Essay on Analyst Herding

Cyrus Aghamolla

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

under the Executive Committee

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

Essay on Analyst Herding

Cyrus Aghamolla

This study investigates a dynamic model of analyst forecasting where the ordering of forecasts and analysts' information endowments are endogenously determined. Analysts are probabilistically informed, potentially biased, and can increase their informedness through information acquisition. I characterize the unique equilibrium which holds for general distributions. The results show that analysts with less bias, greater precision, or a greater likelihood of being informed forecast earlier. Moreover, the main results show (perhaps surprisingly) that analysts always choose to be imperfectly informed, even though information acquisition is costless. This arises from the incentive to induce more timely forecasting by the other analyst. Likewise, analysts choose a positive bias level in equilibrium in order to gain a strategic advantage in their forecast timing. I discuss a number of empirical implications and extend the model to allow analysts to learn over time.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am indebted to my dissertation committee Ilan Guttman, Jonathan Glover, Tim Baldenius, Fabrizio Ferri, and Alon Kalay for their continuous guidance, support, and encouragement. I am particularly grateful to Ilan Guttman for his guidance and for numerous fruitful discussions which improved this dissertation as well as several other projects. I also thank my colleagues Byeongje An, Alejandro Francetich, and Tadashi Hashimoto for many insightful conversations and for their friendship over the course of my doctoral studies. I thank my previous advisors Joe Oppenheimer, John Rust, and Bob Wilson for their continued guidance. Lastly, I am indebted to my family for their support.

Dedicated to the memory of Soraya Aghamolla

Chapter 1

Introduction

A rich empirical literature has documented a well-known regularity that sell-side analysts may strategically time their forecasts for the purpose of observational learning.[1] Indeed, the notion that analysts can learn from each other (or more specifically, from the reports of those who have forecasted before them) is a ubiquitous and pervasive phenomenon persistent through two decades of research.[2] Meanwhile, a disparate yet equally rich literature has shown that analysts tend to upwardly bias their forecasts,[3] and that information acquisition/production plays a crucial role in an analyst's responsibilities (e.g. Ivkovic and Jegadeesh (2004), Chen, Cheng, and Lo (2010), Livnat and Zhang (2012), Keskek, Tse, and Tucker (2014), Altschuler, Chen, and Zhou (2015), Brown et al. (2015)). A natural question thus arises: If analysts can learn from each other, then how does this affect

---

[1]See De Bondt and Forbes (1999), Hong, Kubik, and Solomon (2000), Welch (2000), Cooper, Day, and Lewis (2001), Zitzewitz (2001), Gleason and Lee (2003), Hong and Kubik (2003), Clement and Tse (2005), Jegadeesh and Kim (2010), Clement, Hales, and Xue (2011), Shroff, Venkataraman, and Xin (2014), and Keskek, Tse, and Tucker (2014).

[2]Welch (2000) finds that "the buy or sell recommendations of security analysts have a significant positive influence on the recommendations of the next two analysts" (p. 369). This finding has been reinforced in numerous subsequent studies, namely Cooper et al. (2001), Clement and Tse (2005), Jegadeesh and Kim (2010), Keskek et al. (2014), and Shroff et al. (2014).

Analysts can also potentially access other analysts' reports through their buy-side clients before they have been publicly released (see Irvine, Lipsett, and Puckett (2007)). Anecdotal evidence suggests that this is not unusual.

[3]For example, see De Bondt and Thaler (1990), Francis and Philbrick (1993), Dugar and Nathan (1995), McNichols and O'Brien (1997), Das, Levine, and Sivaramakrishnan (1998), Lin and McNichols (1998), Easterwood and Nutt (1999), Michaely and Womack (1999), Kothari (2001), Lim (2001), Jackson (2005), Brenhardt et al. (2006), Chen and Jiang (2006), and Raedy, Shane, and Yang (2006), among others.

their forecast timing, and, more importantly, their central responsibility of information acquisition/discovery? Although these aspects of analyst behavior have received considerable attention in the empirical literature, few theoretical studies examine the strategic interdependence between analyst forecast timing, information acquisition, and the role of bias.

In this paper, I seek to bridge this gap by providing a unified theory of analyst forecasting where the timing of the forecast, information acquisition, and level of affiliation are endogenously chosen. The results show that the forecast order is determined by the analyst's information quality, information endowment, and potential bias, where the more precise, better informed, or less biased analyst forecasts earlier in expectation. Moreover, due to strategic considerations regarding the subsequent forecast timing, analysts choose to be *imperfectly* informed in equilibrium, even when information acquisition is costless. This result establishes a direct theoretical link between the learning incentives of analysts and their information acquisition, which has heretofore not been captured in the theoretical literature. The results also show that analysts choose to be affiliated in equilibrium, thus providing a theoretical justification for why an analyst may choose employment in which she may encounter a conflict of interest.

The setting is one where each of two analysts must first acquire information and then decide when to issue a forecast of a firm's terminal value. Each analyst has an incentive to delay her forecast in order to observe the report and the corresponding private information of her peer. The additional information from the first analyst's forecast (the "leader") complements the private information of the remaining analyst (the "follower"), and allows for a more accurate prediction by the follower. The key feature of the model is that there may be uncertainty as to the information endowment of each analyst, and this endowment probability is endogenously determined. This assumption differs from virtually all other models of analyst forecast timing, which typically assume that an analyst receives a (potentially noisy) signal of the firm's value with probability one. Although conventional in the theoretical literature, the assumption that an analyst is always endowed with useful information does not seem entirely plausible in practice. Indeed, allowing for the probabilistic endowment of information leads to results that help explain a number of

empirical regularities which have not been captured by previous studies.

Prior to the forecasting decision, each analyst independently chooses the probability for which she receives information in the following stage, and this decision is publicly observable. The likelihood of receiving information can be thought of as an analyst's choice of how many firms or industries to follow; analysts considered specialists who focus on a particular industry are more likely to have insightful information as compared to generalists who follow several industries.[4] The choice of endowment probability can be equivalently represented as an observable effort decision that an analyst makes in each forecasting period, where the amount of effort determines the likelihood that the analyst receives information. Moreover, high informedness can be alternatively interpreted as an analyst's decision to join a firm that has a reputation for extensive information production, such as firms which commonly perform due diligence in the underwriting stage of IPOs, and thus typically receive superior information with a high likelihood. Similarly, high informedness can represent an analyst whose firm specializes in one particular industry or has established lines of communication with certain company executives, whereas less informedness captures an analyst whose firm covers a wide range of industries with a dispersed network of executives, yet with less depth.[5] At the firm level, high informedness can be thought of as a brokerage house or investment-banking firm which emphasizes their research division through a larger allocation of resources.

The analysts' payoff depends on their squared forecast error and the timeliness of their forecast, which is captured by discounting. Several studies document that an analyst's compensation crucially depends on their accuracy.[6] The forecast error in the analyst's

---

[4]This is consistent with the analysis of Clement and Tse (2005), who find that analysts who follow many industries are more likely to herd. Indeed, the authors conclude that this is "due to analysts' inability to develop and use specialized knowledge when they follow many companies or industries" (p. 329).

[5]As noted in Brown et al. (2015), executives may privately disclose immaterial information for the analyst's broader "mosaic." They conclude that "information conveyed in private conversations with management is extremely valuable to sell-side analysts in the post-Reg FD environment" (p. 20). The survey of executives in Soltes (2014) has a similar conclusion that frequent private communication with executives and management plays a significant role for sell-side analysts.

[6]See Stickel (1992), Mikhail, Walther, and Willis (1999), Hong, Kubik, and Solomon (2000), Hong and Kubik (2003), and Wu and Zhang (2009). Additionally, Hong and Kubik (2003) state: "Even if the compensation of analysts does not depend explicitly on forecast accuracy, to generate investment banking business or trading commissions in the longer run, analysts need to cultivate a reputation for forecast expertise among the buy-side" (p. 314).

utility can also indirectly capture the payoff from a subsequent stock recommendation which is more profitable given the analyst's forecast accuracy, as evidenced in Loh and Mian (2006), Ertimur, Sunder, and Sunder (2007), and Brown et al. (2015). Likewise, timeliness is also an important determinant of an analyst's compensation,[7] as expeditious forecasts can generate trading volume and affect promotions (Cooper et al. (2001), Groysberg, Healy, and Maber (2011)). In subsequent analysis, I allow the analysts to be potentially biased in that their compensation may also depend on an endogenously chosen distortion parameter. This represents the analyst's choice of employment, whether at a brokerage house, investment-banking firm, or independent research institution. As shown in a number of studies (e.g. Dugar and Nathan (1995), Lin and McNichols (1997), Michaely and Womack (1999), Mola and Guidolin (2009)), analysts employed by brokerage houses and investment banks, rather than those employed by independent research firms, tend to issue more biased forecasts. While the advantages of working for the former may seem intuitively obvious (e.g. greater resources), this study offers a novel explanation for an analyst's choice of affiliated employment, wherein affiliation arises from a strategic advantage in the forecast timing.

The results concerning the timing of forecasts show that the expected forecast order is determined by (*i*) the quality, or precision, of the analyst's information, (*ii*) the analyst's likelihood of being informed, and (*iii*) the magnitude of the analyst's bias. We see that analysts who have better quality information, are more likely to be informed, or have less potential bias forecast earlier. Moreover, the main result shows that analysts always choose to be imperfectly informed–that is, to be uninformed with positive probability–even when information acquisition is costless. This emerges due to the analyst's incentive to induce the other analyst to forecast earlier and observe their private information. Similarly, analysts choose employment where their compensation potentially relies on issuing a biased forecast in order to induce more timely forecasting by the other analyst.

The equilibrium of the forecasting subgame resembles that of a war of attrition with private information, where each informed analyst mixes between issuing a forecast and

---

[7]See Schipper (1991) for a discussion and Cooper et al. (2001), Hong and Kubik (2003), Jackson (2005), Groysberg et al. (2011), and Loh and Stulz (2011) for empirical support.

remaining quiet at every instant. Mixing occurs in equilibrium as analysts strategically mimic uninformedness in order to induce their peer to forecast earlier. We see that the path of the probabilistic forecast time follows an exponential distribution which reaches one at an endogenously determined terminal time. Because the model allows for heterogeneity in the analysts' precision, distribution, endowment, discount rate, and bias, the mixing probabilities can vary between the two analysts. If this is the case, then the analyst with the "slower" mixing rate will forecast with discrete probability at the beginning of the forecasting subgame, thus leading her expected forecast time to be earlier in equilibrium.

In the information acquisition subgame which precedes the forecasting stage, each analyst simultaneously chooses her likelihood of receiving private information. Since the analyst with a relatively higher likelihood of receiving information forecasts earlier in expectation, each analyst has an incentive to select an endowment probability just below the other's. This leads to an equilibrium choice of endowment probability that is less than one for both analysts.

The results help explain a number of empirical regularities. With regard to herding, the results show that the forecast order is determined by the precision and endowment of the analysts' information. This helps explain the findings of Cooper et al. (2001), Clement and Tse (2005), and Keskek et al. (2014), who document evidence supporting the hypothesis that analysts with superior information forecast earlier. The results also support the conclusions of Shroff, Venkataraman, and Xin (2014), wherein the forecasts of analysts who follow the leader also contain material information. Moreover, the results of the model comport nicely with the empirical findings of Cooper et al. (2001), who conclude that forecasting ability is best measured by the timeliness of the forecast rather than the forecast error. Perhaps most surprisingly, the results help explain why affiliated analysts provide more accurate forecasts than independent analysts, as found in Jacob, Lys, and Neale (1999) and Jacob, Rock, and Weber (2008). The results imply that, rather than due to differences in resources or access to non-public information, analysts who join brokerage or investment-banking firms gain an informational advantage through the strategic timing of forecasts (i.e. the empirical results are driven by analyst selection of employment). Moreover, the results are in line with the findings of Das, Levine, and Sivaramakrishnan

(1998), who document higher levels of bias for analysts' forecasts when there is relatively less public information. In addition, the results correspond to the findings of Gu and Xue (2008), who document that affiliated analysts issue comparatively more accurate earnings forecasts for firms which are also followed by less biased or independent analysts. The results also help to explain the trend for analysts to cover more firms, and sheds light on why some analysts become generalists rather than specialists.

While the model fits nicely with several existing empirical regularities, the results also offer several avenues for future research in terms of empirical predictions. As mentioned previously, the results suggest that relatively earlier forecasts should contain less bias or distortion. Correspondingly, the level of bias or distortion in the forecasts should be comparatively higher when a firm is covered by more analysts. Moreover, we should see more timely forecasts overall when the public information is stronger and when a single highly informed analyst is following the firm. The results of the model also imply that clustered forecasts which occur later should have comparatively higher forecast errors than of clustered forecasts which emerge earlier. The results pertaining to information acquisition indicate that individual forecasts should be relatively less informative, even for earlier forecasts, when there is increased analyst following. The results also imply that high analyst following does not necessarily mean a less opaque informational environment for the firm. Rather, the timeliness of the analysts' forecasts is more indicative of the transparency or strength of a firm's information environment. This follows from the results which imply that (i) increased analyst following may not reduce information asymmetry, (ii) analysts forecast earlier when the precision of their private information is higher and when they are more likely to be informed, and (iii) analysts issue more timely forecasts when there is better public information. The empirical predictions are further discussed in section 5.

1.1   Related literature

This study is related to a number of literatures. I first discuss the current study in the context of the extant literature on forecast timing. Gul and Lundholm (1995) examine a

model where two agents must make a forecast decision and can benefit from observing the announcement of their competitor. They characterize the symmetric equilibrium where agents with good news forecast earlier than agents with bad news. The setting here is similar to Gul and Lundholm (1995) in that both models involve endogenous timing with an interdependent value structure. The current model fundamentally differs from Gul and Lundholm (1995) in four ways: (i) in this setting, agents have uncertainty over the information endowment, whereas Gul and Lundholm assume agents are informed with probability one; (ii) the analysts' payoffs do not depend on the value of the firm, whereas Gul and Lundholm assume the waiting costs are increasing in the value of the firm, thus leading to the aforementioned equilibrium; (iii) analysts may be heterogenous in every dimension, whereas Gul and Lundholm assume agents are homogenous; and (iv) analysts also engage in information acquisition and choice of employment.[8]

Guttman (2010), Aghamolla and Hashimoto (2015), and Xue (2015) also explore analysts' endogenous forecast timing. Guttman (2010) considers a setting where the analysts' precision improves over time, though the payoff from forecasting decreases as the pool of public information improves. Guttman shows how clustering and dispersion of forecasts can emerge. The current setting varies in that the focus is on the equilibrium properties when analysts can learn from the forecasts of other analysts, whereas this feature is absent in Guttman (2010). Moreover, several of the insights of Guttman (2010) also hold in the single-analyst setting of his model, whereas the strategic interaction between analysts is a fundamental feature of the current model. In the extensions, I allow for information arrival to occur stochastically over time, accentuate the salience of timeliness through varied compensation for the leader and follower, as well as show how the payoff function can be endogenized from a market for information as in Guttman (2010).

Xue (2015) investigates a herding model with both an affiliated and independent analyst, where the affiliated analyst receives a more precise signal and the independent analyst

---

[8]In terms of technical assumptions, Gul and Lundholm limit their setting to uniformly distributed signals with a finite support, whereas the results in the forecasting subgame here hold for general and heterogenous distributions. Gul and Lundholm also characterize the unique symmetric pure strategy Nash equilibrium, however, there are also a multiplicity of asymmetric equilibria and potential mixed strategy equilibria. In contrast, the results in the forecasting subgame here are characterized by a unique perfect Bayesian equilibrium without refinements.

can acquire costly information. Xue shows that the independent analyst acquires more information in the presence of the affiliated analyst in order to induce the affiliated analyst to issue a less biased forecast. The model here differs from Xue (2015) in that the strategic timing of forecasts is a fundamental feature of this setting, whereas timing is largely fixed in Xue's model. The current setting also allows for two-sided and costless information acquisition and two-sided endogenous choice of affiliation.[9] Aghamolla and Hashimoto (2015) examine an endogenous timing model with several analysts and show results regarding delay and clustering in analyst forecasting. The current setting differs in that the primary focus is on the interdependence of forecast timing and information acquisition.

This study is also related to the endogenous investment timing literature, particularly Chamley and Gale (1994) and Zhang (1997). These papers show investment cascades following the investment choice of the first agent. In contrast, due to the continuous firm value, the equilibrium here does not involve an information cascade as an informed analyst never disregards her private information. Moreover, the present model does not include an unknown cost (as in Zhang (1997)) and exhibits delay in continuous time (unlike Chamley and Gale (1994)). This paper also builds on the literature concerning reputations in repeated games with two-sided private information, as in Kreps and Wilson (1982), Fudenberg and Tirole (1986), and Abreu and Gul (2000), by endogenizing the ex ante reputation. Kim and Lee (2014) study a concession game where two agents may acquire information about an unknown state variable which affects the payoffs of both agents in the event that the game ends suddenly. Among many differences, Kim and Lee assume that the game terminates with a Poisson arrival and that there is never private information of an agent's informedness–agents pay a fixed cost to acquire information with probability one and this is publicly observed. Kim and Lee rather focus on the effects of disclosure rules of the unknown termination payoff, whereas the focus of this study is on the effects of multiple analysts on information acquisition.

The current setting varies from reputational herding models (Scharfstein and Stein

---

[9]The results in Xue (2015) also critically depend on the discrete structure of the message space (though continuous signal), whereas this is continuous in the current setting.

(1990), Trueman (1994)), in that the forecast order is endogenous and the analysts' utility is indirectly affected by their type through the forecast error, rather than explicitly through a market belief of the type. These models also assume that the agents' payoff does not depend on the value of the realized outcome. This paper is also related to studies which examine analyst information acquisition. These studies consider information acquisition in the context of dynamic cheap talk with unknown alignment (Meng (2015)), varying levels of public information (Fischer and Stocken (2010)), or when the analyst's compensation is based on trading commissions (Hayes (1998)). Among other differences, these studies focus primarily on single-analyst settings, whereas the current model examines information acquisition in the presence of strategic interaction between analysts. This paper also differs from Ottaviani and Sørenson's (2006) reputational cheap talk model as the focus of this study is on the effects of endogenous timing on information acquisition, whereas these are both exogenous in their setting.

The paper proceeds as follows. The following section describes the model, while Section 3 analyzes the equilibrium and presents the results. Section 4 examines the model when analysts may choose their level of affiliation and Section 5 discusses empirical implications and predictions. Section 6 explores a number of extensions and the final section concludes. All proofs are relegated to the Appendix unless otherwise specified.

Chapter 2

Model

The model is composed of two analysts $i \in \{1, 2\}$ who first simultaneously make an information acquisition decision and then decide when to issue a forecast. Time is assumed to be continuous following the analysts' information acquisition decision at date 0. I assume that analysts are uninformed with probability $x_i \in [0, 1]$ and receive information with probability $1 - x_i$. The information acquisition subgame at time 0 consists of three stages (as depicted in Figure 1). In the first stage of $t = 0$, each analyst chooses their probability of receiving information. Correspondingly, in the following stage, each analyst's choice of x_i is publicly revealed and analysts receive private information with probability $1 - x_i$.[1,2]

An analyst is said to be uninformed if she does not receive a signal of the firm value. An informed analyst receives a signal $y_i = v_i + \varepsilon_i$, where v_i is a random variable drawn from some distribution $G_i(v_i)$ which has mean μ_i and variance σ_i^2, and ε_i is a zero mean error term drawn from a distribution Θ_i. I assume that G_i and Θ_i are such that the conditional variance of v_i, denoted by ω_i^2, is constant for all realizations of y_i.[3] The

---

[1]This is equivalent to assuming that analysts make an observable effort decision, $e_i \in [\underline{e}, \overline{e}]$, which maps into the likelihood of receiving information, i.e. $\Upsilon : [\underline{e}, \overline{e}] \mapsto [0, 1]$. An observable effort decision can be thought of as the analyst's acquisition of a useful data set, meetings with management teams or executives, and non-deal road shows where the analyst connects buy-side investors/clients with the executives of a firm they cover.

[2]I assume that information acquisition is costless to emphasize that the result is not driven by acquisition costs but rather from strategic considerations in the forecast timing. Allowing costly information acqusition does not qualitatively affect the results.

[3]All of the results hold if an informed analyst instead receives a perfect signal $y_i = v_i$, however, I allow Θ_i to be non-degenerate for richer comparative statics.

conditional variance ω_i^2 can be thought of as the analyst's experience or prior forecasting ability (see Clement and Tse (2005)). The total firm (terminal) value, V, is the sum of the two components:

$$V = v_1 + v_2.$$

The components v_1 and v_2 are assumed to be independent, however, the results would not qualitatively change if they are correlated. The firm value takes an additive structure to capture complementarities in the informed analysts' information. The additive nature can be thought of as a firm with two segments in disparate industries, where analysts differ in their industry specialities and thus acquire information pertaining to their area of expertise. However, because each analyst must forecast the total expected terminal value of the firm, she benefits from observing the forecast of her peer. Note that *all* of the results continue to hold under an alternative specification where analysts instead receive a joint Normal signal, i.e. where $V = v$ and $y_i = v + \varepsilon_i$, where $v \sim N\left(\mu_v, \sigma_v^2\right)$ and $\varepsilon_i \sim N\left(0, \sigma_i^2\right)$. This is discussed further in Section 6.3. I assume that the value is additively separable for richer comparative statics (e.g. with regard to σ_i^2), and to allow for more general distributions of v.

The forecasting subgame begins in the final stage of $t = 0$, where each analyst decides when to issue a forecast in continuous time. Analyst i's utility is given by:

$$u_i = \left[s - (m_i - V)^2\right] e^{-r_i \tau_i} \cdot \mathbb{I}, \tag{2.1}$$

where m_i is analyst i's forecast, r_i is analyst i's rate of time preference, and $\tau_i \in [0, \infty)$ is the time in which analyst i issues her forecast.[4] The value s is the compensation the analyst receives from issuing a forecast, where $\sigma_i^2 < s < \sigma_1^2 + \sigma_2^2$, and $\mathbb{I}$ is an indicator function equal to one if a forecast is made.[5,6] Each analyst's utility is decreasing in her

---

[4]Section 6.4 discusses endogenizing the payoff function to incorporate the demand for information (as in Guttman (2010)).

[5]An analyst who does not issue a forecast receives a payoff of 0.

[6]Allowing the uninformed type to rather be partially informed would not qualitatively affect the results. In this case, we would instead have the analyst's type be one of two precision levels, $\rho \in \{\rho_H, \rho_L\}$, which are privately observed by each analyst, where $\rho_{H,i} = \frac{1}{\omega_{H,i}^2}$, $\rho_L = \frac{1}{\omega_{L,i}^2}$, and $\omega_{H,i}^2 + \sigma_j^2 < s < \omega_{L,i}^2 + \sigma_j^2$. For notational and expositional ease, I set $\rho_L = 0$.

Figure 2.1: Timeline of the time 0 stage game.

Time 0, Stage 1.
Time 0, Stage 2.
Time 0, Stage 3.

Analysts choose their
endowment probabilities.

Endowment probabilities
are publicly revealed.
Analysts learn if they are
informed and observe the
corresponding private signal,
if informed.

Analysts begin the
forecasting subgame
and decide when
to issue a forecast,
$\tau_i \in [0,\infty)$.
Time proceeds
continuously.
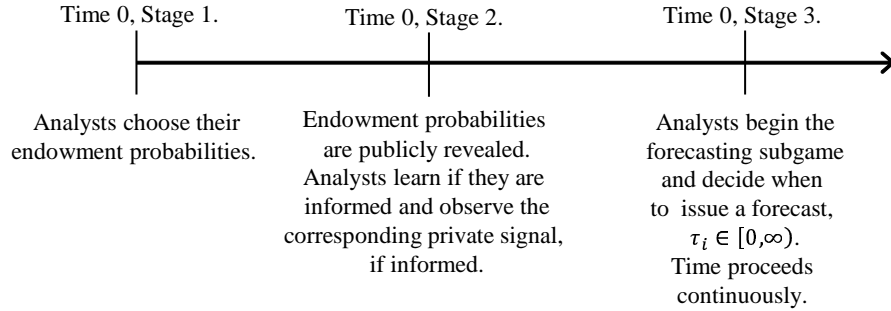
squared forecast error, as illustrated by the disutility $(m_i - V)^2$. The discount rate, $e^{-r_i \tau_i}$, is meant to capture the importance of timeliness of the forecast. Figure 2.1 presents a timeline of the model.

In Section 6.2, I allow for the compensation s to vary in the forecast order to further examine the importance of timeliness. The baseline model also does not include potential biases in the analysts' payoff function; this is explored in Section 4. The equilibrium concept employed is perfect Bayesian equilibrium. The following section characterizes the equilibrium of the baseline model.

Chapter 3

Equilibrium

I first analyze the equilibrium of the forecasting subgame, taking x_i as given. Following the characterization of the equilibrium forecasting strategies, I solve for the endogenous choice of x_i which occurs in the information acquisition subgame.

3.1 Forecasting subgame

In the forecasting subgame, each analyst decides when to issue a forecast. As we can see from the utility structure, each analyst receives a strictly higher payoff if she first observes the announcement of the other analyst before issuing her forecast. Recall that the first analyst to issue a forecast is referred to as the leader, and the second to forecast is referred to as the follower. An informed analyst who forecasts first submits the report:

$$m_i = E\left(v_i | y_i\right) + E\left(v_j\right).$$

We see that the expected utility from an informed leader who forecasts at time t is $Eu_i = \left(s - \omega_i^2 - \sigma_j^2\right) e^{-r_i t}$, and that of an informed follower is $Eu_i = \left(s - \omega_i^2 - \omega_j^2\right) e^{-r_i t}$. An uninformed analyst only issues a forecast when another forecast has been made, and simply repeats the forecast made by the leader. We see that there is a natural incentive to mimic uninformedness, as an analyst will not delay issuing her forecast if she believes her peer to be uninformed. An informed analyst thus trades off the benefit of waiting

to observe the other analyst's forecast with the cost of waiting from a discounted payoff. The equilibrium of the forecasting subgame is akin to a war of attrition with two-sided asymmetric information (e.g. Kreps and Wilson (1982), Ponsati and Sakovics (1995)). The equilibrium is derived in a number of steps and produces a tractable closed-form characterization of the strategies. We see that any equilibrium forecast strategy must have the following properties:

Claim 1: There does not exist an equilibrium in pure strategies.

Suppose we have an equilibrium where analyst 1 forecasts at time t_1 with probability one. In this case, analyst 2 will either forecast immediately after t_1, i.e. at time $t_2 = t_1 + \delta$ for $\delta > 0$, or forecast immediately, i.e. at $t_2 = 0$. If analyst 2 forecasts an instant after t_1, then analyst 1 can do strictly better by saving the discounting costs and forecasting immediately at time 0. However, if analyst 2 does not observe a time 0 forecast from analyst 1, then analyst 2's posterior that analyst 1 is uninformed must go to one, in which case analyst 2 is better off forecasting an instant after time 0 than at some later time. Consequently, this provides an incentive for an informed analyst 1 to delay her forecast time. A similar argument can be made for analyst 2. Hence, any equilibrium must be in mixed strategies.

Let analyst i's forecast strategy be denoted by $\bar{F}_i(t)$, which represents the probability that an informed analyst i forecasts by time t. Let the corresponding hazard rate be denoted by $h_i(t) \equiv \frac{\bar{f}_i(t)}{(1 - \bar{F}_i(t))}$ Similarly, let $F_i(t)$ denote the unconditional probability that analyst i issues a forecast by time t.[1]

Claim 2: Informed analysts will have forecasted with probability one by an endogenously determined, common terminal time T.

To see this, suppose that for an informed analyst 1, her equilibrium path $\bar{F}_1(t)$ reaches one at some time T_1. This implies that an informed analyst 1 would have forecasted with certainty prior to T_1 and the posterior that analyst 1 is uninformed must go to one. Hence, an informed analyst 2 will not delay forecasting if she believes that analyst 1 is uninformed with probability 1, and thus will forecast by T_1 as well. Therefore, $\bar{F}_1(T) = \bar{F}_2(T) = 1$

---

[1] I use the conditional probability in the exposition to better convey the intuition of the equilibrium. All of the claims hold for the corresponding unconditional probability.

at some endogenously determined terminal time T. A similar argument can be made for analyst 2. This implies that the common terminal time must be the minimum of T_1 and T_2.

Claim 3: $\bar{F}_i(t)$ must be continuous for all $t > 0$.

Suppose analyst i's strategy $\bar{F}_1(t)$ had a discontinuity at some time t, then analyst 2 could do strictly better by not forecasting until an instant after time t, i.e. $t + \delta$, rather than issuing a forecast at an instant before time t, i.e. $t - \delta$. Hence, analyst 2 would strictly prefer to wait in the interval $[t - \delta, t]$, but this implies that analyst 1 could do strictly better by forecasting with discrete probability at time $t - \delta$ than at time t.

Claim 4: $\bar{F}_i(t)$ must be strictly increasing until it reaches one.

Suppose $\bar{F}_1(t)$ was constant on some interval, $[a, b]$, which implies that analyst 1 will not forecast in this interval. Analyst 2 can then do strictly better by either forecasting with discrete probability at an instant before time a or an instant after time b. If analyst 2 does not forecast with positive probability until an instant after time b, then analyst 1 can do strictly better by forecasting with discrete probability at time a than at time b, but this violates Claim 3. Likewise, Claim 3 is violated if analyst 2 forecasts with discrete probability an instant before time a.

Claim 5: At most one analyst can have a discrete mass point at time 0.

By Claims 3 and 4, an analyst may forecast with discrete probability only at time 0. However, both analysts cannot have a mass point, since otherwise one analyst can do strictly better by waiting at time 0 with probability one rather than forecasting.

Claims 1-5 indicate that $\bar{F}_i(t)$ must be continuous, strictly increasing, and reaches one at some time T for $i \in \{1, 2\}$. To get a sense of the equilibrium mixing condition, we can examine the local incentive constraint for an informed analyst i.[2] This is given by the following indifference condition:

$$
\begin{aligned}
\left[s - \omega_i^2 - \sigma_j^2\right] e^{-r_i t} = & \beta_j(t) h_j(t) \Delta t \left(s - \omega_i^2 - \omega_j^2\right) e^{-r_i t} \\
& + [1 - \beta_j(t) h_j(t) \Delta t] \left(s - \omega_i^2 - \sigma_j^2\right) e^{-r_i(t + \Delta t)},
\end{aligned}
\tag{3.1}
$$

---

[2]The solution can be alternatively derived using local incentive compatibility conditions, which are sufficient since equation (3.3) is supermodular in t.

where $\beta_j(t)$ is the posterior probability that analyst j is informed by time t and $h_j(t)$ is analyst j's hazard rate. The LHS is analyst i's payoff from forecasting at time t, given that analyst j has not yet forecasted. The RHS of equation (3.1) is analyst i's payoff from waiting one unit of time, Δt,[3] and potentially observing analyst j's forecast. If analyst j issues a forecast in the interval $[t, t + \Delta t]$, then analyst i observes this forecast and issues her own forecast, for a payoff of $\left(s - \omega_i^2 - \omega_j^2\right)e^{-rt}$. However, if analyst j does not forecast in this interval, then analyst i issues a forecast at time $t + \Delta t$, for a payoff of $\left(s - \omega_i^2 - \sigma_j^2\right)e^{-r_i(t+\Delta t)}$, where $\beta_j(t)h_j(t)\Delta t$ captures the probability that an informed analyst j issues a forecast in the interval $[t, t + \Delta t]$. Hence, in equilibrium, each analyst must be indifferent between forecasting immediately and waiting an additional instant. Taking $\Delta t \to 0$ and by L'Hôpital's rule, equation (3.1) becomes:

$$\left(s - \omega_i^2 - \sigma_j^2\right)r_i = \beta_j(t)h_j(t)\left(\sigma_j^2 - \omega_j^2\right). \tag{3.2}$$

Equation (3.2) illustrates the trade-off of waiting for analyst i. The cost of delaying the forecast, which is the loss from discounting (given by the LHS), must exactly offset the benefit of waiting, which is the expected improvement in the error of analyst i's forecast (given by the RHS).

For strategies $F_i(t)$ and $F_j(t)$, an informed analyst i's expected utility is given as:

$$u_i = \max_t \int_0^t \left(s - \omega_i^2 - \omega_j^2\right)e^{-r_i a}f_j(a)\,da + (1 - F_j(t))\left(s - \omega_i^2 - \sigma_j^2\right)e^{-r_i t}. \tag{3.3}$$

Thus, in order for analyst i to be indifferent at every time t, u_i must be constant at all times. Note that the global constraint in equation (3.3) is expressed in terms of the unconditional probability that analyst j forecasts by time t, $F_j(t)$.

Let T_i denote the endogenous terminal time for analyst i, at which an informed analyst i would have forecasted with probability one. Due to the heterogeneity in the analysts' preferences, we may have that $\bar{F}_j(t)$ may reach one earlier than $\bar{F}_i(t)$, and thus implying that $T_j < T_i$. We see how this can arise by considering equation (3.1). Analyst j must set

---

[3]More precisely, intervals of time, Δt, are assumed to be vanishingly small, and can rather be thought of as the limit $\Delta t \to 0$.

$\bar{F}_j(t)$ so that analyst i is indifferent at every instant. If analyst i's LHS payoff is relatively higher than analyst j's, then analyst j will be mixing at a higher rate (i.e. forecasting with greater probability) than analyst i in equilibrium, thus leading $\bar{F}_j(t)$ to reach one at an earlier time than $\bar{F}_i(t)$. Intuitively, if analyst i has a lower option value of waiting (e.g. if $\omega_j^2 > \omega_i^2$), then analyst j must forecast with a relatively higher probability at each instant to keep analyst i indifferent, which leads to $T_j < T_i$. However, by Claim 2, the terminal time for informed analysts i and j must be equal. Hence, we have that both informed analysts must forecast by time $T = \min\{T_1, T_2\}$. In order for this to be the case, analyst i must forecast with discrete probability at $t = 0^+$ to align the path of beliefs, so that T_i is reached at the same time as T_j. The atom at time 0 is uniquely determined by the degree of the heterogeneity in the parameters, thus establishing uniqueness of the forecasting subgame equilibrium. Moreover, this implies that the analyst with the lower option value of waiting forecasts earlier in equilibrium. The solution to the forecasting subgame is characterized as follows:

**Lemma 1** *There is a unique perfect Bayesian equilibrium of the forecasting subgame where (i) an uninformed analyst forecasts if and only if a forecast by the other analyst is made; (ii) both informed types will have forecasted by time $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$; and (iii) the strategy profiles for informed types are given by:*

$$
(F_i(t), F_j(t)) = \begin{cases}
\begin{cases}
F_i(t) = 1 - x_j^{-c_i/c_j} x_i e^{-c_i t} \\
F_j(t) = 1 - e^{-c_j t}
\end{cases} & \text{if } T_j < T_i \\[2em]
\begin{cases}
F_i(t) = 1 - e^{-c_i t} \\
F_j(t) = 1 - e^{-c_j t}
\end{cases} & \text{if } T_j = T_i \\[2em]
\begin{cases}
F_i(t) = 1 - e^{-c_i t} \\
F_j(t) = 1 - x_i^{-c_j/c_i} x_j e^{-c_j t}
\end{cases} & \text{if } T_j > T_i
\end{cases}
$$

where $c_i = \frac{s - \omega_j^2 - \sigma_i^2}{\sigma_i^2 - \omega_i^2} r_j$ and $c_j = \frac{s - \omega_i^2 - \sigma_j^2}{\sigma_j^2 - \omega_j^2} r_i$.

In the symmetric case where the analysts' preferences are homogeneous, neither analyst forecasts with discrete probability at time 0 since it is already the case that $T_j = T_i$. In the heterogeneous case, the above result states that the analyst with the lower option value

Figure 3.1: Forecasting paths and the time 0 mass point.
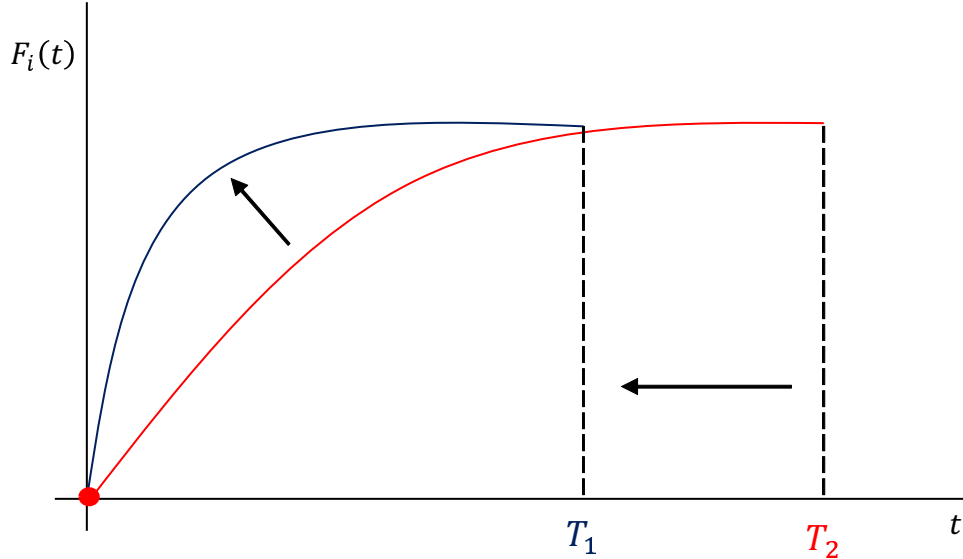
of waiting forecasts at $t = 0^+$ with probability $1 - x_j^{-c_i/c_j} x_i$ or $1 - x_i^{-c_j/c_i} x_j$ if $T_j < T_i$ or $T_j > T_i$, respectively. Lemma 1 also provides a theoretical basis for the use of the leader-follower test statistic, which assumes an exponential arrival of forecasts, by studies that empirically investigate analyst herding (e.g. Cooper et al. (2001), Shroff et al. (2014)). Figure 3.1 illustrates the time 0 mass point.

The option value of waiting is captured by the precision levels, $\frac{1}{\omega_i^2}$, likelihood of endowment, $1 - x_i$, public information, σ_i^2, and the incremental cost of waiting, r_i. We can pin down the expected forecast order according to the analyst who forecasts with discrete probability at time 0:

Corollary 1 *Ceteris paribus, the forecasting time is earlier when analysts have more precise information $\left(T \text{ is decreasing in } \frac{1}{\omega_i^2} \right)$, and the analyst with more precise information forecasts earlier with higher probability.*

The above Corollary states that the analyst with *better* information (higher precision) is expected to forecast first in equilibrium. Moreover, we see that the overall forecasting time improves when just one of the analysts has better information. This corresponds to the findings of Cooper et al. (2001), Clement and Tse (2005), and Keskek et al. (2014)

who find that analysts with superior information issue more timely forecasts. A similar result holds for the endowment probability:

Corollary 2 *Ceteris paribus, the forecasting time is earlier when analysts are more likely to be informed, and the analyst with a higher probability of receiving information forecasts earlier.*

When the probability of being informed, $1 - x_i$, is treated as exogenously given, the result in Corollary 2 states that the analyst with a higher $1 - x_i$ (lower x_i) will forecast earlier in equilibrium. As we will see in the solution of the information acquisition subgame, this incentivizes the analysts to choose a lower endowment probability to induce the other analyst to forecast earlier. The level of public information and the discount rate have a similar effect on the option value of waiting:

Corollary 3 *Ceteris paribus, the forecasting time is earlier when analysts have a higher rate of time preference or have better public information $\left(T \text{ is decreasing in } r_i \text{ and in } \frac{1}{\sigma_i^2} \right)$. The analyst with a higher discount rate forecasts earlier and the analyst whose segment has worse public information $\left(\text{higher } \sigma_i^2 \right)$ forecasts earlier.*

As we expect, discounting speeds up the timing of the forecast and the more impatient analyst forecasts first in equilibrium. Similarly, the better the pool of public information, the more timely the forecasts. This occurs since the option value of waiting is decreasing in the precision of public information, $\frac{1}{\sigma_i^2}$, thus speeding up the forecast times. Surprisingly, we see that the analyst whose segment has worse public information, that is, whose distribution has a higher variance, forecasts earlier. To see this intuitively, suppose that $\max \left\{ \sigma_i^2, \sigma_j^2 \right\} = \sigma_i^2$. This implies that analyst j has a comparatively higher option value of waiting since she can reduce her forecast error more by observing the forecast of analyst i. Hence, the analyst whose segment has relatively worse public information forecasts earlier in equilibrium.

3.2 Information acquisition

Prior to the forecasting subgame, each analyst simultaneously chooses the probability to which she receives information, $1 - x_i$. In the subsequent stage, the endowment likelihoods are publicly revealed and analysts, if informed, privately observe their signals. The choice of x_i is costless and indeed both analysts may choose to receive information with probability one. However, from the equilibrium of the forecasting subgame, analysts have a strategic advantage from being imperfectly informed. As shown in Corollary 2, the analyst with a higher endowment probability forecasts earlier. When choosing x_i, an analyst weighs the cost of being potentially uninformed–not receiving a signal with higher probability–and the benefits from inducing the other analyst to forecast earlier. To find the optimal choice of x_i, it is useful to first examine the simplified case where only one analyst chooses her x_i, keeping the other's fixed. Note that in a single-analyst setting, the analyst will choose to be perfectly informed with probability one.

3.2.1 One-sided choice of x_i

For the analysis in this section, I assume x_j is fixed and known in order to analyze the equilibrium behavior of analyst i. This provides insights that will be helpful when examining the two-sided choice of x_i in the following subsection. We see from Lemma 1 that the equilibrium payoff for an informed analyst i is given as:

$$(1 - F_j(0))\left(s - \omega_i^2 - \sigma_j^2\right) + F_j(0)\left(s - \omega_i^2 - \omega_j^2\right).$$

$F_j(0)$ is the potential atom at time 0, where $F_j(0) = 0$ if $T_i \geq T_j$ and $F_j(0) = 1 - x_i^{-c_j/c_i}x_j$ if $T_i < T_j$. Hence, the equilibrium payoff of an informed analyst i is $s - \omega_i^2 - \sigma_j^2$ when $T_i \geq T_j$. This arises due to the equilibrium condition that the analysts' utility is kept constant at every instant in the forecasting subgame.

First, note that when one, and only one, analyst is known to be informed with probability one, i.e. when $1 - x_j = 1$, the unique equilibrium is where analyst j forecasts immediately. In this case, analyst i's strategy is to issue a forecast only upon observing a forecast by analyst j. To see this for $x_j = 0$ and $x_i > 0$, note that there cannot exist an equilibrium

in which an informed analyst i forecasts with a pure strategy at a time t before analyst j, as analyst i then has a strictly profitable deviation of mimicking uninformedness at time t and withhold forecasting. Also, there cannot be any equilibrium in mixed strategies as in Lemma 1. Mixing breaks down in this case as analyst j would forecast with probability one once her posterior that analyst i is uninformed goes to one, however this violates the condition that analyst i would be indifferent at every instant–as she would strictly prefer to delay forecasting a moment before analyst j forecasts for sure.[4] Hence, any equilibrium must entail a pure strategy of analyst j forecasting prior to analyst i, in which case analyst j then prefers to forecast immediately and save the discounting costs. We can also see this from the indifference condition. The argument implies that $F_i(t) = 0$ for all t, and we can see from equation (3.1) that $\left[s - \omega_j^2 - \sigma_i^2\right] e^{-rt} > \left[s - \omega_j^2 - \sigma_i^2\right] e^{-r(t+\Delta t)}$. Hence, analyst j strictly prefers to forecast immediately than wait in equilibrium when $x_j = 0$ and $x_i > 0$.

Lemma 2 *When $x_j = 0$ and $x_i > 0$, analyst j forecasts at $t = 0$ with probability one, for all ω_i^2, r_i, and σ_i^2.*

This implies that, when x_j is fixed and equal to zero, analyst i has a dominant strategy of choosing arbitrarily small $x_i > 0$ and inducing analyst j to forecast at time 0 with probability one. This is summarized in the following lemma:

Lemma 3 *When $x_j = 0$, then analyst i chooses $x_i = \delta$, where $\delta > 0$ and arbitrarily small, for all ω_i^2, r_i, and σ_i^2.*

When $x_j > 0$, then analyst i must weigh the benefit of inducing analyst j to forecast earlier with the cost of not receiving private information. Note that when $x_j > 0$, analyst i will either choose $x_i > x_j$ or $x_i = 0$, since analyst i receives a strictly higher expected payoff by setting $x_i = 0$ than $x_i \in (0, x_j)$. To focus on the dynamics and economic forces driving information acquisition, I assume that $r_i = r_j$, $\sigma_i^2 = \sigma_j^2$, and $\omega_i^2 = \omega_j^2 = 0$.[5] For

---

[4] This is not a concern when both analysts are potentially uninformed as the posterior that both analysts are uninformed goes to one at the same time. However, this is precluded when one analyst is known to be informed at time 0.

[5] See Section 6.1 for analysis of information acquisition when analysts are heterogeneous.

fixed $x_j > 0$, we see that analyst i maximizes her payoff by setting x_i as a function of x_j only when $x_j \leq \frac{1}{2}$:

Lemma 4 *For fixed $x_j > 0$, analyst i sets x_i such that*

$$x_i = \begin{cases} \left(\frac{x_j \sigma^2}{x_j(s - \sigma^2) + \sigma^2} \right)^{\frac{1}{2}} & \text{if } x_j \in \left(0, \frac{1}{2}\right] \\ 0 & \text{if } x_j \in \left(\frac{1}{2}, 1\right] \end{cases}$$

where $x_i > x_j$ when $x_j \leq \frac{1}{2}$.

Lemmas 3 and 4 imply that analyst i chooses to be imperfectly informed whenever $x_j \leq \frac{1}{2}$, and that the optimal x_i is a function of x_j when $x_j \in \left(0, \frac{1}{2}\right]$. When $c_i = c_j$, we see from Lemma 1 that the size of the atom for analyst j is captured by the expression $1 - \frac{x_j}{x_i}$, and thus the smaller the ratio, the more likely it is that analyst j issues a forecast at time 0. Increasing x_i, however, means that analyst i receives information with lower probability. Absent a signal, analyst i only forecasts by herding with an informed analyst j's forecast. Lemmas 3 and 4 show that there is an interior solution to x_i for a fixed $x_j \leq \frac{1}{2}$. When x_j exceeds $\frac{1}{2}$, the cost of being potentially uninformed with greater probability dominates and analyst i is better off receiving information with probability one and forecasting for sure at time 0. Symmetry of the analysis in the section is also necessary for analytical tractability, however, in Section 6.1 I examine the case where analysts are heterogeneous. These preliminary results prove useful when examining the two-sided choice of x_i in the following section.

3.2.2 Two-sided choice of x_i

I now consider when both analysts simultaneously choose their endowment probabilities. First, we see that there is no equilibrium in pure strategies. For example, consider the case where $x_j > \frac{1}{2}$, then, by Lemma 4, analyst i's best response is $x_i = 0$. However, if this is the case, then by Lemma 3, analyst j can do strictly better by setting $x_j = \delta < \frac{1}{2}$, in which case analyst i can do strictly better by setting $x_i > x_j$. Hence, any equilibrium selection of x_i must include mixing over the interval $[0, q]$, where $q = \left(\frac{\sigma^2}{s + \sigma^2} \right)^{\frac{1}{2}}$. We see that $q \in \left(\frac{1}{2}, 1\right)$ since $\frac{\sigma^2}{s + \sigma^2} < 1$.

Analyst i thus faces the following problem:

$$\max_{x_i} \left\{ \begin{array}{c} E\left[\mathbb{I}_{\{x_i \leq x_j\}}\left[(x_i)(x_j)0 + x_i(1-x_j)(s-\sigma^2) + (1-x_i)(s-\sigma^2)\right]\right] \\ +E\left[\left(1-\mathbb{I}_{\{x_i \leq x_j\}}\right)\left[(x_i)(x_j)0 + x_i(1-x_j)(s-\sigma^2) + (1-x_i)\left[s-\left(\frac{x_j}{x_i}\right)\sigma^2\right]\right]\right] \end{array} \right\}.$$

$$(3.4)$$

The first term in equation (3.4) is analyst i's payoff when she selects an x_i less than x_j, whereas the second term captures analyst i's payoff when x_i exceeds x_j. Let $\Psi(x)$ denote the cumulative distribution function of the analysts' symmetric mixed strategy distribution. Equation (3.4) can thus be expressed as:

$$u_i(x_i) = \int_0^{x_i} \left[x_i(1-x_j)(s-\sigma^2) + (1-x_i)(s-\sigma^2)\right] d\Psi(x_j) \tag{3.5}$$

$$+ \int_{x_i}^q \left[x_i(1-x_j)(s-\sigma^2) + s - x_i s - \left(\frac{x_j}{x_i}\right)\sigma^2 + x_j\sigma^2\right] d\Psi(x_j),$$

We see that the gain from $x_i > x_j$ appears in the second term of equation (3.5), where the extent of the gain is given by the ratio $\frac{x_j}{x_i}$. The potential loss from a higher x_i appears in both terms, however, this is somewhat mitigated by the positive probability that analyst j receives information. A preliminary result concerning the strategy $\Psi(x)$ is given by the following:

Lemma 5 $\Psi(x)$ *is differentiable.*

Lemma 5 states that $\Psi'(y)$ exists and hence the mixing strategy can be expressed in terms of its probability density function $\psi(x)$, i.e. $\Psi(x) = \int_0^x \psi(k)\,dk$. The proof of Lemma 5 also shows that u_i is constant for all $x \in [0,q]$, as is necessary for analyst i to be indifferent. Derivations from equation (3.5) lead to an expression for $\psi(x)$. The following Theorem states one of the central results of the model:

Theorem 1 *In the information acquisition stage, there exists a unique symmetric mixing distribution* $\Psi(x)$, *in which each analyst* $i \in \{1,2\}$ *selects* x_i *according to* $\Psi(x)$. *In the forecasting subgame, each analyst follows the strategy as stated in Lemma 1, with* $\{x_1, x_2\}$ *as specified from the information acquisition stage.*

Theorem 1 states that there is a unique mixing strategy that both analysts follow to

choose their endowment probabilities in the information acquisition stage. This implies that both analysts will always be imperfectly informed, even though they can receive information with probability one at zero cost! We see that this arises due to each analyst's incentive to induce the other analyst to forecast with higher probability at time 0, that is, to give the other analyst the atom. Consequently, both analysts choose to be imperfectly informed in equilibrium. The result implies that the strategic timing of forecasts when there is an opportunity to herd (or observationally learn) leads to inefficient information acquisition, even in the face of zero costs of acquiring information. If one accepts the results from the empirical literature that observational learning among analysts is a ubiquitous and pervasive phenomenon, then the result implies that the very *nature* of analyst forecasting leads to inefficient information acquisition when there are multiple analysts covering a given firm. Theorem 1 also has implications regarding price efficiency, as less information acquisition results in less efficient prices.

The choice of x_i can be considered as an analyst's decision of how many industries or firms to follow; as documented by Clement and Tse (2005), analysts who follow more industries or firms are more likely to herd. In this respect, Theorem 1 helps explain the trend for analysts to cover more firms or become generalists instead of specialists, as this gives them an advantage in the forecast timing. Theorem 1 also implies that analysts have less incentive to acquire information when there is increased analyst following. This implies that individual forecasts will be less informative and that we should see more uninformed herding with greater analyst following. In addition, this suggests that increased analyst following does not necessarily mean a less opaque information environment for the firm. This has implications for proxy variables relying on analyst following as a measure for firm transparency. An alternative measure is to use the timeliness of the forecasts, as Corollary 2 shows that analysts forecast earlier when they are more likely to be informed.

The choice of endowment probability can also be thought of as an effort choice of acquiring information that is made by the analyst at the beginning of each earnings quarter. Observable effort can take the form of scheduled meetings and calls with executives that may increase the likelihood that the analyst has insightful information. This is equivalent to a one-to-one mapping, Υ, of a costless effort choice, $e_i \in [\underline{e}, \overline{e}]$, into the endowment

probability, given as $\Upsilon : [\underline{e}, \overline{e}] \mapsto [0, 1]$, where Υ is strictly increasing in effort. Hence, an analyst who puts in a high level of effort is more likely to receive to useful information. The result in Theorem 1 holds under this alternative formulation as well–analysts choose an inefficient effort level, even though effort is costless, in order to be imperfectly informed. Note that the results are qualitatively unchanged if information acquisition was assumed to be costly; analysts still choose an inefficient level of acquisition when there are multiple analysts.

In addition, the analysts' choice of informedness can be considered in terms of employment, where higher informedness corresponds to an analyst's choice of employment at a boutique research firm which specializes in a particular industry or at a brokerage firm which places significant emphasis on their research division. Hence, this helps to explain why some analysts may join a brokerage house or investment-banking firm whose research division has limited resources. Similarly, we can consider this from a brokerage house or investment-banking firm's perspective, as they may allocate fewer resources into their research division.

Chapter 4

Forecasting with Bias

The previous section examines analyst forecast timing when analysts may be probabilistically informed, receive imprecise information, and can acquire information. Although these are practically relevant features of analyst forecasting, we have also seen through a voluminous literature that analysts generally issue biased forecasts (see Kothari (2001) for a review). However, very few studies to date, to the best of my knowledge, have examined endogenous analyst forecast timing when analysts may be potentially biased.[1] A number of models attribute this bias to arising from trading commissions (e.g. Jackson (2005), Beyer and Guttman (2011)) or to curry favor with management (e.g. Das et al. (1998), Lim (2001)). In this section I show how analyst bias may arise also from the strategic considerations stemming from the endogenous timing of forecasts.

I first examine the model when the potential bias is exogenously given, and then consider the model with an endogenous bias. Suppose that with probability $p_i \in \left[0, \frac{1}{2}\right]$, analyst i's utility function relies on issuing a forecast with an upward bias, and with probability $1 - p_i$, analyst i's utility is the same as in Section 2:

$$u_i = \begin{cases} \left[s - (m_i - b_i - V)^2\right] e^{-r_i \tau_i} \cdot \mathbb{I}, & \text{with probability } p_i \in \left[0, \frac{1}{2}\right] \\ \left[s - (m_i - V)^2\right] e^{-r_i \tau_i} \cdot \mathbb{I}, & \text{with probability } 1 - p_i, \end{cases}$$

---

[1] One exception is Xue (2015), who considers a forecasting model with an analyst who has a known bias and an unbiased analyst. Among several differences, the setting here allows for two-sided bias and bias acquisition. See Section 1.1 for a more detailed discussion of Xue (2015) in relation with this study.

where $b_i > 0$ is a bias term and all other terms are defined as in Section 2. That is, analyst i has bias b_i with probability p_i and is unbiased with probability $1 - p_i$ (similar to Morgan and Stocken (2003) and Meng (2015)), and I assume that p_i and b_i are such that $p_i b_i < \sigma_i$ for $i \in \{1, 2\}$. For simplicity, in this section I assume that $\omega_i^2 = \omega_j^2 = 0$ so that an informed analyst perfectly observes the value of her segment. The analysis follows similarly from Section 3. An informed analyst i's local incentive constraint is given by:

$$
\left[s - \sigma_j^2\right] e^{-r_i t} = \beta_j(t) h_j(t) \Delta t \left(s - p_j^2 b_j^2\right) e^{-r_i t} \tag{4.1}
$$
$$
+ \left[1 - \beta_j(t) h_j(t) \Delta t\right] \left(s - \sigma_j^2\right) e^{-r_i(t+\Delta t)},
$$

where $p_j^2 b_j^2$ is the expected forecast error an informed analyst i incurs after observing analyst j's potentially biased forecast. We see from equation (4.1) that the higher the forecast error from analyst j's potential bias, $p_j^2 b_j^2$, the lower is analyst i's option value of waiting. The derivation of the equilibrium forecasting behavior when analysts are potentially biased is similar to that of Lemma 1. Note that, unlike with x_i, there is no updating on the probability that the analyst is biased as both the biased and unbiased analyst follow the same strategy for their forecast time. We see this in the RHS of equation (4.1), where analyst j, regardless of whether or not she is biased, mixes in accordance with the expected forecast error analyst i retains after observing a forecast from analyst j. The following result presents the equilibrium:

Proposition 1 *There is a unique perfect Bayesian equilibrium of the forecasting subgame when analysts are potentially biased where (i) an uninformed analyst forecasts if and only if a forecast by the other analyst is made; (ii) both informed types will have forecasted by time $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$; and (iii) the strategy profiles for informed types*

are given by:

$$
(F_i(t), F_j(t)) = \begin{cases} \left.\begin{cases} F_i(t) = 1 - x_j^{-c_i/c_j} x_i e^{-c_i t} \\ F_j(t) = 1 - e^{-c_j t} \end{cases}\right\} & \text{if } T_j < T_i \\[2em] \left.\begin{cases} F_i(t) = 1 - e^{-c_i t} \\ F_j(t) = 1 - e^{-c_j t} \end{cases}\right\} & \text{if } T_j = T_i \\[2em] \left.\begin{cases} F_i(t) = 1 - e^{-c_i t} \\ F_j(t) = 1 - x_i^{-c_j/c_i} x_j e^{-c_j t} \end{cases}\right\} & \text{if } T_j > T_i \end{cases}
$$

where $c_i = \frac{s - \sigma_i^2}{\sigma_i^2 - p_i^2 b_i^2} r_j$ *and* $c_j = \frac{s - \sigma_j^2}{\sigma_j^2 - p_j^2 b_j^2} r_i$.

Proposition 1 captures the unique forecasting equilibrium when analysts are potentially biased. We see that the mixing rate, c_i, is characterized by the squared bias, $p_i^2 b_i^2$, of analyst i. As in the equilibrium in Lemma 1, when the final forecast times are misaligned, one analyst must forecast with discrete probability at time 0 to align the path of beliefs. Moreover, the relative biases determine the size of the atom at time 0. When $T_i > T_j$, then analyst i forecasts with probability $1 - x_j^{-c_i/c_j} x_i$ at time 0, and hence the greater is $p_j^2 b_j^2$ relative to $p_i^2 b_i^2$, the greater the probability that analyst i issues a forecast at time 0. We see how this arises from the local incentive constraint in equation (4.1). When $p_j^2 b_j^2$ is high, analyst i has a lower option value of waiting, and thus analyst j's mixing strategy must entail a higher rate of forecasting at every time t to keep analyst i indifferent. The higher mixing rate leads analyst j to reach her terminal time more quickly than analyst i, i.e. $T_j < T_i$. Therefore, to align the terminal times, analyst i has an atom at time 0. This leads to the following ordering according to the potential bias:

Proposition 2 *Ceteris paribus, the forecasting time is earlier when analysts are more biased (T is decreasing in $p_i b_i$), and the analyst who is less biased forecasts earlier with higher probability.*

Proposition 2 shows that the overall forecasting time by informed analysts is earlier when analysts are more biased, and that the terminal time for forecasts speeds up as just one of the analysts becomes more biased. We also see that the less biased analyst

forecasts earlier in equilibrium. This occurs since the less biased analyst forecasts with discrete probability at time 0 to align the path of beliefs. Moreover, the greater the bias of analyst j, the higher the likelihood that analyst i issues a forecast at time 0.

Proposition 2 implies that analysts can benefit from bias if it induces their peer to forecast earlier. As in the preceding section, I now endogenize p_i by allowing analysts to simultaneously choose p_i in the first stage of time 0. In the second stage, p_i becomes publicly revealed and analysts privately learn whether they are informed, their signal (if they are informed), and whether they have a conflict of interest (i.e. if their utility depends on b_i). The analysts then continue in the third stage of time 0 with the forecasting subgame where time proceeds continuously. Moreover, I assume that $x_i = x_j$, $\sigma_i = \sigma_j$, and $r_i = r_j$ to focus on the analyst's acquisition of bias.

The choice of bias probability, p_i, can be interpreted as the analyst's choice of employment, such as a brokerage house, investment banking firm, or independent research institution. Several studies suggest that analysts face a potential conflict of interest when employed at the former two, and document that these analysts typically issue distorted forecasts (e.g. Dugar and Nathan (1995), Lin and McNichols (1997), Michaely and Womack (1999), Mola and Guidolin (2009)). Thus, an analyst who chooses a high p_i can be thought of as affiliated and has joined the research division within a brokerage house, whereas analysts who choose a low p_i can be thought of as independent or whose firms have a reputation for maintaining the "Chinese wall." The equilibrium when p_i is endogenously chosen is stated in the following result:

Theorem 2 *In the bias acquisition stage, the equilibrium choice of p is*
(i) $p_i = 0$, $p_j = \frac{1}{2}$ if $b_i < b_j$,
(ii) $p_i = \frac{1}{2}$, $p_j = 0$ if $b_i > b_j$,
(iii) $p_i = \frac{1}{2}, p_j = 0$ or $p_i = 0$, $p_j = \frac{1}{2}$ if $b_i = b_j$,
and the equilibrium forecasting strategies are given as in Proposition 1.

When analysts are able to choose their level of bias or the likelihood to which their payoff depends on a biased report, we see that analysts have a natural incentive to choose positive levels of each due to the equilibrium effects in the forecasting subgame. However,

the analyst who would receive the compensation with a greater bias, b_i, has an advantage over the other. We see that the unique outcome when $b_i \neq b_j$ is where the analyst with the higher b_i sets $p_i = \frac{1}{2}$ to induce the other analyst to forecast earlier. Moreover, $p_i = \frac{1}{2}$ maximizes the probability of a time 0 forecast by analyst j when $b_i > b_j$. When the bias levels are equal, $b_i = b_j$, we have two asymmetric equilibria where one chooses the maximum bias likelihood and the other sets this likelihood to zero. Note that the results of Theorem 2 would be qualitatively identical if analysts were allowed to choose the bias level b_i while keeping p_i fixed.

The above result also provides a theoretical justification for an analyst's choice of employment for which she may encounter a conflict of interest. Theorem 2 states that the choice of affiliation arises due to the strategic advantage the analyst gains in the forecasting subgame from having a larger potential bias.

Chapter 5

Empirical Implications

5.1 Relation to existing empirical studies

The results help elucidate a number of empirical regularities documented in the literature. As shown in Corollaries 1 and 2, analysts are more likely to be the leading forecasters when they have comparatively more precise information or a higher likelihood of receiving information. This corresponds to the findings of Cooper et al. (2001), Clement and Tse (2005), and Keskek et al. (2014), who find that analysts with superior information or greater forecasting ability issue more timely forecasts. Interestingly, the results of the model concerning the time 0 forecast by the more precise analyst almost exactly correspond to the findings of Keskek et al. (2014). They find "that the timing of individual analysts' forecasts within an information production phase is strongly related to forecast quality" (p. 1507). The time 0 forecast by one of the analysts occurs at the end of the information production/discovery stage of the model, and is made by the analyst with comparatively more precise information. Indeed, the economic forces driving the results of the model–that more precise analysts have a lower option of waiting–seem to govern analysts' behavior in practice.

The corresponding empirical results concerning follower or herding analysts are also captured by the model. As Hong et al. (2000), Clement and Tse (2005), and Jegadeesh and Kim (2010) find, herding analysts are more likely to be of weaker ability. Recall that the likelihood of receiving information corresponds to an analyst's level of specialization

or focus; analysts who are particularly specialized, or who work in specialized research departments, are more likely to receive information than analysts who follow multiple industries or who can be considered generalists. Moreover, the model shows that pure imitation herding occurs only by uninformed analysts, who simply repeat the forecast (or the "consensus") made by the leader. This is in line with the results of Jegadeesh and Kim (2010), and particularly of Clement and Tse (2005), who document this phenomenon as "uninformed herding" (p. 310), and find that analysts who follow more industries are more likely to herd. The results also help to explain the general trend for analysts to cover more firms. The results give an endogenous explanation for why some analysts become generalists and others specialists, as it is driven by strategic considerations in the forecast timing.

The results capture several other important properties of analyst forecasting. As documented in Cooper et al. (2001), timeliness of the forecast is a more informative indication of an analyst's forecasting ability rather than the forecast error. The results of the model confirm Cooper et al. (2001)'s conclusions as the more precise analyst forecasts earlier, even though her resulting forecast error may be higher than the subsequent analysts who issue less timely forecasts. Furthermore, Cooper et al. (2001) find that leaders have a greater market impact than followers. The results of the model imply that an analyst leader will always be informed, thus conveying material information and inciting a market reaction, whereas the follower analyst may sometimes be uninformed and thus issuing a forecast that has little effect. Likewise, a follower may also be informed, albeit with a lower precision, and thus will sometimes elicit a market reaction. The latter effect is evidenced by Shroff et al. (2014), who find that the forecasts of follower analysts can also contain useful information.

This study is also relevant to the empirical papers investigating the properties and interactions of affiliated and independent analysts. We see from Proposition 2 that affiliated or biased analysts are more likely to observe material information of the other analysts before forecasting, which implies that affiliated analysts should have a lower forecast error than independent analysts. This phenomenon is documented by Jacob et al. (1999) and Jacob et al. (2008), who find that affiliated analysts issue more accurate forecasts

than independent analysts. We also see that affiliated analysts have a lower forecast error when independent/unbiased analysts are also following the firm, which comports with the findings of Gu and Xue (2008). The results also correspond to the findings of Das et al. (1998), where analysts issue more biased forecasts when public information is weaker.

5.2 Empirical predictions

A number of empirical predictions emerge from the analysis in this study. The results imply that the posterior probability that an analyst is uninformed increases over time as she delays issuing her forecast. This implies that analysts who forecast after the leader are more likely to exhibit *uninformed herding* at later stages in the forecasting period rather than earlier. Hence, one empirical prediction is that clustering of forecasts that occurs later contains less information than clustering which emerges earlier. Moreover, herding which occurs later will be more in line with an informational cascade rather than herding which occurs earlier, which would rather be closer to observational learning. This also implies that forecasts of follower analysts will induce less market reaction when made in a later stage rather than earlier.

The results also suggest several predictions concerning the overall timing and speed to which forecasts are made. When a single highly informed analyst is following the firm, the forecasting time for all analysts who cover the firm should be earlier. Moreover, as we see in Corollary 1, the timeliness of forecasts crucially depends on the precision of the most informed analyst, rather than on the number of analysts who are covering the firm. Hence, the ability of the lead analysts dictate the overall forecast timing, rather than the number of analysts. Similarly, Corollary 2 implies that the overall timeliness of the forecasts improves with the specialization of the lead analysts. Thus, if one highly specialized analyst is following the firm, this should significantly improve the overall forecasting times of all the analysts who are following the same firm.

Several paradoxical or counter-intuitive predictions also emerge. The results in Theorem 1 imply that individual analysts will have higher forecast errors when more analysts are following the firm. This occurs since the strategic forecast timing induces less informa-

tion acquisition by an analyst when other analysts are following the same firm. Hence, we should see relatively less accurate forecasts and more uninformed herding when the firm has a greater analyst following. The results in Theorem 2 imply that the level of bias or distortion in the forecasts should be relatively higher when more analysts are following the firm. We thus have the empirical prediction that firms who have a higher analyst following can be met with forecasts which entail comparatively greater uninformed herding, greater bias, and less accuracy. This suggests that analyst following may not necessarily be indicative of the transparency of a firm's information environment. As described shortly, an alternative measure is the timeliness of forecasts.

Furthermore, the results imply that independent or unaffiliated analysts issue more timely forecasts. The expected forecasting times for all analysts following a firm should be earlier when the firm is followed by affiliated analysts, and the speed of the overall forecasting times has a monotonic relation with the level of bias of the affiliated analysts. For example, the results imply that forecasts are made in a more timely fashion and information aggregation should occur more quickly when analysts affiliated with the underwriting bank are following a recently public firm after its IPO quiet period.

The results also have implications for the firm's informational environment. The results imply that the timeliness of analysts' forecasts, rather than the number of analysts following the firm, is a more suitable measure for the strength or transparency of a firm's informational environment. This follows from three distinct implications of the model: (*i*) Corollaries 1 and 2 show that the forecasting time is earlier when analysts have more precise information and when they are more likely to be informed, respectively; (*ii*) Theorem 1 implies that increased analyst following may not necessarily lead to a reduction in information asymmetry; and (*iii*) Corollary 3 shows that forecasts will be more timely as public information improves. Taken together, there is a strong theoretical argument that the timeliness of forecasts may be more informative than analyst following of the opaqueness of a firm's informational environment.

There are also predictions concerning the role of public information on the timing of analysts' forecasts. For multi-segmented firms, the results imply that analysts who specialize in the industry/segment with *worse* public information will forecast earlier than

analysts who are following the same firm but whose industry specialization has comparatively stronger public information. However, the overall forecasting time of both analysts in a multi-segmented firm is always expedited as public information improves (in either segment).

Chapter 6

Extensions

6.1   Information acquisition with heterogeneous analysts

In the information acquisition model of section 3.2, I assumed that analysts were ex ante symmetric to highlight the economic forces driving information acquisition as well as to make the analysis analytically tractable for the two-sided acquisition choice. In this section, I allow analysts to be heterogeneous when there is a one-sided information acquisition choice (i.e. where $\omega_i^2 \neq \omega_j^2$, $\sigma_i^2 \neq \sigma_j^2$, or $r_i \neq r_j$). This substantially complexifies the analysis and the two-sided acquisition choice under this setting is no longer tractable, however, even the one-sided acquisition choice shows very interesting properties when there are heterogeneous analysts.

In the one-sided acquisition choice, assume that analyst i must choose her endowment likelihood, $1-x_i$, when analyst j's endowment probability, $1-x_j$, is fixed and within $(0, 1)$. In the homogeneous case, analyst i can induce analyst j to have the atom by setting x_i even slightly higher than x_j. However, under heterogeneity, this no longer holds. For example, if an informed analyst i receives an ex ante more precise signal, she may still have the mass point at time zero even if $x_i > x_j$. Similarly, if her discount rate, r_i or the ex ante variance of her distribution, σ_i^2, is sufficiently higher (see Corollary 3), analyst i may still have the mass point when x_i is above x_j. Hence, under heterogeneity, analyst i's strategy is to choose x_i sufficiently larger than x_j such that this compensates for the difference in, for example, their relative signal precisions so as to induce analyst j to forecast at time

0 with discrete probability. However, if the requisite x_i to induce analyst j's atom is too high so that it becomes unprofitable for analyst i, she would then rather set $x_i = 0$ and forecast immediately. Likewise, if the parameters are such that analyst j has the mass point ex ante (for example, if analyst j's signal is much more precise than analyst i's), then x_i will be determined so as to maximize analyst j's mass point at time 0.

Let $u_i^0 \equiv u_i (x_i = 0)$ be analyst i's expected utility from setting $x_i = 0$ and forecasting immediately, i.e. $x_i = 0 = s - \omega_i^2 - \sigma_j^2$. Similarly, let $u_i^* \equiv u_i (x_i = x_i^*)$ denote expected utility for analyst i when she she sets her endowment probability to x_i^*. When analysts are allowed to be heterogeneous, the optimal information acquisition choice by analyst i is given by the following Proposition:

Proposition 3 *For fixed $x_j \in (0, 1)$, analyst i sets x_i such that:*

$$
x_i = \begin{cases} x_i^* & \text{if } u_i^* \geq u_i^0 \\ 0 & \text{otherwise,} \end{cases}
$$

where x_i^* is the solution to:

$$
x_i^{-\frac{c_j}{c_i}} \left(1 - \frac{c_j}{c_i} + \frac{c_j}{c_i} x_i^{-1} \right) = \frac{1}{\sigma_j^2 - \omega_j^2} \left[\left(s - \omega_j^2 - \sigma_i^2 \right) + \frac{\sigma_i^2 - \omega_i^2}{x_j} \right] \tag{6.1}
$$

where $c_i = \frac{s - \omega_j^2 - \sigma_i^2}{\sigma_i^2 - \omega_i^2} r_j$ and $c_j = \frac{s - \omega_i^2 - \sigma_j^2}{\sigma_j^2 - \omega_j^2} r_i$ for all ω_i^2, σ_i^2, and r_i.

The above Proposition states that analyst i sets x_i according to equation (6.1) when it is profitable to induce analyst j to have the atom at time 0. Otherwise, she chooses to be perfectly informed and forecast immediately with probability one. If analyst i has the mass point ex ante (for example, if $\omega_i^2 < \omega_j^2$), then she will either set $x_i = x_i^* > x_j$ to induce analyst j to have the mass point instead of herself, or set $x_i = 0$ if the ex ante discrepancy is sufficiently large such that it becomes unprofitable (since x_i must be set very high) to induce analyst j to have the mass point.

Proposition 3 can perhaps be more easily understood through the use of a numerical example. I focus on the more interesting case in which analyst i has the mass point at time zero ex ante. For the example, suppose that the analysts are heterogeneous in the precision

Table 6.1: Optimal x_i when $\omega_i^2 = 4$, $\omega_j^2 = 6$, $s = 20$, $r_i = r_j = 0.5$, and $\sigma_i^2 = \sigma_j^2 = 12$

	x_i	T_i	T_j	u_i^*	u_i^0	α_i	α_j
$x_j = 0.01$	0.32	9.03	13.82	6.58	4	0	0.80
$x_j = 0.05$	0.48	5.83	8.99	5.01	4	0	0.65
$x_j = 0.1$	0.57	4.52	6.91	4.17	4	0	0.55
$x_j \geq 0.115$	0	0	≤ 6.49	< 4	4	1	0

of their signals (as denoted by the conditional variance ω_i^2) when informed. Specifically, let $\omega_i^2 = 4$ and $\omega_j^2 = 6$, while the other parameters are symmetric: $s = 20$, $r_i = r_j = \frac{1}{2}$, and $\sigma_i^2 = \sigma_j^2 = 12$.[1] In this example, if $x_i = x_j$, then analyst i would have the mass point at time 0 according to Corollary 1. We have that $\frac{c_j}{c_i} = \frac{8}{3}$ and equation (6.1) is given as $x_i^{-\frac{8}{3}}\left(-\frac{5}{3} + \frac{8}{3}x_i^{-1}\right) = \frac{1}{3} + \frac{8}{(x_j)(6)}$. Let α_i denote the time 0 mass point for analyst i and α_j similarly for analyst j. Table 6.1 shows the optimal x_i for varying values of x_j.

This example highlights very interesting properties of the optimal information acquisition choice with heterogeneous analysts. First, we see that the x_i required to overcome analyst i's precision advantage is always significantly higher than x_j. Analyst i must adopt a substantially higher likelihood of not receiving information so as to reduce the option value of waiting for analyst j. This then increases analyst j's time at which her posterior goes to 1, T_j, thus leading analyst j to have the mass point at time 0 rather than analyst i (since $T_j > T_i$). Second, we see that this strategy of being substantially uninformed is optimal for analyst i. Specifically, when $x_j = 0.1$, analyst i chooses to be uninformed with 57% probability, which gives her an expected payoff higher than that of being informed with probability one. Third, we see that the threshold level at which analyst i chooses to be informed with probability one is reached much more quickly than if analysts were ex ante homogeneous. In the example above, analyst i chooses to be perfectly informed and forecast immediately whenever x_j exceeds 0.1. This is in contrast to the symmetric case in Lemma 4, where there is a much higher threshold $\left(x_j > \frac{1}{2}\right)$ necessary to induce analyst i to set $x_i = 0$. The reason for this is related to the first property of the example–that x_i must significantly exceed x_j in order to compensate for analyst i's ex ante precision ad-

---

[1] This is primarily for expositional ease. The numerical results are qualitatively similar if analysts are also heterogeneous in r and σ^2.

vantage. As x_j increases, analyst i must adopt a comparatively higher probability of being uninformed to induce analyst j to have the mass point, however, this becomes unprofitable at some point for analyst i as it rather lowers her expected payoff.

Lastly, this example shows the downward pattern in analyst i's expected utility as analyst j becomes less informed (as x_j increases). This occurs as analyst i can induce a relatively larger mass point for lower levels of x_j (i.e. α_j is decreasing in x_j), and also since analyst j's relatively lower endowment probability has a direct effect on analyst i's expected utility. It is also interesting to note from this example that the requisite x_i necessary for analyst j to have the mass point becomes larger as analyst i's precision advantage increases. Consequently, this leads to a lower x_j threshold necessary to induce analyst i to choose to be perfectly informed. This is apparent if we rather have $\omega_i^2 = 2$ in the numerical example. In this case, the threshold where $x_i = 0$ is reached is with a much lower threshold x_j. Similar examples can be shown for when analysts are heterogeneous in the variances of their distributions and in their discount rates. Although the analysis in the two-sided acquisition choice is no longer tractable with heterogenous analysts, one-sided acquisition under heterogeneity preserves the equilibrium features of the homogeneous model and also provides additional insights.

6.2 Compensation depends on forecast order

As noted by Groysberg, Healy, and Maber (2011), analysts are rewarded for issuing timely forecasts through promotions and trading commissions from increased volume. In the baseline model, timeliness of the forecast is captured by the discounted compensation the analyst earns when she delays the forecast. An analyst who forecasts earlier saves on the discounting costs, however, they are not necessarily rewarded more for forecasting first. To emphasize the importance of timeliness, I allow for the analyst's salary to vary by the order in which she forecasted and examine the equilibrium effects in the forecasting subgame. Specifically, the analysts' utility function is given by:

$$u_i = \left[s_k - (m_i - V)^2 \right] e^{-r_i \tau_i} \cdot \mathbb{I},$$

where $k \in \{L, F\}$. The leader receives a salary s_L, and the follower receives s_F, where $s_L > s_F > \sigma_i^2$ and $s_L < \sigma_i^2 + s_F < \sigma_1^2 + \sigma_2^2$. I assume here that $\omega_i^2 = \omega_j^2 = 0$, so that informed analysts receive a perfect signal of the value of their segment v_i, however, the results in this section would not be qualitatively affected if signals were imperfect and with different variances. The remaining structure is the same as in the baseline model.

Each analyst now has a lower option value of waiting given the difference in salaries. The local incentive constraint for an informed analyst i is given as:

$$\left[s_L - \sigma_j^2 \right] e^{-r_i t} = \beta_j(t) h_j(t) \Delta t (s_F) e^{-r_i t} \tag{6.2}$$
$$+ \left[1 - \beta_j(t) h_j(t) \Delta t \right] \left(s_L - \sigma_j^2 \right) e^{-r_i(t+\Delta t)},$$

We see from equation (6.2) that analyst j must forecast with a higher rate since the LHS is now relatively larger. The equilibrium of the forecast subgame is given as:

Proposition 4 *When the analysts' compensations are determined by the forecast order, the equilibrium of the forecasting subgame consists of strategies $(F_i(t), F_j(t))$ as defined in Lemma 1, except where $c_j = \frac{s_L - \sigma_j^2}{\sigma_j^2 - (s_L - s_F)} r_i$ and $c_i = \frac{s_L - \sigma_i^2}{\sigma_i^2 - (s_L - s_F)} r_j$. The overall forecasting terminal time is decreasing in the difference $s_L - s_F$.*

Because the option value of waiting is lower, each analyst must mix with a higher rate of forecasting to keep the other analyst indifferent in equilibrium. This consequently lowers the overall forecasting time. Moreover, the forecasts are more quickly issued as the reward for forecasting first increases. This is expected as the option value of waiting is decreasing in the difference $s_L - s_F$. We see that delay continues to persist in the forecast time, as the option value of waiting is positive. However, delay vanishes in the limit where $s_L - s_F \to \sigma_i^2$, since the value of waiting then tends to zero. The equilibrium of the information acquisition subgame remains qualitatively unaffected.

Note that this alternative specification is equivalent to a setting where an analyst who forecasts first is punished less severely in terms of the disutility they receive from the forecast error, i.e. if the utility function was rather:

$$u_i = \left[s - \alpha_k (m_i - V)^2 \right] e^{-r_i \tau_i} \cdot \mathbb{I},$$

where $\alpha_k \in (0,1)$ for $k \in \{L, F\}$, $\alpha_L < \alpha_F$, and $\alpha_L \sigma_i^2 < \alpha_F \sigma_i^2 < s < \alpha_L \left(\sigma_1^2 + \sigma_2^2\right)$. In this case, the leader's forecast error has a coefficient α_L, whereas the coefficient on the forecast error is α_F for the follower. This specification conveys the intuitive notion that analysts who forecast early may be punished less severely for inaccuracies as relatively less information is available at the time of their forecast. This setting is equivalent to the preceding specification where the analysts' compensation depends on the forecast order and yields qualitatively identical results to Proposition 4 and to the results in Section 3.

6.3 Alternative value function

The baseline model assumes that the total value of the firm, V, is the sum of the values of two segments: $V = v_1 + v_2$. I impose this structure to allow for general and heterogenous distributions of v_i, however, the results in both the forecasting subgame and information acquisition subgame are qualitatively unaffected under alternative value structures. Specifically, if we have a single segment, $V = v$, where each analyst receives an imperfect signal, $y_i = v + \varepsilon_i$, with some probability $1 - x_i$, then the results go through as long as the conditional variance of v is decreasing in the number of signals and does not depend on the value of y_i.

For example, suppose that $v \sim N\left(\mu_v, \sigma_v^2\right)$ and $\varepsilon_i \sim N\left(0, \sigma_{\varepsilon_i}^2\right)$, where $\sigma_{\varepsilon_i}^2 > 0$. As in the baseline case, each analyst is informed with probability $1 - x_i$ and has a payoff function as in equation (2.1). Let ω_i^2 denote the conditional variance for an informed analyst i who observes $y_i = v + \varepsilon_i$, i.e. $\omega_i^2 = \left(E\left(v|y_i\right) - v\right)^2$, where $\omega_i^2 < s < \sigma_v^2$. Let ϕ^2 denote the conditional variance for an informed analyst who observes both y_1 and y_2, i.e. $\phi^2 = \left(E\left(v|y_1, y_2\right) - v\right)^2$. We see that each analyst has an incentive to delay her forecast to observe the forecast of her peer, as $\phi^2 < \omega_i^2$. I assume that the signals are independent, however, the results would not be qualitatively affected if there is correlation. The indifference condition for an informed analyst i is then given by:

$$\left[s - \omega_i^2\right] e^{-r_i t} = \beta_j(t) h_j(t) \Delta t \left(s - \phi^2\right) e^{-r_i t} \tag{6.3}$$
$$+ \left[1 - \beta_j(t) h_j(t) \Delta t\right] \left(s - \omega_i^2\right) e^{-r_i(t+\Delta t)}.$$

The analysis is qualitatively similar to the baseline case. Moreover, the properties of the forecast order are preserved in this setting as well. The equilibrium of the forecasting subgame is given in the following result:

Proposition 5 *When the firm value and signals are joint normal, the equilibrium of the forecasting subgame consists of strategies $(F_i(t), F_j(t))$ as defined in Lemma 1, except where $c_j = \frac{s - \omega_i^2}{\omega_i^2 - \phi^2} r_i$ and $c_i = \frac{s - \omega_j^2}{\omega_j^2 - \phi^2} r_j$. The corresponding results in Corollaries 1, 2, 3, and Theorem 1 continue to hold.*

The proof of Proposition 5 proceeds similarly from the results in Section 3. We see that the option value of waiting is given by the decrease in the forecast error from observing the other analyst's forecast, $\omega_i^2 - \phi^2$, and that this determines the mixing rate. Proposition 5 also states that the forecast order is still determined by the analysts' relative precision and likelihoods of receiving information; all else equal, the analyst with a lower ω_i^2 forecasts earlier and the analyst with a higher $1 - x_i$ forecasts earlier. The equilibrium of the information acquisition stage is qualitatively unaffected as well, since analysts continue to have an incentive to receive information with probability less than one in order to induce the other analyst to forecast earlier. The results concerning forecasting with a potential bias and bias acquisition as shown in Section 4 continue to hold under this alternative specification as well. The normal framework is used in the following subsection to explore the model's robustness when informational content of the forecast is included in the analysts' payoff function.

6.4 Compensation depends on informational content of forecast

The baseline model captures the importance of accuracy through the squared error of the analyst's forecast. This depicts an analyst's incentive to disseminate accurate information to their buy-side clients and retail investors. It can also be argued that an analyst's compensation or payoff also relies on the informational content of the report, that is, the new information (or interpretation/insight) that an analyst brings to the market. As

mentioned previously (see footnote 2), analysts may obtain useful information from other analysts in other non-public channels before the forecasts have been widely distributed (e.g. through their common buy-side clients). However, it may also be the case that analysts are rewarded on the market impact/reaction of their report or on completely new information that they contribute.

To investigate this further, assume that, as in Section 6.3, $V = v$, $v \sim N\left(\mu_v, \sigma_v^2\right)$, and $\varepsilon_i \sim N\left(0, \sigma_{\varepsilon_i}^2\right)$, where $\sigma_{\varepsilon_i}^2 > 0$, and each analyst, if informed, receives a signal $y_i = v + \varepsilon_i$. The information content of the forecast can be captured as in Admati and Pfleiderer (1986) and Guttman (2010). Let $\pi_i \equiv \frac{1}{\omega_i^2}$ denote the conditional precision of an informed analyst i's forecast, and let $\rho_t \equiv \frac{1}{\varphi_t^2}$ denote the precision of publicly available information. The additional component of the utility function that depends on new information is represented as:

$$C_t = \log\left(\frac{\pi_i + \rho_t}{\rho_t}\right). \tag{6.4}$$

Note that ρ_t is constant except when a forecast is released by one of the analysts. Since there is no further private information gathering by either analyst after time 0, π_i is fixed but ρ_t may increase after a forecast is released. To examine the effects of information content-based compensation, consider the alternative utility function which incorporates C_t:

$$u_i = \left[s + C_t - (m_i - V)^2\right] e^{-r_i \tau_i} \cdot \mathbb{I}. \tag{6.5}$$

In this case, there is an additional benefit in forecasting early, as C_t is relatively larger for the analyst that forecasts first. However, the benefit of learning from the other analyst is maintained through the disutility in the forecast error. Hence, the payoff structure in equation (6.5) is qualitatively equivalent to that in Section 6.2, as there is an additional benefit to the leader. It is straightforward to show that the qualitative equilibrium properties of Section 2 are preserved under this alternative utility structure as well. Thus, compensation partly based on the informational contribution of the forecast does not qualitatively affect the results when analysts prefer to have a lower forecast error.

6.4.1 Endogenous payoff function

As shown by Admati and Pfleiderer (1986) and Guttman (2010), equation (6.4) can be derived endogenously from the payoff that an analyst receives from selling her forecast to an uninformed investor. The same argument can be applied here to endogenously generate the analyst's payoff function, discounted to the time in which she issues her forecast:

$$u_i = C_t e^{-r_i \tau_i} \cdot \mathbb{I}, \tag{6.6}$$

where the risk aversion parameter in Guttman (2010) is set to $\gamma = \frac{1}{2}$ and where signals and values are normally distributed as in Section 6.3 (i.e. an informed analyst receives a signal $y_i = v + \varepsilon_i$, where $v \sim N\left(\mu_v, \sigma_v^2\right)$, $\varepsilon_i \sim N\left(0, \sigma_{\varepsilon_i}^2\right)$, where $\sigma_{\varepsilon_i}^2 > 0$). Under the payoff structure in equation (6.6), an additional assumption on the information structure is necessary to capture analysts' observational learning from their peers. For example, there may exist complementarities in the analysts' information sets such that the precision of analyst i's signal is improved upon observing the forecast and learning the information of analyst j. This may happen, for instance, when analysts have different comparative advantages and backgrounds such that the information released by one analyst may help with the discovery, analysis, or interpretation of the information of another analyst. For example, suppose analyst 1 specializes in firm-specific expertise and analyst 2's advantage is with overall industry expertise. A report/forecast by analyst 2 on the financial soundness of major suppliers or customers may allow analyst 1 to better assess the firm's long-run strategy and direction, thus allowing her to make a more accurate forecast.

To capture complementarity in the information sets of the two analysts, we can assume that the errors, ε_1 and ε_2, covary so that there is additional benefit to observing the forecast, and thus the signal, of the other analyst, i.e. by setting $cov\left(\varepsilon_1, \varepsilon_2\right) = c \neq 0$.[2]

---

[2]For c to result in a reduction of the conditional variance, the following condition must be satisfied:

$$\frac{\sigma_1^2 \sigma_2^2}{\sigma_v^2 \left(\sigma_1^2 + \sigma_2^2\right) + \sigma_1^2 \sigma_2^2} > \frac{\sigma_1^2 \sigma_2^2 - c^2}{\sigma_v^2 \left(\sigma_1^2 + \sigma_2^2\right) + \sigma_1^2 \sigma_2^2 - 2\sigma_v^2 - c^2}$$

In the notationally less cumbersome symmetric case where $\sigma_1^2 = \sigma_2^2 \equiv \sigma_i^2$, c must satisfy:

$$c > \frac{\sigma_i^2 - \sigma_i^4 + 2\sigma_v^2 - 2\sigma_v^2 \sigma_i^2}{\sigma_i^2}$$

In this case, the results of the baseline model are qualitatively preserved as long as there is sufficient reduction in the conditional variance of analyst i's signal upon observing the forecast of analyst j. Alternatively, complementarities can be depicted as in Kim and Verrechia (1994, 1997), where each informed analyst receives a private signal of the firm's value or earnings, $y_i = v + \varepsilon_i$, as well as an additional signal regarding the error of the other analyst's information, $O_i = \varepsilon_j - \delta_i$, where $\delta_i \sim N\left(0, \sigma_{\delta_i}^2\right)$. In this case, the signal O_i is only helpful to analyst i if she observes the forecast of analyst j, thus giving her additional information beyond the signal of analyst j, represented as $y_{i,O} \equiv y_j - O_i = v + \delta_i$. For example, analyst i has private information regarding the random error in analyst j's report due to her understanding of analyst j's firm or information discovery/analysis methods (e.g. the model that analyst j uses to generate her forecasted numbers). This plausible alternative information structure also generates benefits to observational learning by the analysts when $\sigma_{\delta_i}^2$ is sufficiently low (and similarly for $\sigma_{\delta_j}^2$). The qualitative equilibrium properties under this utility and information structure are also preserved as the economic forces driving the model–the benefit to waiting and from inducing more timely forecasting by the other analyst–continue to hold.

One should note, however, that the above specifications do not generate significant insights beyond the utility representation in the baseline model and also cause the analysis to be far more algebraically cumbersome. Nevertheless, these alternative utility and information structures allow the payoff function to be endogenized from an investor's demand for information and demonstrate that the results are robust to alternative modeling specifications.

6.5 Commitment to forecasting time

The baseline model assumes that analysts are not committed to any particular forecast time and can issue their forecast as time proceeds. However, it is plausible that an analyst may announce the date in which she expects to release her forecast, and thus, in some

---

or, in the case of negative covariance,

$$c < -\sigma_i^2 - 2\sigma_v^2.$$

sense, committing ex ante to a forecast time. In this extension, I consider the alternative model in which analysts publicly announce a forecast time to which they are committed and cannot issue a forecast at any other time. I first examine the case in which the selection of forecast times is sequential, where analyst j selects her forecast time τ_j and then, having observed τ_j, analyst i selects her forecast time τ_i. Under this setting, analyst j will choose her forecast time τ_j in order to induce analyst i to forecast immediately. That is, she will set τ_j such that analyst i finds it more profitable to set $\tau_i = 0$ rather than an instant after analyst j's forecast time, $\tau_i = \tau_j + \varepsilon$. This is given by:

$$e^{-r_i \tau_j^*} \left(s - \omega_i^2 - \omega_j^2 \right) \leq s - \omega_i^2 - \sigma_j^2, \tag{6.7}$$

which is an equilibrium strategy for analyst i if this is more profitable for analyst j than from forecasting immediately:

$$e^{-r_j \tau_j^*} \left(s - \omega_i^2 - \omega_j^2 \right) \geq s - \sigma_i^2 - \omega_j^2.$$

If the above condition does not hold, then analyst j sets $\tau_j = 0$ and analyst i sets $\tau_i = \varepsilon$, as any $\tau_j < \tau_j^*$ will induce analyst i to set $\tau_i = \tau_j + \varepsilon$. In this case, analyst j can do strictly better by forecasting immediately. This analysis assumes that both analysts are perfectly informed, however, endogenizing information acquisition leads to a similar result as the main result of the baseline model. If there is some probability x_i that analyst i is uninformed, analyst j's payoff must satisfy

$$e^{-r_j \tau_j^*} \left(\left(s - \omega_i^2 - \omega_j^2 \right) (1 - x_i) + \left(s - \sigma_i^2 - \omega_j^2 \right) (x_i) \right) \geq s - \sigma_i^2 - \omega_j^2.$$

Similarly, if analyst j was also probabilistically uninformed, a lower forecast time would be required to satisfy equation (6.7). In this case, the equilibrium information acquisition strategy again involves mixing over x_i.

In the case where the selection of the committed forecast time is simultaneous, the equilibrium entails mixing over the selection of τ_i. This is similar to the argument made in Claim 1–no two pure strategy forecast times can be mutual best responses. Information

acquisition is less clear in this context though it seems to be the case that it remains to be inefficient if it leads to more density over lower values of τ_i in the forecast time selection stage.

6.6 Learning over time

In the baseline model, analysts can only learn their private information at time 0, at which point they are either informed or uninformed. However, it may be the case that analysts receive information at different points in time and can also learn new private information by delaying their forecast. In this section, I extend the baseline setting by allowing analysts to learn information over time. Suppose information arrives for each analyst according to a Poisson process with intensity λ. The remaining structure is the same as in the baseline model, where $V = v_1 + v_2$, analysts have utility as in equation (2.1), and $y_i = v_i$. The signal and the arrival time T are privately observed by each analyst. The arrival time is thus exponentially distributed, where the probability of being informed for each analyst is:

$$\Omega(t) = 1 - e^{-\lambda t}.$$

We see that delay by an informed analyst will be present in this setting as well. I conjecture that an analyst who learns information at time T will wait an additional amount of time $t(T)$ before forecasting, where $t(T)$ is the delay strategy based on the arrival time T. The time the forecast is issued is thus $T + t(T)$. There are two simplifying conditions regarding a symmetric pure strategy $t(T)$. The first is that an analyst who receives information at time 0 will issue her forecast immediately, i.e. $t(0) = 0$. The benefit of delaying the forecast is the option value of observing the other analyst's forecast. If $t(T) \geq t(T')$, then an analyst can only observe the forecast of those who received information before them. Thus, an informed analyst at time 0 can only learn information if the other analyst was also informed at time 0. However, the option value in this case is zero since the event has zero probability. Alternatively, if $t(0) > t(T')$, where $T' > 0$, this would contradict the single-crossing condition necessary for a symmetric pure strategy equilibrium. This implies that $t(T)$ is increasing, at least initially, and is never strictly decreasing for all T.

According to the strategy $t(T)$, an analyst who receives information at time T can only observe the forecast of an analyst who received information between time $T - t(u)$ and time T. An analyst who received information prior to time $T - t(u)$ would have already forecasted, and an analyst who receives information after time T will delay their forecast until some time after $T + t(T)$. Therefore, the relevant interval for an informed analyst i at time T is the probability of information arrival for analyst j in the interval $[u + t(u), T]$. Let $k(u) = u + t(u) \equiv T$, and so $k^{-1}(T) = u$. For notational ease, let $\xi(T) \equiv k^{-1}(T)$. The informed analyst's payoff can be expressed as:

$$V_i = \max_h \int_{\xi(T)}^T s \cdot e^{-r(u+t(u)-T)} \cdot \lambda e^{-\lambda(u-\xi(T))} du + (1 - \Omega(T - \xi(T))) e^{-r(t(T))} \left(s - \sigma_j^2\right)$$

$$(6.8)$$

The first term in equation (6.8) is analyst i's expected payoff from waiting an additional time $t(T)$ before forecasting and observing the other analyst's forecast. The second term is analyst i's expected payoff from waiting until time $T + t(T)$ and not observing analyst j's forecast.

The solution to the problem involves solving for the functional form of the delay period, $t(T)$. However, this is unfortunately analytically intractable without imposing further conditions. Suppose we can set a blocked interval, where analysts may learn information but are prohibited from issuing a forecast. This can be thought of as the early stages of an IPO process, where analysts affiliated with the underwriting banks learn material information but disclosure is prohibited. Let B denote the end point of the blocked interval, $[0, B)$. We can thus choose B which induces the strategy $t(T)$ to be constant for all arrival times T:

Proposition 6 *Suppose there is a time interval where analysts are not able to forecast, then the equilibrium forecasting strategy follows a constant delay time given that $B = t^*$, where*

$$t^* = \frac{\ln\left(\sigma^2(\lambda + r) - rs\right)}{\lambda + r}.$$

The above Proposition states that the delay time is constant and equal to B when analysts may learn information but cannot forecast in the interval $[0, B)$. This is a special case since analysts who learn information within the blocked interval have the same

incentives as those who learn information after time B.

Chapter 7

Conclusion

This study offers a unified theory for analyst forecast timing, information acquisition, and choice of employment. I characterize the unique equilibrium of the forecasting game to show that the forecast order is determined by the precision of the analysts' signals, the likelihood the analysts' privately receive information, the level of the analysts' biases, the public information in the analysts' segments, and the analysts' discount rates. The results capture several significant and interesting properties of the forecast order and allow for endogenous information acquisition. I show that there exists a unique symmetric equilibrium for which analysts choose to be imperfectly informed. This occurs due to the analysts' incentive to gain a strategic advantage in the forecasting subgame. Similarly, when the choice of employment is endogenized, one of the analysts chooses to be affiliated in equilibrium.

The key novelty of the model is that I allow analysts to be probabilistically informed and analysts choose this probability prior to their forecasting decision. The uncertain nature of information acquisition is of practical relevance and departs from the extant analyst herding literature. Allowing for uncertainty over the information endowment leads to number of interesting results that shed light on empirical regularities which have heretofore not been captured in the theoretical literature. Moreover, the results provide several avenues for future research through empirical predictions. The main result captures the link between an analyst's primary responsibility of information acquisition/production and the strategic nature of the forecast timing. The results also provide a theoretical justification

for an analyst's choice of affiliated employment.

Suggestions for future theoretical work that build on the model here include endogenizing the market's perception of the analyst's type in the forecasting game as part of her compensation. The model assumes that an analyst's compensation only depends on her forecast error and timeliness, however, it may be the case that an analyst's perceived ability (informed or uninformed) may also have some impact on her compensation. Another possible extension includes an exogenous shock to the public information that occurs either at some known or unknown time. It seems that this would induce further delay by analysts, or push to immediate forecasting if an analyst's compensation depends on their perceived ability. This work provides useful insights on the economic forces which drive analyst behavior and offers several avenues for future theoretical and empirical research.

Bibliography

[1] Abreu, D., and F. Gul. 2000. "Bargaining and reputation." *Econometrica* 60 (1): 85-117.

[2] Admati, A.R., and P. Pfleiderer. 1986. "A monopolistic market for information." *Journal of Economic Theory* 39 (2): 400-438.

[3] Aghamolla, C., and T. Hashimoto. 2015. "Endogenous timing, clustering, and delay in analyst herding." Working paper.

[4] Altschuler, D., G. Chen, and J. Zhou. 2015. "Anticipation of management forecasts and analysts' private information search." *Review of Accounting Studies* 20 (2): 803-838.

[5] Bernhardt, D., M. Campello, and E. Kutsoati. 2006. "Who herds?" *Journal of Financial Economics* 80 (3): 657-675.

[6] Beyer, A., and I. Guttman. 2011. "The effect of trading volume on analysts' forecast bias." *The Accounting Review* 86 (2): 451-481.

[7] Brown, L.D., A.C. Call, M.B. Clement, and N.Y. Sharp. 2015. "Inside the 'black box' of sell-side financial analysts." *Journal of Accounting Research* 53 (1): 1-47.

[8] Chamley, C., and D. Gale. 1994. "Information revelation and strategic delay in a model of investment." *Econometrica* 62 (5): 1065-1085.

[9] Chen, Q., and W. Jiang. 2005. "Analysts' weighting of private and public information." *Review of Financial Studies* 19 (1): 319–355.

[10] Chen, X., Q. Cheng, and K. Lo. 2010. "On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation." *Journal of Accounting and Economics* 49 (3): 361-381.

[11] Clement, M.B., 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 (3): 279-299.

[12] Clement, M.B. and S.Y. Tse. 2005. "Financial analyst characteristics and herding behavior in forecasting." *Journal of Finance* 60 (1): 307-341.

[13] Cooper, R.A., T.E. Day, and C.M. Lewis. 2001. "Following the leader: A study of individual analysts' earnings forecasts." *Journal of Financial Economics* 61 (3): 383-416.

[14] Das, S., C.B. Levine, and K. Sivaramakrishnan. 1998. "Earnings predictability and bias in analysts' earnings forecasts." *The Accounting Review* 73 (2): 277-294.

[15] De Bondt, W.F.M., and W.P. Forbes. 1999. "Herding in analyst earnings forecasts: Evidence from the United Kingdom." *European Financial Management* 5 (2): 143-163.

[16] De Bondt, W., and R. Thaler. 1990. "Do security analysts overreact?" *American Economic Review* 80 (2): 52-57.

[17] Dugar, A., and S. Nathan. 1995. "The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations." *Contemporary Accounting Research* 12 (1): 131-160.

[18] Easterwood, J., and S. Nutt. 1999. "Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism?" *Journal of Finance* 54 (5): 1777-1797.

[19] Ertimur, Y., J. Sunder, and S.V. Sunder. 2007. "Measure for measure: The relation between forecast accuracy and recommendation profitability of analysts." *Journal of Accounting Research* 45 (3): 567-606.

[20] Evans, G.C. 1910. "Volterra's integral equation of the second kind, with discontinuous kernel." *Transactions of the American Mathematical Society* 11 (4): 393-413.

[21] Fischer, P.E., and P.C. Stocken. 2010. "Analyst information acquisition and communication." *The Accounting Review* 85 (6): 1985-2009.

[22] Francis, J., and D. Philbrick. 1993. "Analysts' decisions as products of a multi-task environment." *Journal of Accounting Research* 31 (2): 216-230.

[23] Fudenberg, D., and J. Tirole. 1986. "A theory of exit in duopoly." *Econometrica* 54 (4): 943-960.

[24] Gleason, C.A., and C.M.C. Lee. 2003. "Analyst forecast revisions and market price discovery." *The Accounting Review* 78 (1): 193-225.

[25] Groysberg, B., P. Healy, and D. Maber. 2011. "What drives sell-side analyst compensation at high-status investment banks?" *Journal of Accounting Research* 49 (4): 969-1000.

[26] Gu, Z. and J. Xue. 2008. "The superiority and disciplining role of independent analysts." *Journal of Accounting and Economics* 45: 289-316.

[27] Gul, F., and R. Lundholm. 1995. "Endogenous timing and the clustering of agents' decisions." *Journal of Political Economy* 103 (5): 1039-1066.

[28] Guttman, I. 2010. "The timing of analysts' earnings forecasts." *The Accounting Review* 85 (2): 513-545.

[29] Hayes, R. 1998. "The impact of trading commission incentives on analyst's stock coverage decision and earnings forecasts." *Journal of Accounting Research* 36 (2): 299-320.

[30] Hong, H., and J. Kubik. 2003. "Analyzing the analysts: Career concerns and biased earnings forecasts." *Journal of Finance* 58 (1): 313-351.

[31] Hong, H., J. Kubik, and A. Solomon. 2000. "Security analysts' career concerns and herding of earnings forecasts." *RAND Journal of Economics* 31: 121-144.

[32] Ivkovic, Z., and N. Jegadeesh. 2004. "The timing and value of forecast and recommendation revisions." *Journal of Financial Economics* 73: 433-463.

[33] Irvine, P., M. Lipson, and A. Puckett. 2007. "Tipping." *Review of Financial Studies* 20 (3): 741-768.

[34] Jackson, A. 2005. "Trade generation, reputation, and sell-side analysts." *Journal of Finance* 60 (2): 673-717.

[35] Jacob, J., T.Z. Lys, and M.A. Neale. 1999. "Expertise in forecasting performance of security analysts." *Journal of Accounting and Economics* 28: 51-82.

[36] Jacob, J., S. Rock, and D. Weber. 2008. "Do non-investment bank analysts make better earnings forecasts?" *Journal of Accounting, Auditing and Finance* 23: 23-60.

[37] Jegadeesh, N., and W. Kim. 2010. "Do analysts herd? An analysis or recommendations and market reactions." *Review of Financial Studies* 23 (2): 901-937.

[38] Keskek, S., S.Y. Tse, and J.W. Tucker. 2014. "Analyst information production and the timing of annual earnings forecasts." *Review of Accounting Studies* 19: 1504-1531.

[39] Kim, K., and F.Z.X. Lee. 2014. "Information acquisition in a war of attrition." *American Economic Journal: Microeconomics* 6 (2): 37-78.

[40] Kim, O., and R.E. Verrecchia. 1994. "Market liquidity and volume around earnings announcements." *Journal of Accounting and Economics* 17 (1-2): 41-67.

[41] Kim, O., and R.E. Verrecchia. 1997. "Pre-announcement and event-period private information." *Journal of Accounting and Economics* 24 (3): 395-419.

[42] Kothari, S. 2001. "Capitals research in accounting." *Journal of Accounting and Economics* 31 (1-3): 105-231.

[43] Kreps, D., and R. Wilson. 1982. "Reputation and imperfect information." *Journal of Economic Theory* 27 (2): 253-279.

[44] Lim, T. 2001. "Rationality and analysts forecast bias." *The Journal of Finance* 56 (1): 369-385.

[45] Lin, H., and M. McNichols. 1998. "Underwriting relationships, analysts' earnings forecasts and investment recommendations." *Journal of Accounting and Economics* 25 (1): 101-127.

[46] Livnat, J., and Y. Zhang. 2012. "Information interpretation or information discovery: Which role of analysts do investors value more?" *Review of Accounting Studies* 17: 612-641.

[47] Loh, R.K., and G.M. Mian. 2006. "Do accurate earnings forecasts facilitate superior investment recommendations?" *Journal of Financial Economics* 80 (2): 455-483.

[48] Loh, R.K. and R.M. Stulz. 2011. "When are analyst recommendation changes influential?" *Review of Financial Studies* 24 (2): 593–627.

[49] McNichols, M., and P. O'Brien. 1997. "Self-selection and analyst coverage." *Journal of Accounting Research* 35 (Supplement): 167-199.

[50] Meng, X. 2015. "Analyst reputation, communication, and information acquisition." *Journal of Accounting Research* 53 (1): 119–173.

[51] Michaely, R., and K. Womack. 1999. "Conflict of interest and credibility of underwriter analyst recommendations." *Review of Financial Studies* 12 (4): 653–686.

[52] Mikhail, M., B. Walther, and R. Willis. 1999. "Does forecast accuracy matter to security analysts?" *The Accounting Review* 74 (2): 185-200.

[53] Mola, S., and M. Guidolin. 2009. "Affiliated mutual funds and analyst optimism." *Journal of Financial Economics* 93 (1): 108-137.

[54] Morgan, J., and P. Stocken. 2003. "An analysis of stock recommendations." *RAND Journal of Economics* 34 (1): 183–203.

[55] Ottaviani, M., and P.N. Sørenson. 2006. "Reputational Cheap Talk." *RAND Journal of Economics* 37 (1): 155–175.

[56] Ponsati, C., and J. Sakovics. 1995. "The war of attrition with incomplete information." *Mathematical Social Sciences* 29 (3): 239-254.

[57] Raedy, J., P. Shane, and S. Yang. 2006. "Horizon-dependent underreaction in financial analysts' earnings forecasts." *Contemporary Accounting Research* 23 (1): 291-322.

[58] Scharfstein, D. S., and J.C. Stein. 1990. "Herd behavior and investment." *American Economic Review* 80 (3): 465-479.

[59] Schipper, K. 1991. "Analysts' forecasts." *Accounting Horizons* 5 (4): 105-121.

[60] Shroff, P., R. Venkataraman, and B. Xin. 2014. "Timeliness of analysts' forecasts: The information content of delayed forecasts." *Contemporary Accounting Research* 31 (1): 202-229.

[61] Soltes, E. 2014. "Private interaction between firm management and sell-side analysts." *Journal of Accounting Research* 52 (1): 245-272.

[62] Stickel, S. 1992. "Reputation and performance among security analysts." *The Journal of Finance* 47 (5): 1811–1836.

[63] Trueman, B. 1994. "Analyst forecasts and herding behavior." *Review of Financial Studies* 7 (1): 97-214.

[64] von Wolfersdorf, L. 2007. "On the theory of convolution equations of the third kind." *Journal of Mathematical Analysis and Applications* 331 (2): 1314-1336.

[65] Welch, I. 2000. "Herding among security analysts." *Journal of Financial Economics* 57: 369-396.

[66] Wu, J., and A. Zang. 2009. "What determines financial analysts' career outcomes during mergers?" *Journal of Accounting and Economics* 47: 59-86.

[67] Xue, H. 2015. "Independent and affiliated analysts: Disciplining and herding." Working paper.

[68] Zhang, J. 1997. "Strategic delay and the onset of investment cascades." *RAND Journal of Economics* 28 (1): 188-205.

[69] Zitzewitz. E. 2001. "Measuring herding and exaggeration by equity analysts and other opinion sellers." Working paper.

Appendix A

Appendix of Proofs

Proof of Lemma 1. By Claims 1-5, $F_j(t)$ must be such that an informed analyst i's utility is constant for all t. This is given by:

$$u_i = \int_0^t \left(s - \omega_i^2 - \omega_j^2\right) e^{-r_i a} f_j(a)\, da + \left(1 - F_j(t)\right) \left(s - \omega_i^2 - \sigma_j^2\right) e^{-r_i t}$$

Taking the FOC with respect to t, we have

$$\left(s - \omega_i^2 - \omega_j^2\right) e^{-r_i t} f_j(t) - f_j(t) \left(s - \omega_i^2 - \sigma_j^2\right) e^{-r_i t} - \left(1 - F_j(t)\right) \left(s - \omega_i^2 - \sigma_j^2\right) e^{-r_i t} r_i = 0$$

$$f_j(t) \left(\sigma_j^2 - \omega_j^2\right) - \left(1 - F_j(t)\right) \left(s - \omega_i^2 - \sigma_j^2\right) r_i = 0$$

$$f_j(t) + \frac{F_j(t) \left(s - \omega_i^2 - \sigma_j^2\right) r_i}{\sigma_j^2 - \omega_j^2} = \frac{\left(s - \omega_i^2 - \sigma_j^2\right) r_i}{\sigma_j^2 - \omega_j^2}$$

Let $c_j = \frac{\left(s - \omega_i^2 - \sigma_j^2\right) r_i}{\sigma_j^2 - \omega_j^2}$ and multiplying the above equation by $e^{c_j t}$, we have

$$f_j(t) e^{c_j t} + F_j(t) e^{c_j t} \cdot c_j = c_j e^{c_j t}$$

which is equivalent to

$$\frac{d\left(F_j\left(t\right)e^{c_jt}\right)}{dt} = c_je^{c_jt}$$

$$F_j\left(t\right)e^{c_jt} = e^{c_jt} - \alpha$$

$$F_j\left(t\right) = 1 - \alpha e^{-c_jt}$$

We see that $F_j\left(t\right)$ follows an exponential path, which implies that the constant hazard rate is equal to c_j. Thus,

$$\frac{d}{dt}\ln\left(1 - F_j\left(t\right)\right) = -c_j$$

$$\ln\left(1 - F_j\left(t\right)\right) - \ln\left(1 - F_j\left(0\right)\right) = -c_jt$$

$$F_j\left(t\right) = 1 - \left(1 - F_j\left(0\right)\right)e^{-c_jt}$$

Thus, $\alpha = 1 - F_j\left(0\right)$, and is analyst j's potential atom at $t = 0$. By similar argument for analyst i, we have

$$F_i\left(t\right) = 1 - \left(1 - F_i\left(0\right)\right)e^{-c_it}$$

where

$$c_i = \frac{s - \omega_j^2 - \sigma_i^2}{\sigma_i^2 - \omega_i^2}r_j$$

The endogenous terminal time to the game is given by $T = \min\left\{T_1, T_2\right\}$. To determine T, we have that

$$1 - e^{-c_iT_i} = 1 - x_i$$

$$-c_iT_i = \ln\left(x_i\right)$$

$$T_i = -\frac{1}{c_i}\ln\left(x_i\right)$$

Hence, $T = \min\left\{-\frac{1}{c_i}\ln(x_i), -\frac{1}{c_j}\ln(x_j)\right\}$. Suppose that $T_i > T_j$. Then $F_j(0) = 0$ and $F_i(0)$ is determined by

$$1 - (1 - F_i(0)) e^{-c_i T_j} = 1 - x_i$$

$$(1 - F_i(0)) e^{-c_i T_j} = x_i$$

$$F_i(0) = 1 - x_i \exp\left(c_i\left(-\frac{1}{c_j}\ln(x_j)\right)\right)$$

$$F_i(0) = 1 - x_j^{-\frac{c_i}{c_j}} x_i$$

And hence $F_i(t) = 1 - \left(x_j^{-\frac{c_i}{c_j}} x_i\right) e^{-c_i t}$. ∎

Proof of Corollary 1. Recall that $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$, which is

$$T = \min\left\{-\frac{\sigma_j^2 - \omega_j^2}{\left(s - \omega_i^2 - \sigma_j^2\right)r_i}\ln(x_j), -\frac{\sigma_i^2 - \omega_i^2}{\left(s - \omega_j^2 - \sigma_i^2\right)r_j}\ln(x_i)\right\}$$

which we see is decreasing in $\frac{1}{\omega_i^2}$ and $\frac{1}{\omega_j^2}$, which implies the first part of the claim. Assuming $x_i = x_j$, $r_i = r_j$, and $\sigma_i^2 = \sigma_j^2$, then if $T_i > T_j$,

$$-\frac{1}{c_i}\ln(x_i) > -\frac{1}{c_j}\ln(x_j)$$

$$\frac{\sigma_i^2 - \omega_i^2}{s - \omega_j^2 - \sigma_i^2}\ln(x_i) < \frac{\sigma_j^2 - \omega_j^2}{s - \omega_i^2 - \sigma_j^2}\ln(x_j)$$

$$\frac{\sigma^2 - \omega_i^2}{s - \omega_j^2 - \sigma^2} > \frac{\sigma^2 - \omega_j^2}{s - \omega_i^2 - \sigma^2}$$

Then,

$$\left(\sigma^2 - \omega_i^2\right)\left(s - \omega_i^2 - \sigma^2\right) > \left(\sigma^2 - \omega_j^2\right)\left(s - \omega_j^2 - \sigma^2\right)$$

$$\sigma^2 s - \omega_i^2 s - \omega_i^2 \sigma^2 + \omega_i^4 - \sigma^4 + \sigma^2\omega_i^2 > \sigma^2 s - \omega_j^2 s - \omega_j^2\sigma^2 + \omega_j^4 - \sigma^4 + \sigma^2\omega_j^2$$

$$-\omega_i^2 s + \omega_i^4 > -\omega_j^2 s + \omega_j^4$$

$$s\left(\omega_j^2 - \omega_i^2\right) > \left(\omega_j^2 - \omega_i^2\right)\left(\omega_j^2 + \omega_i^2\right)$$

If $\omega_j^2 - \omega_i^2 > 0$, then

$$s > \omega_j^2 + \omega_i^2$$

which holds since $s > \omega_j^2 + \sigma_i^2 > \omega_j^2 + \omega_i^2$. If $\omega_j^2 - \omega_i^2 < 0$, then

$$s < \omega_j^2 + \omega_i^2$$

However, this can never be the case since $s > \omega_j^2 + \sigma_i^2$. Thus, if $T_i > T_j$, then it must be that $\omega_j^2 > \omega_i^2$, and so $\frac{1}{\omega_i^2} > \frac{1}{\omega_j^2}$, and analyst i has the atom at time 0. Hence, the analyst with more precise information forecasts earlier. ∎

Proof of Corollary 2. Recall that $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$, which is

$$T = \min\left\{-\frac{\sigma_j^2 - \omega_j^2}{\left(s - \omega_i^2 - \sigma_j^2\right)r_i}\ln(x_j), -\frac{\sigma_i^2 - \omega_i^2}{\left(s - \omega_j^2 - \sigma_i^2\right)r_j}\ln(x_i)\right\}.$$

Assuming $\omega_i^2 = \omega_j^2$, $r_i = r_j$, and $\sigma_i^2 = \sigma_j^2$, then if $T_i > T_j$,

$$-\frac{1}{c_i}\ln(x_i) > -\frac{1}{c_j}\ln(x_j)$$

$$\frac{\sigma_i^2 - \omega_i^2}{s - \omega_j^2 - \sigma_i^2}\ln(x_i) < \frac{\sigma_j^2 - \omega_j^2}{s - \omega_i^2 - \sigma_j^2}\ln(x_j)$$

$$\frac{\sigma^2 - \omega^2}{s - \omega^2 - \sigma^2}\ln(x_i) < \frac{\sigma^2 - \omega^2}{s - \omega^2 - \sigma^2}\ln(x_j)$$

$$\ln(x_i) < \ln(x_j)$$

$$x_i < x_j$$

Which is equivalent to $1 - x_i > 1 - x_j$. ∎

Proof of Corollary 3. Recall that $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$, which is

$$T = \min\left\{-\frac{\sigma_j^2 - \omega_j^2}{\left(s - \omega_i^2 - \sigma_j^2\right)r_i}\ln(x_j), -\frac{\sigma_i^2 - \omega_i^2}{\left(s - \omega_j^2 - \sigma_i^2\right)r_j}\ln(x_i)\right\}$$

which we see is decreasing in r_i and r_j. Differentiating T_i with respect to σ_j^2, we have that

$$\frac{-\ln(x_j)\left(s - \omega_i^2 - \sigma_j^2\right)r_i + r_i\left(\omega_j^2 - \sigma_j^2\right)\ln(x_j)}{\left(\left(s - \omega_i^2 - \sigma_j^2\right)r_i\right)^2}$$

$-\ln(x_j)\left(s - \omega_i^2 - \sigma_j^2\right)r > 0$, $\left(\left(s - \omega_i^2 - \sigma_j^2\right)r_i\right)^2 > 0$, and $r_i\left(\omega_j^2 - \sigma_j^2\right)\ln(x_j) > 0$ since $\omega_j^2 - \sigma_j^2 < 0$ and $\ln(x_j) < 0$. A similar calculation can be done for σ_i^2. Hence T is increasing in σ_i^2 and thus decreasing in $\frac{1}{\sigma_i^2}$. Assuming $\omega_i^2 = \omega_j^2$, $x_i = x_j$, and $\sigma_i^2 = \sigma_j^2$, then if $T_i > T_j$,

$$-\frac{1}{c_i}\ln(x_i) > -\frac{1}{c_j}\ln(x_j)$$

$$\frac{\sigma_i^2 - \omega_i^2}{\left(s - \omega_j^2 - \sigma_i^2\right)r_j}\ln(x_i) < \frac{\sigma_j^2 - \omega_j^2}{\left(s - \omega_i^2 - \sigma_j^2\right)r_i}\ln(x_j)$$

$$\frac{1}{r_j} > \frac{1}{r_i}$$

$$r_j < r_i$$

Since analyst i has the atom at $t = 0$ and $r_i > r_j$, this implies that the less patient analyst forecasts earlier. Similarly, if $\omega_i^2 = \omega_j^2$, $x_i = x_j$, and $r_i = r_j$, then if $T_i > T_j$,

$$-\frac{1}{c_i}\ln(x_i) > -\frac{1}{c_j}\ln(x_j)$$

$$\frac{\sigma_i^2 - \omega_i^2}{\left(s - \omega_j^2 - \sigma_i^2\right)r_j}\ln(x_i) < \frac{\sigma_j^2 - \omega_j^2}{\left(s - \omega_i^2 - \sigma_j^2\right)r_i}\ln(x_j)$$

$$\left(\sigma_i^2 - \omega^2\right)\left(s - \omega^2 - \sigma_j^2\right) > \left(\sigma_j^2 - \omega^2\right)\left(s - \omega^2 - \sigma_i^2\right)$$

$$\sigma_i^2 s - \omega^2 \sigma_i^2 - \sigma_i^2 \sigma_j^2 - \omega^2 s + \omega^4 + \omega^2 \sigma_j^2 > \sigma_j^2 s - \omega^2 \sigma_j^2 - \sigma_i^2 \sigma_j^2 - \omega^2 s + \omega^4 + \omega^2 \sigma_i^2$$

$$\sigma_i^2\left(s + 2\omega^2\right) > \sigma_j^2\left(s + 2\omega^2\right)$$

$$\sigma_i^2 > \sigma_j^2$$

∎

Proof of Lemma 4. Assuming x_j is fixed, analyst i faces the following maximization

problem:

$$\max_{x_i} \left\{ (1 - x_i) \left[\left(x_i^{-1} x_j \right) \left(s - \sigma^2 \right) + \left(1 - x_i^{-1} x_j \right) (s) \right] + x_i \left(1 - x_j \right) \left(s - \sigma^2 \right) \right\}$$

$$= \max_{x_i} \left\{ (1 - x_i) s - (1 - x_i) \left(\frac{x_j}{x_i} \right) \sigma^2 + x_i \left(1 - x_j \right) \left(s - \sigma^2 \right) \right\}$$

$$= \max_{x_i} \left\{ (1 - x_i) s - \left(\frac{x_j}{x_i} \right) \sigma^2 + x_j \sigma^2 + x_i \left(1 - x_j \right) \left(s - \sigma^2 \right) \right\}$$

Taking the FOC, this becomes

$$-s + \left(\frac{x_j}{x_i^2} \right) \sigma^2 + (1 - x_j) \left(s - \sigma^2 \right) = 0$$

$$\frac{x_j}{x_i^2} = \frac{x_j s + \sigma^2 - x_j \sigma^2}{\sigma^2}$$

$$x_i^2 = x_j \frac{\sigma^2}{x_j \left(s - \sigma^2 \right) + \sigma^2}$$

$$x_i = (x_j)^{\frac{1}{2}} \left(\frac{\sigma^2}{x_j \left(s - \sigma^2 \right) + \sigma^2} \right)^{\frac{1}{2}}$$

We see that

$$(x_j)^{\frac{1}{2}} \left(\frac{\sigma^2}{x_j \left(s - \sigma^2 \right) + \sigma^2} \right)^{\frac{1}{2}} > x_j$$

$$\frac{\sigma^2}{x_j \left(s - \sigma^2 \right) + \sigma^2} > x_j$$

$$\sigma^2 - \sigma^2 x_j > x_j^2 \left(s - \sigma^2 \right)$$

$$s < \frac{\sigma^2 \left(1 - x_j \right)}{x_j^2} + \sigma^2$$

which requires that

$$\frac{1 - x_j}{x_j^2} \geq 1.$$

To see that $x_i = 0$ when $x_j > \frac{1}{2}$, limit the choice of x_i to be either $x_i = x_j$ or $x_i > x_j$.

Then analyst i will choose $x_i > x_j$ if and only if:

$$\left\{ \begin{array}{c} (x_i)(x_j)\,0 + x_i(1-x_j)(s-\sigma^2) \\ + (1-x_i)\left[(1-F_j(0))(s-\sigma^2)+F_j(0)(s)\right] \end{array} \right\} > \left\{ \begin{array}{c} (s-\sigma^2)(1-\hat{x}_i) \\ +x_i\left(s-\sigma_j^2\right)(1-x_j)+(\hat{x}_i)(x_j)\,0 \end{array} \right\}$$

$$(1-x_i)\left[(1-F_j(0))(s-\sigma^2)+F_j(0)(s)\right] > (s-\sigma_j^2)(1-\hat{x}_i)$$

where $x_i > \hat{x}_i$ on the RHS and $\hat{x}_i = x_j$ on the LHS. Then,

$$(1-x_i)\left[\left(\frac{x_j}{x_i}\right)(s-\sigma^2)+\left(1-\frac{x_j}{x_i}\right)(s)\right] > (s-\sigma^2)(1-x_j)$$

$$(1-x_i)\left[s-\left(\frac{x_j}{x_i}\right)\sigma^2\right] > (s-\sigma^2)(1-x_j)$$

$$x_i s + \left(\frac{x_j}{x_i}\right)\sigma^2 < \sigma^2 + x_j s$$

$$s(x_i - x_j) + \sigma^2\left(\frac{x_j - x_i}{x_i}\right) < 0$$

Let $x_i = x_j + \kappa$, then

$$s(\kappa) - \sigma^2\left(\frac{\kappa}{x_i}\right) < 0$$

$$s < \frac{\sigma^2}{x_i}$$

$$s < \frac{\sigma^2}{x_j + \kappa}$$

$$s(x_j + \kappa) < \sigma^2$$

Since $\frac{s}{2} < \sigma^2$, then this holds as long as $x_j \le \frac{1}{2}$. This implies that analyst i's payoff is greater by setting $x_i = x_j$ when $x_j > \frac{1}{2}$ then by setting $x_i > x_j$, given that x_i is restricted to the interval $[x_j, 1]$. However, we know that analyst i can do strictly better by setting $x_i = 0$ when $x_i = x_j$ which implies that $x_i = 0$ when $x_j > \frac{1}{2}$. ∎

Proof of Lemma 5. Analyst i's expected utility is given as

$$u_i = \int_0^x \left[x(1-y)(s-\sigma^2) + (1-x)(s-\sigma^2) \right] d\Psi(y) \tag{A.1}$$

$$+ \int_x^q \left[x(1-y)(s-\sigma^2) + s - xs - \left(\frac{y}{x}\right)\sigma^2 + y\sigma^2 \right] d\Psi(y).$$

Where $x_i = x$ and $x_j = y$. It follows from equation $(A.1)$ that u_i is continuous on $[0,q]$ since $\Psi(x)$ is continuous by definition. Define $\mathcal{F}_i := \{x_i | u_i(x_i) = \max u_i\}$. From Lemma 4 and since $\Psi(x)$ is nondecreasing, $\mathcal{F}_i$ is dense in $[0,q]$. Hence, u_i is constant for $x \in [0,q]$ and so $\mathcal{F}_i = [0,q]$. This implies that u_i is differentiable. The fundamental theorem of calculus and differentiability of u_i imply that $\Psi(x)$ is differentiable on $[0,q]$. ∎

Proof of Theorem 1. Analyst i's maximization problem is:

$$\max_{x_i} \left\{ \begin{array}{c} E\left[\mathbb{I}_{\{x_i \leq x_j\}} \left[(x_i)(x_j)0 + x_i(1-x_j)(s-\sigma^2) + (1-x_i)(s-\sigma^2) \right] \right] \\[2mm] + E\left[(1-\mathbb{I}_{\{x_i \leq x_j\}}) \left[\begin{array}{c} (x_i)(x_j)0 + x_i(1-x_j)(s-\sigma^2) \\ + (1-x_i)\left[(s-\sigma^2)\left(1-\left(1-\frac{x_j}{x_i}\right)\right) + s\left(1-\frac{x_j}{x_i}\right) \right] \end{array} \right] \right] \end{array} \right\}$$

$$= \max_{x_i} \left\{ \begin{array}{c} E\left[\mathbb{I}_{\{x_i \leq x_j\}} + x_i(1-x_j)(s-\sigma^2) + (1-x_i)(s-\sigma^2) \right] \\[2mm] + E\left[(1-\mathbb{I}_{\{x_i \leq x_j\}}) \left[x_i(1-x_j)(s-\sigma^2) + (1-x_i)\left[s - \left(\frac{x_j}{x_i}\right)\sigma^2 \right] \right] \right] \end{array} \right\}$$

$$= \max_{x_i} \left\{ \begin{array}{c} E\left(\mathbb{I}_{\{x_i \leq x_j\}} \left[x_i(1-x_j)(s-\sigma^2) + s - x_i s - \sigma^2 + x_i \sigma^2 \right] \right) \\[2mm] + E\left((1-\mathbb{I}_{\{x_i \leq x_j\}}) \left[x_i(1-x_j)(s-\sigma^2) + s - x_i s - \left(\frac{x_j}{x_i}\right)\sigma^2 + x_j \sigma^2 \right] \right) \end{array} \right\}$$

Hence, setting the equation above equal to u_i and taking expectation, we have:

$$u_i = \int_0^x \left[x(1-y)(s-\sigma^2) + (1-x)(s-\sigma^2) \right] \psi(y) \, dy \tag{A.2}$$

$$+ \int_x^q \left[x(1-y)(s-\sigma^2) + s - xs - \left(\frac{y}{x}\right)\sigma^2 + y\sigma^2 \right] \psi(y) \, dy$$

Where $x_i = x$ and $x_j = y$ for ease of exposition. By Lemma 5, $\Psi(y)$ is differentiable and analyst i's utility must be constant over all choices of x_i. Taking the FOC of equation

$(A.2)$ with respect to x, we have:

$$\left[\begin{array}{c} \left[x\left(1-x\right)\left(s-\sigma^2\right)+\left(1-x\right)\left(s-\sigma^2\right)\right]\psi\left(x\right)+\left(\sigma^2-s\right)\int_0^x y\psi\left(y\right)dy \\ -\left[x\left(1-x\right)\left(s-\sigma^2\right)+s-xs-\left(\frac{x}{x}\right)\sigma^2+x\sigma^2\right]\psi\left(x\right) \\ +\int_x^q\left[-ys-\sigma^2+y\sigma^2+\left(\frac{y}{x^2}\right)\sigma^2\right]\psi\left(y\right)dy \end{array}\right]=0$$

$$\left[\begin{array}{c} \left[s-xs-\sigma^2+\sigma^2x\right]\psi\left(x\right)+\left(\sigma^2-s\right)\int_0^x y\psi\left(y\right)dy-\left[s-xs-\sigma^2+x\sigma^2\right]\psi\left(x\right) \\ +\int_x^q\left[-ys-\sigma^2+y\sigma^2+\left(\frac{y}{x^2}\right)\sigma^2\right]\psi\left(y\right)dy \end{array}\right]=0$$

$$\left(\sigma^2-s\right)\int_0^x y\psi\left(y\right)dy+\int_x^q\left[-ys-\sigma^2+y\sigma^2+\left(\frac{y}{x^2}\right)\sigma^2\right]\psi\left(y\right)dy=0$$

So we have

$$\left(\sigma^2-s\right)\int_0^x y\psi\left(y\right)dy+\int_x^q\left[y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)-\sigma^2\right]\psi\left(y\right)dy=0$$

Which becomes

$$\left[\begin{array}{c} \left(\sigma^2-s\right)\int_0^x y\psi\left(y\right)dy+\int_0^q\left[y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)-\sigma^2\right]\psi\left(y\right)dy \\ -\int_0^x\left[y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)-\sigma^2\right]\psi\left(y\right)dy \end{array}\right]=0$$

$$\left[\begin{array}{c} \int_0^x\left[\left(\sigma^2-s\right)y-\left(y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)-\sigma^2\right)\right]\psi\left(y\right)dy \\ +\int_0^q y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)\psi\left(y\right)dy-\int_0^q\sigma^2\psi\left(y\right)dy \end{array}\right]=0$$

$$\left[\begin{array}{c} \int_0^x\left[\left(\sigma^2-s\right)y-\left(y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)-\sigma^2\right)\right]\psi\left(y\right)dy \\ +\int_0^q y\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)\psi\left(y\right)dy-\int_0^q\sigma^2\psi\left(y\right)dy \end{array}\right]=0$$

Which is

$$\int_0^x\left[\sigma^2-y\frac{\sigma^2}{x^2}\right]\psi\left(y\right)dy+\left(\sigma^2-s+\frac{\sigma^2}{x^2}\right)\mu-\sigma^2=0 \tag{A.3}$$

Multiplying the entire equation by x^3, we have

$$\int_0^x\sigma^2\left[x^3-xy\right]\psi\left(y\right)dy-\left(x^2\left(s-\sigma^2\right)-\sigma^2\right)\mu x-x^3\sigma^2=0$$

Differentiating the above with respect to x, we have

$$\sigma^2 \left(x^3 - x^2\right) \psi\left(x\right) + \int_0^x \sigma^2 \left[3x^2 - y\right] \psi\left(y\right) dy - 3x^2 \left(s - \sigma^2\right) + \sigma^2 \mu - 3x^2 \sigma^2 = 0$$

which is

$$\psi\left(x\right) - \int_0^x \frac{3x^2 - y}{x^2 - x^3} \psi\left(y\right) dy = \frac{3x^2 s - \sigma^2 \mu}{\sigma^2 \left(x^3 - x^2\right)}$$

Which is a Volterra equation of the second kind. $K\left(x,y\right) = \frac{3x^2-y}{x^2-x^3}$ and $h\left(x\right) = \frac{3x^2 s - \sigma^2 \mu}{\sigma^2 (x^3 - x^2)}$ for $x, y \in [0, q]$ are discontinuous at $x = 0$. We have that $\int_0^x |K\left(x,y\right)| dy$ converges for all x except for $x = 0$ in $[0, q]$ and is finite in any region $[\epsilon, \epsilon + q]$ except for $x = 0$ in $[0, q]$. Note that we can divide $[0, q]$ into k smaller intervals with bounds (a_i, a_{i+1}) such that $\int_{a_i}^x |K\left(x,y\right)| dy \leq 1$ since we can find a_i such that $x^2 - x^3 > \left|3x^2 \left(x - a_i\right) - \frac{1}{2} \left(x^2 - a_i^2\right)\right|$, by setting a_i close to x. Hence, by Evans (1910), there exists a unique solution to $\psi\left(x\right)$.[1]

■

Proof of Proposition 1. By Claims 1-5, $F_j\left(t\right)$ must be such that an informed analyst i's utility is constant for all t. This is given by:

$$u_i = \int_0^t \left(s - p_j^2 b_j^2\right) e^{-r_i a} f_j\left(a\right) da + \left(1 - F_j\left(t\right)\right) \left(s - \sigma_j^2\right) e^{-r_i t}$$

Taking the FOC with respect to t, we have

$$\left(s - p_j^2 b_j^2\right) e^{-r_i t} f_j\left(t\right) - f_j\left(t\right) \left(s - \sigma_j^2\right) e^{-r_i t} - \left(1 - F_j\left(t\right)\right) \left(s - \sigma_j^2\right) e^{-r_i t} r_i = 0$$

Proceeding as in the proof of Lemma 1 yields the result. ■

---

[1] Evans (1910) shows this by dividing $K\left(\cdot\right)$ into two components, one constant and the other a function of the arguments. For the solution, an approximating series can be constructed in which each term is constructed by first solving a singular differential equation and then solving an integral for the inhomogeneity of the subsequent differential equation. Alternatively, equation $(A.3)$ can be expressed as a Volterra equation of the third kind:

$$\psi\left(x\right) x - \int_0^x \frac{2x}{1 - x} \psi\left(y\right) dy = \frac{2xs\mu}{\sigma^2 \left(x - 1\right)}$$

in which we can apply von Wolfersdorf (2007).

Proof of Proposition 2. Recall that $T = \min\left\{-\frac{1}{c_j}\ln(x_j), -\frac{1}{c_i}\ln(x_i)\right\}$, which is

$$T = \min\left\{-\frac{\sigma_j^2 - p_j^2 b_j^2}{\left(s - \sigma_j^2\right)r_i}\ln(x_j), -\frac{\sigma_i^2 - p_i^2 b_i^2}{\left(s - \sigma_i^2\right)r_j}\ln(x_i)\right\}$$

which we see is decreasing in $p_i^2 b_i^2$ and thus $p_i b_i$. Analyst i has an atom at time 0 if $T_i > T_j$, and so

$$-\frac{1}{c_i}\ln(x_i) > -\frac{1}{c_j}\ln(x_j)$$

$$\frac{\sigma_i^2 - p_i^2 b_i^2}{s - \sigma_i^2}\ln(x_i) < \frac{\sigma_j^2 - p_j^2 b_j^2}{s - \sigma_j^2}\ln(x_j)$$

When $x_i = x_j$, then

$$\frac{\sigma_i^2 - p_i^2 b_i^2}{s - \sigma_i^2} > \frac{\sigma_j^2 - p_j^2 b_j^2}{s - \sigma_j^2}$$

We see from the above that the bias $p_i^2 b_i^2$ only lowers T, thus reducing the overall time until the posterior is reached and strengthening the analyst's position. This becomes

$$\left(\sigma_i^2 - p_i^2 b_i^2\right)\left(s - \sigma_j^2\right) > \left(\sigma_j^2 - p_j^2 b_j^2\right)\left(s - \sigma_i^2\right)$$

$$\sigma_i^2 s + p_i^2 b_i^2 \sigma_j^2 - p_i^2 b_i^2 s - \sigma_i^2 \sigma_j^2 > \sigma_j^2 s - p_j^2 b_j^2 s + p_j^2 b_j^2 \sigma_i^2 - \sigma_i^2 \sigma_j^2$$

$$\sigma_i^2 s + p_i^2 b_i^2 \sigma_j^2 - p_i^2 b_i^2 s > \sigma_j^2 s - p_j^2 b_j^2 s + p_j^2 b_j^2 \sigma_i^2$$

When $\sigma_i^2 = \sigma_j^2$, then

$$p_i^2 b_i^2 \sigma_j^2 - p_i^2 b_i^2 s > -p_j^2 b_j^2 s + p_j^2 b_j^2 \sigma_i^2$$

$$-p_i^2 b_i^2 s + p_j^2 b_j^2 s > p_j^2 b_j^2 \sigma_i^2 - p_i^2 b_i^2 \sigma_j^2$$

$$s\left(p_j^2 b_j^2 - p_i^2 b_i^2\right) > \sigma^2\left(p_j^2 b_j^2 - p_i^2 b_i^2\right)$$

If $p_j^2 b_j^2 - p_i^2 b_i^2 > 0$, then $s > \sigma^2$. If $p_j^2 b_j^2 - p_i^2 b_i^2 < 0$, then $s < \sigma^2$, but this cannot happen. Thus we must have that $p_i^2 b_i^2 < p_j^2 b_j^2$ if $T_i > T_j$. Hence, the player who is less biased forecasts earlier. ∎

Proof of Theorem 2. We see from Proposition 1 that the equilibrium payoff for an informed analyst i is given as

$$(1 - F_j(0))\left(s - \sigma_j^2\right) + F_j(0)\left(s - p_j^2 b_j^2\right).$$

From Proposition 2, when $x_i = x_j$, $\sigma_i = \sigma_j$, and $r_i = r_j$, if $p_j^2 b_j^2 < p_i^2 b_i^2$, then $F_j(0) > 0$, and $F_j(0) = 0$ otherwise. Assume that in the case of indifference, the analyst sets p_i to be the lowest value which achieves the same payoff. When $b_i > b_j$, then analyst i can induce an atom by analyst j by setting $p_i = \frac{1}{2}$, since then $p_i^2 b_i^2 > p_j^2 b_j^2$ for any choice of p_j. In this case, $F_j(0) > 0$ and $F_i(0) = 0$, and so an informed analyst j has an equilibrium payoff of $s - \sigma_i^2$ for any choice of p_j, and thus sets $p_j = 0$. $F_j(0)$ is maximized when $p_i = \frac{1}{2}$ and hence is a best response to any choice of p_j. A similar argument holds for $b_i < b_j$ which gives $p_i = 0$, $p_j = \frac{1}{2}$. When $b_i = b_j$, any choice of $p_i < \frac{1}{2}$ is met with a $p_j > p_i$, and vice versa. Hence, at least one of p_j or p_i must be equal to $\frac{1}{2}$. If $p_i = \frac{1}{2}$, then analyst j is indifferent to any choice of $p_j \in \left[0, \frac{1}{2}\right]$, since if $p_j = \frac{1}{2}$, then neither analyst has an atom, and if $p_j < \frac{1}{2}$, then only analyst j has an atom. In either case, analyst i will never have an atom, and analyst j receives the same expected utility, thus setting $p_j = 0$. ■

Proof of Proposition 3. Analyst i's program is given as:

$$\max_{x_i} \left\{ \begin{array}{c} (1 - x_i)\left[\left(x_i^{-\frac{c_j}{c_i}} x_j\right)\left(s - \omega_i^2 - \sigma_j^2\right) + \left(1 - x_i^{-\frac{c_j}{c_i}} x_j\right)\left(s - \omega_i^2 - \omega_j^2\right)\right] \\ + x_i(1 - x_j)\left(s - \omega_j^2 - \sigma_i^2\right) \end{array} \right\}$$

$$= \max_{x_i} \left\{ \begin{array}{c} (1 - x_i)\left[\begin{array}{c} x_i^{-\frac{c_j}{c_i}} x_j s - x_i^{-\frac{c_j}{c_i}} x_j \sigma_j^2 - x_i^{-\frac{c_j}{c_i}} x_j \omega_i^2 \\ + s - x_i^{-\frac{c_j}{c_i}} x_j s - \omega_i^2 + x_i^{-\frac{c_j}{c_i}} x_j \omega_i^2 - \omega_j^2 + x_i^{-\frac{c_j}{c_i}} x_j \omega_j^2 \end{array}\right] \\ + x_i(1 - x_j)\left(s - \omega_j^2 - \sigma_i^2\right) \end{array} \right\}$$

$$= \max_{x_i} \left\{ (1 - x_i)\left[-x_i^{-\frac{c_j}{c_i}} x_j \sigma_j^2 + s - \omega_i^2 - \omega_j^2 + x_i^{-\frac{c_j}{c_i}} x_j \omega_j^2\right] + x_i(1 - x_j)\left(s - \omega_j^2 - \sigma_i^2\right) \right\}$$

$$= \max_{x_i} \left\{ \begin{array}{c} \left[-x_i^{-\frac{c_j}{c_i}} x_j \left(\sigma_j^2 - \omega_j^2\right) + s - \omega_i^2 - \omega_j^2 + x_i^{1-\frac{c_j}{c_i}} x_j \left(\sigma_j^2 - \omega_j^2\right) - x_i s + \omega_i^2 x_i + \omega_j^2 x_i\right] + \\ x_i(1 - x_j)\left(s - \omega_j^2 - \sigma_i^2\right) \end{array} \right\}$$

Taking the FOC with respect to x_i, we have

$$
\begin{bmatrix}
\frac{c_j}{c_i}x_i^{-\frac{c_j}{c_i}-1}x_j\left(\sigma_j^2-\omega_j^2\right)+\left(1-\frac{c_j}{c_i}\right)x_i^{-\frac{c_j}{c_i}}x_j\left(\sigma_j^2-\omega_j^2\right) \\
-s+\omega_i^2+\omega_j^2+(1-x_j)\left(s-\omega_j^2-\sigma_i^2\right)
\end{bmatrix}=0
$$

$$
\begin{bmatrix}
\frac{c_j}{c_i}x_i^{-\frac{c_j}{c_i}-1}x_j\left(\sigma_j^2-\omega_j^2\right)+\left(1-\frac{c_j}{c_i}\right)x_i^{-\frac{c_j}{c_i}}x_j\left(\sigma_j^2-\omega_j^2\right) \\
-s+\omega_i^2+\omega_j^2+s-x_js-\sigma_i^2+\sigma_i^2x_j-\omega_j^2+x_j\omega_j^2
\end{bmatrix}=0
$$

$$
\begin{bmatrix}
\frac{c_j}{c_i}x_i^{-\frac{c_i}{c_j}-1}x_j\left(\sigma_j^2-\omega_j^2\right)+\left(1-\frac{c_j}{c_i}\right)x_i^{-\frac{c_j}{c_i}}x_j\left(\sigma_j^2-\omega_j^2\right) \\
+\omega_i^2-x_js-\sigma_i^2+\sigma_i^2x_j+x_j\omega_j^2
\end{bmatrix}=0
$$

which is

$$
\frac{c_j}{c_i}x_i^{-\frac{c_j}{c_i}-1}+\left(1-\frac{c_j}{c_i}\right)x_i^{-\frac{c_j}{c_i}}=\frac{x_js+\sigma_i^2-\sigma_i^2x_j-x_j\omega_j^2-\omega_i^2}{x_j\left(\sigma_j^2-\omega_j^2\right)}
$$

$$
x_i^{-\frac{c_j}{c_i}}\left(1-\frac{c_j}{c_i}+\frac{c_j}{c_i}x_i^{-1}\right)=\frac{s-\sigma_i^2-\omega_j^2}{\left(\sigma_j^2-\omega_j^2\right)}+\frac{\sigma_i^2-\omega_i^2}{x_j\left(\sigma_j^2-\omega_j^2\right)}
$$

where $c_i=\frac{s-\omega_j^2-\sigma_i^2}{\sigma_i^2-\omega_i^2}r_j$ and $c_j=\frac{s-\omega_i^2-\sigma_j^2}{\sigma_j^2-\omega_j^2}r_i$.

When analyst i has the atom ex ante, then she sets $x_i=x_i^*$, where x_i^* is the solution to equation (6.1) when $u_i^*\geq u_i(x_i=0)=s-\omega_i^2-\sigma_j^2$. Otherwise, she can do strictly better by receiving information with certainty and forecasting immediately. When analyst j has the atom ex ante, then she sets x_i according to equation (6.1) to maximize the size of the mass point for analyst j. $\blacksquare$

Proof of Proposition 4. By Claims 1-5, $F_j(t)$ must be such that an informed analyst i's utility is constant for all t. This is given by:

$$
u_i=\int_0^t(s_F)e^{-r_ia}f_j(a)\,da+(1-F_j(t))\left(s_L-\sigma_j^2\right)e^{-r_it}
$$

Taking the FOC with respect to t, we have

$$
(s_F)e^{-r_it}f_j(t)-f_j(t)\left(s_L-\sigma_j^2\right)e^{-r_it}-(1-F_j(t))\left(s_L-\sigma_j^2\right)e^{-r_it}r_i=0
$$

Proceeding as in the proof of Lemma 1 yields the result. The terminal time is given as

$$T = \min \left\{ -\frac{\sigma_j^2 - (s_L - s_F)}{\left(s_L - \sigma_j^2\right) r_i} \ln(x_j), -\frac{\sigma_i^2 - (s_L - s_F)}{\left(s_L - \sigma_i^2\right) r_j} \ln(x_i) \right\}$$

which is decreasing in $s_L - s_F$. ∎

Proof of Proposition 5. We have that

$$u_i = \int_0^t \left(s - \phi^2\right) e^{-r_i a} f_j(a) \, da + (1 - F_j(t)) \left(s - \omega_i^2\right) e^{-r_i t}$$

Taking the FOC with respect to t, we have

$$\left(s - \phi^2\right) e^{-r_i t} f_j(t) - f_j(t) \left(s - \omega_i^2\right) e^{-r_i t} - (1 - F_j(t)) \left(s - \omega_i^2\right) e^{-r_i t} r_i = 0$$

Proceeding as in the proof of Lemma 1 yields the result. The terminal time is given as

$$T = \min \left\{ -\frac{\omega_i^2 - \phi^2}{\left(s - \omega_i^2\right) r_i} \ln(x_j), -\frac{\omega_j^2 - \phi^2}{\left(s - \omega_j^2\right) r_j} \ln(x_i) \right\}$$

which we see is decreasing in $\frac{1}{\omega_i^2}$ and $\frac{1}{\omega_j^2}$. Assuming $x_i = x_j$ and $r_i = r_j$, then if $T_i > T_j$,

$$-\frac{1}{c_i} \ln(x_i) > -\frac{1}{c_j} \ln(x_j)$$

$$\frac{\omega_j^2 - \phi^2}{s - \omega_j^2} \ln(x) < \frac{\omega_i^2 - \phi^2}{s - \omega_i^2} \ln(x)$$

$$\frac{\omega_j^2 - \phi^2}{s - \omega_j^2} > \frac{\omega_i^2 - \phi^2}{s - \omega_i^2}$$

$$\omega_j^2 s - \phi^2 s - \omega_i^2 \omega_j^2 + \phi^2 \omega_i^2 > \omega_i^2 s - \phi^2 s - \omega_j^2 \omega_i^2 + \omega_j^2 \phi^2$$

$$\omega_j^2 > \omega_i^2$$

Thus, if $T_i > T_j$, then it must be that $\omega_j^2 > \omega_i^2$, and so $\frac{1}{\omega_i^2} > \frac{1}{\omega_j^2}$, and analyst i has the atom at time 0. Hence, the analyst with more precise information forecasts earlier. Results for Corollaries 2 and 3 follow similarly. Analyst i's expected utility from the information

acquisition stage is given as Theorem 1 holds as the only change is

$$u_i = \int_0^x \left[x \left(1 - y\right) \left(s - \omega_j^2\right) + \left(1 - x\right) \left(s - \omega_i^2\right) \right] \psi\left(y\right) dy \tag{A.4}$$

$$+ \int_x^q \left[x \left(1 - y\right) \left(s - \omega_j^2\right) + \left(1 - x\right) \left(s - \phi\right) + \left(\phi - \omega_i\right) \frac{y}{x} - \left(\phi - \omega_i\right) y \right] \psi\left(y\right) dy$$

where $q \in (0, 1)$ is a fixed parameter as a determined by the Normal analog for Lemma 4. Equation $(A.4)$ is qualitatively identical to equation $(A.2)$ and hence the proof of Theorem 1 can be replicated. ∎

Proof of Proposition 6. When $t\left(T\right)$ is constant for all T, $\xi\left(T\right) = T - t$, and so equation (6.8) becomes:

$$V_i = \max_t \int_{T-t}^T s \cdot e^{-r(u+t-T)} \cdot \lambda e^{-\lambda(u-T+t)} du + \left(1 - \Omega\left(t\right)\right) e^{-rt} \left(s - \sigma^2\right).$$

The FOC with respect to t is

$$- \left(r + \lambda\right) \int_{T-t}^T \lambda s e^{-r(u+t-T)-\lambda(u-T+t)} du + \lambda s + e^{-t(\lambda+r)} \left(- \left(r + \lambda\right)\right) \left(s - \sigma^2\right) = 0$$

$$- \left(r + \lambda\right) \lambda s \left[\frac{e^{-r(u+t-T)-\lambda(u-T+t)}}{- \left(r + \lambda\right)} \right]_{T-t}^T + \lambda s + e^{-t(\lambda+r)} \left(- \left(r + \lambda\right)\right) \left(s - \sigma^2\right) = 0$$

$$\lambda s \left[e^{-t(r+\lambda)} - 1 \right] + \lambda s + e^{-t(\lambda+r)} \left(- \left(r + \lambda\right)\right) \left(s - \sigma^2\right) = 0$$

$$e^{-t(\lambda+r)} \left[\lambda s - \left(r + \lambda\right) \left(s - \sigma^2\right) \right] = 0$$

Thus,
$$t^* = \frac{\ln\left(\sigma^2\left(\lambda + r\right) - rs\right)}{\lambda + r}.$$

Which is an equilibrium delay strategy when $B = t^*$. ∎