

Forecast Timing and Reputational Concerns: Theory and Evidence

Mark A. Chen *

April 26, 2007

Preliminary: comments welcome

*Robert H. Smith School of Business, University of Maryland, and the Office of Economic Analysis, U.S. Securities and Exchange Commission. E-mail: machen@rhsmith.umd.edu. Phone: (301) 405-2171. I am grateful to seminar participants at the Office of Economic Analysis for their helpful comments and feedback on an early version of this work. The Securities and Exchange Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. The views expressed herein are those of the author and do not necessarily reflect the views of the Commission, the Commissioners, or members of the staff.

Abstract

I develop a model in which analysts strategically time and exaggerate their forecasts to convince the public that they are skilled. The model's key predictions are that an analyst will tend to forecast later if he has a better ex ante reputation, and he will tend to forecast earlier if the quality of public information is lower. Using a database of individual analysts' earnings forecasts, I empirically test and find support for both of these predictions.

1 Introduction

Among the most widely-used research outputs of stock analysts are their forecasts of corporate earnings. Clearly, the extent to which such forecasts are directly useful to market participants depends not only on accuracy, but also on timeliness. Empirical studies have shown that forecast timing varies systematically across individual analysts and is related to observable analyst characteristics such as forecasting experience, past accuracy, or employer prestige (see, e.g., Lys and Sohn(1990), Stickel (1992), Mikhail, Walther, and Willis (1999), Hong, Kubik, and Solomon (2000), Hong and Kubik (2003), and Ivkovich and Jegadeesh (2004)). Later forecasts appear to be more accurate (Brown (1991)). In addition, forecast timeliness seems to be related to stock price impact, and timeliness can serve as an effective means of ranking analysts (Cooper, Day, and Lewis (2000)). To date, however, there has been little direct investigation of the underlying causes of forecast timing, and key questions remain. For instance, why would a rational analyst choose to forecast earlier despite an expected loss in accuracy? How does forecast timing vary with key aspects of the analyst’s information environment?

This paper develops and analyzes a model in which reputational concerns are the chief determinant of analysts’ timing decisions. I consider a stylized, multiperiod setting in which one or more analysts attempt to forecast a firm’s uncertain earnings. I assume that analysts may differ in their skill level, i.e., their ability to obtain private signals about earnings-relevant information. Initially, the public is uninformed about an analyst’s skill level. Because an analyst cares not only about the intrinsic accuracy of a forecast but also about the public’s perception that he is skilled, he may time his forecast or bias it strategically in equilibrium.

Over time, information arrives to the public about the components of earnings, and some of this information is correlated with analysts’ private information. The main result of the paper is that forecast timing, in addition to forecast bias, can serve as a credible means of conveying skill levels. In particular, I show that while exaggeration (relative to posterior beliefs) is one means by which a skilled analyst can distinguish himself, issuing forecasts earlier is often an even more effective way for an analyst to convince the public that he is skilled. The key idea underlying this result is that forecasting earlier entails a cost in terms of foregone accuracy, but the skilled individual—relative to his unskilled counterpart—faces a lower cost from moving earlier due to the fact that he already possesses more of the information.

The model illustrates that the optimal forecast delay of the skilled analyst will depend on the fundamental tradeoff between two considerations. On the one hand, waiting longer gives an analyst an expected improvement in his information due to the information that arrives publicly about the systematic component of earnings. On the other hand, by waiting he also faces the hazard that his informational advantage about the firm-specific earnings component will be impaired, thereby prohibitively increasing the cost of separation. I characterize the optimal forecast delay in equilibrium as a function of the parameters of the analyst's objective function, the quality of the public information flow over time, and the analyst's initial reputation (i.e., the market's prior belief about analyst skill).

The characterization of equilibrium permits a direct examination of how forecast timing depends cross-sectionally on the key parameters of the model. The comparative statics lead to two main testable implications. First, the analyst's ex ante reputation should be directly related to the optimal forecast delay. This result can be seen as a reflection of the idea that analysts who already enjoy a favorable assessment by the market will not want to "go out on a limb" if there is little to be gained by forecasting early. Second, the forecast delay should be greater if the quality of public information is higher. An increase in the relative informational advantage of the skilled analyst leads him to forecast earlier as the expected per-period costs of losing his informational advantage become higher.

I test the main cross-sectional predictions of the model for forecast timing with a large sample of individual analysts' quarterly earnings-per-share (EPS) forecasts from 1995 to 2003. To obtain an empirical proxy for individual analysts' ex ante reputation, I use information gathered from Institutional Investor magazine's All-America Research Team survey. Published annually, the All-America survey recognizes a select group of analysts for their general research and forecasting ability based on the responses of a large number of money managers and institutional investors. I conduct a duration analysis to examine the effects of different covariates on the hazard rate of individual analysts' forecasts. The tests reveal that, controlling for salient firm, employer, and analyst characteristics (including workload, employer size, and forecasting experience), ex ante reputation is positively and significantly related to forecast delay. In addition, forecast delay is positively related to measures of the quality of information flow and negatively related to measures of a skilled analyst's informational advantage. Overall,

the results strongly support the cross-sectional predictions, particularly the reputation-based prediction, which would be difficult to explain with a story devoid of reputational concerns.

While the overall modeling approach taken in this paper appears to be novel, a number of related aspects have been studied in the literature on strategic timing of actions. For instance, Gul and Lundholm (1995) consider a continuous-time model in which forecasters have private information. In their model, equilibrium delay is determined by the tradeoff between the benefits of information externalities and the cost of delay. However, in contrast to the present model, their model does not consider reputation, and it assumes exogenous preferences over the extent of delay. Gale and Chamley (1994) are generally concerned with social learning. In their model, they show how agents with private information will strategically delay in order to learn from the private information of others. This model differs from the one in the current paper in that it deals with a discrete action space, and reputational incentives do not play an explicit role. Moreover, unlike the model of the current paper, it does not readily deliver implications about the cross-sectional determinants of delay.

A number of other papers consider the issue of strategic behavior among forecasters and other agents. Much of this literature has focused on how reputational concerns can give rise to herding incentives (see, e.g., Scharfstein and Stein (1990), Trueman (1994), and Graham (1999)).¹ Herding per se is not a central feature of the present model since analysts' action spaces are continuous. However, the reputational herding literature contains the idea (upon which the current paper builds) that rational agents will strategically attempt to influence the market's posterior assessments about ability. A separate branch of the literature has studied in detail how forecasters may bias their forecasts in settings of reputational cheap talk. For example, Ottaviani and Sorensen (2005a, 2005b) show, in a static, one-shot forecasting game with a somewhat richer private information structure than the one considered here, that forecasters will tend to exaggerate when they are well-informed. This result parallels the forecast exaggeration that arises in the current paper if the general model is restricted to a one-period game.

¹ See also Graham (1996) for empirical evidence on herding among investment newsletters and Welch (2000) and Jegadeesh et al. (2007) for empirical evidence on herding in security analysts' stock recommendations. Hong, Kubik, and Solomon (1999) document that analyst forecasts tend to herd in ways that vary systematically with forecasting experience. Bernhardt, Campello, and Kutsoati (2005) also study herding among forecasts (not necessarily arising from reputational concerns).

Finally, other papers have considered how the actions and forecasts of privately-informed agents may exhibit boldness or bias over time in response to dynamic reputational concerns (see, e.g., Prendergast and Stole (1996), Chemmanur and Fulghieri (1995), and Clarke and Subramanian (2006)). The current model differs in that it pertains mainly to short-term reputation formation. Extending the results of the current paper to study dynamic reputation formation would seem to be an important avenue for future inquiry.

The rest of the paper is organized as follows. Section 2 outlines a basic version of the model for a single analyst, solves for the equilibrium, and examines comparative statics properties of the equilibrium. Section 3 discusses the extension of the analysis to the case of multiple analysts. Section 4 discusses and implements several empirical tests of the model’s cross-sectional predictions. Section 5 provides a summary and conclusion.

2 The Model

2.1 The Basic Model

Consider a setting in which an analyst faces the task of forecasting a company’s upcoming earnings result. The forecasting season extends over periods $1, 2, \dots, T$ where $T > 2$, and the company reports its earnings publicly in period $T + 1$. Earnings are equal to the sum of two unobserved components:

$$\pi = x + y.$$

The y component captures the influence of systematic factors that arise from economy-wide sources, while the x component reflects the firm-specific part of earnings. At the start of period 0 (the period preceding the start of the forecasting season), market participants share a common prior belief that x and y are zero-mean, independently normally distributed variables with variances of σ_x^2 and σ_y^2 , respectively.

During period 0, the analyst can, depending on his type, acquire additional information about the firm-specific earnings component x . There are two possible types of analyst: “skilled” (S) and “unskilled” (U). Whereas an unskilled analyst receives no private information about x , a skilled analyst, through his research, is able to obtain a private, noisy signal $v_0 = x + \epsilon_0$, where $\epsilon_0 \sim N(0, \sigma_\epsilon^2)$ and ϵ_0 is independent of x .²

²This signal may be thought of as the cumulation of all insights that the analyst has derived about

Whether an analyst is skilled or not is known only to him at the start of the game, but other market participants share a common prior belief that, with probability $\theta \in (0, 1)$, he is skilled. This can be thought of as the analyst's *ex ante* reputation.

At the beginning of each period $t = 1, 2, \dots, T$, information about one or both components of earnings arrives publicly. First, there is a noisy signal $w_t = y + \eta_t$ about the systematic earnings component, where $\eta_t \sim N(0, \sigma_\eta^2)$, where η_t is independent of x , y , and ϵ_0 , and where the η_t s are serially independent. Second, with probability $\gamma \in (0, 1)$, a public signal arrives that is informative about the firm-specific earnings component x . For simplicity, I assume that this signal, if it arrives, completely reveals x .³

At the end of period t , after his beliefs have been updated in light of the new information that has arrived, the analyst can issue a forecast $F \in \mathbf{R}$ if he has not already done so.⁴ It is assumed throughout that the analyst and investors use observed signals and forecasts to update their prior beliefs in a Bayesian fashion.

In this model, a key idea is that the analyst cares about both his reputation (i.e., the market's posterior beliefs about whether he is skilled) as well as about his intrinsic accuracy in forecasting π . To capture these two objectives in a simple fashion, I assume that the analyst, regardless of type, is an expected utility maximizer and seeks to maximize his expected final payoff:

$$U = aPr(\text{Analyst is skilled}|\Omega_T, F, t, \pi) - b(F - \pi)^2$$

where $a, b > 0$, F is the analyst's forecast, t is the period in which the forecast was issued, and Ω_T is the sequence of all earnings signals that have arrived over the forecasting season. (Throughout the analysis, we shall denote by Ω_t the sequence of public earnings-relevant signals that have arrived through period t .) The first term in the objective function is a constant multiple of the market's posterior beliefs about the probability that the analyst is skilled. It captures the notion, central to the literature on reputational

the firm prior to the current forecasting quarter.

³The key property is that the signal substantially decreases the disparity between the two types of analysts. For example, the signal might correspond to management-issued earnings guidance, a pre-announcement, or an earnings warning. Allowing in the model for imperfect signals about x would complicate the analysis somewhat, but qualitatively similar conclusions could be drawn in a probabilistic sense.

⁴The assumption that the analyst only issues a single forecast during the quarter is consistent with real-world forecasting in which it is uncommon for an analyst to issue more than one forecast in a quarter.

herding, that analysts may be subject to career concerns and may want to convince the labor market or investors that they are skilled, whether or not they actually are. The second term reflects a concern for absolute accuracy of the forecast. Note that the functional form does not preclude dependence of ex post reputation on forecast accuracy.⁵

2.2 Equilibrium Forecasting

In this section, we characterize equilibrium forecasting behavior in this multiperiod game. We begin by examining how the expected error of a forecast depends upon the information available to the analyst and upon the nature of the forecast itself. Let x^S and σ_x^2 denote the skilled analyst's private posterior beliefs about the mean and variance of x at the end of period 0, i.e., after the private signal v_0 has arrived (this will continue to be the analyst's posterior beliefs in subsequent periods as long as x has not been revealed). From the well-known Bayesian updating property of normal distributions (see, e.g., DeGroot (1970)), we have that $x^S = \sigma_x^2 v_0 / (\sigma_x^2 + \sigma_\epsilon^2)$ and $\sigma_x^2 = \sigma_x^2 \sigma_\epsilon^2 / (\sigma_x^2 + \sigma_\epsilon^2)$.

Since there is no private information concerning the systematic earnings component, in each period the posterior beliefs about y will be the same across the two analyst types. The posterior will be normally distributed. Let y_t and $\sigma_{y,t}^2$ denote the common posterior mean and variance of y , respectively, at the end of period $t = 1, \dots, T$. In general, the posterior mean y_t will depend on the signal realizations and the previous period's posterior mean.⁶ The posterior variance, however, will be independent of the

⁵As a practical matter, analysts may have explicit or implicit incentives that reward them for intrinsic accuracy. For instance, institutional clients that use an analysts' forecasts may value accuracy independently of the analyst's reputation. Also, brokerage houses may provide incentives to encourage forecast accuracy. Indeed, the Global Settlement on Research Conflicts of Interest reached in 2003 between securities regulators and large Wall Street firms stipulated that lead analysts should have a substantial portion of their compensation tied to the accuracy of their research.

⁶The posterior mean at the end of period t is given by a weighted average of the period- t signal and the previous period's posterior mean:

$$y_t = \left(\frac{\sigma_y^2}{t\sigma_y^2 + \sigma_\eta^2} \right) w_t + \left(\frac{(t-1)\sigma_y^2 + \sigma_\eta^2}{t\sigma_y^2 + \sigma_\eta^2} \right) y_{t-1}.$$

signal realizations. Indeed, from successive bayesian updating, we have that

$$\sigma_{y,t}^2 = \frac{\sigma_y^2}{t\sigma_y^2 + \sigma_\eta^2}. \quad (1)$$

Note that, absent any reputational concerns, the analyst would simply wait as long as possible and issue a forecast equal to his posterior mean on π (such a forecast would minimize the expected forecast error). In this circumstance, Equation (1) implies that the expected error will depend only on whether or not x is revealed before period T , and not on the signal realizations y_1, y_2, \dots, y_T . In general, however, reputational concerns may cause an analyst to choose earlier forecasts or biased ⁷ forecasts to optimally influence the market’s perception of skill.

To characterize the equilibrium of the multiperiod game, we use the concept of Perfect Bayesian Equilibrium (PBE), which requires that (i) wherever possible, participants update their beliefs rationally in a Bayesian fashion given the history of play; and (ii) each participant’s equilibrium strategies are optimal at each stage of play, given each participant’s beliefs. We will focus on pure-strategy PBEs in which out-of-equilibrium beliefs held by the market satisfy “reasonable” restrictions. In particular, we require that the market discount the possibility that one type of analyst was responsible for an out-of-equilibrium action if that type would never benefit from the action but the other type of analyst might conceivably benefit. This requirement is an implication of the standard “intuitive criterion” (Cho and Kreps (1986)) that has been widely used in the analysis of signaling-game equilibria.

A priori, the set of candidate equilibria appears quite large: the two analyst types may conceivably select pooling and/or separating strategies that depend in various ways on the entire history of signals. Fortunately, as a first step we can establish the following key result, which substantially narrows down the potential set of equilibria.

Lemma 1 *In any equilibrium, if $v_0 \neq 0$, then the skilled and unskilled types will not issue the same forecast in period t unless x has been revealed before the end of period t .*

Lemma 1 shows that pooling of types can only occur in a period t if either the skilled analyst’s private signal is exactly 0 or if x has been revealed to the public. Put differently, pooling in a period cannot be sustained if the skilled analyst’s private

⁷Throughout, “bias” will pertain to how far away is an analyst’s forecast from his posterior, not how far away it is from realized earnings or from a consensus forecast.

posterior belief differs from the unskilled (public) posterior. Intuitively, what makes separation a credible alternative is the continuity of the skilled analyst's action space in conjunction with his informational advantage. When he has a more precise estimate of the mean level of earnings than his unskilled counterpart, he can do better than pooling by biasing his forecast in the same direction as his private information, thus taking on an inaccuracy penalty that the unskilled analyst is unable to profitably bear. If, however, x is revealed, then the private informational advantage is lost, and as a result the skilled type is forced to pool with the unskilled.⁸

We can now describe equilibrium in the basic multiperiod model. As the following result shows, the timing and forecasting behavior of each type of agent are uniquely characterized in terms of the skilled analyst's signal v_0 and the underlying parameters of the model. (For simplicity, we set aside the zero-probability case where $v_0 = 0$.)

⁸As it happens, multiple, Pareto-ranked pooling equilibria may arise in subgames following the revelation of x . In these equilibria, the skilled and unskilled types become equally well informed about x and y . We will assume throughout the analysis that when multiple, strictly-pareto-ranked pooling equilibria of this type are possible, then the Pareto-dominant one occurs.

Proposition 1 *Suppose that $v_0 \neq 0$. Then the equilibrium forecasting behavior of each type of analyst is given by the following:*

Type U:

The analyst waits until the end of period T to forecast. If the firm-specific earnings component x has not been revealed by that time, then the analyst issues a forecast equal to

$$F^U = y_T. \quad (2)$$

Otherwise, if x has been revealed before the end of period T , then the analyst issues a forecast equal to

$$F^U = y_T + x. \quad (3)$$

Type S:

For $t = 1, \dots, T$ define

$$D_t \equiv \sqrt{\frac{(1 - [1 - (1 - \gamma)^{T-t}] \theta) a}{b} + \sigma_{y,T}^2 - \sigma_{y,t}^2 + [(1 - \gamma)^{T-t} - 1] \sigma_x^2} \quad (4)$$

and let m be the smallest positive integer between 1 and T such that

$$\begin{aligned} a(\gamma - \gamma\theta) &> b(\sigma_m^2 - \gamma\sigma_T^2 - (1 - \gamma)\sigma_{m+1}^2 + \gamma\sigma_S^2) \\ &+ b(\max\{D_m - |x^S|, 0\})^2 - b(1 - \gamma)(\max\{D_{m+1} - |x^S|, 0\})^2. \end{aligned} \quad (5)$$

The analyst waits until the end of period m . If x has not been revealed by that time, the analyst issues a forecast at the end of period m equal to

$$F^S = \begin{cases} y_m + x^S & \text{if } |x^S| > D_m \\ y_m + D_m & \text{if } 0 < x^S \leq D_m \\ y_m - D_m & \text{if } -D_m \leq x^S < 0 \end{cases}$$

Otherwise, if x has been revealed before the end of period m , then the analyst waits until the end of period T and issues a forecast equal to

$$F^S = y_T + x. \quad (6)$$

Proposition 1 shows that the two types of analyst will typically exhibit very different behavior. The unskilled analyst will always wait as late as possible to forecast. Intuitively, Lemma 1 shows that the unskilled analyst's type is bound to be learned in equilibrium (unless x is revealed at some point), and so waiting until period T is always best because not only does this maximize the number of signals about y , but it also maximizes the chance that he will be able to pool and get a reputational payoff of θ instead of zero.

The skilled analyst, on the other hand, will generally forecast early and distinguish his type. According to Proposition 1, there is an increasing sequence of cutoff levels D_1, D_2, \dots, D_T that describes the minimum distance away from the public posterior that a skilled forecast would need to be in periods $1, \dots, T$ to prevent mimicry. (Note that the D_t s are independent of the signal realizations v_0 and w_1, w_2, \dots, w_T ; only the posterior variances matter for the D_t s.) When the skilled analyst's information v_0 is extreme, he can separate simply by efficiently forecasting his posterior. However, when his private posterior is close to the public posterior, he will choose to exaggerate in order to meet the no-mimicry condition.

Finally, the equilibrium forecasting time for the skilled analyst is governed by the arbitrage condition (5), which provides a necessary and sufficient condition for the analyst to prefer to forecast immediately in a period. Whereas necessity follows automatically, sufficiency is shown in the proof of Proposition 1 via backward induction arguments (see Appendix for details). This arbitrage condition reflects a fundamental tradeoff in the decision of whether to move immediately or wait until later. The tradeoff can be described as follows. Consider a period $t \leq T$. On the one hand, so long as x has not been revealed yet, forecasting today allows the skilled analyst to separate for sure; he does not have to run the risk that his informational advantage will be lost via revelation of x . On the other hand, forecasting today causes the skilled analyst to forego additional signals on y . Over time, the incremental benefit of getting an additional signal on y decreases, and so at some point the skilled analyst will find it advantageous not to wait.

2.3 Comparative Statics: Equilibrium Forecast Delay

In this section we focus on how the skilled analyst's equilibrium delay depends on the key model parameters. From Proposition 1, it is clear that the delay depends only on the D_t s, the minimum distances from the public posteriors needed to separate, and on

the stringency of inequality (5) in Proposition 1, which captures the skilled analyst's tradeoff between forecasting or waiting in period t . Clearly, some parameter values (e.g., small values of θ or extreme values of v_0) dictate corner solutions in which the analyst forecasts in period 1 or period T . We therefore restrict attention to the region of the parameter space in which there is an interior delay value. Some caution is warranted in interpreting the comparative statics since the main endogenous variable of interest—delay—will not vary continuously in the underlying parameters. Therefore, we implicitly assume that T is sufficiently large that the comparative statics can be given a meaningful interpretation.

Proposition 2 *On average, the equilibrium forecast delay of the skilled analyst*

- (i) *increases in θ , the analyst's ex ante reputation;*
- (ii) *decreases in a/b ;*
- (iii) *decreases in γ when $|x^S| \geq D_T$;*
- (iv) *increases in σ_ϵ^2 .*

The proof of Proposition 2 follows primarily from an inspection of Proposition 1 and from the fact that the optimal delay depends only on the sequence of D_t s and the forecast-or-wait condition (5). A rough intuition for the results can be given as follows. For result (i), an increase in θ reduces the skilled analyst's opportunity cost of separating. In other words, pooling becomes more advantageous as θ increases. Since the risk of pooling is inherent in waiting, a rise in θ tilts the per-period tradeoff in favor of waiting. In addition, there is a smaller, indirect effect whereby an increase in θ increases the opportunity cost to the unskilled analyst of moving earlier, and this alters the cutoff levels D_1, D_2, \dots, D_T . Result (ii) is an expression of the fact that separation becomes relatively more valuable compared to pooling as a increases relative to b , and moving earlier minimizes the chance that the skilled analyst will be required to pool with the unskilled analyst.

Result (iii) captures the idea that the greater the chance that x is revealed next period, the costlier it becomes for the skilled analyst to wait as the reputational payoff to pooling is always dominated by the reputational payoff to separating. Finally, (iv) might appear at first to be a somewhat counterintuitive result. It states that delay is decreasing in the quality of the skilled analyst's private information. However, this result follows simply from the fact that the private signal v_0 confers a forecasting advantage to the skilled analyst in the face of uncertainty. The larger is this advantage, the costlier

it becomes to wait due to the risk that the advantage is lost upon revelation of the firm-specific component x .

3 The Case of Multiple Analysts

We now consider the generalization of the basic model to the case of multiple analysts. The main difference from before is that now analysts can independently observe private signals and possibly learn from each others' public forecasts. Thus, suppose that each of $J > 1$ analysts can be either skilled or unskilled. For $j = 1, \dots, J$, analyst j , if skilled, receives a signal $v_0^j = x + \epsilon_0^j$, where $\epsilon_0^j \sim N(0, \sigma_\epsilon^2)$ with the ϵ_0^j s independent. As before, each analyst must issue a forecast by period T or before.

This generalization of the basic model gives rise to two issues that complicate the characterization of equilibrium. First, an analyst can learn not only from the forecasts of other analysts, but also from others' timing decisions. For example, if an analyst's equilibrium strategy involves forecasting immediately upon receiving an extreme signal, then the analyst's decision to wait will cause the other analysts to truncate their beliefs about his posterior. A second complicating factor is that the forecasts themselves will in general be nonlinear functions of private information. Even in the single-analyst model, as Proposition 1 showed, forecasts may involve bias and exaggeration. In the general case, individual analysts update their posteriors based on others' potentially distorted forecasts, and hence individual posterior beliefs over x will in general be non-Gaussian.

The additional complications notwithstanding, we can still establish results similar to those in Section 2.2. The following result shows that the equilibrium timing behavior of any unskilled analyst is to wait as long as possible, and equilibrium timing of a skilled analyst is governed by an arbitrage condition parallel to that of Proposition 1:

Proposition 3 *Suppose that $v_0^j \neq 0$ for $j = 1, \dots, J$. Then the equilibrium forecasting behavior of each type of analyst satisfies the following:*

Type U, Analyst j:

The analyst will wait until the end of period T to forecast. If the firm-specific earnings component x has not been revealed by that time, then the analyst issues a forecast equal to

$$F^{U,j} = y_T + E[x|h_T]. \tag{7}$$

where h_T denotes the public history of forecasts through period T . Otherwise, if x has been revealed before the end of period T , then the analyst issues a forecast equal to

$$F^{U,j} = y_T + x. \quad (8)$$

Type S, Analyst j:

For $t = 1, \dots, T$ define

$$D_t \equiv \sqrt{\frac{(1 - [1 - (1 - \gamma)^{T-t}] \theta) a}{b} + \sigma_{y,T}^2 - \sigma_{y,t}^2 + (1 - \gamma)^{T-t} E[\text{Var}[x|h_T]|h_t] - \text{Var}[x|h_t]}. \quad (9)$$

The analyst will wait until the first period $m \in 1, \dots, T$ such that

$$\begin{aligned} a(\gamma - \gamma\theta) &> b(\sigma_m^2 - \gamma\sigma_T^2 - (1 - \gamma)\sigma_{m+1}^2 + \gamma\text{Var}[x|h_m, v_0^j]) \\ &+ b(\max[D_m - |E[x|h_m, v_0^j]|, 0])^2 \\ &- E[b(1 - \gamma)(\max\{D_{m+1} - |E[x|h_m, v_0^j]|, 0\})^2 | h_m, v_0^j]. \end{aligned} \quad (10)$$

If x has not been revealed by that time, the analyst issues a forecast at the end of period m equal to

$$F^{S,j} = \begin{cases} y_m + E[x|h_m, v_0^j] & \text{if } |E[x|h_m, v_0^j]| > D_m \\ y_m + D_m & \text{if } 0 < E[x|h_m, v_0^j] \leq D_m \\ y_m - D_m & \text{if } -D_m \leq E[x|h_m, v_0^j] < 0 \end{cases}$$

Otherwise, if x has been revealed before the end of period m , then the analyst waits until the end of period T and issues a forecast equal to

$$F^{S,j} = y_T + x. \quad (12)$$

I conjecture that comparative statics results, parallel to those in Proposition 2, can be shown to hold in the general case. A property that is sufficient to ensure such results hold is that a change in one of the underlying parameters (say θ) will, for a given set of initial private signal realizations, shift all skilled analysts' forecasts temporally in the same direction (either earlier or later).

4 Empirical Tests

The comparative statics results outlined in Proposition 2, in particular results (i) and (iv), give rise to two empirical implications that are directly testable. First, the more favorable is the market's initial assessment of an analyst's skill, the later in the forecasting season the analyst will tend to forecast. Second, the lower the quality of information that is publicly available about firm-specific earnings, the earlier an analyst will tend to forecast. In this section, I test these two implications using a panel dataset of individual analysts' quarterly earnings-per-share (EPS) forecasts. Focusing on quarterly EPS forecasts has the distinct advantage of yielding a large number of forecasting seasons of similar length in which (a) information arrival is most likely uncertain; and (b) analysts' forecasts are most likely to have information content.

4.1 Data

To construct the sample of forecasts, I use the Institutional Brokers Estimate System (I/B/E/S) Detail History File. The I/B/E/S Detail History File identifies forecasts at the level of the individual analyst and includes information on the dates of forecast issuance. To facilitate combining the forecast data with other variables used in the analysis, I limit the sample to current-quarter EPS forecasts issued during 1995-2003 for which associated I/B/E/S stock price data and earnings realizations data are available. I also require that a forecast be issued after the previous quarter's earnings report but during the three months constituting the current reporting quarter. Finally, I follow much of the literature and include only forecast quarters that end in March, June, September, or December. This leaves a total of 298,129 observed forecasts.

Testing the two main implications of the model requires measures of analysts' ex ante reputation and measures of the potential informational advantage to being skilled. For the reputation measure, I gather data from Institutional Investor magazine's All-America Research Survey over the 1995-2003 period. The survey, published annually in October, recognizes several hundred sell-side analysts who were voted as top stock researchers in a subjective poll of institutional money managers. Since only analysts' last names and first initials are available through I/B/E/S, I use the NASD BrokerCheck database (available online at www.nasd.com) to resolve ambiguities in matching I/B/E/S names with All-America analysts. Thus, I am able to construct a

variable that indicates whether a forecasting analyst was named to the first, second, third, or honorable mention team in the All-America survey.

In order to capture the potential informational advantage to being skilled, I use firms' market capitalization, which has been used as a proxy for the difficulty with which earnings can be forecasted. In some of the tests I also use analyst coverage, which should be directly related to the total amount of information produced about idiosyncratic earnings.

The empirical tests also control for other characteristics of an analyst's research environment that may be related to his propensity to delay, such as his research experience, his workload intensity, or his employer's resources (see, e.g., Clement (1999), Jacob, Lys, and Neale (1999)). I construct a variable that measures an analysts' forecasting experience in terms of the length of time since he first appeared in the I/B/E/S database.⁹ This experience measure takes into account all types of I/B/E/S forecasts made by an analyst, including long-term earnings growth forecasts, quarterly forecasts, and annual forecasts. Analysts' workloads are measured as the number of different companies or, alternatively, as the number of different I/B/E/S industry classifications (2-digit SIG codes) for which an analyst issued forecasts in a given year. To capture the resources that are available to an analyst via his brokerage firm, I use the total number of forecasters employed by the brokerage firm in a given year.

4.2 Results

Previous studies have documented that analysts often revise their earnings forecasts within the first few days after the previous quarter's earnings release to mechanically reflect the new earnings information (see, e.g., Cooper, Day, and Lewis (2000)). Since such forecasts are likely to contain little of an analyst's private information, I exclude from the sample any forecasts occurring within the three days after the previous quarter's earnings report. This reduces the sample size by 53,671 observations, leaving a total sample of 244,458 forecasts.

⁹The fact that I/B/E/S forecast data before January 1981 are generally not available leads to the well-known truncation problem wherein forecasting experience may be understated for a subset of analysts. This problem does not appear to be a severe one, however: the main qualitative results are unchanged when I exclude analysts whose observed experience dates back to 1981.

Table 1 presents some basic descriptive statistics for the sample of forecasts. On average, analysts in the sample have over 6 years' worth of forecasting experience and cover over 15 companies, albeit only in a few industries. Only about 4 percent of all forecasts are issued by first-team All-Americans, while 11 percent are issued by analysts named to the first, second, or third All-America teams. The average and median delay between the last earnings report and the current quarter's forecast is approximately one month. Across all forecasts, the typical company has coverage from 8 analysts and has a market capitalization of about 1.5 billion dollars. As Panel C indicates, the average forecast delay appears to have decreased over the sample period by about 5 percent from 31.57 days in 1995 to 29.83 days in 2003.

To test the models' two predictions pertaining to forecast delay, I estimate Cox proportional hazards models in which the main covariates are the logarithm of a company's market capitalization and a binary variable indicating whether an analyst was a member of the All-America first team in the latest annual survey. The regressions also control for the extent of analyst coverage, the log of brokerage firm size (as measured by total number of I/B/E/S analysts employed), and the forecasting season length, i.e., the number of days between the previous earnings report and the end of the current quarter. In these regressions, only an analyst's first forecast during the valid forecasting season is included (this requirement further reduces the sample by 49,733 observations).

Table 2 reports the results of these hazard-rate regressions. As shown in Column (1), the hazard rate is significantly negatively related to All-America status as well as to firm size. In other words, better ex ante reputation is associated with a greater delay, as is larger firm size. These relationships are highly statistically significant. In column (2), replacing the first-team All-America measure with the more inclusive measure of All-America status does not change the qualitative conclusions. Columns (3) and (4) show that when the regressions are augmented with controls for the analyst's workload and forecasting experience, the coefficient estimates for firm size and All-America status remain negative. The coefficient on season length is, as expected, significantly negative in all regressions since the unit of measurement for the hazard rate is constant across all observations.

Table 3 reports the results of hazard-rate regressions for subsamples based on analyst coverage. Included as covariates in these regressions are first-team All-America status, firm size, brokerage size, workload, and forecasting experience. As is evident from the

table, the coefficients on All-America status and firm size remain significantly negative across each of the 10 strata. Overall, then, the results strongly support the view that ex ante reputation and the quality of public information are both positively related to delay.

5 Conclusion

This paper develops a stylized, multiperiod model in which analysts with private information about their forecasting skill face reputational concerns that drive them to forecast strategically. The (essentially) unique equilibrium of this game is characterized as one in which a skilled analyst will choose to forecast early and forego the benefit of public information in order to credibly convey his skill. Two key predictions emerge from the comparative statics analysis: first, analysts will tend to forecast later when their ex ante reputations are better; and second, analysts will tend to forecast earlier when their potential informational advantage is greater. Empirical tests using proxies for ex ante reputation and the quality of public information strongly support the model's predictions. While other explanations for observed forecast timing patterns may exist, the theory presented here explains cross-sectional patterns that a different theory without reputational concerns would be hard-pressed to explain.

Appendix

Proof of Lemma 1.

Consider any PBE in pure strategies. Suppose by way of contradiction that, for some period t , the following are true: (1) by the end of period t , x has not been revealed yet; (2) the skilled analyst has updated his posterior mean on x to $x^S \neq 0$ after having observed the private signal $v_0 \neq 0$; and (3) both types issue the same forecast \tilde{F} in period t . We proceed to argue that the skilled analyst can do strictly better by deviating to a strategy whereby he issues a forecast \tilde{G} in period t instead of issuing \tilde{F} .

Let $\hat{\theta}_{T+1}(F, t, \pi, \Omega_T)$ denote the market's final posterior belief of the probability that the analyst is skilled after having observed realized earnings π , forecast F in period t , and the sequence of public signals Ω_T . Thus, the reputational payoff that the skilled analyst expects from forecasting F in period t is given by

$$\mu_F^S \equiv E[\hat{\theta}_{T+1}(F, t, \pi, \Omega_T) | \Omega_t, v_0] \quad (13)$$

and the reputational payoff that the unskilled analyst expects from forecasting F in period t is given by

$$\mu_F^U \equiv E[\hat{\theta}_{T+1}(F, t, \pi, \Omega_T) | \Omega_t]. \quad (14)$$

Observe that since both types are using pure strategies and pooling their forecasts in equilibrium, the market must assign positive probability weight to the unskilled type. In other words, μ_F^S must be strictly less than 1. Now define functions

$$d_1(z) = a(1 - \mu_F^S) + b(\tilde{F}^2 - z^2) + 2by_t(z - \tilde{F}) \quad (15)$$

$$d_2(z) = a(1 - \mu_F^S) + b(\tilde{F}^2 - z^2) + 2b(x^S + y_t)(z - \tilde{F}) \quad (16)$$

Note that $d_1(\tilde{F}) = d_2(\tilde{F}) = a(1 - \mu_F^S) > 0$. Furthermore, we also have that $d_1'(z) = -2bz + 2by_t$ and $d_2'(z) = -2bz + 2bx^S + 2by_t$. Thus, it follows that if $x^S > 0$ ($x^S < 0$), then there exists a sufficiently positive (sufficiently negative) number \tilde{G} such that $d_1(\tilde{G}) < 0$ and $d_2(\tilde{G}) > 0$.

Consider now the expected payoff to each type of analyst from deviating to the out-of-equilibrium forecast \tilde{G} at time t . The expected net benefit to the unskilled analyst from deviating to \tilde{G} is equal to

$$B^U \equiv a\mu_{\tilde{G}}^U - E[b(\tilde{G} - \pi)^2|\Omega_t] - [a\mu_{\tilde{F}}^U - E[b(\tilde{F} - \pi)^2|\Omega_t]] \quad (17)$$

$$= a(\mu_{\tilde{G}}^U - \mu_{\tilde{F}}^U) - b(\tilde{G}^2 - 2\tilde{G}y_t + \sigma_x^2 + \sigma_{y,t}^2 + y_t^2) \quad (18)$$

$$+ b(\tilde{F}^2 - 2\tilde{F}y_t + \sigma_x^2 + \sigma_{y,t}^2 + y_t^2) \\ = a(\mu_{\tilde{G}}^U - \mu_{\tilde{F}}^U) + b(\tilde{F}^2 - \tilde{G}^2) + 2by_t(\tilde{G} - \tilde{F}) \quad (19)$$

$$\leq d_1(\tilde{G}) \quad (20)$$

$$< 0. \quad (21)$$

Thus the unskilled analyst can never gain by deviating to \tilde{G} . However, it is possible for the skilled analyst to gain from such a deviation. In particular, if $\mu_{\tilde{G}}^S = 1$, then the expected net benefit to the skilled analyst from deviating to \tilde{G} in period t is given by

$$B^S \equiv a - E[b(\tilde{G} - \pi)^2|\Omega_t, v_0] - [a\mu_{\tilde{F}}^S - E[b(\tilde{F} - \pi)^2|\Omega_t, v_0]] \quad (22)$$

$$= a(1 - \mu_{\tilde{F}}^S) - b(\tilde{G}^2 - 2\tilde{G}(x^S + y_t) + \sigma_S^2 + \sigma_{y,t}^2 + y_t^2) \quad (23)$$

$$+ b(\tilde{F}^2 - 2\tilde{F}(x^S + y_t) + \sigma_S^2 + \sigma_{y,t}^2 + y_t^2) \\ = a(1 - \mu_{\tilde{F}}^S) + b(\tilde{F}^2 - \tilde{G}^2) + 2b(x^S + y_t)(\tilde{G} - \tilde{F}) \quad (24)$$

$$\leq d_2(\tilde{G}) \quad (25)$$

$$< 0. \quad (26)$$

Satisfaction of the Intuitive Criterion requires that the public believe that the analyst forecasting \tilde{G} is skilled with certainty, implying that $\hat{\theta}_{T+1}(\tilde{G}, t, \pi, \Omega_T) = 1$ and hence that $\mu_{\tilde{G}}^S = 1$. Thus, the skilled analyst would indeed have an incentive to deviate from \tilde{F} , which contradicts the purported equilibrium.

Proof of Proposition 1.

We proceed in three steps. In the first step, we characterize the unskilled analyst's equilibrium behavior. In the second step, we derive a necessary and sufficient condition that must be satisfied for a forecast by the skilled analyst in period t to be an equilibrium forecast. Finally, we build on the results of step 2 to solve for the skilled analyst's equilibrium forecast delay.

Step 1: Unskilled analyst's behavior

In any period $t \in 1, \dots, T$, if x has been revealed and the analyst has not yet forecasted, then the skilled and unskilled analysts have the same information about x and y and have the same preferences. Thus, equilibrium entails pooling on the same forecast at the same time. The unique, strictly pareto dominant outcome is for both types to pool on an unbiased forecast in period T of $F^U = F^S = E[\pi|\Omega_T, x] = y_T + x$. This results in an expected payoff of $a\theta - b\sigma_{y,T}^2$.

Now suppose that x has not already been revealed by the last period of the forecasting season (Period T). Also suppose the unskilled analyst has not forecasted beforehand. Then Lemma 1 implies that type separation will occur. Thus, the unskilled analyst does not receive any reputational payoff in equilibrium, and he will simply issue an unbiased forecast $F^U = E[\pi|\Omega_T] = y_T$ to minimize mean squared error.

Now consider the unskilled analyst's decision at any period $t \in 1, \dots, T - 1$ where he has not already forecasted and where x has not been revealed yet. By Lemma 1, if the unskilled analyst forecasts immediately, then his type is revealed. Therefore, if he does forecast in period t , then his optimal choice is to minimize mean squared error with an unbiased forecast $F_t^U = y_t$, which results in an expected payoff of $-b(\sigma_x^2 + \sigma_{y,t}^2)$. If he chooses instead to wait until period T , however, he can obtain a strictly higher expected payoff of $[1 - (1 - \gamma)^{T-t}](a\theta - b\sigma_{y,T}^2) + (1 - \gamma)^{T-t}[-b(\sigma_x^2 + \sigma_{y,T}^2)]$. Therefore, forecasting before period T cannot be part of an equilibrium strategy.

Step 2: Critical levels D_1, D_2, \dots, D_T

We proceed to characterize, for each period $t \in 1, \dots, T$, a critical number D_t that is used to describe how the skilled analyst forecasts if indeed he forecasts in period t . First, note from the above arguments that if x is revealed in any period t , then the unique pareto-dominant equilibrium has the skilled analyst pooling with the unskilled type by forecasting in period T . Assume then that, by the end of period T , x has not been revealed and S has not already forecasted. Lemma 1 implies type separation, and so we look for a necessary condition to prevent mimicry. Define

$$D_T \equiv \sqrt{\frac{a}{b}} \tag{27}$$

and observe that the skilled analyst's forecast must be at least a distance D_T apart from y_T (the public posterior mean on π) to prevent mimicry. To see why this is true, recall from above that the unskilled type forecasts $F^U = y_T$ in equilibrium. Suppose for some realized v_0 the skilled analyst issues a period T forecast of F_T^S where $|F_T^S - y_T| < D_T$.

Then the unskilled type's payoff from mimicking with a forecast of F_T^S is $a - bE[(F_T^S - \pi)^2|\Omega_T]$, and hence his net expected gain from deviating to F_T^S is given by

$$M_T^U = a - bE[(F_T^S - \pi)^2|\Omega_T] - [-bE[(y_T - \pi)^2|\Omega_T]] \quad (28)$$

$$= a - bE[(y_T - \pi - (y_T - F_T^S))^2|\Omega_T] - [-bE[(y_T - \pi)^2|\Omega_T]] \quad (29)$$

$$= a - bE[(y_T - F_T^S)^2|\Omega_T] \quad (30)$$

$$> a - b(a/b) \quad (31)$$

$$= 0 \quad (32)$$

which is a contradiction to the purported equilibrium. Hence, we must have that $|F_T^S - y_T| \geq D_T$.

Next, we characterize the skilled analyst's optimal forecasts. There are three cases to consider, depending on the value of x^S , the skilled analyst's private posterior mean for x (we continue to maintain the assumption that $v_0 \neq 0$ and that $x^S \neq 0$).

Case 1: $0 < x^S < D_T$.

Equilibrium requires that the skilled analyst issues forecast $F_T^S = y_T + D_T$. To see why this is true, suppose to the contrary that S forecasts $F_T^S = y_T + M$ for some $M > D_T$. The skilled analyst could strictly gain from deviating to a forecast of D_T if the market would interpret such a forecast as coming from a skilled analyst for certain. (Indeed, $y_T + D_T$ is closer to the skilled analyst's private posterior mean than is $y_T + M$ and thus leads to lower expected error.) At the same time, it is easy to check that the unskilled analyst could never profitably deviate to $y_T + D_T$, regardless of the ensuing beliefs. Thus, the Intuitive Criterion implies that forecasting $y_T + D_t$ will indeed lead to a market belief that the analyst is skilled for certain, and so the skilled analyst would want to deviate, which is a contradiction. Thus, $F^S = y_T + D_T$ in this case.

Case 2: $-D_T < x^S < 0$.

By an argument similar to that used in Case 1, $F_T^S = y_T - D_T$.

Case 3: $|x_S| > D_T$.

In this case, the skilled analyst simply forecasts $F_T^S = y_T + x^S$ as this minimizes expected error and maximizes expected reputation (the no-mimicry constraint is satisfied.) In addition, it is easy to check that any other forecast $F_T^S \neq y_t + x^S$ is ruled out by the Intuitive Criterion.

Assume finally that, for a given period $t < T$, x has not been revealed and that the

skilled analyst has not yet forecasted. Define

$$D_t \equiv \sqrt{\frac{(1 - [1 - (1 - \gamma)^{T-t}]\theta)a}{b} + \sigma_{y,T}^2 - \sigma_{y,t}^2 + [(1 - \gamma)^{T-t} - 1]\sigma_x^2}. \quad (33)$$

By Lemma 1, any equilibrium with a forecast in the current period must entail separation. This implies that, in equilibrium, the skilled analyst cannot issue a forecast F_t^S where $|F_t^S - y_t| \geq D_t$. Indeed, suppose to the contrary that $|F_t^S - y_t| \leq D_t$. As argued above, the unskilled analyst in equilibrium will wait until period T and issue an unbiased forecast F^U ($= x + y_T$ if x is revealed and y_T otherwise). The expected benefit from this is $B_{t,wait}^u = [1 - (1 - \gamma)^{T-t}]a\theta - (1 - \gamma)^{T-t}b\sigma_x^2 - b\sigma_{y,T}^2$ where the first term reflects the expected reputational benefit from having x revealed and being able to pool with the skilled type; the second term is the expected loss in accuracy from not having x revealed; and the last term is the baseline expected inaccuracy associated with the systematic component y .

Now, by construction of D_t , $B_{t,wait}^U = a - b\sigma_x^2 - b\sigma_{y,t}^2 - bD_t^2$. If the unskilled analyst forecasts today and mimics, then he gets $B_{t,mimic}^U = a - b(\sigma_x^2 + \sigma_{y,t}^2) - b(F_t^S - y_t)^2$. So $B_{t,wait}^U < B_{t,mimic}^U$, which is a contradiction.

If it is indeed optimal for the analyst to forecast in period t , then by reasoning similar to that in cases 1-3 for period T above, we have that, depending on the skilled analyst's posterior x^S for the idiosyncratic component, he will forecast

$$F^S = \begin{cases} y_m + x^S & \text{if } |x^S| > D_m \\ y_m + D_m & \text{if } 0 < x^S \leq D_m \\ y_m - D_m & \text{if } -D_m \leq x^S < 0 \end{cases}$$

Step 3: Equilibrium forecast delay

From the above arguments, for any period $t \in 1, \dots, T$, if x has been revealed, then the skilled analyst will wait until period T and pool by issuing $y_T + x$, obtaining expected payoff of $a\theta - b\sigma_{y,T}^2$. We use backward induction to characterize the equilibrium timing of the skilled analyst's forecast when x has not been revealed. Define, for $t = 1, \dots, T$,

$$V_t^S = a - b\sigma_{y,t}^2 - b\sigma_S^2 - b[\max(D_t - |x^S|, 0)]^2, \quad (34)$$

$$W_t^S = \gamma(a\theta - \sigma_{y,T}^2) + (1 - \gamma)[a - b\sigma_{y,t+1}^2 - b\sigma_S^2 - b[\max(D_{t+1} - |x^S|, 0)]^2]. \quad (35)$$

The expression V_t^S represents the expected payoff to the skilled analyst, as of period t , if x has not been revealed and he forecasts immediately; W_t^S represents the expected payoff from waiting until period $t + 1$ and either forecasting immediately (if x has still not arrived) or pooling in period T (if x has been revealed). Note that in computing these payoffs, we have used the fact that expected quadratic loss is equal to variance plus the square of bias.

Consider first the skilled analyst's decision at period T . Suppose he has not forecasted, and suppose x has not been revealed. Then, by Lemma 1 and the preceding arguments, he separates with his forecast and obtains expected payoff of V_T^S . Now consider the skilled analyst's decision at period $T - 1$, assuming that he has not forecasted already and that x has not been revealed. If he forecasts today, by Lemma 1 he separates and obtains expected payoff of V_{T-1}^S . If he waits, his expected payoff is W_{T-1}^S . Thus, he forecasts today if and only if $V_{T-1}^S - W_{T-1}^S > 0$.

Now consider the skilled analyst's decision at the end of any period $r \in 1, \dots, T - 2$. Assume we have shown that the following induction hypothesis is true: if the end of period $r + 1$ arrives with x not having been revealed and S not having forecasted, then S forecasts immediately if and only if $V_{r+1}^S - W_{r+1}^S > 0$.

Suppose S has not yet forecasted, and x has not yet been revealed. We proceed to show that S forecasts today if and only if $V_r^S > W_r^S$. Now, forecasting today (period r) yields an expected payoff of V_r^S , while waiting until next period (period $r + 1$) yields at least the payoff from forecasting then, W_r^S (and possibly more).

First, suppose that

$$V_r^S - W_r^S < 0 \tag{36}$$

Then clearly S prefers to wait. Now suppose that

$$V_r^S - W_r^S \geq 0 \tag{37}$$

We will show in fact that Condition (37) implies $V_{r+1}^S - W_{r+1}^S \geq 0$. The induction hypothesis would then imply that, in period $r + 1$, S would forecast immediately, and hence it would follow from (37) that S prefers forecasting today to waiting.

Notice that $D_r < D_{r+1} < \dots < D_T$. There are therefore four cases to consider.

Case 1: $|x^S| \geq D_{r+2}$.

In this case, we have the following simplified expressions:

$$V_r^S = a - b\sigma_{y,r}^2 - b\sigma_S^2 \quad (38)$$

$$W_r^S = \gamma(a\theta - \sigma_{y,T}^2) + (1 - \gamma)[a - b\sigma_{y,r+1}^2 - b\sigma_S^2] \quad (39)$$

$$V_{r+1}^S = a - b\sigma_{y,r+1}^2 - b\sigma_S^2 \quad (40)$$

$$W_{r+1}^S = \gamma(a\theta - \sigma_{y,T}^2) + (1 - \gamma)[a - b\sigma_{y,r+2}^2 - b\sigma_S^2]. \quad (41)$$

Thus, we have that

$$V_{r+1}^S - W_{r+1}^S = a(\gamma - \gamma\theta) - b(\sigma_{y,r+1}^2 - (1 - \gamma)\sigma_{y,r+2}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) \quad (42)$$

$$\geq a(\gamma - \gamma\theta) - b(\sigma_{y,r}^2 - (1 - \gamma)\sigma_{y,r+1}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) \quad (43)$$

$$= V_R^S - W_R^S \quad (44)$$

$$\geq 0. \quad (45)$$

Case 2: $D_{r+1} \leq |x^S| \leq D_{r+2}$.

In this case,

$$V_{r+1}^S - W_{r+1}^S = a(\gamma - \gamma\theta) - b(\sigma_{y,r+1}^2 - (1 - \gamma)\sigma_{y,r+2}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) \quad (46)$$

$$+ b(1 - \gamma)[\max(D_{r+2} - |x^S|, 0)]^2$$

$$> a(\gamma - \gamma\theta) - b(\sigma_{y,r+1}^2 - (1 - \gamma)\sigma_{y,r+2}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) \quad (47)$$

$$\geq V_r^S - W_r^S \quad (48)$$

$$\geq 0.$$

Case 3: $D_r \leq |x^S| < D_{r+1}$.

For this case, we have that

$$V_{r+1}^S - W_{r+1}^S = a(\gamma - \gamma\theta) - b(\sigma_{y,r+1}^2 - (1 - \gamma)\sigma_{y,r+2}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) \quad (50)$$

$$- b(D_{r+1} - |x^S|)^2 + b(1 - \gamma)(D_{r+2} - |x^S|, 0)^2.$$

Notice from this equation that $V_{r+1}^S - W_{r+1}^S$ is strictly concave in $|x^S|$. Thus, it suffices to show that $V_{r+1}^S - W_{r+1}^S$ is positive at $|x^S| = D_{r+1}$ and at $|x^S| = 0$ (i.e., at the boundary points of the relevant range). Direct computations and some algebraic simplification show that, at $|x^S| = 0$, $V_{r+1}^S - W_{r+1}^S = b\gamma(\sigma_x^2 - \sigma_S^2)$. Also, at $|x^S| = D_{r+1}$, $V_{r+1}^S - W_{r+1}^S = a(\gamma - \gamma\theta) - b(\sigma_{y,r+1}^2 - (1 - \gamma)\sigma_{y,r+2}^2 - \gamma\sigma_{y,T}^2 + \gamma\sigma_S^2) + b(1 - \gamma)(D_{r+2} - D_{r+1})^2 > V_r^S - W_r^S > 0$.

Case 4: $|x^S| < D_r$.

In this case, we have that

$$V_{r+1}^S = a - b\sigma_{y,r+1}^2 - b\sigma_S^2 - b(D_{r+1} - |x^S|)^2 \quad (51)$$

$$W_{r+1}^S = \gamma(a\theta - \sigma_{y,T}^2) + (1 - \gamma)[a - b\sigma_{y,r+2}^2 - b\sigma_S^2 - b(D_{r+2} - |x^S|)^2]. \quad (52)$$

Multiplying out the squared binomials in each equation, using the definition of D_{r+1} and D_{r+2} , and some algebraic simplifications reveals that

$$V_{r+1}^S - W_{r+1}^S = b(\gamma\sigma_x^2 - \gamma\sigma_S^2) + 2b|x^S|(D_{r+1} - (1 - \gamma)D_{r+2}) - b\gamma|x^S|^2 \quad (53)$$

$$> b(\gamma\sigma_x^2 - \gamma\sigma_S^2) \quad (54)$$

$$> 0 \quad (55)$$

Proof of Proposition 2:

This follows directly from inspecting the expression D_t in Proposition 1 and the condition (5) in Proposition 1.

Proof of Proposition 3 (sketch):

The proof of this proposition parallels the proof of Proposition 1. First, observe that the analogue of Lemma 1 holds for the two types of each analyst j since the main arguments in the proof apply to a fixed period t . Next, note that all of the unskilled types will wait to forecast until period T because, in each period $t = 1, \dots, T - 1$, the expected variance reduction from waiting another period is positive, and hence the expected squared error is lower. Third, in each period a skilled analyst j will have to overcome a no-mimicry condition D_t as defined in the statement of the proposition. This characterizes analyst j 's optimal forecast if indeed he forecasts in period t . Finally, the optimal timing is determined by an arbitrage condition similar to that in the single-analyst case, except now the analyst will compute $Var[x|h_m, v_0^j]$ instead of σ_S^2 and $E[x|h_m, v_0^j]$ instead of x^S . Note that the ex ante forecasting time can no longer be calculated in a simple manner since variances will depend on signal realizations and prior forecasts. Nonetheless, arguments similar to those used in Step 3 of Proposition 1 go through with V_t^S and W_t^S defined appropriately.

References

- Bernhardt, D., M. Campello, and M. Kutsoati, 2006, Who herds? *Journal of Financial Economics* 80, 657-675.
- Brown, L., 1991, Forecast selection when all forecasts are not equally recent, *International Journal of Forecasting* 7, 349-356.
- Chamley, C. and D. Gale, 1994, Information revelation and strategic delay in a model of investment, *Econometrica* 62, 1065-1085.
- Cho, I. and D. Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 102, 179-222.
- Clarke, J. and A. Subramanian, 2006, Dynamic forecasting behavior: Theory and evidence, *Journal of Financial Economics* 80, 81-113.
- Clement, M., 1999, Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27, 285-303.
- Clement, M. and S.Y. Tse, 2005, Financial analyst characteristics and herding behavior in forecasting, *Journal of Finance* 60, 307-341.
- Cooper, R., T. Day, and C. Lewis, 2000, Following the leader: A study of individual analysts' forecasts, *Journal of Financial Economics* 61, 383-416.
- Ehrbeck, T. and R. Waldmann, 1996, Why are professional forecasters biased? Agency versus behavioral explanations, *Quarterly Journal of Economics* 111, 21-40.
- Graham, J., 1999, Herding among investment newsletters: theory and evidence, *Journal of Finance* 54, 237-268.
- Gul, F. and R. Lundholm, 1995, Endogenous timing and the clustering of agents' decisions, *Journal of Political Economy* 103, 1039-1066.
- Guttman, I., 2005, The timing of analysts' earnings forecasts, working paper, Stanford University Graduate School of Business.

- Hong, H., and J. Kubik, 2003, Analyzing the analysts: Career concerns and biased earnings forecasts, *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.
- Ivkovic, Z., and N. Jegadeesh, 2004, The timing and value of forecast and recommendation revisions, *Journal of Financial Economics* 73, 433-463.
- Jackson, A., 2005, Trade generation, reputation, and sell-side analysts, *Journal of Finance* 60, 673-717.
- Keane, M. and M. Runkle, Rationality and Forecast bias, *Journal of Political Economy*
- Lys, T. and S. Sohn, 1990, The association between revisions of analysts' earnings forecasts and security-price changes, *Journal of Accounting and Economics* 13, 341-363.
- Mikhail, M., B. Walther, and R. Willis, 1999, Does forecast accuracy matter to security analysts? *The Accounting Review* 74, 185-200.
- Morgan, J. and P. Stocken, 2003, An analysis of stock recommendations, *RAND Journal of Economics* 34, 183-203.
- Ottaviani, M. and P. Sorensen, 2005a, Reputational cheap talk, *Rand Journal of Economics* 37, 155-175.
- Ottaviani, M. and P. Sorensen, 2005b, The strategy of professional forecasting, *Journal of Financial Economics* 81, 441-466.
- Prendergast, C. and L. Stole, 1996, Impetuous youngsters and jaded old-timers: acquiring a reputation for learning, *Journal of Political Economy* 104, 1105-1134.
- Scharfstein, D. and J. Stein, 1990, Herd behavior and investment, *American Economic Review* 80, 465-479.
- Stickel, S., 1992, Reputation and performance among security analysts, *Journal of Finance* 47, 1811-1836.

Trueman, B., 1994, Analyst forecasts and herding behavior, *Review of Financial Studies* 7, 97-124.

Welch, I., 2000, Herding among security analysts, *Journal of Financial Economics* 58, 369-396.

Table 1
Descriptive Statistics

The sample consists of current-quarter EPS forecasts in I/B/E/S made during March, June, September, and December quarters from 1994-2003. Excluded are forecasts issued within three days after the previous quarter's earnings report. Panel A reports analyst, forecast, and broker characteristics by individual forecast. Panel B shows company characteristics by forecast. Finally, Panel C reports the mean and median forecast delay and forecasting season length according to each year in the sample period. Forecasting season length is the number of days between the last earnings report and the end of the current quarter.

Panel A: Analyst, forecast, and broker characteristics						
	N	mean	median	1 st Q	3 rd Q	Std. dev.
# of companies covered by analyst in current year	294,785	17.40	15	11	21	13.33
# of industries covered by analyst (2-digit SIG) in current year	294,785	2.98	3	2	4	2.02
Analyst forecasting experience (years)	296,950	6.41	5.28	2.06	9.81	5.19
Size of forecasting analyst's employer (# analysts employed)	297,477	74.09	51	20	97	76.78
Days between forecast and prior quarter's earnings report	298,129	30.90	29	9	50	21.45
Days between forecast and end of current quarter	298,129	35.96	36	17	55	21.43
Length of forecasting season	298,129	66.86	68	62	74	9.48
Institutional Investor All-American 1 st team	298,129	0.039	0	0	0	0.194
Institutional Investor All-American team (1 st , 2 nd , or 3 rd)	298,129	0.111	0	0	0	0.314

Table 1, cont'd.

Panel B: Company characteristics						
	N	mean	median	1 st Q	3 rd Q	Std. dev.
# of analysts covering company	298,129	9.98	8	4	14	7.04
Market capitalization	284,426	8,626.05	1,550.67	451.24	5,435.7	26,179.39
Panel C: Forecasting season length and forecast delay, by year						
Year	N	Length of forecasting season (days)		Days between prior quarter earnings report and forecast		
		Mean	Median	Mean	Median	
1995	30,206	68.87	71	31.57	29	
1996	36,767	67.85	69	31.62	29	
1997	38,088	67.14	69	31.22	29	
1998	42,424	66.45	69	31.92	30	
1999	38,358	66.05	69	31.88	31	
2000	5,040	63.67	65	17.01	13	
2001	39,057	67.70	68	30.78	28	
2002	33,940	66.38	68	30.07	28	
2003	34,249	65.14	68	29.83	28	

Table 2
Analysis of Forecast Delay

The sample is drawn from I/B/E/S and consists of current-quarter EPS forecasts made during the March, June, September, and December quarters from 1995-2003. The sample only includes an analyst's earliest forecast issued within three days after the previous quarter's earnings report. Reported in the table are estimated coefficients from Cox proportional hazards regression models. Firm size is measured as the log or market capitalization; broker size is measured as the log of one plus the number of employees covered. Forecasting season length is the number of days between the last earnings report and the end of the current quarter. Z-statistics appear in parentheses below coefficient estimates.

Explanatory Variable	(1)	(2)	(3)	(4)
Firm size	-0.0250 (-15.57)	-0.0243 (-15.09)	-0.0253 (-15.60)	-0.0246 (-15.21)
Analyst coverage	-0.0042 (-10.09)	-0.0044 (-10.44)	-0.0042 (-9.95)	-0.0044 (-10.27)
Institutional Investor 1 st team All-American	-0.0831 (-7.65)		-0.0892 (-8.15)	
Institutional Investor 1 st , 2 nd , or 3 rd -team All-American		-0.0873 (-12.66)		-0.0957 (-13.67)
Number of companies covered by analyst			-0.0005 (-3.08)	-0.0005 (-2.66)
Number of industries covered by analyst			0.0019 (1.65)	0.0015 (1.31)
Forecasting experience			0.0027 (6.29)	0.0033 (7.61)
Brokerage size	-0.0099 (-4.73)	-0.0050 (-2.32)	-0.0110 (-5.18)	-0.0057 (-2.65)
Length of forecasting season	-0.0283 (-123.82)	-0.0283 (-123.68)	-0.0282 (-122.59)	-0.0282 (-122.46)
Number of observations	236,045	236,045	234,630	234,630
Model P-value	0.000	0.000	0.000	0.000

Table 3
Forecast Delay: Separate Regressions by Coverage

The sample is drawn from I/B/E/S and consists of current-quarter EPS forecasts made during the March, June, September, and December quarters from 1995-2003. The sample only includes an analyst's earliest forecast issued within three days after the previous quarter's earnings report. Reported in the table are estimated coefficients from Cox proportional hazards regression models. Firm size is measured as the log of market capitalization; broker size is measured as the log of one plus the number of employees covered. Forecasting experience is measured in years. Z-statistics appear in parentheses below coefficient estimates.

Analyst Coverage	N	II 1 st team All-American		Firm Size		Brokerage Size		Workload (# of companies)		Forecasting Experience		P-value
		Coeff	z-stat	Coeff	z-stat	Coeff	z-stat	Coeff	z-stat	Coeff	z-stat	
1	8,434	-0.1736	-3.38	-0.0487	-4.98	0.0915	7.67	-0.0052	-6.43	0.0103	4.42	0.000
2	14,547	-0.0735	-1.96	-0.0371	-4.84	0.0773	8.84	-0.0033	-5.46	0.0064	3.69	0.000
3	17,656	-0.0824	-2.65	-0.0236	-3.53	0.0418	5.46	-0.0024	-4.70	0.0060	3.87	0.000
4	17,756	-0.1091	-3.86	-0.0186	-2.93	0.0251	3.30	-0.0015	-2.79	0.0051	3.39	0.000
5	17,012	-0.0994	-3.59	-0.0225	-3.51	0.0096	1.23	-0.0006	-1.05	0.0050	3.28	0.000
6	15,946	-0.0902	-3.43	-0.0456	-6.91	0.0190	2.31	-0.0008	-1.43	0.0057	3.52	0.000
7	15,665	-0.1159	-4.44	-0.0279	-4.15	0.0102	1.23	-0.0012	-2.17	0.0028	1.71	0.000
8	14,084	-0.1289	-4.87	-0.0425	-6.38	0.0026	0.29	-0.0007	-1.12	0.0053	3.13	0.000
9	12,487	-0.1281	-4.61	-0.0510	-7.12	-0.0165	-1.75	-0.0006	-0.81	0.0041	2.28	0.000
10	11,799	-0.0845	-3.02	-0.0499	-6.46	-0.0140	-1.43	-0.0015	-1.79	0.0050	2.68	0.000