to as the "false consensus effect" (henceforth "FCE"). I provide a brief overview of the literature on the false consensus effect in Section 2. The purpose of this paper is to determine whether and how a false consensus affects financial analysts' earnings forecasts in the days following firms' earnings announcements.

In Section 3, I formulate a model of financial analysts in which the analysts have a false consensus. In the model, when a firm publicly announces its quarter t earnings, all analysts covering the firm simultaneously receive signals about the firm's quarter $t+1$ earnings. The analysts correctly perceive their signals' precisions, but they overestimate the correlation of their signal errors. This assumption, that analysts overestimate the correlation of their signal errors, formalizes the false consensus effect: agents overestimate their similarity to others, and hence, the likelihood that other analysts are focusing on the same information within the firm's quarterly earnings announcement as they are. Consider the following extreme example. Suppose a firm just announced its quarter t earnings, and there are two analysts, A and B, forecasting the firm's quarter $t+1$ earnings. Analyst A focuses on the firm's quarter t revenues when analyzing the announcement, while B focuses on the firm's quarter t costs, perhaps because A (B) is skilled at determining which revenue (cost) components are persistent but unskilled at determining which cost (revenue) components are persistent. A and B both update their beliefs about the firm's quarter $t+1$ earnings based on the information they analyze. If A and B overestimate their similarity to each other, A will overestimate the likelihood that B focuses on revenue information when B updates his beliefs, and B will overestimate the likelihood that A focuses on cost information when A updates his beliefs. In this example, it is likely that A and B would overestimate the correlation of their signal errors, because they underestimate the likelihood that they are focusing on different parts of the earnings announcement.

Since the informativeness of a collection of signals is decreasing in the correlation of the signal errors, an analyst who overestimates the correlation of signal errors underestimates the precision of *collections* of signals observed by groups of analysts. Hence, he places too little weight on the aggregate information content of his signal and other analysts' signals, and his posterior mean is too close to his prior mean. I show that if financial analysts overestimate the correlation of their private signal errors, the likelihood of an analyst's revised forecast being too close to his earlier forecast is increasing in the number of other analysts issuing forecasts between the time of the analyst's initial forecast and the time of his revised forecast. (See Proposition 1.) To clarify the proposition's prediction, consider the following example. Analysts A and B are each forecasting earnings for two firms, Y and Z. Suppose analyst A forecasts firm Y's earnings to be \$2.00 per share, and analyst B forecasts firm Z's earnings to be \$5.00. Analysts C, D, and E then issue forecasts for firm Y's earnings, and analyst F issues a forecast for firm Z's earnings. After observing C's, D's, and E's forecasts, A upgrades his forecast for Y's earnings to \$2.10 per share. After observing F's forecast, B upgrades his forecast for Z's earnings to \$5.10 per share. The testable prediction of my model is that the likelihood of Y's true earnings being greater than \$2.10 per share is greater than the likelihood of Z's true earnings being greater than \$5.10 per share. To see this, note that three analysts issued forecasts between A's initial forecast and his revised forecast, whereas only one analyst issued a forecast between B's initial forecast and his revised forecast. Moreover, ex post, A's revised forecast of \$2.10 is too close to his earlier forecast if and only if Y's true earnings are greater than \$2.10 per share, and B's revised forecast of \$5.10 is too close to his earlier forecast of

\$5.00 per share if and only if Z's true earnings are greater than \$5.10.

To understand the intuition behind this prediction, note that overestimation of signal error correlations only causes an analyst's posterior beliefs to be biased when he updates his beliefs based on a collection of signals, not when he updates his beliefs based only on his own signal. More specifically, because the informativeness of a collection of signals is decreasing in the correlation of the signal errors, if analysts overestimate the correlation of their signal errors, they will underestimate the precision of collections of signals. If no analysts issue forecasts between the times of his earlier forecast and his revised forecast, the bias will not affect the analyst's beliefs about the firm's future earnings. For these revisions, it should be equally likely that the true earnings are greater than the revised forecast as it is that the true earnings are less than the revised forecast. If, on the other hand, many analysts issue forecasts between his earlier forecast and his revised forecast, the bias should be apparent, and the bias worsens as the analyst observes more analysts' forecasts.

To empirically test my proposition, I obtain earnings forecasts issued by sell-side analysts. Because my theory concerns how analysts update their beliefs upon observing other analysts' signals (and not other news), I restrict my sample to forecast revisions that are likely to have been precipitated by other analysts' reports rather than the arrival of new information about the firm's earnings. Specifically, I restrict my sample to forecast revisions in which (i) the revised forecast is issued shortly after the analyst's earlier forecast, (ii) the firm does not have an earnings announcement between the earlier forecast and the revised forecast, and (iii) the stock has low cumulative abnormal returns (in absolute value) between the analyst's earlier forecast and his revised forecast.

In Section 4, I show that the likelihood that an analyst's revised forecast is too close (ex post) to his earlier forecast is significantly positively related to how many other analysts issue forecasts between the time of the analyst's earlier forecast and the time of his revised forecast, as my model predicts. This relationship holds unconditionally as well as after controlling for the direction of the forecast revision (i.e., up revision versus down revision), the relationship between the direction of the revision and the stock's cumulative abnormal returns (i.e., whether the revision is in the same direction as the stock's CARs or in the opposite direction), the analyst's experience and accuracy, and the firm's size, book-to-market, and analyst coverage. In terms of economic significance, when the control variables are set to their means, the model predicts there is a 49.0% chance that an analyst's revised forecast will be too close to his earlier forecast if no other analysts issue forecasts between his earlier forecast and his revised forecast, and a 59.0% chance if 11 or more analysts issue forecasts between his earlier forecast and his revised forecast.

I consider alternative explanations for my findings in Section 5. In Section 5.1, I consider the possibility that my findings are driven by rational analysts' career concerns: in Section 5.1.1, I derive the implications of a rational model of strategic forecasting and show that its predictions are not consistent with my findings, and in Section 5.1.2, I show that the analysts who exhibit the bias I document do not seem to experience favorable career outcomes relative to those who do not. I show that my findings are not explained by analyst overconfidence in Section 5.2.

In Section 6, I examine the relationship between the market's reaction to forecast revisions

and the analyst's tendency to exhibit the bias I document. I find that the market does not appear to respond differently based on the analyst's prior tendency to exhibit the bias, even though the analyst's prior tendency to exhibit the bias is related to his future forecast errors.

Section 7 concludes.

2 Psychological Foundation

Research on the false consensus effect dates back to Katz and Allport (1931), who found that admitted cheaters' estimates of the prevalence of cheating is significantly higher than non-cheaters' estimates. Ross, Greene, and House (1977) also investigated the phenomenon, and they were the first to use the term "false consensus effect" to refer to it. Since Ross, Greene, and House (1977), hundreds of studies have been conducted on the FCE.

Ross, Greene, and House (1977) asked students a series of two types of questions. One type asked whether the students engage in a certain behavior (e.g. watch television at least 30 hours a month), have a particular preference (e.g. prefer wheat or white bread), struggle with a certain "problem" (e.g. whether they feel they can sufficiently control their temper), have a certain belief (e.g. that nuclear weapons will be used in warfare in the next 20 years), etc. The students were also asked to estimate the percentage of other students who engage in the same behavior, have the same preferences, struggle with the same problems, have the same beliefs, etc. A clear finding arose: the estimates provided by the students who engaged in the behavior, had the preference, etc., differed significantly from those who did not engage in the behavior, who did not have the preference, etc. For example, people who watched television at least 30 hours a month provided an average estimate of 49.2% for the percent of their fellow students who watch television at least 30 hours a month, while the average estimate provided by those who did not watch television at least 30 hours a month was 40.9%. The discrepancies were even larger for the other topics mentioned above (58.8% versus 31.2% for whether nuclear weapons would be used in the next 20 years, 52.5% versus 37.4% for the wheat versus white bread preference, and 42.1% versus 27.9% for the difficulty in controlling one's temper question). In 32 out of the 34 questions, the average estimates provided for the percent of students who believe (participate in, prefer, etc.) X were higher among those students who believe (participate in, prefer, etc.) X than among those who do not believe (participate in, prefer, etc.) X, supporting the hypothesis that people overestimate the degree to which they are similar to others. Since then, hundreds of studies have been conducted and the vast majority support the FCE as defined by Ross, Greene, and House (1977).

Early researchers (including Ross, Greene, and House (1977)) interpreted such findings as evidence that people have an egocentric bias. Dawes and Mulford (1996) recognized that this conclusion is flawed because it is rational for Bayesians to update their beliefs based on a sample of one (e.g., their own type). In response to this objection, psychologists have modified the criterion for bias and defined a "truly false consensus effect" (TFCE) to occur if respondents weight their own responses more heavily than the response of a randomly chosen individual (Engelmann and Strobel

(2000)) or if the dummy for whether they engage in a certain behavior (i.e., their "endorsement") is positively correlated with their forecast error for the percent of people who actually engage in the behavior (Krueger and Zeiger (1993)). Both types of TFCE are supported by the literature.[1] Egan, Merkle, and Weber (2010) find support for this latter measure of a TFCE by analyzing investors' survey responses to questions about others' beliefs about future stock returns. Krueger (1998) and Krueger (2000) provide more comprehensive reviews of this literature.

3 A Model of Forecasting Following Earnings Announcements

I model a group of financial analysts forecasting a company's quarter $t+1$ earnings following the company's quarter t earnings announcement.

3.1 Analysts' Information

For a generic firm-quarter (j,t), let X denote the firm's quarter $t+1$ earnings. Prior to j's quarter t earnings announcement, each analyst believes X is distributed

$$X \sim N\left(\mu, \frac{1}{\tau_x}\right). \tag{1}$$

In my model, analysts share a common prior about the firm's earnings, but my testable hypothesis will only require that the analysts are able to infer other analysts' signals when they release their research reports. Analysts might infer other analysts' signals by reading their reports to see how they justify their forecasts. In the case where other analysts are revising their forecasts, the analyst can compare other analysts' initial forecasts and their revised forecasts.

Earnings announcements appear to provide analysts with useful information about the firm's future earnings, as analyst forecast activity increases substantially immediately following such events. In Figure 1, I plot analyst forecasting activity in event time, where the event date (day 0) is a firm's earnings announcement date. 28.3% of all forecasts issued within 30 days of an earnings announcement date are issued on the day immediately after one of the firm's earnings announcement dates. My findings reported in Figure 1 are similar to those reported in Keskek, Tse, and Tucker (2011).

[INSERT FIGURE 1 HERE]

I assume that at the time of the quarter t earnings announcement, all the analysts following the firm simultaneously receive signals about the firm's quarter $t+1$ earnings. Formally, each analyst

[1] Engelmann and Strobel (2000) are an exception in that they argue the TFCE does not arise when monetary rewards are at stake. In a later study, Engelmann and Strobel (2004) show that the TFCE does arise even when monetary rewards are at stake as long as mental effort is required to determine the responses of other subjects.

i ($i = 1, \ldots, N$) receives a private signal about the value of X:

$$s_i = X + \varepsilon_i. \tag{2}$$

The information released by firms around earnings announcements is public. However, it is reasonable to assume that analysts receive different signals from the announcement. Firms generally announce their earnings via press releases containing unaudited, condensed income statements and balance sheets. In addition, they often hold conference calls to discuss their operating performance and future outlook, and they field questions from analysts. My assumption that analysts receive heterogeneous signals just captures the idea that analysts focus on different pieces of information during the earnings announcement (e.g., due to limited attention) or simply have different interpretations of the information that is released.

Because earnings announcements are public information events that may have differential interpretations, I allow for the possibility that the signal error has a systematic component (η) common to all agents and a private component (ξ_i) that is unique to each analyst $i = \{1, 2, \ldots, N\}$. More formally, the signal error in (2) can be decomposed as

$$\varepsilon_i = \left(\sqrt{\frac{\rho}{\tau}}\right)\eta + \left(\sqrt{\frac{1-\rho}{\tau}}\right)\xi_i, \tag{3}$$

where $(\sqrt{\tau_x}(X-\mu), \eta, \xi_1, \ldots, \xi_N) \sim N(0, I_{N+2})$, where I_{N+2} denotes the (N+2)-dimensional identity matrix. Note that is simply the standard differential information model (e.g., Grossman (1976)) extended to allow for correlation of signal errors and applied in the context of financial analysts and firms' earnings.[2]

The parameters τ_x and τ represent the precisions of the analysts' priors and signals, respectively, while ρ measures the correlation of the analysts' signal errors. Increasing ρ has no effect on the precision of the analysts' *individual* signals, but it increases the precision of the signals about the *other analysts'* signals and it decreases the precision of any (non-trivial) *collection* of signals, e.g., $\{s_1, \ldots, s_N\}$.

3.2 Analysts' Beliefs

Analyst i *believes* that his signal error is given by the equation:

$$\varepsilon_i \;=^i\; \left(\sqrt{\frac{\hat{\rho}}{\hat{\tau}}}\right)\eta + \left(\sqrt{\frac{1-\hat{\rho}}{\hat{\tau}}}\right)\xi_i, \tag{4}$$

where $(\sqrt{\tau_x}(X-\mu), \eta, \xi_1, \ldots, \xi_N) \sim^i N(0, I_{N+2})$. Throughout this paper, I use the notation $x =^i y$ to represent analyst i's belief that $x = y$. Similarly, $\sim^i$ denotes i's belief about the distribution of a random variable.

[2] Holthausen and Verrecchia (1990) also allow for signal errors to have public and private components. In their model, the variable of interest is an asset's fundamental value, whereas in mine, it is a firm's future earnings.

It is straightforward to verify the following equations concerning analysts' precisions, the correlation of their errors, and their beliefs about these parameters. I use the notation $\mathrm{Var}_i(\cdot)$, $\mathrm{Corr}_i(\cdot,\cdot)$, and $\mathbf{E}_i[\cdot]$ to represent the variance, correlation, and expectation taken with respect to i's subjective probability distribution.

$$\begin{aligned}
\mathrm{Var}(\varepsilon_i) &= \frac{1}{\tau} && (5)\\
\mathrm{Var}_i(\varepsilon_i) &= \frac{1}{\hat{\tau}} && (6)\\
\mathrm{Corr}(\varepsilon_j, \varepsilon_k) &= \rho, \quad j \neq k\ , && (7)\\
\mathrm{Corr}_i(\varepsilon_j, \varepsilon_k) &= \hat{\rho}, \quad j \neq k\ . && (8)
\end{aligned}$$

There are several key parameters in the model: the analysts' signal precision (τ), the analysts' perceived signal precision ($\hat{\tau}$), the correlation of the analysts' signal errors (ρ), and the perceived correlation of the analysts' errors ($\hat{\rho}$).

Researchers have examined the effects of biases in perceived signal precisions. Research in psychology suggests that people are overconfident, i.e., that they overestimate their abilities. Behavioral economists have modelled this by assuming that $\hat{\tau} > \tau$. See, e.g., Daniel, Hirshleifer, and Subrahmanyam (1998) and Odean (1998).

Biases in signal error correlations have significantly different consequences than biases in signal precisions. Suppose a person has a biased perception of his signal precision, but he correctly perceives how correlated his signal error is with others'. This person's bias will be apparent if he acts only on his own signal—he will overreact to the signal if he overestimates its precision, and he will underreact if he underestimates its precision. If he is able to observe information about others' signals, the bias will generally be mitigated—even if he overestimates his own signal's precision, he correctly assesses the precisions and correlations of other peoples' signals, so he realizes that if he observes sufficiently many signals from others, his own signal should be assigned very little weight compared to the other agents' aggregate signal. With biased perceptions of signal error correlations, the pattern is different. Suppose a person correctly perceives the precision of his private signal, but he incorrectly perceives how correlated his signal error is with others'. If he acts only on his private signal, he will appear rational since he correctly perceives the precision of his signal. However, when he starts acting on *collections* of people's signals, the bias will become apparent. Hence, compared to people's perceptions of their signal precisions, people's perceptions of their signal error correlations should have a greater impact on information aggregation. Because of this distinction, I focus on biases in people's perceptions of signal error correlations rather than biases in their perceptions of signal precisions. As discussed in Section 2, research in social psychology suggests that people tend to overestimate their similarity to others, a bias referred to as the false consensus effect (FCE). I model the false consensus effect by assuming that $\hat{\rho} > \rho$. To my knowledge, I am the first to analyze biases in signal error correlations.

3.3 Analysts' Forecasts

I assume analysts seek to maximize their forecast accuracy by issuing forecasts that are consistent with their true beliefs. This can be sustained in equilibrium if brokerages can commit ex ante to compensating analysts based on their forecast accuracy. In the case of macroeconomic forecasters, Engelberg, Manski, and Williams (2009) provide evidence that most forecasters' point forecasts are close to their subjective means. Moreover, empirical evidence suggests there should be some limits to how far analysts' forecasts deviate from their true beliefs—Mikhail, Walther, and Willis (1999) and Hong, Kubik, and Solomon (2000) provide evidence that forecast accuracy is an important determinant of analysts' career outcomes. Hence, forecasts should generally be near analysts' subjective means even if they are not equal to them.

It is worth noting, however, that if brokerages cannot commit ex ante to compensating analysts based on their accuracy, and analysts are instead compensated based on brokerages' ex post assessments of their abilities, truthful forecasting can generally not be supported in equilibrium. For example, Trueman (1990) develops an information structure such that in equilibrium, low skilled analysts (knowingly) insufficiently revise their forecasts following public news events, and Trueman (1994) develops an information structure such that in equilibrium, analysts (knowingly) tend to issue forecasts that are closer to the consensus than their private information warrants. I return to the issue of strategic forecasting in Section 5.1, where I examine the theoretical predictions of a model of non-truthful strategic forecasting, and I empirically examine the career outcomes of the analysts whose forecasts are most consistent with my model's predictions.

The analysts are labelled based on the order they issue forecasts following the previous quarter's earnings announcement. That is, analyst 1 is the first to issue an earnings forecast, followed by analyst 2, and so on. When analyst 1 issues his forecast, the forecast is based entirely on his prior and his signal, s_1. Recalling that $\hat{\tau} = \tau$, i.e., the analysts are not overconfident:

$$\begin{aligned} f_1 &= \hat{\mathbf{E}}[X|s_1] \\ &= \frac{\tau_x}{\tau_x + \hat{\tau}}\mu + \frac{\hat{\tau}}{\tau_x + \hat{\tau}}s_1 \\ &= \frac{\tau_x}{\tau_x + \tau}\mu + \frac{\tau}{\tau_x + \tau}s_1, \end{aligned} \tag{9}$$

where $\hat{\mathbf{E}}[\cdot]$ denotes the expectation with respect to analysts' beliefs. If analysts were "rational," $\hat{\mathbf{E}}[\cdot]$ and $\mathbf{E}[\cdot]$ would be equivalent. Since analysts correctly assess the precisions of their signals, they rationally update their beliefs in response to observing a single signal: $\hat{\mathbf{E}}[X|s_i] = \mathbf{E}[X|s_i]$ for all i. More generally, though, analysts have biased beliefs, so $\hat{\mathbf{E}}[\cdot]$ and $\mathbf{E}[\cdot]$ are different operators. Their biased perception of signal error correlations results in biased posteriors when they incorporate signals from multiple analysts. Specifically, $\hat{\mathbf{E}}[X|s_1, \ldots, s_n] \neq \mathbf{E}[X|s_1, \ldots, s_n]$ for all $n > 1$.

The second analyst to issue a forecast observes the first analyst's forecast, f_1, as well as his

own private signal, s_2. His forecast is thus

$$f_2 = \hat{\mathbf{E}}[X|f_1, s_2]. \tag{10}$$

Because $f_1(\cdot)$ is an injective (or "one-to-one") function of s_1, other analysts can infer s_1 based on the first analyst's forecast. I assume throughout my analysis that each analyst knows the other analysts' true signal precisions (τ) and perceived precisions ($\hat{\tau}$). In practice, this assumption corresponds to the idea that an analyst can infer how another analyst updated his beliefs in response to a firm's earnings announcement by reading that other analyst's research report.

Since the other analysts realize that s_1 is an informative signal, they will use this information when issuing their forecasts. Letting $\overline{s_n}$ represent the average of the first n analysts' signals,

$$\overline{s_n} = \frac{1}{n}\sum_{i=1}^{n} s_i, \tag{11}$$

and letting $\hat{\tau}_{\overline{s_n}}$ denote the analysts' beliefs about the informativeness of $\overline{s_n}$, Equation (10) simplifies to[3]

$$\begin{aligned} f_2 &= \hat{\mathbf{E}}[X|s_1, s_2] \\ &= \frac{\tau_x}{\tau_x + \hat{\tau}_{\overline{s_2}}}\mu + \frac{\hat{\tau}_{\overline{s_2}}}{\tau_x + \hat{\tau}_{\overline{s_2}}}\overline{s_2}. \end{aligned} \tag{12}$$

The process repeats. Since f_2 is an injective function of $\overline{s_2}$, all analysts are able to infer $\overline{s_2}$. The third analyst then issues the forecast

$$\begin{aligned} f_3 &= \hat{\mathbf{E}}[X|f_1, f_2, s_3] \\ &= \hat{\mathbf{E}}[X|s_1, s_2, s_3] \\ &= \frac{\tau_x}{\tau_x + \hat{\tau}_{\overline{s_3}}}\mu + \frac{\hat{\tau}_{\overline{s_3}}}{\tau_x + \hat{\tau}_{\overline{s_3}}}\overline{s_3}. \end{aligned} \tag{13}$$

Continuing,

$$f_4 = \hat{\mathbf{E}}[X|f_1, f_2, f_3, s_4] \tag{14}$$

$$\vdots$$

$$f_n = \hat{\mathbf{E}}[X|f_1, f_2, \ldots, f_{n-1}, s_n]$$

$$\vdots$$

$$\begin{aligned} f_{n+k} &= \hat{\mathbf{E}}[X|f_1, f_2, \ldots, f_{n+k-1}, s_{n+k}] \\ &= \frac{\tau_x}{\tau_x + \hat{\tau}_{\overline{s_{n+k}}}}\mu + \frac{\hat{\tau}_{\overline{s_{n+k}}}}{\tau_x + \hat{\tau}_{\overline{s_{n+k}}}}\overline{s_{n+k}}. \end{aligned} \tag{15}$$

[3] Given my assumptions, it is a standard result in Bayesian statistics that $\mathbf{E}[X|s_1, \ldots, s_n] = \mathbf{E}[X|\overline{s_n}]$. This can also easily be proven by taking conditional expectations of multivariate normal distributions.

Later, after analysts $n+1$ through $n+k$ issue forecasts, analyst n revises his forecast from f_n to f_n^R:

$$f_n^R \quad = \quad \hat{\mathbf{E}}[X|f_1, \ldots, f_{n-1}, s_n, f_{n+1}, \ldots, f_{n+k}].$$

Note that this model assumes each analyst's forecast following an earnings announcement is based on (i) his beliefs about the quarter $t+1$ earnings before observing the quarter t earnings announcement (his prior), (ii) his interpretation of the quarter t earnings announcement (his signal), and (iii) other analysts' reports following the quarter t earnings announcement (other analysts' signals).

3.4 Testable Prediction

Recall my assumptions about analysts' beliefs: $\hat{\tau} = \tau$ and $\hat{\rho} > \rho$. Since analysts correctly perceive the precisions of their priors and signals, a single analyst acting in isolation (e.g., covering a firm that no other analysts cover) will appear rational. His biased perception of signal error correlations will not affect his forecasts—since he cannot observe other analysts' signals, he cannot make mistakes combining other analysts' signals with his own.

Biased perceptions of signal error correlations can only become apparent when analysts aggregate their information with other analysts' information. Since the precision of a collection of signals is decreasing in the correlation of the signal errors, an analyst who overestimates the correlation of his signal errors with others' will tend to underestimate the precision of the collection of signals consisting of his signal and others' signals. To develop a testable prediction, I therefore examine the relationship between f_n, the nth analyst's forecast, and f_n^R, his revised forecast, as a function of the number of analysts who issue forecasts between the times of the initial forecast and the revised forecast. For example, in Section 3.3, k denotes the number of analysts who issue forecasts between n's initial forecast and his revised forecast. Mathematically, I show that the likelihood of the revised forecast, f_n^R, being too close (ex post) to his earlier forecast, f_n, increases in the number of other analysts (k) issuing forecasts between the times f_n and f_n^R are issued. I examine ex post probabilities of underreaction, and not the size of forecast errors, because even with a false consensus, forecast errors will generally become smaller (in absolute value) as an analyst observes more signals. However, the *rate* at which they reduce is slower, and the *limit* of the resulting mean squared error (as $n \to \infty$) is larger, if they have a false consensus than if they correctly assess the signal error correlations. Ex post underreaction probabilities, on the other hand, are monotonically related to the number of other analysts' signals an analyst has observed.

Proposition 1. *If $\hat{\tau} = \tau$ and $\hat{\rho} > \rho$, then for all n, the probability of the event*

$$\left\{ sgn(X - f_n^R) = sgn(f_n^R - f_n) \right\}$$

is increasing in the number of analysts who issue forecasts between f_n and f_n^R, where $sgn(\cdot)$ is the sign function (which maps positive numbers to $+1$ and negative numbers to -1).

Proposition 1 provides the foundation for my empirical analysis. (See Appendix C for proofs of all propositions.)

4 Empirical Analysis of the Model's Prediction

Proposition 1 predicts that the likelihood of ex post underreaction increases in the number of other analysts issuing forecasts between the time of the analyst's earlier forecast and the time of his revised forecast. When $n = 0$, the proposition applies to analysts who issue forecasts *before* an earnings announcement and revise the forecasts following an earnings announcement.[4] Table 10 from Clement, Hales, and Xue (2011) confirms this prediction. Hence, Proposition 1 provides an explanation for their findings. In addition, the proposition provides testable predictions for analysts whose initial forecasts are issued *following* an earnings announcement ($n > 0$). In this section, I test the proposition among {initial forecast, revised forecast} pairs where each forecast is issued between consecutive earnings announcements. I define *NUM* as the number of other analysts who issue forecasts between the analyst's earlier forecast and his revised forecast. According to Proposition 1, ex post underreact probabilities should increase in *NUM*.

Figure 2 illustrates the timing of earnings announcements and forecasts, as well as the construction of the variable *NUM*. Figure 3 graphically depicts the testable prediction.

[INSERT FIGURES 2 AND 3 HERE]

4.1 Data

My data consist of earnings forecasts issued by sell-side financial analysts, as supplied by I/B/E/S. I start with the 4,103,517 distinct observations from the unadjusted file with forecast period indicator equal to 6 or 7. I get earnings information from the I/B/E/S unadjusted actuals file. 3,913,937 of the 4,103,517 forecasts can be linked to earnings actuals with non-missing earnings announcement dates. 3,675,720 of these forecasts can be matched to CRSP, and 3,582,911 of these observations have non-missing I/B/E/S split adjustment factors. From these forecasts, there are 3,558,388 distinct (analyst, firm, forecast date, quarter end date). I restrict my sample to these distinct forecasts. 1,672,607 of these forecasts are forecast revisions. 516,354 of these revisions satisfy the condition that the analyst's previous forecast was also issued after the firm's earnings announcement. 493,325 satisfy the following conditions: (i) the quarter end date for the quarter being forecasted is no more than 105 days later than the previous quarter's quarter end date, (ii) the earnings announcement date for the quarter being forecasted is no more than 90 days after the quarter end date, and (iii) the previous quarter's earnings announcement date is no more than 90 days after its quarter end

[4] My proof applies for the $n = 0$ case. Since $\hat{\omega}_0 = 0$, $f_0 = \mu$, and all references to $\overline{s_0}$ can be ignored (since they are multiplied by $\hat{\omega}_0 = 0$).

date. 442,172 of these forecasts can be assigned to size and book-to-market quintiles based on NYSE cutoffs.

In my model, the likelihood of an analyst's revised forecast equalling his earlier forecast or the firm's true earnings is 0, because all the distributions are continuous. In practice, forecasts and announced earnings are necessarily discrete, due to rounding. Therefore, I exclude forecast revisions that are equal to the analyst's earlier forecast or the firm's reported earnings. I exclude these observations because I am unable to determine whether the analyst underreacted (ex post) in his revision in these cases. For a discussion of the assumptions needed so that my results are not biased by excluding these observations, see Appendix B. This filter reduces my sample to 386,697 observations.

I apply several non-standard filters to my sample so that it is as consistent as possible with my model's assumptions. My model assumes there is no extraneous news after the firm's earnings announcement and prior to the analysts' forecasts. Analysts' beliefs continue to evolve, but they evolve in response to other analysts' forecasts, not in response to other news about the firm's next quarter earnings. In practice, analysts often revise their forecasts following news that is unrelated to the other analysts' forecasts. To maximize the likelihood that the analyst revised his forecast based on other analysts' reports, I restrict my sample to forecast revisions that (i) are issued within 30 days of the analyst's earlier forecast and (ii) are for companies whose stocks' cumulative abnormal returns (CARs) are less than 2% (in absolute value) between the analyst's earlier forecast and the revised forecast (both forecast days are included in the CAR computation).[5] These filters increase the likelihood that the forecast revisions are based on other analysts' forecasts, but they do not rule out the possibility that other news causes the analyst to revise his forecast. I also exclude forecast revisions that occur on the day immediately following the analyst's earlier forecast because for these revisions, there is no time for other analysts to issue forecasts between the analyst's earlier forecast his revised forecast.[6] Finally, since the timing of the forecasts relative to the earnings announcement is critical, I require that I/B/E/S and Compustat earnings announcement dates agree. After these restrictions, my sample consists of 26,100 forecast revisions.

To summarize, my sample consists of the 26,100 $\{f^0, f^R\}$ pairs (i.e., {initial forecast, revised forecast} pairs) for a firm's quarter $t+1$ earnings such that: (i) f^0 was issued after the firm's quarter t earnings announcement, (ii) $f^R \notin \{EPS, f^0\}$, where EPS is the firm's reported quarter $t+1$ earnings (which the analyst is forecasting), (iii) f^R was issued between 2 and 30 days from the date f^0 was issued, and (iv) the cumulative abnormal return between the date f^0 is issued and the date that f^R is issued is between -2% and 2%.

The sample spans from 1985 to 2011. Earnings for 3,405 distinct firms are forecasted by 4,796 distinct analysts. The average time between the previous quarter's earnings announcement date and the analyst's earlier forecast is 30.88 days (std = 26.14), the average time between the analyst's

[5] I compute daily abnormal returns (ARs) as the difference between a stock's daily return and its size and book-to-market matched return, where the stocks are sorted into quintiles based on NYSE cutoff values. See Appendix A for a more detailed description of these variable definitions.

[6] In unreported regressions, I find that all of my results are qualitatively similar whether I remove these forecasts from my sample or not.

earlier forecast and the revised forecast is 15.69 days (std = 8.63), and the average time between the revised forecast and the forecasted quarter's earnings announcement date is 45.28 days (std = 26.96). The average *Coverage*, defined as the number of distinct analysts issuing a forecast for the firm in the given quarter, is 13.3 (std = 7.4). These summary statistics are reported in Table 1.

[INSERT TABLE 1 HERE]

NUM, the number of distinct analysts issuing forecasts between the times of the analyst's earlier forecast and his revised forecast, is my independent variable of interest. The distribution of *NUM* for my sample of 26,100 forecast revisions is plotted in Figure 4. In my probit analysis, I reassign all values of *NUM* that are greater than 11 to take the value of 11, to minimize the effects of outliers. My qualitative results do not depend on this reassignment.

[INSERT FIGURE 4 HERE]

Large firms with high analyst coverage are heavily represented in my sample of 26,100 forecast revisions. 12,593 of the observations lie in the largest quintile of stocks based on NYSE cutoffs, while only 1,617 of the observations lie in the smallest quintile based on NYSE cutoffs. The disparity is more striking when firms are sorted based on their level of analyst coverage: 18,727 of the observations lie in the quintile of stocks with the highest analyst coverage, while only 273 lie in the quintile with the lowest analyst coverage.[7] These facts are not surprising—small firms tend to have few analysts covering them, and firms with few analysts covering them have little chance of appearing in my sample.

[INSERT TABLE 2 HERE]

The sample also consists of a large number of growth stocks that are easy to forecast, where the measure of forecast difficulty is the average absolute value of price-scaled forecast errors one to eight quarters ago. Moreover, more of the revisions in my sample are made by analysts with low accuracy than high accuracy, and more of them are issued by analysts with high experience than low experience. See Appendix A for a detailed definition of each of the variables in Table 2

4.2 Results

To test Proposition 1 (on page 11), I begin by running a probit regression. The sample consists of the 26,100 observations described in Section 4.1. The dependent variable is the dummy for whether the

[7] Here, the quintiles are based on the entire set of stocks being forecasted in the I/B/E/S dataset in a given quarter.

analyst's revised forecast is too close (ex post) to his earlier forecast. More formally, using the notation from Proposition 1 on page 11, the dependent variable is the dummy $\mathbf{1}_{\{\text{sgn}(X-f_n^R)=\text{sgn}(f_n^R-f_n)\}}$. The independent variable is *NUM*, which is the number of forecasts issued by other analysts between the date of the analyst's earlier forecast and his revised forecast. If the analysts have a false consensus, the coefficient of *NUM* should be significantly positive. The results of this regression are reported in Column 1 of Table 3.

[INSERT TABLE 3 HERE]

The likelihood of underreaction is strongly related to the number of analysts issuing forecasts between the analyst's earlier forecast and his revised forecast, as my theory predicts.

Of course, the probit model with only one independent variable is simplistic. Endogeneity is an obvious concern. It is possible that an omitted variable is actually driving the analyst underreaction that I document, and *NUM* is simply correlated with this omitted variable. For example, it is possible that analysts underreact to news about companies' earnings, and *NUM* simply proxies for the presence of news. Though I cannot completely dismiss this possibility, it is worth noting that the sample was specifically designed to minimize this risk. For each forecast revision in my sample, the analyst's revised forecast must be issued within 30 days of his earlier forecast, there cannot be an earnings announcement between the analyst's earlier forecast and his revised forecast, and the firm's cumulative abnormal returns between the analyst's earlier forecast and his revised forecast must be less than 2% in absolute value.

To reduce the possibility that omitted variables are driving my result, I add control variables to my probit regressions. First, I control for the fact that the likelihood of underreaction is strongly related to whether the revision is an upward revision or a downward revision. In my sample of forecast revisions, analysts underreact (ex post) in 65% percent of their upward revisions, while they underreact for only 42% percent of their downward revisions. This discrepancy might be due to analysts seeking to please management by issuing beatable forecasts: Kang, O'Brien, and Sivaramakrishnan (1994) document that short-term forecasts are often pessimistic, and Richardson, Teoh, and Wysocki (2004) provide evidence that this is due to analysts' incentives to issue short-term forecasts that management can beat. Because of the disparity in underreaction probabilities between upward and downward revisions, I include the dummy variable *UpRevision* which equals 1 if and only if the revision is an upward revision. As expected, the coefficient for this variable is positive and highly significant. Moreover, *NUM* remains statistically significant after inclusion of this variable. The results of the probit regression including this dummy are reported in Column 2 of Table 3.

Next, I control for whether the revision is in the same direction as the stock's abnormal returns. Elliott, Philbrick, and Wiedman (1995) have documented that analysts tend to underreact to information contained in a company's stock returns. I define a dummy variable *ReviseWithCARs* which equals 1 if and only if the forecast revision is in the same direction as the cumulative abnormal

returns (CARs) for the company's stock between the date of the analyst's earlier forecast and his revised forecast. For example, if a stock has had positive CARs and the analyst issues a revised forecast that is greater than his earlier forecast, the variable takes the value 1. However, if a stock had positive CARs and the analyst issues a revised forecast that is lower than his earlier forecast, the variable takes the value 0. Because it has been documented that analysts underreact to market information, I predict (and find) that the coefficient of this variable is positive and highly significant. Moreover, *NUM* remains highly significant after including this additional control variable. The results of this probit are reported in Column 3 of Table 3.

Another potential source of an omitted variables bias is that the firms, or analysts, associated with high *NUM* revisions might differ fundamentally from the firms, or analysts, associated with low *NUM* revisions. It is possible that *NUM* is simply capturing differences in firms or analysts, and not biases in how analysts incorporate the information content of other analysts' forecasts. For example, high *NUM* revisions tend to be for firms with high analyst coverage. So it's possible that analysts are just more prone to underreacting in their revisions for firms with high analyst coverage than they are for firms with low analyst coverage.[8]

Though I cannot dismiss stories such as these, I address the issue by defining control variables to capture the type of firm associated with each revision as well as the type of analyst that issued the revision. My firm-type control variables include the firm's size, book-to-market ratio, analyst coverage, and forecast difficulty (as measured by the absolute value of previous price-scaled forecast errors). My analyst-type control variables include the analyst's previous accuracy (both for the firm being forecasted and his overall accuracy) and his experience (both for the firm being forecasted and his overall experience). These controls are motivated in part by Mikhail, Walther, and Willis (1997) and Brown (2001), who show that accuracy is increasing in firm-level experience and prior accuracy, respectively. Because there is no reason to expect a particular functional relationship between these controls and underreaction probabilities, I create quintile dummies for each of the controls. For example, the variable *SizeDummy1* is the dummy that the firm is in the smallest size quintile. See Appendix A for a detailed description of the definitions of my controls.

In Column 4 of Table 3, I include firm-type controls but not analyst-type controls. The addition of these controls increases the coefficient of *NUM* and the pseudo-R^2. In Column 5 of Table 3, I include the analyst-type controls but not the firm-type controls. The addition of these controls increases the pseudo-R^2 and has no effect on the coefficient of *NUM*. In Column 6 I include both the firm-type and analyst-type controls.

While my firm-type and analyst-type controls can account for much of the heterogeneity across firms and analysts, there is undoubtedly some unobservable heterogeneity that is not captured by these controls. Firm and analyst fixed effects can control for such unobserved heterogeneity. Since fixed effects are problematic in probit regressions (due to the incidental parameter problem), I

[8] Another way of ensuring that my results are not driven by differences in coverage between high *NUM* observations and low *NUM* observations is to simply restrict the sample to the firms with the highest analyst coverage. In unreported probits, I restrict attention to the 18,727 observations in the highest *Coverage* quintile, and all my results are qualitatively similar.

estimate OLS with fixed effects to account for unobserved heterogeneity among analysts and firms.

In Column 7, I report the results of a regression using the same variables that were used in Column 6, the only difference being that in Column 7 I report the results of an OLS regression, whereas in Column 6 I report the results of a probit. In Column 8, I include firm fixed effects but not analyst fixed effects. In Column 9, I include analyst fixed effects but not firm fixed effects. Finally, in Column 10, I include both firm and analyst fixed effects.

In all of the regressions, the coefficient of *NUM* is highly significant, confirming my model's prediction.

4.3 Robustness

In this section, I show that my testable prediction is confirmed when my filters are less restrictive, and that it broadly holds regardless of the direction of the revision (up versus down, and its position relative to the earlier forecast and the consensus forecast).

4.3.1 Less Restrictive Filters

In my model, analysts update their beliefs based only on other analysts' forecasts. To empirically test the model, I chose a sample of forecast revisions in which there is likely little news about the firm between the analyst's forecasts. Specifically, to obtain my sample of 26,100 forecast revisions in Table 3, I restricted the sample to forecast revisions that were (i) issued within a month of the analyst's earlier forecast and (ii) for firms whose stock had cumulative abnormal returns (CARs) less than 2% (in absolute value) between the day of the analyst's earlier forecast and his revised forecast. To ensure that my results are not sensitive to these restrictions, I run the probit regression from Column 6 of Table 3 for samples of forecast revisions that are less restrictive.

I begin by relaxing the requirement that the cumulative abnormal returns be less than 2% in absolute value. In Column 2 of Table 4, I do not place any restrictions on the CARs of the firm being forecasted. When I remove that restriction, the sample grows to 116,443 forecast revisions. The economic significance of *NUM* is lower for this less restrictive sample, but *NUM* is still significant at the 1% level. Next, I relax the requirement that the revised forecast be issued within one month of the analyst's earlier forecast. In Column 3 of Table 4, I do not place any restrictions on the timing of the analyst's revised forecast, but I continue requiring that the stock being analyzed has a CAR that is less than 2% in absolute value. When doing this, the sample grows to 53,938 forecast revisions, and as before, the economic significance of *NUM* drops, but the coefficient of *NUM* is still significant at the 1% level. Finally, I loosen both of the restrictions by simply requiring that the CARs between the analyst's earlier forecast and his revised forecast be less than 10%, regardless of the amount of time that elapses between the analyst's earlier forecast and his revised forecast. This sample contains 202,326 forecast revisions. Again, the economic significance of *NUM* is lower than it is for my original sample of 26,100 observations, but it is still significant at the 1% level.

[INSERT TABLE 4 HERE]

Taken together, the results presented in Table 4 show that my qualitative results are not dependent on the specific choice of sample. Moreover, it is noteworthy the economic significance of *NUM* is highest in the sample for which my model should be most applicable.

4.3.2 Direction of Revision

In my model, all of the probability distributions are symmetric. If the conditional distributions of forecasts and earnings are asymmetric, and if analysts do not provide subjective medians as their point forecasts, then it is possible that my empirical findings are driven by a relationship between *NUM* and skewness of the distributions.

To examine this, I divide the forecasts into separate samples depending on whether the revision was greater than or less than the analyst's earlier forecast. In Column 1 of Table 5, I run the same regression as in Column 6 of Table 3, except that the sample is restricted to the 11,317 revised forecasts that are greater than the analyst's earlier forecast. In Column 2 of Table 5, I run the same regression, except that the sample consists of the 14,783 revised forecasts that are less than the analyst's earlier forecast. In both of these regressions, I drop the control variable *UpRevision* because all of the revisions are in the same direction. I do, however, include all of the other control variables from Column 6 of Table 3. The coefficient of *NUM* is statistically significant in both regressions.

[INSERT TABLE 5 HERE]

There is a large literature examining the relationship between analysts' forecasts and the consensus forecast. Analysts are said to "herd" when they underweight their private information and issue forecasts near the consensus. It is worth noting that the bias I model is fundamentally different than herding. With herding, analysts change the weighting of their private signal relative to others'. Analysts with a false consensus, on the other hand, weight their signal equally to others', but they improperly weight the total collection of signals consisting of their own signal and others'.

I divide the forecast revisions into three disjoint groups: 6,545 "Away" revisions consisting of the revisions that are in the opposite direction of the consensus forecast, 5,943 "Herding" revisions consisting of the revisions that lie between the analyst's earlier forecast and the consensus forecast, and 10,807 "Past" revisions consisting of the revisions that are in the same direction as the consensus, but that go beyond the consensus forecast. I define the consensus forecast on a given date by taking the mean of each analyst's most recent forecast, where the forecasts are restricted to those that are at most 30 days old.

For each of the three samples, I run the regressions from Column 6 of Table 3. The results of these regressions are reported in Columns 3 through 5 of Table 5. In each of the three samples, the coefficient of *NUM* is positive, and it is highly significant in the Herding and Past samples. Hence, the relationship between *NUM* and ex post underreaction is not simply a by-product of analyst herding.

5 Alternative Explanations

In this section, I examine whether my results can be explained by analysts' career concerns or analyst overconfidence

5.1 Career Concerns

5.1.1 A Model of Strategic Forecasting

When developing my model in Section 3, I assumed that analysts' forecasts equal their subjective means. This can be an equilibrium outcome if brokerages can commit (ex ante) to compensating analysts based on their mean squared forecast errors. If, however, analysts are compensated according to the market's ex post assessment of the analyst's ability, analysts' forecasts will not equal their subjective means (see, e.g., Trueman (1994) and Ottaviani and Sørensen (2006)). Moreover, if analysts are evaluated relative to their peers (e.g., in a contest-like fashion), their equilibrium forecasts may differ from their true beliefs (Ottaviani and Sørensen (2006)). Such career concerns can lead to predictable biases in analysts' forecasts even though the analysts are rational. In this section, I follow Marinovic, Ottaviani, and Sørensen (2012) in developing a forecasting model that incorporates analysts' reputational concerns and possible contest-like payouts.

As in Section 3, analysts have a common prior over a firm's quarter $t+1$ earnings, X, and when the firm announces its quarter t earnings, analysts simultaneously receive private signals for X: $s_i = X + \varepsilon_i$. The precisions of the prior and the private signals are denoted by τ_x and τ, respectively. Analysts believe their signal errors have correlation $\hat{\rho}$, whereas in reality the correlation equals ρ. Marinovic, Ottaviani, and Sørensen (2012) assume analysts release their forecasts simultaneously, but I continue to assume that they are released sequentially.

The key difference between this model and the one in Section 3 is that in Section 3, analysts' forecasts equal their true beliefs, but in this section they do not. I follow Marinovic, Ottaviani, and Sørensen (2012): letting $\mathbf{E}_i[\cdot]$ denote i's expectations, $\mathbf{E}_b[\cdot]$ denote brokerages' expectations, and Z_i denote public information available at the time i issues his forecast, I assume analyst i issues the forecast f_i to maximize the function

$$
\begin{aligned}
U_i &= u_r^{\delta} u_c^{1-\delta}, \quad \text{where} && (16)\\
u_r &= \exp\{-\mathbf{E}_i[\mathbf{E}_b(X - s_i|f_i, X)^2|s_i, Z_i]\}, && (17)\\
u_c &\propto \frac{g_{X|s_i}(f_i)}{h_{f|X=f_i}(f_i)}, \quad \text{where} && (18)\\
g_{X|s_i}(\cdot) &= \text{conditional probability density of } X \text{ (given } s_i \text{ and } Z_i\text{)}, \quad \text{and} && (19)\\
h_{f|X=f_i}(f_i) &= \text{fraction of others' forecasts that equal } f_i\text{, conditional on } X = f_i \text{ (and } Z_i\text{)}. && (20)
\end{aligned}
$$

As Marinovic, Ottaviani, and Sørensen (2012) note, $h_{f|X=f_i}(\cdot)$ is not a probability density because the conditioning event varies with the function's argument.

The function u_r captures analyst i's reputational concerns: i wants brokerages to believe the

signal he received was close to the true earnings realization. The function u_c captures contest-like payouts when i is evaluated relative to his peers: i "wins" the contest if he correctly forecasts the firm's earnings, in which case he splits the prize equally with the other analysts who correctly forecast earnings. The parameter θ denotes the intensity of reputational concerns relative to the contest-like payouts the analysts face:

$$\theta = \frac{\delta}{1-\delta}. \tag{21}$$

Z_i, the public information available at the time i issues his forecast, consists of the forecasts $f_1, \ldots, f_{i-1}$ as well as public information about the firm's quarter $t+1$ earnings prior to its quarter t earnings announcement.

I consider the limiting case in which infinitely many analysts issue forecasts at the same time. This assumption may seem unnatural in my sequential forecasting environment: on the one hand, each analyst ($i = 1, \ldots, N$) issues a forecast as though infinitely many others are simultaneously issuing forecasts, but on the other hand, each analyst i only observes finitely many forecasts ($f_1, \ldots, f_{i-1}$) before issuing his forecast. The model, however, makes sense in an environment in which there are many private/internal forecasts that can be observed by brokerages (ex post) when evaluating the analysts who issue public forecasts, but the analysts issuing the public forecasts cannot observe these private/internal forecasts (until possibly after they issue their public forecasts).

Proposition 2. *For each $n = 1, \ldots,$ let D_n be defined as*

$$D_n = \tau_x[(n-1)\hat{\rho}+1] + n\tau. \tag{22}$$

If

$$\theta < \theta_n^+ \equiv \frac{\tau D_n[(n-2)\hat{\rho}+1]}{2(D_n-\tau)(n\hat{\rho}+1)(1-\hat{\rho})}, \tag{23}$$

then the quadratic equation

$$\begin{aligned} 0 = {} & \alpha_n^2\left(\frac{D_n}{(n-1)\hat{\rho}+1} - \frac{\tau(1-\hat{\rho})}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right) \\ & +\alpha_n\left(\frac{2\tau}{n\hat{\rho}+1} - \frac{\tau}{(n-1)\hat{\rho}+1} - \frac{\tau\hat{\rho}}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right) \\ & +\left(2\theta - \frac{\tau[(n-1)\hat{\rho}+1]}{(n\hat{\rho}+1)(1-\hat{\rho})} - \frac{2\theta\tau}{D_n} + \frac{\tau\hat{\rho}}{(n\hat{\rho}+1)(1-\hat{\rho})}\right). \end{aligned} \tag{24}$$

has a unique positive solution. Let α_n^* denote this unique positive solution. If in addition to (23),

$$\frac{[(n+1)\hat{\rho}+1]\alpha_n^*}{[(n-1)\hat{\rho}+1](n\hat{\rho}+1)} + \frac{2\theta}{D_n} - \frac{\hat{\rho}}{(n\hat{\rho}+1)(1-\hat{\rho})} > 0, \tag{25}$$

then

$$f_n^* \quad = \quad (1-\alpha_n^*)\left(\frac{(n-1)\tau}{D_{n-1}}\right)\overline{s_{n-1}} + \alpha_n^* s_n \tag{26}$$

is the unique symmetric linear equilibrium forecasting strategy in which analysts place positive weight on their private signals.

The condition in (23) guarantees the uniqueness of the positive weight the analysts place on their private signals. The condition in (25) guarantees that f_n^* maximizes (rather than minimizes) analyst n's objective function. The proposition generalizes Proposition 1 in Marinovic, Ottaviani, and Sørensen (2012) in two ways: it applies to a sequential (rather than simultaneous) forecasting environment, and it allows for analysts to believe their signal errors are correlated.

In Section 3, I showed that if analysts have a false consensus and their forecasts equal their beliefs, then $\mathbf{P}\Big(\text{sgn}(X - f_n^R) = \text{sgn}(f_n^R - f_n)\Big)$ increases in k, where f_n is analyst n's initial forecast, f_n^R is his revised forecast, and k is the number of analysts issuing forecasts (chronologically) between the dates f_n and f_n^R are issued. In Section 4, I empirically confirmed this prediction. Here, I examine whether the strategic forecasting model developed in this section makes a similar prediction when analysts are rational.

As in Section 3, analysts' forecasts are injective functions of their private signals and public information; this, when combined with the common prior assumption, implies that $f_n^R = f_{n+k}$.[9] By Dharmadhikari and Joag-Dev (1984), I can investigate the effect of k on $\mathbf{P}\Big(\text{sgn}(X - f_n^R) = \text{sgn}(f_n^R - f_n)\Big)$ by analyzing the relationship between $\text{corr}(X - f_{n+k}, f_{n+k} - f_n)$ and k, which I do numerically.

Without loss of generality, I normalize the precision of the prior to equal 1: $\tau_x = 1$. I let the parameters τ, $\hat{\rho}$, δ n, and k take the following values:

$$\begin{aligned}
\tau &\in \bigcup_{m=1,\ldots,10} \left\{m, \frac{1}{m}, \frac{10+m}{10}, \frac{10}{10+m}\right\} \\
\hat{\rho} &\in \bigcup_{m=0,\ldots,9} \left\{\frac{m}{10}\right\} \\
\delta &\in \bigcup_{m=1,\ldots,99} \left\{\frac{m}{100}\right\} \\
n &\in \{1,\ldots,5\} \\
k &\in \{1,\ldots,5\}.
\end{aligned}$$

To analyze the model's predictions with rationality, I begin by taking the cartesian product of these parameters. This results in 183,150 unique $(\tau, \rho, \hat{\rho}, \delta, n)$ combinations with the property that $\hat{\rho} = \rho$. Of these 183,150 combinations, 42,049 have the property that (23) and (25) are satisfied for all

[9] Recall that in practice, f_{n+k} often differs from f_n^R. This is likely due to differences in analyst n and analyst $n+k$'s priors, which is why I analyze forecast revisions in my empirical analysis.

$k = 1, \ldots, 5$. The majority (32,093) of these 42,049 parameter combinations have the property that $\text{corr}(X - f_{n+k}, f_{n+k} - f_n)$ is *less* for $k = 5$ than for $k = 1$. In other words, for most of these parameter values, the likelihood of ex post underreaction *decreases* in k, contrary to my empirical findings in Section 4. Moreover, 9,955 of the 9,956 parameter values that predict the correlation to increase in k have the property that $\text{corr}(X - f_{n+5}, f_{n+5} - f_n) < 0$. Negative correlations between $X - f_{n+k}$ and $f_{n+k} - f_n$ imply that ex post underreaction probabilities are less than 50%, which is not consistent with the empirical evidence (Elliott, Philbrick, and Wiedman (1995)). Hence, only *one* of the 183,150 parameter values predicts that (i) ex post underreaction probabilities increase in *NUM* (an empirical finding I document in Section 4) and (ii) ex post underreactions are greater than 50% (an empirical finding documented by Elliott, Philbrick, and Wiedman (1995)).

Of course, the fact that this model does not explain my empirical findings does not disprove all career concerns explanations. It is possible that my results are explained by a different information structure or a career concerns objective function that differs from (16). Exploring this possibility is beyond the scope of this paper and is an avenue of future research.

5.1.2 Empirical Analysis of Career Outcomes

If my empirical findings from Section 4 are due to career concerns, then one might expect that analysts who exhibit the bias should experience better career outcomes than ones who do not. Moreover, the bias should be strongest for the analysts who have the strongest incentives to exhibit the bias. If, on the other hand, my findings are due to a behavioral bias (e.g., the false consensus effect), then analysts should learn over time and correct their mistakes. In other words, the bias should be most severe for analysts with little experience.

To determine whether analysts who exhibit the bias experience favorable career outcomes, I start with the 96,190 unique analyst-year pairs, (i, t), from the unadjusted I/B/E/S detail history file. Then, I restrict the sample to (i, t) pairs such that (i) $t < 2012$, (ii) $i \neq 000000$, and (iii) i has non-missing *Total Accuracy* as of January 1 of year t.[10] This leaves me with 47,035 (analyst, year) pairs. To estimate *NUM* coefficients at the analyst-year level, I need the analyst to issue many forecast revisions. I require that i issue at least 20 forecast revisions in years $t - 5$ through $t - 1$ satisfying:

(i) the revised forecast is not equal to the analyst's earlier forecast or the true earnings per share

(ii) the revised forecast was issued between 2 and 30 days after the analyst's earlier forecast

(iii) both the initial forecast and the revised forecast are issued after the firm's quarter q earnings announcement, where $q + 1$ denotes the quarter being forecasted

(iv) the CARs between the analyst's earlier forecast and his revised forecast are less than 10% in absolute value[11]

[10] I exclude 2012 forecasts so that I can examine whether the analyst exits the sample in the subsequent year.

[11] For my primary analysis in Section 4.2, I require that the CARs be less than 2% in absolute value. Here, I loosen

(v) the earnings announcement dates from I/B/E/S and Compustat agree

After applying these filters, I am left with 2,645 analyst-year pairs. For each of these 2,645 (i, t) pairs, I take all of i's forecast revisions in years $t-5$ through $t-1$ satisfying (i)-(v) above, and run a probit, where the dependent variable is the dummy for whether the revision was too close to the earlier forecast, and the independent variables are *NUM*, *UpRevision*, and *ReviseWithCARs*. The probits converged for 2,461 of these (i, t) pairs, leaving me with 2,461 *NUM* coefficient estimates at the analyst-year level. To address the question of whether analysts have incentives to have positive *NUM* coefficients, I compare the career outcomes of analysts with high *NUM* coefficients to those with low *NUM* coefficients. Since this sample is small relative to my other samples, and since each analyst's *NUM* coefficient is estimated with noise, this test will likely have less power than my other tests. In particular, a lack of relationship between analysts' *NUM* coefficients and their subsequent career outcomes will suggest, but not prove, that my results are due to behavioral biases rather than career concerns.

Regarding career outcomes, it is generally accepted that analysts prefer working for large brokerage houses (see, for example, the arguments presented in Hong and Kubik (2003)). I classify an analyst as having a career advancement in year t if he (i) begins the year working for a brokerage house that is not one of the ten largest, (ii) changes brokerages during year t, and (iii) ends the year working for one of the largest ten brokerages. I classify an analyst as having a career demotion in year t if he begins the year working for one of the ten largest brokerage houses, changes brokerages during year t, and ends the year working for a brokerage that is not one of the ten largest. I also track the likelihood of an analyst exiting the I/B/E/S dataset: I classify an analyst as having exited the profession in year t if he issues a forecast in year t but does not issue a forecast in year $t+1$.

Given my definitions of advancements and demotions, it is impossible for an analyst at a top ten brokerage to have an advancement or for an analyst at a non-top ten brokerage to have a demotion. Hence, I separately analyze the outcomes of analysts who begin the year at a non-top ten brokerage (Panels A and C of Table 6) and those who begin the year at a top ten brokerage (Panel B). I control for *Experience* by sorting the (i, t) pairs into quintiles based on i's *Experience* at the beginning of year t; I conduct this sort within each year and within the analyst's brokerage's status (top ten or non-top ten). Within each year, starting brokerage status, and *Experience* quintile, I sort analysts into two groups, high and low, based on whether their *NUM* coefficient is greater than the median for that {year, brokerage status, *Experience* quintile} triple. Finally, I report the number and percentage of advancements, demotions, and exits for the high and low *NUM* coefficient groups. These results are presented in Panels A and B. In Panel C, I report the frequency of career advancements when analysts not working for top ten brokerages are sorted on *Total Accuracy* within each year and *Experience* quintile.

[INSERT TABLE 6 HERE]

this restriction because otherwise, my sample is too small to reasonably analyze. The qualitative results in Section 4.2 still hold with the looser restrictions. (See Table 4.)

2.1% of analysts with low *NUM* coefficients who begin a year at a non-top ten brokerage advance to a top ten brokerage within the year, whereas 2.5% of analysts with high *NUM* coefficients advance. (See Panel A.) This difference is not statistically significant. Regarding exits, 11.1% (11.5%) of analysts with low (high) *NUM* coefficients who begin a year at a non-top ten brokerage exit the I/B/E/S dataset. This difference is also statistically insignificant.

Though there is scant evidence that *NUM* coefficients affect advancement probabilities, there is some evidence that analysts with higher *NUM* coefficients are less likely to be demoted. 3.7% of analysts with low *NUM* coefficients who begin the year at a top ten brokerage get demoted, whereas only 2.0% of analysts with high *NUM* coefficients get demoted. See Panel B. This difference, however, is not statistically significant, and the data on exit probabilities suggests the opposite: analysts with high *NUM* coefficients are more likely to exit (15.4%) than ones with low *NUM* coefficients (13.2%).[12]

Of course, lack of statistical significance does not prove that *NUM* coefficients and career advancements are independent of one another. It is possible there is simply not enough data to detect the relationship. However, in Panel C, I show that there is enough data to establish a significant relationship between an analyst's *Total Accuracy* and the likelihood of advancement: 3.3% of high *Total Accuracy* analysts have career advancements, whereas only 1.5% of low *Total Accuracy* analysts do. This difference is significant at the 5% level. Regarding exits, there is no significant difference between the likelihoods (11.1% versus 11.5%).

The results in Table 6 suggest that if analysts receive benefits from having positive *NUM* coefficients, it is that high *NUM* coefficients reduce the likelihood of being fired from a top ten brokerage. Hence, if career concerns explain my findings, the bias I document should be strongest among analysts who work for top ten brokerages.

To test this, I revisit the sample of 26,100 forecast revisions described in Section 4.1. I define the dummy *TopTen* to equal 1 if the analyst is employed by one of the ten largest brokerages (on the day the revised forecast is issued). *TopTen* $\times$ *NUM* is the interaction of *TopTen* and *NUM*. When these variables are added to my probit in Table 7, *TopTen* $\times$ *NUM* is significantly *negative*, suggesting that the bias is *least* severe for analysts working at a top ten brokerage.

[INSERT TABLE 7 HERE]

Interpreting coefficients of interaction terms in probits (and logits) can be problematic because the magnitude of the interaction effect does not equal the marginal effect of the interaction term (Ai and Norton (2003)). Since OLS does not suffer from this problem, In Column 2, I run the regression using OLS instead of a probit. The interaction is significantly negative, again suggesting that the bias is *least* severe for analysts working at a top ten brokerage.[13]

[12] Disappearing from I/B/E/S is not necessarily a negative career outcome; some of these analysts are promoted. See, e.g., Wu and Zang (2009).

[13] One problem with using OLS to model probabilities is that the predicted values need not lie in the unit interval. In this regression, however, this is not a concern since the fitted values all lie within the interval $[0.3, 0.8]$.

The results in Table 6 and Panel A of Table 7 do not support the idea that career incentives are driving the positive *NUM* coefficient I document. If the bias is a result of behavioral biases, the bias should be strongest among inexperienced analysts who have not had time to learn from their mistakes and correct their biases. To test this, I define the dummy *NewAnalyst* to equal 1 if the analyst is in the lowest *Total Experience* quintile, and the interaction *NewAnalyst* $\times$ *NUM* as the product of *NewAnalyst* and *NUM*. I include these variables in a probit and report the results in Panel B of Table 7. As predicted, the coefficient of the interaction is positive and significant, suggesting that the bias is most severe among analysts with the least experience. In the right column, I estimate the coefficient via OLS, and I again find the interaction to have a positive and significant coefficient.[14]

Summarizing, I find that

(i) there appears to little relationship between an analyst's *NUM* coefficient and his career outcomes,

(ii) the strongest empirical support for the career concerns hypothesis is that analysts at top ten brokerages are less likely to be demoted if they have a high *NUM* coefficient (though this is not statistically significant), but

(iii) analysts at top ten brokerages are *least* prone to the bias, and

(iv) the bias is most pronounced for analysts who lack forecasting experience.

Taken together, these findings suggest that behavioral biases, and not career concerns, are driving my results. However, it is worth recalling that the lack of statistical significance between analysts' *NUM* coefficients and subsequent career outcomes does not disprove the career concerns hypothesis; it is possible that the insignificance is due to the relatively small sample.

5.2 Overconfidence

In the previous section, I argued that my empirical results are likely due to behavioral biases rather than strategic forecasting. Overconfidence is a psychological bias that has received much attention in the economics literature. Is analyst overconfidence consistent with my empirical findings?

Recall the model: each analyst receives a signal, s_i, which is equal to the variable of interest, X, plus an error term, ε_i. The analysts' signal precision is τ, the error correlation is ρ, and analysts *believe* the signal precision and error correlation are $\hat{\tau}$ and $\hat{\rho}$, respectively. So far, I have focused on the case in which $\hat{\rho} > \rho$ and $\hat{\tau} = \tau$. To my knowledge, I am the first to examine the implications of this potential bias.

Overconfidence is generally modelled by assuming $\hat{\tau} > \tau$ and $\hat{\rho} = \rho = 0$. In the case of analysts, overconfidence clearly predicts that analysts would weight their own signals more strongly than they should. Empirically, Bernhardt, Campello, and Kutsoati (2006)) and Chen and Jiang

[14] In this regression, all of the fitted values lie in the interval $[0.22, 0.91]$.

(2006) document that analysts tend to issue forecasts too far from the consensus in the sense that they could lower their forecast error by issuing forecasts closer to the consensus. I examine whether my findings are consistent with analyst overconfidence by determining the model's predictions when $\hat{\tau} > \tau$. I follow the overconfidence literature by assuming $\hat{\rho} = \rho = 0$.

The following proposition will be used to produce a testable hypothesis for overconfidence.

Proposition 3. *If $\hat{\tau} > \tau$ and $\hat{\rho} = \rho = 0$, then the probability of the event*

$$\left\{ sgn(X - f_n^R) = sgn(f_n^R - f_n) \right\}$$

is increasing in NUM for $NUM < n - 1 + \frac{\hat{\tau}+\tau_x}{\tau}$, and decreasing in NUM for $NUM > n - 1 + \frac{\hat{\tau}+\tau_x}{\tau}$, where $sgn(\cdot)$ is the sign function which maps positive numbers to $+1$ and negative numbers to -1.

For low *NUM*, overconfidence and the false consensus effect make similar predictions: that $\mathbf{P}(\text{sgn}(X - f_n^R) = \text{sgn}(f_n^R - f_n))$ is increasing in *NUM*. However, whereas the probability is monotonic in *NUM* when analysts overestimate the correlation of their signal errors, when they are overconfident, it rises up to a point and then decreases. As $NUM \to \infty$, analysts place significantly more weight on the other analysts' signals (which are correctly weighted) than their own (which is overweighted), and the underreaction probability converges to 0.5.[15] To get a sense of where the maximum probability occurs, consider the case where $\hat{\tau} = 2\tau$, $\tau_x = \tau$, $\rho = 0$, and $n = 1$. In this case, the *NUM* that maximizes the underreaction probability is 3. Figure 5 depicts $\mathbf{P}(\text{sgn}(X - f_n^R) = \text{sgn}(f_n^R - f_n))$ as a function of *NUM* when analysts overestimate the precision of their signals.

[INSERT FIGURE 5 HERE]

To test this hypothesis, I run the probits from Column 6 of Table 3, but I restrict the sample to the first analyst to issue a forecast following an earnings announcement. Moreover, I restrict the sample to forecast revisions with $NUM > k$ for various k. If overconfidence is driving the relationship between *NUM* and ex post underreaction, then the coefficient of *NUM* should decrease as *NUM* is increased, and for larger values of k (certainly for $k \geq \frac{\hat{\tau}+\tau_x}{\tau}$), the coefficient should be negative.

The evidence does not support this prediction. When the sample is not restricted based on *NUM* (Column 1), the coefficient on *NUM* is 0.012. As k rises, the coefficient of *NUM* tends to rise as well, especially for relatively large k. When $k = 5$, the coefficient of *NUM* is 0.029. For $k = 6$, the coefficient is 0.048, which is significant at the 5% level. As k rises, the coefficient of *NUM* continues to rise, but the statistical significance declines as the sample size declines.

[INSERT TABLE 8 HERE]

[15] This can be seen by inspecting equations (83) and (84) in Appendix C and noting that as $n \to \infty$, $\tau_B \to \infty$.

Another way to test the predictions of a false consensus versus the predictions of overconfidence is to plot the predicted probabilities of ex post underreaction as a function of *NUM* when all of the control variables are set to their means. If the plot resembles Figure 3 but not Figure 5, that is strong evidence for a false consensus. If the plot resembles Figure 5 but not Figure 3, that is strong evidence for overconfidence.

I test this by running the regression from Column 6 of Table 3, with the only change being that I replace the variable *NUM* with dummy variables for each of the values of *NUM*. I then compute the probit's predicted probability of underreaction as a function of *NUM*, when all of the control variables are set to their means, and the appropriate *NUM* dummy is set to 1 (all the other *NUM* dummies are set to 0).

[INSERT FIGURE 6 HERE]

Clearly, Figure 6 more closely resembles Figure 3 than Figure 5, suggesting that a false consensus, and not overconfidence, is driving the relationship between analyst underreaction and *NUM*.

6 Market Reaction to Forecast Revisions

An analyst who overestimates the correlation of his signal error with others' will tend to insufficiently revise his forecasts (Section 3). A natural question is whether the market's reaction to an analyst's forecast revision is related to the analyst's tendency to exhibit the bias.[16] Consider two hypothetical analysts: analyst i, who is subject to the bias, and analyst j, who is not. If the market is aware of this, then (ceteris paribus) i's revisions should be associated with *larger* price responses, since i's bias causes him to underreact when revising his earnings forecasts.

To test this, I take the 3,558,388 distinct (analyst, firm, forecast date, quarter end date) 4-tuples from the unadjusted I/B/E/S file satisfying the following requirements: the forecast period indicator is equal to 6 or 7 (i.e., short term quarterly earnings forecasts), the split adjustment factor is non-missing, and the company can be linked to CRSP and the earnings actuals file.

I restrict this sample to the most recent forecast revision issued by each analyst at least 30 days before the forecast period end date. Since Clement and Tse (2003) document that an analyst's prior accuracy in forecasting a firm's earnings is related to the market's future response to his earnings forecasts for that firm, I use the variable *Firm Error* (described in Table 9) as a control variable in my analysis, and I require that it be defined for each forecast in my sample. Applying these filters results in a sample of 680,415 forecast revisions.

There is an important caveat when analyzing returns around forecast revisions. It is debatable whether the market responds to financial analysts' earnings forecasts and recommendations. Altinkilic and Hansen (2009) examine intraday price movements in the minutes surrounding analyst

[16] I thank an anonymous referee for this suggestion.

recommendations and find little evidence that the market significantly responds to the recommendations. They also document that recommendations are often issued following important news events about the firm such as earnings announcements, and they argue that much of the price movements attributed to recommendations are actually driven by the news event preceeding the recommendations.

To reduce the likelihood that I am capturing returns to non-analyst related news events, I restrict my sample so that the revised forecast is at least three trading days from an earnings announcement date. This ensures that the three day window surrounding the forecast revision does not overlap with the three day window surrounding an earnings announcement. Though this obviously does not ensure that the market is responding to the analyst's forecast revision, it mitigates the concern. This reduces my sample to 245,885 revisions. 211,780 of these forecast revisions are for firms that can be matched to size and book to market portfolios and whose stock price is at least \$5 per share two days prior to the revised forecast.

Clement and Tse (2003) analyze returns around forecast revisions and document that returns are related to the following factors: the amount of time that has elapsed since the previous forecast for the firm's earnings, the forecast horizon, the analyst's prior accuracy, and the size of the analyst's brokerage. Motivated by their findings, I define the following control variables: *Raw Elapsed Time* is the number of days between the forecast and the next most recent forecast for the firm's earnings, *Raw Horizon* is the number of days between the date the forecast is issued and the forecast period end date, *Raw Firm Error* is the average price-scaled forecast error among i's forecasts for the firm's earnings issued between one and eight quarters before the quarter the revision is issued (where the forecasts have *Raw Horizon* between 31 and 90 days), and *Raw Brokerage Size* is the number of analysts employed by the analyst's brokerage in the year the forecast revision is issued. I follow Clement and Tse (2003) by normalizing these four raw control variables as follows:

$$\text{Control}_{ijt} = \frac{\text{Raw Control}_{ijt} - \text{Raw Control}_{\min_{jt}}}{\text{Raw Control}_{\max_{jt}} - \text{Raw Control}_{\min_{jt}}}, \tag{27}$$

where Raw $\text{Control}_{\min_{jt}}$ (Raw $\text{Control}_{\max_{jt}}$) is the minimum (maximum) of the raw control among all analysts issuing forecasts for the firm that quarter. Clearly, these controls are only defined if Raw $\text{Control}_{\max_{jt}} \neq$ Raw $\text{Control}_{\min_{jt}}$, which can only be satisfied if at least two analysts are covering the firm. An advantage of this scaling methodology is that it removes any time or firm effects (e.g., size, book-to-market, coverage, etc.) from the variables—in each firm quarter, all of the variables take a minimum value of 0 and a maximum value of 1.

To distinguish between analysts who exhibit a false consensus from those that do not, I estimate *NUM* coefficients at the analyst-year level following the methodology described in Section 5.1.2. I define the variable *FCE* by applying the scaling methodology described in (27) to analysts' *NUM* coefficients. This ensures that my measure of analysts' prior FCE tendencies is not correlated with firm size or analyst coverage. I require that the four control variables and *FCE* be defined. After applying this filter, my sample consists of 15,529 forecast revisions.

For each of these forecast revisions, I define the cumulative abnormal return (CAR) as the sum of the abnormal returns in the three day window centered around the revision date, where a stock's abnormal return is the difference between the stock's return and its size/book-to-market matched portfolio. To test whether the market reacts more to forecast revisions issued by analysts exhibiting the bias, I regress the three day CARs around the forecast revision onto *REVP*, which is defined as the (price-scaled) difference between the analyst's revised forecast and his earlier forecast, and the interaction $REVP \times FCE$. I present the results of this regression in Column 1 of Table 9. In Column 2, I include the interactions that Clement and Tse (2003) document are associated with CARs.[17] The coefficient of $REVP \times FCE$ is neither positive nor significantly significant in either regression, suggesting that for a given level of *REVP*, the market does not respond more to analysts with high *NUM* coefficients.

[INSERT TABLE 9 HERE]

One possible explanation for these findings is that there are few significant cross-sectional differences in analysts' propensity to exhibit a false consensus. In other words, differences in analysts' estimated *NUM* coefficients are due to random sampling and have no predictive power in predicting each analyst's future forecast errors. Another possibility is that analysts with high *NUM* coefficients are not more likely to underreact unconditionally—an analyst's *NUM* coefficient captures the relationship between the analyst's likelihood of underreacting (ex post) and the number of analysts issuing forecasts between his forecasts, *not* the likelihood of the analyst underreacting unconditionally. For example, an analyst who always overreacts (i.e., never underreacts ex post) when *NUM* is low and issues rational forecasts (i.e., underreacts ex post 50% of the time) when *NUM* is high would have a positive *NUM* coefficient.

I rule out these possibilities in Columns 3 and 4 of Table 9, where I regress price-scaled forecast errors (of the revised forecast) onto the independent variables in Columns 1 and 2. Analysts' *NUM* coefficients (based on their forecasts in the five years prior to the year of the revised forecast) are positively associated with their forecast errors. This suggests that there is heterogeneity in analysts' propensities to exhibit a false consensus, and that analysts who exhibit the bias are more likely to underreact unconditionally in their forecast revisions.

It is unclear why the market does not react differently based on analysts' *NUM* coefficients. One possibility is that traders, like analysts, overestimate the public component of agents' (including financial analysts and traders) error terms. This could explain why the market seems to underreact to analysts' forecast revisions (e.g., Gleason and Lee (2003)). Another possibility is that the market pays less attention to the earnings forecasts of analysts who most exhibit psychological biases.

[17] The insignificance of these control variables does not contradict Clement and Tse (2003) because our samples are quite different—I restrict my sample to forecasts not issued near earnings announcements and issued by analysts with estimated *NUM* coefficients.

7 Conclusion

For financial analysts to efficiently aggregate their information, it is important that they have unbiased perceptions of their similarity to other analysts. If they overestimate their similarity to other analysts, as the social psychology literature suggests, the likelihood of ex post underreaction in forecast revisions will increase in the number of other analysts that issue forecasts between the time of the analyst's earlier forecast and the time of his revised forecast. Empirically, I confirm this prediction using earnings forecasts issued by sell-side analysts.

There are several avenues for future research. Trueman (1990) and Trueman (1994) show that some properties of analysts' forecast errors that appear to be due to behavioral biases are consistent with analysts shading their forecasts away from their true beliefs due to career concerns. I considered this possibility in Section 5.1, but my theoretical analysis was restricted to a particular type of objective function, and my empirical analysis was somewhat inconclusive due to the small sample. Future researchers may develop better ways to test my behavioral explanation versus alternative explanations consistent with analyst rationality. Also, overestimation of signal error correlations undoubtedly has implications in other areas of economics, accounting, and finance. Testing my differential information model with a false consensus to other environments is another avenue for future research.

A Variable Definitions

Note: for each of the variables below, I divide the stocks/analysts into quintiles each quarter based on the values of the variables. I create dummies for each of the quintiles. When controlling for analyst type and firm type, I include the quintile dummies rather than the underlying variable because there is no reason to expect a specific functional relationship between the variables and underreaction probabilities. My results are not sensitive to this technique.

- *Size* is defined as the company's market cap in the June preceding the forecast revision. *Size Quintile* is defined by assigning stocks to a quintile based on the NYSE cutoff values that quarter.

- *Book-to-market* is defined as $\frac{\text{BE}}{\text{market}}$, where BE is the book equity at the end of the latest fiscal year ending before the latest June preceding the forecast date, and market is defined as the size of the company in the December preceding the latest June preceding the date of the forecast. More formally, book equity is defined as shareholder equity minus preferred stock plus investment tax credit (TXDITC) minus post-retirement benefit assets (PRBA). If total stockholder equity (SEQ) is non-missing, I set shareholder equity to equal it. Otherwise, if total common/ordinary equity (CEQ) and total preferred stock (PSTK) are non-missing, I set shareholder equity to be the sum of these variables. Otherwise, I define shareholder equity as total assets (AT) minus total liabilities (LT) minus minority interest (MIB). I define preferred stock to be the redemption value of preferred stock (PSTKRV) if this is non-missing. Otherwise, I define it to be the liquidating value of preferred stock (PSTKL). If both PSTKRV and PSTKL are missing, I define preferred stock to be total preferred stock (PSTK).

- For each forecast f, I defined the analyst's *Experience* as the number of firm-quarters the analyst has forecasted prior to f, where the sample of firm-quarters is restricted to ones in which the firm reported its earnings prior to f. In other words, an analyst does not have positive *Experience* until after (i) he issues an earnings forecast for a firm-quarter and (ii) the firm announces its earnings for that quarter. *Experience Quintile* is defined by ranking the analysts on their *Experience* each quarter, where the cutoff values are based on all analysts who issued a forecast in the quarter.

- *Firm Experience* is defined analogously to *Experience*, except that the analyst's previous forecasts are restricted to the firm being forecasted. *Firm Experience Quintile* is defined by ranking the analysts on their *Firm Experience* each quarter, where the cutoff values are based on all analysts who issued a forecast in the quarter.

- My (inverse) measure of an analyst's *Total Accuracy* in a given quarter is the average value of $\frac{|f-\text{EPS}|}{P}$ for all the forecasts the analyst issued between one and eight quarters before the quarter for which the measure is defined, where f is analyst's forecast, EPS is the actual earnings, and P is the company's price on the earnings announcement date. The forecasts f

used to define *Total Accuracy* are restricted to the analyst's forecasts with horizons between 31 and 90 days, where horizon is the number of days between the forecast and the quarter end date. Analysts are assigned to quintiles based on *Total Accuracy* by ranking them, each quarter, on their *Total Accuracy*. For my regressions, I create dummies for each quintile. Since analysts that have not issued forecasts in the previous eight quarters (e.g., those with no *Experience*) have undefined *Total Accuracy*, I create an additional dummy for *Total Accuracy* being undefined.

- *Firm Accuracy* is defined analogously to *Total Accuracy*, except the forecasts are restricted to the firm being forecasted.
- A firm's *Difficulty* in a given quarter is defined as the average $\frac{|f-\text{EPS}|}{P}$ for all forecasts issued for the firm between one and eight quarters before the quarter being analyzed, where f is a forecast, EPS is the actual earnings, and P is the company's price on the earnings announcement date. The forecasts used to define *Difficulty* are restricted to have horizons between 31 and 90 days, where horizon is the number of days between the forecast and the quarter end date. Each quarter, firms are ranked on their *Difficulty* and are assigned to quintiles.
- A company's *Coverage* in a given quarter is defined as the number of distinct analysts who issued a forecast for the company's earnings (at any horizon) during the quarter. Firms are ranked on their *Coverage* each quarter and are assigned to quintiles.

B Rounding Assumptions

Letting f^0 denote the analyst's initial reported (and rounded) forecast, $f^{*,0}$ denote the analyst's unrounded initial forecast (which is unreported and hence unobservable—it can be thought of as the forecast the analyst would have issued if he were not forced to round), f^R denote the analyst's revised (and rounded) forecast, $f^{*,R}$ denote the analyst's unreported, unrounded revised forecast (which is unobserved), $\text{EPS}^{\text{reported}}$ denote the firm's reported (rounded) earnings, and EPS^* the firm's "true" (unrounded) earnings per share (which is unobserved), and Z the control variables, I assume:

$$\mathbf{P}\big(\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,0}) \mid \{\text{EPS}^{\text{report}} = f^R\} \cup \{f^R = f^0\}, Z\big) \tag{28}$$

$$=$$

$$\mathbf{P}\big(\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,R}) \mid \{\text{EPS}^{\text{report}} \neq f^R\} \cap \{f^R \neq f^0\}, Z\big) \tag{29}$$

$$=$$

$$\mathbf{P}\big(\text{sgn}(\text{EPS}^{\text{report}} - f^R) = \text{sgn}(f^R - f^0) \mid \{\text{EPS}^{\text{report}} \neq f^R\} \cap \{f^R \neq f^0\}, Z\big) \tag{30}$$

Note that the event $\{\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,0})\}$ in expression (28) is what the theory makes predictions about—namely that the probability of this event occurring is increasing in the number of forecasts issued by other analysts between the analyst's earlier forecast and his revised forecast. The conditioning event in (28)—the event $\{\text{EPS}^{\text{report}} = f^R\} \cup \{f^R = f^0\}$—is simply the event that I am unable to observe the event $\{\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,0})\}$ due to rounding.

The conditioning event in expression (29), $\{\text{EPS}^{\text{report}} \neq f^R\} \cap \{f^R \neq f^0\}$, corresponds to the event that I am able to observe the event of interest, namely

$$\{\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,0})\}.$$

Finally, the event $\{\text{sgn}(\text{EPS}^{\text{report}} - f^R) = \text{sgn}(f^R - f^0)\}$ in expression (30) is the event mirroring $\{\text{sgn}(\text{EPS}^* - f^{*,R}) = \text{sgn}(f^{*,R} - f^{*,0})\}$, the difference being that the terms in (30), $\text{EPS}^{\text{report}}$, f^R, and f^0 are rounded and observable, whereas the terms in (28), EPS^*, $f^{*,R}$, and $f^{*,0}$ are not.

The first equality says the probability of an analyst's unreported unrounded revised forecast being too close (ex post) to his unreported unrounded initial forecast, conditional on the control variables and the event that the revised forecast was equal to either the earlier forecast or the reported earnings, is equal to the probability of an analyst's unreported unrounded revised forecast being too close (ex post) to his unreported unrounded initial forecast, conditional on the control variables and the event that the revised forecast does not equal the earlier forecast or the reported earnings. The last equality holds as long as analysts' forecasts are their subjective means, rounded to the nearest penny. Together, these assumptions are equivalent to the assumption that the probability of an analyst underreacting in his forecast revision, conditional on me being able to observe whether he underreacts, is equal to the probability of an analyst underreacting in his forecast revision, conditional on me not being able to observe whether he underreacts.

C Proofs

Proof. (**Proposition 1**)

Let k denote the number of other analysts issuing forecasts between analyst n's initial forecast and his revised forecast. Since f_{n+k} and f_n^R both equal $\hat{\mathbf{E}}[X|s_1, \ldots, s_{n+k}]$, it trivially follows that in my model, $f_n^R = f_{n+k}$.[18] Hence, to prove the theorem, it suffices to show that for all n,

$$\mathbf{P}(\text{sgn}(X - f_{n+k}) = \text{sgn}(f_{n+k} - f_n)) \tag{31}$$

is increasing in k, where $\mathbf{P}(\cdot)$ is the probability measure. To show that $\mathbf{P}(\text{sgn}(X - f_{n+k}) = \text{sgn}(f_{n+k} - f_n))$ is increasing in k, it suffices to show that $\text{corr}(X - f_{n+k}, f_{n+k} - f_n)$ increases in k. (See Dharmadhikari and Joag-Dev (1984).)

Recalling (2), (3), and (11), it follows that the average signal, $\overline{s_n}$, received by the first n analysts is given by the equation

$$\overline{s_n} \;=\; X + \left(\sqrt{\frac{\rho}{\tau}}\right)\eta + \left(\sqrt{\frac{1-\rho}{\tau}}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right). \tag{32}$$

Recalling (2), (4), (11), and my assumption that $\hat{\tau} = \tau$, analysts' beliefs about $\overline{s_n}$ are given by the equation

$$\overline{s_n} \;=^i\; X + \left(\sqrt{\frac{\hat{\rho}}{\tau}}\right)\eta + \left(\sqrt{\frac{1-\hat{\rho}}{\tau}}\right)\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right). \tag{33}$$

For notational convenience, for each n, let D_n, $\overline{D}_n$, $\hat{\tau}_{\overline{s_n}}$, and $\hat{\omega}_n$ be defined by

$$D_n \;=\; \tau_x[(n-1)\hat{\rho} + 1] + n\tau \tag{34}$$

$$\overline{D}_n \;=\; \tau_x[(n-1)\rho + 1] + n\tau \tag{35}$$

$$\hat{\tau}_{\overline{s_n}} \;=\; \frac{n\tau}{(n-1)\hat{\rho} + 1} \tag{36}$$

$$\hat{\omega}_n \;=\; \frac{n\tau}{D_n} \tag{37}$$

It is easily verified that agents believe the aggregate signal, $\overline{s_n}$, has precision $\hat{\tau}_{\overline{s_n}}$, and that $\hat{\omega}_n = \frac{\hat{\tau}_{\overline{s_n}}}{\tau_x + \hat{\tau}_{\overline{s_n}}}$. Hence, by (15), $f_n = (1 - \hat{\omega}_n)\mu + \hat{\omega}_n\overline{s_n}$. Without loss of generality, let $\mu = 0$, so that f_n can be expressed as

$$f_n \;=\; \hat{\omega}_n\overline{s_n}. \tag{38}$$

[18] In practice, this prediction is clearly false. I interpret differences between the two forecasts to be attributed to differences in analyst n's and analyst $n+k$'s priors rather than differences in the weightings of the signals the analysts have received. My testable prediction only requires that analysts be able to back-out the other analysts' signals; the analysts need not actually have the same priors. Only the forecasts f_n and f_n^R need be based on the same prior, and since these forecasts are issued by the same analyst, this assumption is innocuous.

Recalling (1), (3), (32), (35), (37), and (38),

$$\begin{aligned}
\mathrm{Cov}(X, \overline{s_{n+k}}) &= \frac{1}{\tau_x} && (39)\\
\mathrm{Var}(\overline{s_{n+k}}) &= \frac{1}{\tau_x} + \frac{\rho}{\tau} + \frac{1-\rho}{(n+k)\tau} \\
&= \frac{\overline{D}_{n+k}}{(n+k)\tau\tau_x} && (40)\\
\mathrm{Cov}(\overline{s_n}, \overline{s_{n+k}}) &= \frac{1}{\tau_x} + \frac{\rho}{\tau} + \frac{1-\rho}{(n+k)\tau} \\
&= \frac{\overline{D}_{n+k}}{(n+k)\tau\tau_x} && (41)\\
X - f_{n+k} &= X - \hat{\omega}_{n+k}\overline{s_{n+k}} && (42)\\
f_{n+k} - f_n &= \hat{\omega}_{n+k}\overline{s_{n+k}} - \hat{\omega}_n\overline{s_n} && (43)\\
\hat{\omega}_{n+k} - \hat{\omega}_n &= \frac{(n+k)\tau}{D_{n+k}} - \frac{n\tau}{D_n} \\
&= \frac{k\tau\tau_x(1-\hat{\rho})}{D_n D_{n+k}}. && (44)
\end{aligned}$$

Hence,

$$\mathrm{Cov}(X - f_{n+k}, f_{n+k} - f_n) = \frac{k\tau\tau_x(n+k-1)(1-\hat{\rho})(\hat{\rho}-\rho)}{D_n D_{n+k}^2} \quad (45)$$

Recalling (1) and (37)-(44),

$$\mathrm{Var}(X - f_{n+k}) = \frac{D_{n+k}^2 - (n+k)\tau(2D_{n+k} - \overline{D}_{n+k})}{\tau_x D_{n+k}^2} \quad (46)$$

Recalling (37)-(44),

$$\mathrm{Var}(f_{n+k} - f_n) = \left(\frac{\tau}{\tau_x D_n^2 D_{n+k}^2}\right)\left[(n+k)\overline{D}_{n+k}D_n^2 - 2nD_nD_{n+k}\overline{D}_{n+k} + n\overline{D}_n D_{n+k}^2\right] \quad (47)$$

Let ϱ denote $\mathrm{corr}(X - f_{n+k}, f_{n+k} - f_n)$, and let δ_1 and δ_2 be given by the equations:

$$\begin{aligned}
\delta_1 &= (n+k)\overline{D}_{n+k}D_n^2 - 2nD_nD_{n+k}\overline{D}_{n+k} + n\overline{D}_n D_{n+k}^2 && (48)\\
\delta_2 &= D_{n+k}^2 - (n+k)\tau(2D_{n+k} - \overline{D}_{n+k}) && (49)
\end{aligned}$$

Combining (45), (46), (47), (48), and (49)

$$\begin{aligned}
\varrho &= \frac{\mathrm{Cov}(X - f_{n+k}, f_{n+k} - f_n)}{\sqrt{\mathrm{Var}(X - f_{n+k}) * \mathrm{Var}(f_{n+k} - f_n)}} \\
&= \frac{k\tau_x^2\sqrt{\tau}(n+k-1)(1-\hat{\rho})(\hat{\rho}-\rho)}{\sqrt{\delta_1\delta_2}}. && (50)
\end{aligned}$$

Let Δ_1 and Δ_2 be defined as

$$\Delta_1 = \tau(n+k)(\rho[n+k-1]+1) + \tau_x(\hat{\rho}[n+k-1]+1)^2 \tag{51}$$

$$\Delta_2 = k\Big((1-\hat{\rho})^2\tau_x(\tau+\rho\tau_x) + n(1-\rho)(\tau+\hat{\rho}\tau_x)^2\Big) + (1-\rho)\Big(n(\tau+\hat{\rho}\tau_x) - \hat{\rho}\tau_x + \tau_x\Big)^2 \tag{52}$$

From (48), (49), (51), and (52),

$$\delta_1\delta_2 = k\tau_x^2\Delta_1\Delta_2. \tag{53}$$

(Note: I have omitted steps because the proof is very cumbersome. It can easily be verified using computer software.)

Combining (50) and (53),

$$\varrho = \tau_x\sqrt{\tau}(1-\hat{\rho})(\hat{\rho}-\rho)\left(\frac{\sqrt{k}(n+k-1)}{\sqrt{\Delta_1\Delta_2}}\right). \tag{54}$$

To complete the proof, it suffices to show that $\frac{\partial\varrho}{\partial k} > 0$. For notational simplicity, I let Δ_1' and Δ_2' denote $\frac{\partial\Delta_1}{\partial k}$ and $\frac{\partial\Delta_2}{\partial k}$, respectively. Taking the partial of ϱ with respect to k,

$$\frac{\partial\varrho}{\partial k} = \tau_x\sqrt{\tau}(1-\hat{\rho})(\hat{\rho}-\rho)\left(\frac{\sqrt{\Delta_1\Delta_2}\left(\sqrt{k}+\frac{n+k-1}{2\sqrt{k}}\right) - \frac{\sqrt{k}(n+k-1)(\Delta_1\Delta_2'+\Delta_1'\Delta_2)}{2\sqrt{\Delta_1\Delta_2}}}{\Delta_1\Delta_2}\right) \tag{55}$$

so

$$\left(\frac{2\sqrt{k}(\Delta_1\Delta_2)^{\frac{3}{2}}}{\tau_x\sqrt{\tau}(1-\hat{\rho})(\hat{\rho}-\rho)}\right)\frac{\partial\varrho}{\partial k} = \Delta_1\Delta_2(3k+n-1) - k(n+k-1)(\Delta_1\Delta_2'+\Delta_1'\Delta_2). \tag{56}$$

Since $\rho\in[0,1)$, $\hat{\rho}\in(\rho,1)$, $\tau>0$, $\tau_x>0$, $k>0$, $\Delta_1>0$, and $\Delta_2>0$, the equation above implies

$$\operatorname{sgn}\left(\frac{\partial\varrho}{\partial k}\right) = \operatorname{sgn}\Big(\Delta_1\Delta_2(3k+n-1) - k(n+k-1)(\Delta_1\Delta_2'+\Delta_1'\Delta_2)\Big). \tag{57}$$

Differentiating (51) and (52),

$$\Delta_1' = \tau(2[n+k]\rho+[1-\rho]) + 2\tau_x\hat{\rho}([n+k]\hat{\rho}+[1-\hat{\rho}]) \tag{58}$$

$$\Delta_2' = (1-\hat{\rho})^2\tau_x(\tau+\rho\tau_x) + n(1-\rho)(\tau+\hat{\rho}\tau_x)^2 \tag{59}$$

Plugging (58) and (59) into (57), it can be shown that

$$\operatorname{sgn}\left(\frac{\partial\varrho}{\partial k}\right) = \operatorname{sgn}\Big(P\Big[k^2(Q+n[1-\rho]R+[1-\hat{\rho}]^2\tau_xS) + k(1-\rho)TU + (n-1)(1-\rho)VW\Big]\Big),$$

where P-W are given by the equations

$$\begin{aligned}
P &= (k+n)(\hat{\rho}\tau_x+\tau)-\hat{\rho}\tau_x+\tau_x \\
Q &= n^2(1-\rho)(\hat{\rho}\tau_x+\tau)(\rho\tau+\hat{\rho}^2\tau_x) \\
R &= \tau\tau_x(-\rho\hat{\rho}+2\rho+3\hat{\rho})+(\rho+1)\tau^2+2(2-\hat{\rho})\hat{\rho}^2\tau_x^2 \\
S &= \tau_x(-\rho\hat{\rho}+2\rho+\hat{\rho})+(\rho+1)\tau \\
T &= \tau_x(2[n-1]\hat{\rho}+3)+(2n+1)\tau \\
U &= n\tau([n-1]\rho+1)+\tau_x([n-1]\hat{\rho}+1)^2 \\
V &= n(\hat{\rho}\tau_x+\tau)-\hat{\rho}\tau_x+\tau_x \\
W &= n\tau([n-1]\rho+1)+\tau_x([n-1]\hat{\rho}+1)^2.
\end{aligned}$$

It is easily verified that each of P-W are all positive, so ϱ is increasing in k, completing the proof. □

Proof. (**Proposition 2**)

First note that (24) is equivalent to $E_1(\alpha_n)=E_2(\alpha_n)$, where

$$\begin{aligned}
E_1(\alpha_n) &= \alpha_n^2\left(\frac{D_n}{(n-1)\hat{\rho}+1}-\frac{\tau(1-\hat{\rho})}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right)+\alpha_n\left(\frac{2\tau}{n\hat{\rho}+1}\right)+\left(2\theta-\frac{\tau[(n-1)\hat{\rho}+1]}{(n\hat{\rho}+1)(1-\hat{\rho})}\right) \\
E_2(\alpha_n) &= \alpha_n\left(\frac{\tau}{(n-1)\hat{\rho}+1}+\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right)+\left(\frac{2\theta\tau}{D_n}-\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)(1-\hat{\rho})}\right).
\end{aligned}$$

It can easily be verified that $E_1(0)<E_2(0)$ if and only if $\theta<\theta_n^+$. It can also be easily verified that for all $\alpha_n>0$,

$$\frac{\partial E_1(\alpha_n)}{\partial\alpha_n}>\frac{\partial E_2(\alpha_n)}{\partial\alpha_n},$$

from which it trivially follows that if (23) is satisfied, then (24) has a unique positive solution, which I denote α_n^*.

Consider symmetric forecasting strategies that are linear combinations of the available public information, $\hat{\mathbf{E}}[X|s_1,\ldots,s_{n-1}]$, and s_n:

$$f_n = (1-\alpha_n)\hat{\mathbf{E}}[X|\overline{s_{n-1}}]+\alpha_n s_n. \tag{60}$$

When solving for the equilibrium weights, I can assume without loss of generality that $\mu=0$ and

$\overline{s_{n-1}} = 0$.[19] Given these assumptions, (60) simplifies to

$$f_n = \alpha_n s_n. \tag{61}$$

The equilibrium forecast, f_n^*, satisfies

$$\begin{aligned} f_n^* &= \arg\max_{f_n} \frac{\ln U_n}{1-\delta} \\ &= \arg\max_{f_n} -\theta \hat{\mathbf{E}}\left\{[X - f_n^{-1}(f_n)]^2 | \overline{s_n}\right\} \\ &\quad + \left(-\frac{\hat{\tau}(X|\overline{s_n})}{2}(f_n - \mathbf{E}[X|\overline{s_n}])^2 + \frac{\hat{\tau}(f|X=f_n,\overline{s_n})}{2}(f_n - \hat{\mathbf{E}}[f|X=f_n,\overline{s_n}])^2\right), \end{aligned} \tag{62}$$

where $f_n^{-1}(\cdot)$ maps forecasts to signals according to (61): $f_n^{-1}(f) = \frac{f}{\alpha_n}$. The first order condition for (62) is

$$\begin{aligned} 0 &= \left(\frac{-\theta 2}{-\alpha_n}\right) \hat{\mathbf{E}}\left\{\left[x - \frac{f_n}{\alpha_n}\right] | \overline{s_n}\right\} - \hat{\tau}(x|\overline{s_n})(f_n - \hat{\mathbf{E}}[x|\overline{s_n}]) \\ &\quad + \hat{\tau}(f|x=f_n,\overline{s_n})(f_n - \hat{\mathbf{E}}[f|x=f_n,\overline{s_n}])\left(1 - \frac{\partial \hat{\mathbf{E}}[f|x=f_n,\overline{s_n}]}{\partial f_n}\right). \end{aligned} \tag{63}$$

The following equations can easily be verified:

$$\hat{\mathbf{E}}[X|\overline{s_n}] = \frac{n\tau\overline{s_n}}{D_n} \tag{64}$$

$$\hat{\tau}(X|\overline{s_n}) = \frac{D_n}{(n-1)\hat{\rho}+1} \tag{65}$$

$$\hat{\tau}(f|X=f_n,\overline{s_n}) = \frac{\tau[(n-1)\hat{\rho}+1]}{\alpha_n^2(n\hat{\rho}+1)(1-\hat{\rho})}. \tag{66}$$

Lemma 1. *If analysts issue forecasts according to (61), then*

$$\hat{\mathbf{E}}[f|X=f_n,\overline{s_n}] = \frac{\alpha_n[(1-\hat{\rho})f_n + n\hat{\rho}\overline{s_n}]}{(n-1)\hat{\rho}+1}. \tag{67}$$

[19] At the time analyst n issues his forecast, $\overline{s_{n-1}}$ has already been impounded into both the analyst's and the market's expectations. Analyst n does not care about the market's beliefs about how far his signal (s_n) is from others' signals; all he cares about are the market's beliefs about how far s_n was from X (captured by the function u_r) and the expected contest-like payout when competing against other analysts issuing with the same public information set as him (captured by the function u_c). (Recall that n is not competing against analysts 1 through $n-1$, but rather against analysts issuing forecasts simultaneous to him.)

When I numerically analyze the properties of the forecast errors later in this section, I *do not* assume that $\hat{\mathbf{E}}[x|\overline{s_{n-1}} = 0]$. This assumption is only invoked to simplify the solution for finding the equilibrium weighting on the analysts' private signals.

It trivially follows from the lemma that

$$\frac{\partial \hat{\mathbf{E}}[f|x=f_n,\overline{s_n}]}{\partial f_n} = \frac{\alpha_n(1-\hat{\rho})}{(n-1)\hat{\rho}+1}. \tag{68}$$

Combining (64)-(68), (63) becomes:

$$\begin{aligned} 0 &= \frac{2\theta}{\alpha_n}\left\{\frac{n\tau\overline{s_n}}{D_n}-\frac{f_n}{\alpha_n}\right\}-\left(\frac{D_n}{(n-1)\hat{\rho}+1}\right)\left(f_n-\frac{n\tau\overline{s_n}}{D_n}\right) \\ &\quad +\frac{\tau[(n-1)\hat{\rho}+1]}{\alpha_n^2(n\hat{\rho}+1)(1-\hat{\rho})}\left(f_n-\frac{\alpha_n[(1-\hat{\rho})f_n+n\hat{\rho}\overline{s_n}]}{(n-1)\hat{\rho}+1}\right)\left(1-\frac{\alpha_n(1-\hat{\rho})}{(n-1)\hat{\rho}+1}\right). \end{aligned} \tag{69}$$

Recalling $\overline{s_{n-1}}=0$, it follows that $n\overline{s_n}=s_n$, and the equation above becomes

$$\begin{aligned} 0 &= \frac{2\theta}{\alpha_n}\left\{\frac{\tau s_n}{D_n}-\frac{f_n}{\alpha_n}\right\}-\left(\frac{D_n}{(n-1)\hat{\rho}+1}\right)\left(f_n-\frac{\tau s_n}{D_n}\right) \\ &\quad +\frac{\tau[(n-1)\hat{\rho}+1]}{\alpha_n^2(n\hat{\rho}+1)(1-\hat{\rho})}\left(f_n-\frac{\alpha_n[(1-\hat{\rho})f_n+\hat{\rho}s_n]}{(n-1)\hat{\rho}+1}\right)\left(1-\frac{\alpha_n(1-\hat{\rho})}{(n-1)\hat{\rho}+1}\right). \end{aligned} \tag{70}$$

This implies

$$\begin{aligned} &f_n\left[\frac{2\theta}{\alpha_n^2}+\frac{D_n}{(n-1)\hat{\rho}+1}-\frac{\tau[(n-1)\hat{\rho}+1]}{\alpha_n^2(n\hat{\rho}+1)(1-\hat{\rho})}\left(1-\frac{\alpha_n(1-\hat{\rho})}{(n-1)\hat{\rho}+1}\right)^2\right]= \\ &s_n\left[\frac{2\theta\tau}{\alpha_n D_n}+\frac{\tau}{(n-1)\hat{\rho}+1}-\frac{\tau[(n-1)\hat{\rho}+1]}{\alpha_n^2(n\hat{\rho}+1)(1-\hat{\rho})}\left(1-\frac{\alpha_n(1-\hat{\rho})}{(n-1)\hat{\rho}+1}\right)\left(\frac{\alpha_n\hat{\rho}}{(n-1)\hat{\rho}+1}\right)\right]. \end{aligned} \tag{71}$$

Recalling $f_n=\alpha s_n$,

$$\begin{aligned} &\alpha_n^2\left(\frac{D_n}{(n-1)\hat{\rho}+1}-\frac{\tau(1-\hat{\rho})}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right)+\alpha_n\left(\frac{2\tau}{n\hat{\rho}+1}\right)+\left(2\theta-\frac{\tau[(n-1)\hat{\rho}+1]}{(n\hat{\rho}+1)(1-\hat{\rho})}\right) \\ &\qquad = \\ &\alpha_n\left(\frac{\tau}{(n-1)\hat{\rho}+1}+\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right)+\left(\frac{2\theta\tau}{D_n}-\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)(1-\hat{\rho})}\right). \end{aligned} \tag{72}$$

which implies

$$\begin{aligned} 0 &= \alpha_n^2\left(\frac{D_n}{(n-1)\hat{\rho}+1}-\frac{\tau(1-\hat{\rho})}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right) \\ &\quad +\alpha_n\left(\frac{2\tau}{n\hat{\rho}+1}-\frac{\tau}{(n-1)\hat{\rho}+1}-\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)[(n-1)\hat{\rho}+1]}\right) \\ &\quad +\left(2\theta-\frac{\tau[(n-1)\hat{\rho}+1]}{(n\hat{\rho}+1)(1-\hat{\rho})}-\frac{2\theta\tau}{D_n}+\frac{\tau\hat{\rho}}{(n\hat{\rho}+1)(1-\hat{\rho})}\right), \end{aligned} \tag{73}$$

which, as I have already shown, has a unique positive solution, α_n^*.

All that remains is to prove that $f_n^* \equiv \alpha_n^* s_n$ maximizes (rather than minimizes) n's objective function. It easily follows from (63) and (69) that the condition in (25) guarantees that f_n^* maximizes n's objective function, completing the proof.

□

Proof. (**Lemma 1**)

Since analysts' forecasts are given by $f_n = \alpha_n s_n$, it suffices to show that

$$\hat{\mathbf{E}}[s|x = f_n, \overline{s_n}] = \frac{(1-\hat{\rho})f_n + n\hat{\rho}\overline{s_n}}{(n-1)\hat{\rho}+1}. \tag{74}$$

Since each analyst believes $(x, \overline{s_n}, s)$ is distributed

$$\begin{bmatrix} x \\ \overline{s_n} \\ s \end{bmatrix} \sim^i N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\tau_x} & \frac{1}{\tau_x} & \frac{1}{\tau_x} \\ \frac{1}{\tau_x} & \frac{D_n}{n\tau\tau_x} & \frac{1}{\tau_x} + \frac{\hat{\rho}}{\tau} \\ \frac{1}{\tau_x} & \frac{1}{\tau_x} + \frac{\hat{\rho}}{\tau} & \frac{1}{\tau_x+\tau} \end{bmatrix}\right),$$

it follows that

$$\hat{\mathbf{E}}[s|x = f_n, \overline{s_n}] = \left(\frac{n\tau\tau_x}{(n-1)\hat{\rho}+1}\right) \begin{bmatrix} \frac{1}{\tau_x} & \frac{\tau+\hat{\rho}\tau_x}{\tau\tau_x} \end{bmatrix} \begin{bmatrix} \frac{D_n}{n\tau\tau_x} & -\frac{1}{\tau_x} \\ -\frac{1}{\tau_x} & \frac{1}{\tau_x} \end{bmatrix} \begin{bmatrix} f_n \\ \overline{s_n} \end{bmatrix},$$

i.e.,

$$\hat{\mathbf{E}}[s|x = f_n, \overline{s_n}] = \frac{(1-\hat{\rho})f_n + n\hat{\rho}\overline{s_n}}{(n-1)\hat{\rho}+1}, \tag{75}$$

completing the proof. □

Proof. (**Proposition 3**)

Note that given my assumptions, the framework is identical to one in which an agent observes a signal, s_A (the collection of signals $\{s_1, \ldots, s_n\}$), whose precision he overestimates, issues a forecast based on s_A, and then observes a second signal, s_B (the collection of signals $\{s_{n+1}, \ldots, s_{n+k}\}$), whose precision he correctly estimates, and then issues a second forecast based on s_A and s_B. Because I assume that $\rho = \hat{\rho} = 0$, s_A and s_B are independent signals for X. Let τ_A and τ_B denote their true precisions, and let $\hat{\tau}_A$ and $\hat{\tau}_B$ denote the agent's beliefs about their precisions. Based on my assumptions, $\hat{\tau}_A > \tau_A$ and $\hat{\tau}_B = \tau_B$.

Let f_A be the first forecast issued by the analyst (after observing s_A), and let f_B be the revised forecast which is issued after the agent observes the signals s_A and s_B:

$$f_A = \frac{\tau_x}{\tau_x + \hat{\tau}_A}\mu + \frac{\hat{\tau}_A}{\tau_x + \hat{\tau}_A}s_A \tag{76}$$

$$f_B = \frac{\tau_x}{\tau_x + \hat{\tau}_A + \tau_B}\mu + \frac{\hat{\tau}_A}{\tau_x + \hat{\tau}_A + \tau_B}s_A + \frac{\tau_B}{\tau_x + \hat{\tau}_A + \tau_B}s_B. \tag{77}$$

Lemma 2. $\mathbf{P}(sgn(X - f_B) = sgn(f_B - f_A))$ *is increasing in* τ_B *for* $\tau_B < \hat{\tau}_A + \tau_x$*, and is decreasing in* τ_B *for* $\tau_B > \hat{\tau}_A + \tau_x$*.*

Recall that in this (overconfidence) framework, τ_B refers to the precision of the collection of signals $\{s_{n+1}, \ldots, s_{n+k}\}$, and $\hat{\tau}_A$ refers to the perceived precision of the collection of signals $\{s_1, \ldots, s_n\}$. From the lemma, the probability of underreaction is maximized when $\tau_B = \hat{\tau}_A + \tau_x$, i.e., when $k\tau = \hat{\tau} + (n-1)\tau + \tau_x$. Solving for k completes the proof.

□

Proof. (**Lemma 2**)

From (76) and (77),

$$
\begin{aligned}
X - f_B &= \Big(\frac{\tau_x}{\tau_x + \hat{\tau}_A + \tau_B}\Big)X - \frac{\hat{\tau}_A}{\tau_x + \hat{\tau}_A + \tau_B}\varepsilon_A - \frac{\tau_B}{\tau_x + \hat{\tau}_A + \tau_B}\varepsilon_B - \frac{\tau_x}{\tau_x + \hat{\tau}_A + \tau_B}\mu && (78)\\
f_B - f_A &= \Big(\frac{\tau_B\tau_x}{(\tau_x + \hat{\tau}_A)(\tau_x + \hat{\tau}_A + \tau_B)}\Big)X - \Big(\frac{\hat{\tau}_A\tau_B}{(\tau_x + \hat{\tau}_A)(\tau_x + \hat{\tau}_A + \tau_B)}\Big)\varepsilon_A + \Big(\frac{\tau_B}{\tau_x + \hat{\tau}_A + \tau_B}\Big)\varepsilon_B \\
&\quad - \frac{\tau_x\tau_B}{(\tau_x + \hat{\tau}_A)(\tau_x + \hat{\tau}_A + \tau_B)}\mu && (79)\\
\mathrm{Cov}(X - f_B, f_B - f_A) &= \frac{\tau_B\hat{\tau}_A(\hat{\tau}_A - \tau_A)(\tau_x + \hat{\tau}_A)}{\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_x + \hat{\tau}_A + \tau_B)^2} && (80)\\
\mathrm{Var}(X - f_B) &= \frac{\tau_B[\tau_A(\tau_x + \hat{\tau}_A)^2] + (\tau_A\tau_x + \hat{\tau}_A{}^2)(\tau_x + \hat{\tau}_A)^2}{\tau_A(\tau_x + \hat{\tau}_A + \tau_B)^2(\tau_x + \hat{\tau}_A)^2} && (81)\\
\mathrm{Var}(f_B - f_A) &= \frac{\tau_B^2(\tau_x\tau_A + \hat{\tau}_A{}^2) + \tau_B(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2)}{(\tau_x + \hat{\tau}_A)^2(\tau_x + \hat{\tau}_A + \tau_B)^2\tau_A} && (82)
\end{aligned}
$$

Let $\Gamma_{\mathrm{OC}}(\cdot)$ be the correlation of the random variables $X - f_B$ and $f_B - f_A$ as a function of τ_B.

$$
\Gamma_{\mathrm{OC}}(\tau_B) = \frac{\tau_B\hat{\tau}_A(\hat{\tau}_A - \tau_A)(\tau_x + \hat{\tau}_A)}{\sqrt{\Delta(\tau_B)}}, \qquad (83)
$$

where

$$
\begin{aligned}
\Delta(\tau_B) &= \Big(\tau_B[\tau_A(\tau_x + \hat{\tau}_A)^2] + (\tau_A\tau_x + \hat{\tau}_A{}^2)(\tau_x + \hat{\tau}_A)^2\Big)\Big(\tau_B^2(\tau_x\tau_A + \hat{\tau}_A{}^2) + \tau_B(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2)\Big)\\
&= \tau_B^3\Big[\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_x\tau_A + \hat{\tau}_A{}^2)\Big]\\
&\quad + \tau_B^2\Big[\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2) + (\tau_x\tau_A + \hat{\tau}_A{}^2)^2(\tau_x + \hat{\tau}_A)^2\Big]\\
&\quad + \tau_B\Big[(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2)(\tau_A\tau_x + \hat{\tau}_A{}^2)(\tau_x + \hat{\tau}_A)^2\Big]. \qquad (84)
\end{aligned}
$$

(To justify (83), it is helpful to let C_N, X_N, F_N, C_D, X_D, and F_D be defined as the numerators and

denominators of $\mathrm{Cov}(X - f_B, f_B - f_A)$, $\mathrm{Var}(X - f_B)$, and $\mathrm{Var}(f_B - f_A)$ in equations (80)-(82), so that

$$\begin{aligned}
\mathrm{Cov}(X - f_B, f_B - f_A) &= \frac{C_N}{C_D} \\
\mathrm{Var}(X - f_B) &= \frac{X_N}{X_D} \\
\mathrm{Var}(f_B - f_A) &= \frac{F_N}{F_D}.
\end{aligned}$$

Then clearly, $\Gamma_{\mathrm{OC}}(\tau_B) = \frac{C_N\sqrt{X_D F_D}}{C_D\sqrt{X_N F_N}}$, so to justify (83), it suffices to show that

$$\frac{C_N\sqrt{X_D F_D}}{C_D} = \tau_B \hat{\tau}_A(\hat{\tau}_A - \tau_A)(\tau_x + \hat{\tau}_A),$$

which can be verified.)

Differentiating (83),

$$\Gamma'_{\mathrm{OC}}(\tau_B) = \hat{\tau}_A(\hat{\tau}_A - \tau_A)(\tau_x + \hat{\tau}_A)\Big[\frac{\sqrt{\Delta(\tau_B)} - \frac{1}{2}\tau_B[\Delta(\tau_B)]^{-\frac{1}{2}}\Delta'(\tau_B)}{\Delta(\tau_B)}\Big]$$

Hence, since the agent is overconfident ($\hat{\tau}_A > \tau_A$),

$$\begin{aligned}
\mathrm{sgn}(\Gamma'_{\mathrm{OC}}(\tau_B)) &= \mathrm{sgn}\Big(\sqrt{\Delta(\tau_B)} - \frac{\tau_B\Delta'(\tau_B)}{2\sqrt{\Delta(\tau_B)}}\Big) \\
&= \mathrm{sgn}\Big(2\Delta(\tau_B) - \tau_B\Delta'(\tau_B)\Big) \qquad (85)
\end{aligned}$$

Differentiating (84),

$$\begin{aligned}
\Delta'(\tau_B) &= 3\tau_B^2\Big[\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_x\tau_A + \hat{\tau}_A{}^2)\Big] \\
&\quad +2\tau_B\Big[\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2) + (\tau_x\tau_A + \hat{\tau}_A{}^2)^2(\tau_x + \hat{\tau}_A)^2\Big] \\
&\quad +\Big[(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2)(\tau_A\tau_x + \hat{\tau}_A{}^2)(\tau_x + \hat{\tau}_A)^2\Big] \qquad (86)
\end{aligned}$$

Hence, from (84) and (86)

$$\begin{aligned}
2\Delta(\tau_B) - \tau_B\Delta'(\tau_B) &= (\tau_B^3)\Big[-\tau_A(\tau_x + \hat{\tau}_A)^2(\tau_x\tau_A + \hat{\tau}_A{}^2)\Big] \\
&\quad +\tau_B\Big[(\tau_A\tau_x^2 + 2\tau_A\hat{\tau}_A\tau_x + \tau_A\hat{\tau}_A{}^2)(\tau_A\tau_x + \hat{\tau}_A{}^2)(\tau_x + \hat{\tau}_A)^2\Big] \\
&= \tau_A\tau_B(\tau_x + \hat{\tau}_A)^2(\tau_A\tau_x + \hat{\tau}_A{}^2)\Big[\tau_x^2 + 2\hat{\tau}_A\tau_x + \hat{\tau}_A{}^2 - \tau_B^2\Big] \qquad (87)
\end{aligned}$$

From (85) and (87), with overconfidence the probability of underreacting is maximized when

$$\begin{aligned} \tau_B^2 &= \hat{\tau}_A{}^2 + 2\hat{\tau}_A\tau_x + \tau_x^2 \\ &= (\hat{\tau}_A + \tau_x)^2, \quad \text{i.e., when} \\ \tau_B &= \hat{\tau}_A + \tau_x. \end{aligned} \tag{88}$$

□

References

Ai, C., and E. Norton (2003): "Interaction Terms in Logit and Probit Models," *Economics Letters*, 80, 123–129.

Altinkilic, O., and R. Hansen (2009): "On the Information Role of Stock Recommendation Revisions," *Journal of Accounting and Economics*, 48, 17–36.

Bernhardt, D., M. Campello, and E. Kutsoati (2006): "Who Herds?," *Journal of Financial Economics*, 80, 657–675.

Brown, L. (2001): "How important is past analyst forecast accuracy?," *Financial Analysts Journal*, 57, 44–49.

Chen, Q., and W. Jiang (2006): "Analysts' Weighting of Private and Public Information," *The Review of Financial Studies*, 19, 319–355.

Clement, M., J. Hales, and Y. Xue (2011): "Understanding Analysts' use of Stock Returns and other Analysts' Revisions when Forecasting Earnings," *Journal of Accounting and Economics*, 51, 279–299.

Clement, M., and S. Tse (2003): "Do Investors Respond to Analysts' Forecast Revisions as If Forecast Accuracy is All that Matters?," *The Accounting Review*, 78 (1), 227–249.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998): "Investor Psychology and Security Market Under- and Overreactions," *The Journal of Finance*, 53, 1839–1885.

Dawes, R., and M. Mulford (1996): "The False Consensus Effect and Overconfidence: Flaws in Judgment or Flaws in How We Study Judgment?," *Organizational Behavior and Human Decision Processes*, 65, 202–211.

Dharmadhikari, S., and K. Joag-Dev (1984): "Regions Whose Probabilities Increase with the Correlation Coefficient and Slepian's Theorem," *Lecture Notes-Monograph Series*, 5, 156–164.

Egan, D., C. Merkle, and M. Weber (2010): "The Beliefs of Others - Naive Realism and Investment Decisions," *Working Paper*.

Elliott, J., D. Philbrick, and C. Wiedman (1995): "Evidence from Archival Data on the Relation Between Security Analysts' Forecast Errors and Prior Forecast Revisions," *Contemporary Accounting Research*, 11, 919–938.

Engelberg, J., C. Manski, and J. Williams (2009): "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business and Economic Statistics*, 165, 146–158.

Engelmann, D., and M. Strobel (2000): "The False Consensus Effect Disappears if Representative and Monetary Incentives are Given," *Experimental Economics*, 3, 241–260.

——— (2004): "Deconstruction and Reconstruction of an Anomaly," *Working Paper*.

Gleason, C., and C. Lee (2003): "Analyst Forecast Revisions and Market Price Discovery," *The Accounting Review*, 78, 193–225.

Grossman, S. (1976): "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information," *The Journal of Finance*, 31, 573–585.

Holthausen, R., and R. Verrecchia (1990): "The Effect of Informedness and Consensus on Price and Volume Behavior," *The Accounting Review*, 65, 191–208.

Hong, H., and J. Kubik (2003): "Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts," *The Journal of Finance*, 58, 313–351.

Hong, H., J. Kubik, and A. Solomon (2000): "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31, 121–144.

Kang, S.-H., J. O'Brien, and K. Sivaramakrishnan (1994): "Analysts' Interim Earnings Forecasts: Evidence on the Forecasting Process," *Journal of Accounting Research*, 32, 103–112.

Katz, D., and F. Allport (1931): *Students' Attitudes*. Syracuse: Craftsman Press.

Keskek, S., S. Tse, and J. W. Tucker (2011): "Analyst Information Production and Forecast Timing Around Earnings Announcements," *Working Paper*.

Krueger, J. (1998): *Advances in Experimental Social Psychology*vol. 30, pp. 163–240. Academic Press.

——— (2000): *Handbook of Social Comparison: Theory and Research*pp. 323–351. Plenum Publishers.

Krueger, J., and J. Zeiger (1993): "Social Categorization and the Truly False Consensus Effect," *Journal of Personality and Social Psychology*, 65, 670–680.

Marinovic, I., M. Ottaviani, and P. N. Sørensen (2012): "Forecasters' Objectives and Strategies," *Working Paper*.

Mikhail, M., B. Walther, and R. Willis (1997): "Do Security Analysts Improve their performance with experience?," *Journal of Accounting Research*, 35, 131–157.

——— (1999): "Does Forecast Accuracy Matter to Security Analysts?," *The Accounting Review*, 74, 185–200.

Odean, T. (1998): "Volume, Volatility, Price, and Profit When All Traders Are Above Average," *The Journal of Finance*, 53, 1887–1934.

Ottaviani, M., and P. N. Sørensen (2006): "The Strategy of Professional Forecasting," *Journal of Financial Economics*, 81, 441–466.

Richardson, S., S. Teoh, and P. Wysocki (2004): "The Walk-down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading," *Contemporary Accounting Research*, 21, 885–924.

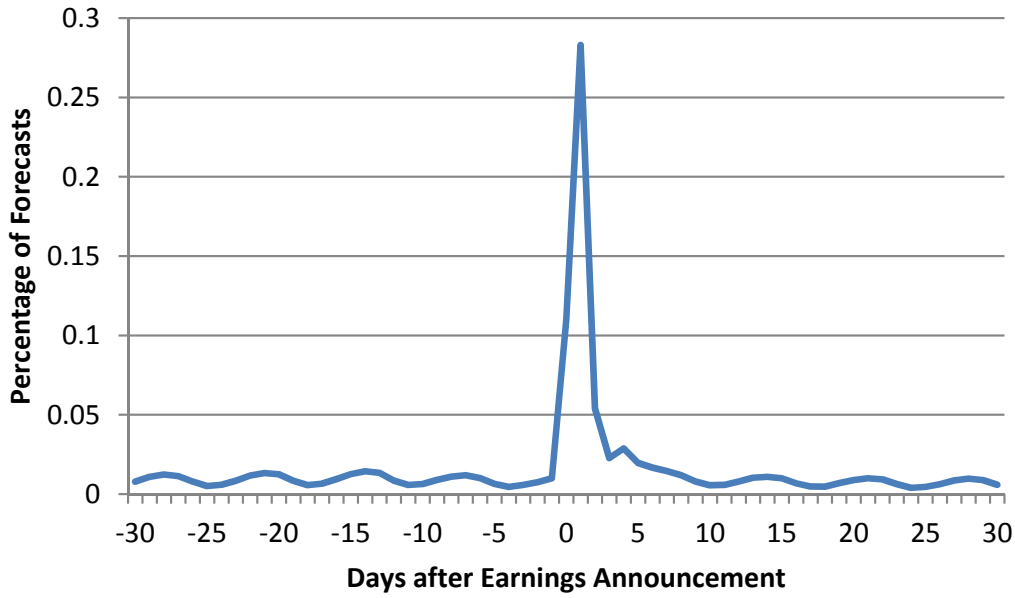
Ross, L., D. Greene, and P. House (1977): "The "False Consensus Effect": An Egocentric Bias in Social Perception and Attribution Processes," *Journal of Experimental Social Psychology*, 13, 279–301.

Trueman, B. (1990): "On the Incentives for Security Analysts to Revise their Earnings Forecasts," *Contemporary Accounting Research*, 7, 203–222.

——— (1994): "Analyst Forecasts and Herding Behavior," *The Review of Financial Studies*, 7, 97–124.

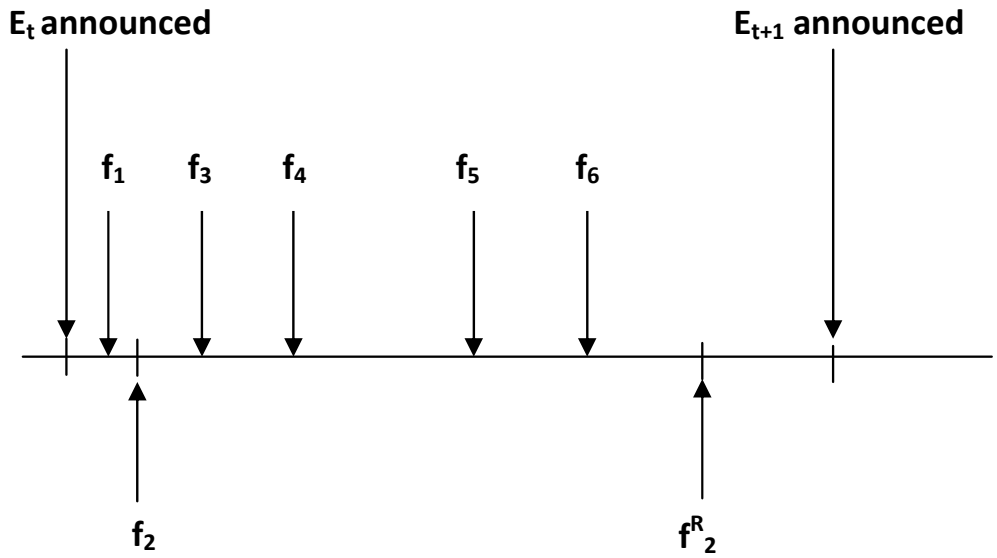
Wu, J. S., and A. Zang (2009): "What Determine Financial Analysts' Career Outcomes during Mergers?," *Journal of Accounting and Economics*, 47, 59–86.

Figure 1: Analyst Forecasting Activity Around Earnings Announcements



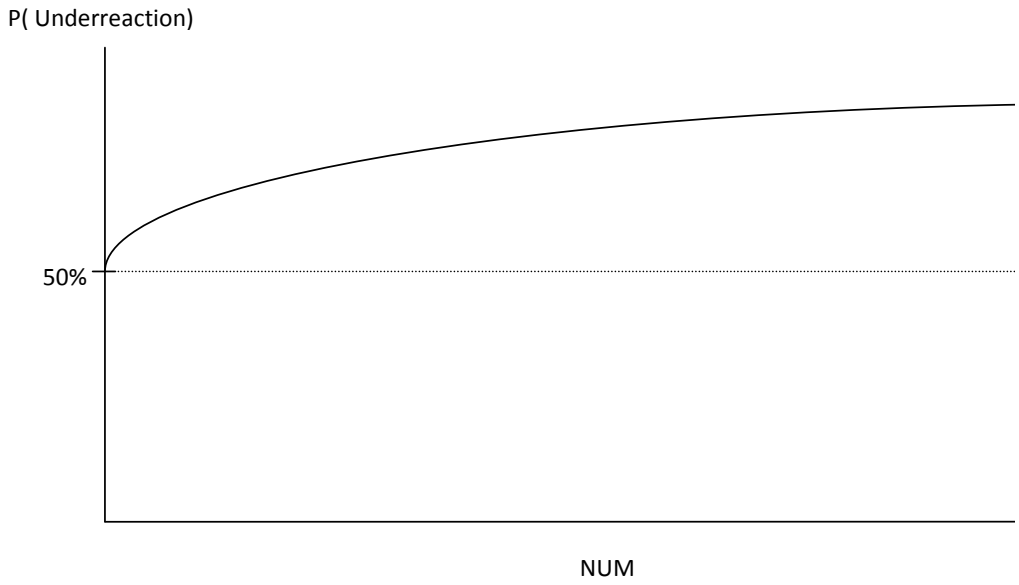
I take the 3,296,837 I/B/E/S quarterly earnings forecasts that are issued between a firm's earnings announcement dates, where the earnings announcement dates differ between 61 and 120 days. From this sample, I take the 2,687,500 forecasts that are within 30 days of an earnings announcement. I plot the percentage of these forecasts that lie in each of the 61 days surrounding the earnings announcement date. Event day 0 corresponds to the earnings announcement date.

Figure 2: Illustration of Sample and *NUM* Variable Construction



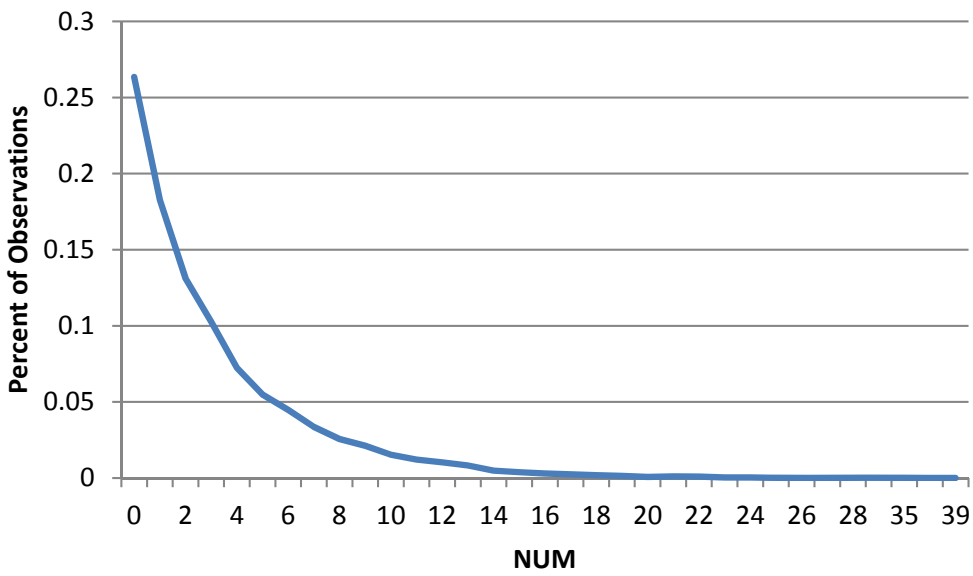
I illustrate the construction of my sample and *NUM*, my main variable of interest. For a given firm-quarter (j, t), I consider the analysts who issue forecasts and forecast revisions for firm j's quarter $t+1$ earnings following j's quarter t earnings announcement. For each such (f_n, f_n^R) pair, I count the number of analysts issuing forecasts between f_n and f_n^R (chronologically). In this example, four analysts (3, 4, 5, and 6) issue forecasts for firm j's quarter $t+1$ earnings between analyst 2's initial forecast and his revised forecast. Hence, in this example, $NUM = 4$ for the pair (f_2, f_2^R).

Figure 3: False Consensus Prediction



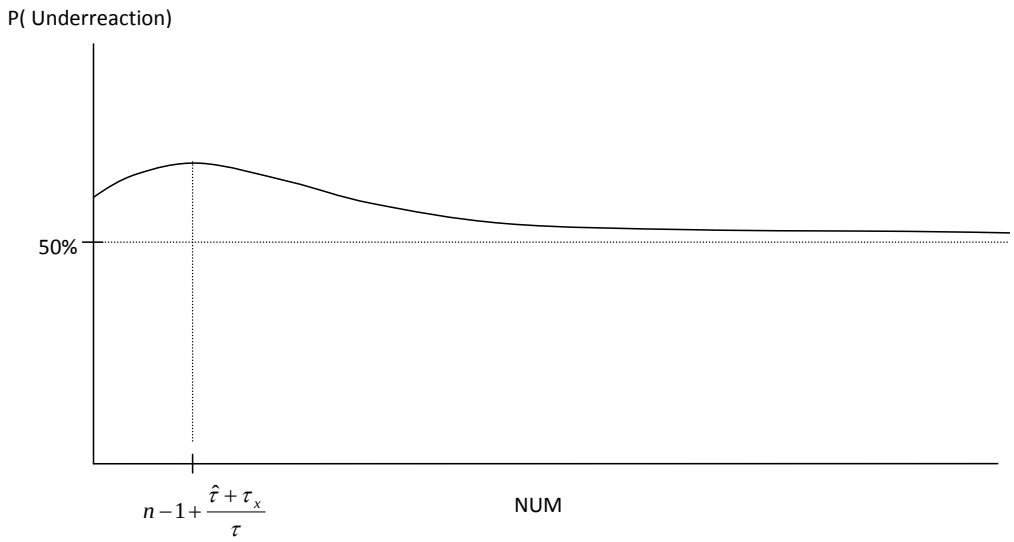
Based on my model of sequential forecasting with a false consensus effect ($\hat{\rho} > \rho$), I plot the predicted relationship between the probability of forecast revision underreaction and *NUM*.

Figure 4: Distribution of *NUM*



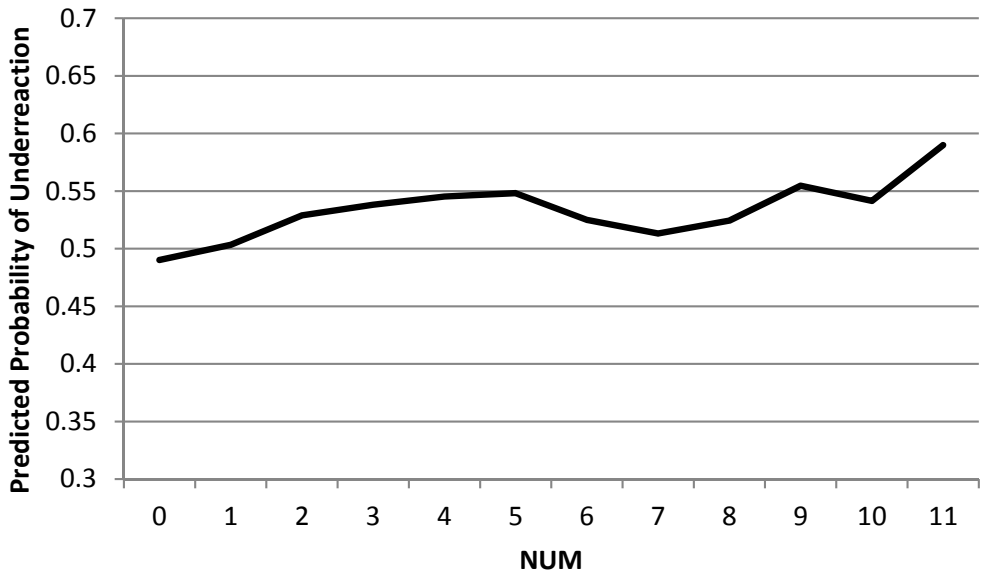
From my sample of 26,100 forecast revisions, I plot the percentage of observations for each of the values of *NUM*. In my probit analysis, I reassign all values of *NUM* that are greater than 11 to take the value of 11, to avoid the effects of outliers. My qualitative results do not depend on this reassignment.

Figure 5: Overconfidence Prediction



Based on my model of sequential forecasting with overconfidence (see Section 5.2), I plot the relationship between the probability of forecast revision underreaction and *NUM* that is predicted by analyst overconfidence.

Figure 6: Probit Probabilities



I run a probit regression that is similar to the one in column (6) of Table 3, except that the variable *NUM* is replaced by dummies for the 12 possible values of *NUM* ($NUM = 0, \ldots, NUM = 11$). I then set all of the control variables to their means, and plot the predicted probabilities of forecast revision underreaction by *NUM* based on the probit model. All revised forecasts in which at least 11 different analysts issued forecasts between the analyst's earlier forecast and his revised forecast are assigned a *NUM* dummy of 11.

Table 1: Summary Statistics

		Mean	Std Dev
Number of Observations	26,100		
Number of Firm-quarters	16,433		
Number of Analysts	4,796		
Number of Firms	3,405		
Time Span	1985-2011		
Coverage		13.29	7.43
Days since last EPS announcement		30.88	26.14
Days between forecasts		15.69	8.63
Days until EPS announcement		45.28	26.96

I report summary statistics for my sample consisting of 26,100 forecast revisions. "Coverage" is defined as the number of distinct analysts who issued a forecast for the company's earnings (for any quarter's earnings) during the quarter the forecast is issued. "Days since last EPS announcement" is defined as the number of days between the previous quarter's earnings announcement date and the analyst's earlier forecast. "Days between forecasts" is defined as the number of days between the analyst's earlier forecast and his revised forecast. "Days until EPS announcement" is defined as the number of days between the date the company announces its quarterly earnings and the date the analyst's revised forecast is issued.

Table 2: Number of Observations, by Quintile

Variable / Quintile	1	2	3	4	5
Size	1,617	2,564	3,666	5,660	12,593
BM	7,003	6,162	5,487	4,607	2,841
Coverage	273	798	1,719	4,583	18,727
Difficulty	5,422	6,228	6,302	5,158	2,990
Firm Accuracy	4,107	5,054	4,418	3,676	2,850
Total Accuracy	4,440	6,005	6,047	4,787	3,275
Firm Experience	3,391	4,137	4,826	5,620	8,126
Total Experience	1,672	3,139	4,946	6,547	9,796

I show the number of observations from my sample of 26,100 forecast revisions that lie in each of the control variable quintiles. The number of observations for *Firm Accuracy* (*Total Accuracy*) do not sum to 26,100 because those variables are undefined if the analyst has not issued a forecast in the 8 quarters before the quarter the forecast is issued. When defining quintile dummies, I create a sixth dummy for each these two variables to capture whether the variables are undefined. See Appendix A for a detailed description for how these variables are defined.

Table 3: Underreaction Probits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
NUM	0.015***	0.016***	0.016***	0.017***	0.016***	0.017***	0.006***	0.005***	0.003**	0.004***
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.001]	[0.001]	[0.001]	[0.001]
UpRevision		0.588***	0.589***	0.589***	0.589***	0.589***	0.230***	0.237***	0.226***	0.236***
		[0.042]	[0.042]	[0.042]	[0.042]	[0.042]	[0.016]	[0.016]	[0.016]	[0.017]
ReviseWithCARs			0.056***	0.056***	0.054***	0.055***	0.021***	0.020***	0.023***	0.023***
			[0.015]	[0.015]	[0.015]	[0.015]	[0.006]	[0.006]	[0.007]	[0.007]
Constant	0.013	-0.241***	-0.270***	-0.214***	-0.316***	-0.236***	0.364***			
	[0.012]	[0.022]	[0.023]	[0.041]	[0.040]	[0.065]	[0.038]			
R^2/Pseudo R^2	0.001	0.039	0.039	0.041	0.041	0.041	0.056	0.205	0.235	0.358
N	26,100	26,100	26,100	26,100	26,100	26,100	26,100	26,100	26,100	26,100
Regression Type	Probit	Probit	Probit	Probit	Probit	Probit	OLS	OLS	OLS	OLS
Firm-type controls	No	No	No	Yes	No	Yes	Yes	No	No	No
Analyst-type controls	No	No	No	No	Yes	Yes	Yes	No	No	No
Firm Fixed Effects	No	No	No	No	No	No	No	Yes	No	Yes
Analyst Fixed Effects	No	No	No	No	No	No	No	No	Yes	Yes

The dependent variable in these regressions is the dummy for whether the analyst underreacted (ex post) in his forecast revision. More specifically, the dependent variable is defined as $\mathbf{1}_{\{\text{sgn}(\text{EPS}-f^R)=\text{sgn}(f^R-f^0)\}}$, where EPS is the reported earnings for the quarter being forecasted, f^R is the analyst's revised forecast, and f^0 is the analyst's earlier forecast. This dummy takes the value 1 if and only if the analyst would have had a lower forecast error (in absolute value) if he had moved one penny further in the direction of his forecast revision. The sample is restricted to forecast revisions that are not equal to the reported earnings or the analyst's earlier forecast. *NUM*, the independent variable of interest, is defined as the number of distinct analysts who issue forecasts after the analyst's earlier forecast and before his revised forecast. See Figure 2 for an illustration. *UpRevision* is a dummy for whether the revised forecast is greater than the analyst's earlier forecast. *ReviseWithCARs* is a dummy for whether the revised forecast is in the same direction as the company's stock's cumulative abnormal returns (CARs) between the analyst's earlier forecast and his revised forecast. E.g., if the stock had positive CARs and the analyst issued a revised forecast that was greater than (less than) his earlier forecast, this variable takes the value 1 (0). Firm-type controls consist of quintile dummies for the variables *Size*, *BM*, *Coverage*, and *Difficulty*. Analyst-type controls consist of quintile dummies for the variables *Firm Accuracy*, *Total Accuracy*, *Firm Experience*, and *Total Experience*. For definitions of these variables, see Appendix A. Robust standard errors are clustered by quarter and are presented in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 4: Robustness checks

	Original Sample	1 month, any CAR	$\lvert CAR \rvert \leq 2\%$	$\lvert CAR \rvert \leq 10\%$
NUM	0.017***	0.014***	0.013***	0.007***
	[0.004]	[0.002]	[0.002]	[0.001]
UpRevision	0.589***	0.557***	0.725***	0.732***
	[0.042]	[0.039]	[0.038]	[0.033]
ReviseWithCARs	0.055***	0.164***	0.034***	0.095***
	[0.015]	[0.013]	[0.012]	[0.010]
Constant	-0.236***	-0.270***	-0.337***	-0.378***
	[0.065]	[0.039]	[0.050]	[0.036]
Pseudo R^2	0.041	0.039	0.06	0.062
N	26,100	116,443	53,938	202,326

The dependent variable in these probit regressions is same as in Table 3: it is the dummy for whether the analyst would have had a lower forecast error (in absolute value) if he had moved one penny further in the direction of his forecast revision. Column (1) repeats the regression from Column (6) of Table 3: the sample consists of forecast revisions that satisfy the following two conditions: (i) the revised forecast is made within a month of the analyst's earlier forecast, and (ii) the cumulative abnormal returns (CARs) between the analyst's earlier forecast and his revised forecast is less than 2% (in absolute value). Columns (2)-(4) examine less restrictive samples. In Column (2), I include all forecast revisions that occur within a month of the analyst's earlier forecast, regardless of the CARs between the analyst's earlier forecast and his revised forecast. In Column (3), I include all forecast revisions such that the CARs between the analyst's earlier forecast and his revised forecast are less than 2%, regardless of the amount of time between the analyst's earlier forecast and his revised forecast. In Column (4), I include all forecast revisions such that the CARs between the analyst's earlier forecast and his revised forecast are less than 10%, regardless of the amount of time between the analyst's earlier forecast and his revised forecast. *NUM*, the independent variable of interest, is defined as the number of forecasts issued by other analysts between the analyst's earlier forecast and his revised forecast. All of these regressions contain both the analyst-type and firm-type controls from Table 3. For definitions of these variables, see Appendix A. Robust standard errors are clustered by quarter and presented in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 5: Underreaction Probits, by Direction of Revision

	Up	Down	Away	Herding	Past
NUM	0.029***	0.009*	0.006	0.017**	0.011**
	[0.005]	[0.005]	[0.005]	[0.007]	[0.005]
ReviseWithCARs	0.073***	0.042**	0.057	0.059*	0.067***
	[0.024]	[0.020]	[0.035]	[0.035]	[0.025]
UpRevision			0.513***	0.471***	0.729***
			[0.060]	[0.059]	[0.049]
Constant	0.137	-0.089	-0.388***	-0.106	-0.236**
	[0.113]	[0.083]	[0.114]	[0.137]	[0.102]
Pseudo R^2	0.014	0.01	0.036	0.032	0.061
N	11,317	14,783	6,545	5,943	10,807

The dependent variable in these probit regressions is same as in Table 3: it is the dummy for whether the analyst would have had a lower forecast error (in absolute value) if he had moved one penny further in the direction of his forecast revision. The samples for these probits are divided based on the direction of the forecast revisions: the samples in Columns (1) and (2) consist of all upwards and downwards revisions, respectively. Columns (3)-(5) divide the sample based on whether the revision was (i) away from the consensus forecast ("Away"), (ii) between the analyst's earlier forecast and the consensus forecast ("Herding"), or (iii) in the same direction as the consensus forecast, but past the consensus forecast ("Past"). The number of observations in Columns 3-5 do not sum to 26,100 because revisions in which the analyst's earlier forecast or his revised forecast equal the consensus forecast are not included in any of the three samples. *NUM*, the independent variable of interest, is defined as the number of forecasts issued by other analysts between the analyst's earlier forecast and his revised forecast. All of these regressions contain the analyst-type and firm-type controls from Table 3. For definitions of these variables, see Appendix A. Robust standard errors are clustered by quarter and presented in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 6: Employment Outcomes of Analysts

Panel A: Outcomes of Analysts outside of Top 10, by *NUM* Coefficient					
Num Coefficient	Obs	Advancements	Pct Advanced	Exits	Pct Exits
Low	702	15	2.1%	78	11.1%
High	750	19	2.5%	86	11.5%
High Minus Low			0.4%		0.4%

Panel B: Outcomes of Analysts at Top 10, by *NUM* Coefficient					
Num Coefficient	Obs	Demotions	Pct Demoted	Exits	Pct Exits
Low	462	17	3.7%	61	13.2%
High	506	10	2.0%	78	15.4%
High Minus Low			-1.7%		2.2%

Panel C: Outcomes of Analysts outside of Top 10, by *Accuracy*					
Accuracy	Obs	Advancements	Pct Advanced	Exits	Pct Exits
Low	750	11	1.5%	86	11.5%
High	702	23	3.3%	78	11.1%
High Minus Low			1.8%**		-0.4%

I analyze the career outcomes of analysts in a given year as a function of their *NUM* coefficient, their *Total Accuracy*, and whether or not they began the year at a top 10 brokerage (based on size). I define a career advancement as moving from a non-top ten brokerage to a top ten brokerage. I define a career demotion as moving from a top ten brokerage to a non-top ten brokerage. I consider an analyst to exit the profession in year t if he issues a forecast in year t but not in year $t+1$. Each year, I sort analysts into quintiles based on their *Total Experience*. Within each year and *Total Experience* quintile, I sort analysts into "high" and "low" groups based on whether their *NUM* coefficients are greater than the median. The career outcomes for these analysts are reported in Panels A and B. Within each year and *Total Experience* quintile, I sort analysts into "high" and "low" groups based on whether their *Total Accuracy* measures are greater than the median. The career outcomes for these analysts employed in non-top ten brokerages are reported in Panel C. My sample is restricted to 2,420 analyst-year pairs, (i, t), such that i issued at least 20 (forecast, revised forecast) pairs within consecutive quarterly earnings announcements. *, **, and *** represent statistical significance of t-tests at the 10%, 5% and 1% level, respectively.

Table 7: Testing Career Concerns versus Behavioral Biases

Panel A: Top Ten Dummy and *NUM* Interaction (Career Concerns)		
	Probit	OLS
NUM	0.022***	0.008***
	[0.004]	[0.001]
TopTen × *NUM*	-0.012**	-0.005**
	[0.005]	[0.002]
TopTen	0.040**	0.016*
	[0.020]	[0.009]
UpRevision	0.590***	0.235***
	[0.042]	[0.006]
ReviseWithCARs	0.055***	0.020***
	[0.015]	[0.006]
Constant	-0.255***	0.376***
	[0.066]	[0.037]
R^2	0.042	0.064
N	26,100	26,100
Panel B: Low *Experience* Dummy and *NUM* Interaction (Career Concerns)		
	Probit	OLS
NUM	0.017***	0.006***
	[0.004]	[0.001]
LowExp × *NUM*	0.077*	0.030*
	[0.044]	[0.016]
LowExp	-0.342***	-0.129***
	[0.105]	[0.047]
UpRevision	0.588***	0.235***
	[0.042]	[0.006]
ReviseWithCARs	0.056***	0.021***
	[0.015]	[0.006]
Constant	-0.256***	0.430***
	[0.063]	[0.037]
R^2	0.042	0.064
N	26,100	26,100

In each panel, I report the results of probit regressions (left column) and OLS regressions (right column). The dependent variable in each regression is the dummy for whether the analyst's revised forecast is too close (ex post) to his earlier forecast. *TopTen* is a dummy for whether or not the analyst is employed by one of the ten largest brokerages at the time the revised forecast is issued. *LowExp* is a dummy for whether or not the analyst is in the bottom quintile of analysts based on *Experience*, where the *Experience* rankings are done annually. As in Column (6) of Table 3, I include analyst-type and firm-type dummies. In Panel B, I omit *Firm Experience* quintile dummies, and the only *Total Experience* quintile dummy I include is the *Total Experience* quintile = 1 dummy (i.e., *LowExp*). Robust standard errors are clustered by quarter and presented in brackets. *, **, and *** represent statistical significance of t-tests at the 10%, 5% and 1% level, respectively.

Table 8: Underreaction Probits, Overconfidence Tests

	All *NUM*	*NUM*> 0	*NUM*> 1	*NUM*> 2	*NUM*> 3	*NUM*> 4	*NUM*> 5	*NUM*> 6	*NUM*> 7	*NUM*> 8	*NUM*> 9
NUM	0.012**	0.008	0.006	0.002	0.002	0.014	0.029*	0.048**	0.046	0.092	0.105
	[0.005]	[0.006]	[0.007]	[0.007]	[0.011]	[0.015]	[0.017]	[0.023]	[0.036]	[0.060]	[0.128]
UpRevision	0.499***	0.532***	0.570***	0.603***	0.661***	0.684***	0.677***	0.744***	0.804***	0.876***	0.971***
	[0.049]	[0.053]	[0.055]	[0.057]	[0.057]	[0.061]	[0.065]	[0.070]	[0.089]	[0.110]	[0.124]
ReviseWithCARs	0.047	0.048	0.074*	0.090**	0.066	0.076	0.104	0.103	0.132	0.087	0.154
	[0.029]	[0.032]	[0.042]	[0.045]	[0.050]	[0.057]	[0.066]	[0.077]	[0.097]	[0.105]	[0.113]
Constant	-0.148	-0.111	-0.187	0.01	0.012	0.033	-0.174	-0.015	0.104	-0.121	-0.413
	[0.106]	[0.113]	[0.131]	[0.141]	[0.168]	[0.217]	[0.239]	[0.307]	[0.465]	[0.726]	[1.429]
Pseudo R^2	0.031	0.035	0.042	0.048	0.058	0.066	0.073	0.087	0.097	0.107	0.126
N	8410	6184	4755	3729	2888	2297	1811	1392	1070	838	662

The dependent variable in these probit regressions is same as in Table 3: it is the dummy for whether the analyst would have had a lower forecast error (in absolute value) if he had moved one penny further in the direction of his forecast revision. The samples are restricted based on whether *NUM* is sufficiently high. Column (1) consists of all the observations from Table 3 in which *NUM*> 0, Column (2) consists of all the observations from Table 3 in which *NUM*> 1, etc. *NUM*, the independent variable of interest, is defined as the number of forecasts issued by other analysts between the analyst's earlier forecast and his revised forecast. For definitions of the control variables, see Appendix A. Robust standard errors are clustered by quarter and presented in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table 9: Market Reaction to Forecast Revisions

	(1)	(2)	(3)	(4)
	CAR regressions		$\frac{\text{EPS}-f^R}{P}$ regressions	
REVP	2.43***	1.07**	0.43***	0.29***
	[0.27]	[0.46]	[0.04]	[0.06]
REVP × *FCE*	-0.15	-0.20	0.14**	0.13**
	[0.42]	[0.42]	[0.06]	[0.06]
REVP × *Firm Error*		0.25		0.03
		[0.42]		[0.06]
REVP × *Elapsed Time*		-0.16		0.09*
		[0.40]		[0.05]
REVP × *Horizon*		2.18***		-0.05
		[0.40]		[0.05]
REVP × *Brokerage Size*		0.53		0.27***
		[0.41]		[0.06]
Intercept	-0.00	-0.00	-0.00	-0.00
	[0.00]	[0.00]	[0.00]	[0.00]
R^2	0.01	0.01	0.03	0.03
N	15,529	15,529	15,529	15,529

In Columns 1 and 2, I regress three day CARs around forecast revisions onto various characteristics of the revisions. In Columns 3 and 4, I regress price scaled forecast errors onto these same variables. *REVP* is defined as (winsorized) $\frac{f^R-f}{P}$, where f^R is the analyst's revised forecast, f is his earlier forecast, and P is the firm's stock price two days prior to the day the revised forecast is issued. *FCE* is the analyst's *NUM* coefficient based on his forecast revisions in the five years prior to the year in which the revised forecast is issued, scaled using the procedure described in (27) from Section 6. The methodology for estimating *NUM* coefficients at the analyst-year level is described in Section 5.1.2. *Raw Horizon* is the number of days between the date the forecast is issued and the forecast period end date, *Raw Firm Error* is the average value of $\frac{|f-\text{EPS}|}{P}$ among i's forecasts for the firm's earnings issued between one and eight quarters before the quarter the revision is issued, where the forecasts have *Raw Horizon* between 31 and 90 days, *Raw Elapsed Time* is the number of days between the forecast and the next most recent forecast for the firm's earnings, and *Raw Brokerage Size* is the number of analysts employed by the analyst's broker in the year the forecast revision is issued. I follow Clement and Tse (2003) in using (27) (from Section 6) to scale the raw variables to obtain the (scaled) variables *Firm Error*, *Horizon*, *Elapsed Time*, and *Brokerage Size*. My results are qualitatively similar when I include (the non-interacted variables) *FCE*, *Firm Error*, *Elapsed Time*, *Horizon*, and *Brokerage Size* as additional controls (along with the interaction variables).