

Analyst reputation, communication and information acquisition

Xiaojing Meng

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

©2012

Xiaojing Meng

All Rights Reserved

Abstract

Analyst Reputation, Communication and Information Acquisition

Xiaojing Meng

Strategic information transmission models, also called cheap talk models, have become increasingly popular in accounting, as they have successfully brought new insights to various accounting topics. This dissertation consists of two chapters, each analyzes a model of strategic information transmission between an expert and a decision maker.

In the first chapter, I study how reputational concerns affect analysts' incentives to invest in information acquisition, and subsequently, their strategic communication with investors in form of "repeated cheap talk". In a setting where analysts' incentives may be misaligned with the investors in a particular fashion (i.e. biased towards issuing optimistic reports), an equilibrium exists in which only aligned analysts will acquire information. As a result, investors may favorably update their beliefs about the analysts' type (as being aligned) when the report is consistent with the realized state. Hence reputational concerns serve as a disciplining device to curb analysts' opportunistic behavior, consistent with economic intuition. This is in sharp contrast to earlier studies that have treated information as exogenous and identical, in which case reputational concerns may work against informative communication.

The second chapter is based on joint work with Tim Baldenius and Nahum Melumad. In this work, we study the optimal board composition—of monitoring and advisory "types"—

within a framework of strategic communication between the CEO and the board when the CEO is an empire builder. The board of directors performs the dual role of monitoring and advising the firm's management. At times, it makes certain key decisions itself. A major concern regarding the effectiveness of boards is CEO power, in particular as it relates to the board nomination process and CEO entrenchment. Monitoring types on the board aim to uncover information known to the CEO, whereas advisors aim to uncover incrementally decision-relevant information. Successful board monitoring allows for selective intervention even if authority is formally delegated to the CEO. Counter to conventional wisdom, we show that powerful CEOs, who influence the board nomination process, may in fact prefer more monitors on the board than do shareholders. Regulatory interventions (such as the Sarbanes-Oxley Act) that attempt to strengthen the monitoring role of boards may thus be harmful in precisely those cases where agency problems are severe. Lastly, to prevent that CEOs entrench themselves by choosing "complex" projects, shareholders may want to commit to an advisor-heavy board.

Contents

List of Figures	iii
List of Tables	v
Acknowledgements	vi
Dedication	viii
1 Analyst Reputation, Communication and Information Acquisition	1
1.1 Introduction	1
1.2 Model Setup	9
1.3 The Repeated Communication Game — Exogenous and Commonly Known Precision	16
1.3.1 The Second Period Communication Game	16
1.3.2 The First Period Communication Game	20
1.4 Unobservable Choice of Precision	26

1.5	Policy Implications	37
1.6	Empirical Implications	40
1.7	Conclusion	42
2	Board Composition and CEO power	45
2.1	Introduction	45
2.2	Model	52
2.2.1	Payoffs for Given Allocation of Decision Rights	54
2.2.2	Communication Equilibrium	58
2.3	Board Composition and Allocation of Decision Rights	59
2.3.1	Exogenous Board Composition	60
2.3.2	Exogenous Allocation of Decision Rights	62
2.4	Equilibrium Choice of Board Composition and Decision Rights	64
2.5	CEO Power I: Influence Over Board Composition	68
2.6	CEO Power II: Entrenchment	74
2.6.1	Endogenous Project Complexity	76
2.6.2	Preempting CEO Entrenchment	80
2.7	Concluding Remarks	89
A	Proofs for Chapter1	91
A.1	Formal definition of the equilibrium	91
A.2	Properties of informative communication in the first period	95

A.3	Proof of Corollary 1	109
A.4	Proof of Proposition 1	114
A.5	Proof of Proposition 2	116
A.6	Proof of Proposition 3	126
A.7	Proof of Corollary 2	130
B	Proofs for Chapter 2	136
B.1	Preliminaries.	136
B.2	Proof of Lemma 3.	137
B.3	Proof of Lemma 4.	138
B.4	Proof of Lemma 5.	139
B.5	Proof of Proposition 4.	141
B.6	Proof of Proposition 5.	143
B.7	Proof of Lemma 7	146
B.8	Proof of Proposition 6	147

List of Figures

1.1	Timeline	11
1.2	The impact of m_1 on $a_1(m_1)$	33
1.3	The impact of m_2 and λ_2 on $a_2(m_2, \lambda_2)$	33
1.4	Social welfare as a function of c	39
2.1	Timeline I—Shareholders/board choose x and S	65
2.2	Board composition as a function of CEO bias under Timeline I	67
2.3	Timeline II—CEO influences board composition	69
2.4	Board composition as a function of CEO bias under Timeline II	73
2.5	Timeline III—CEO chooses A endogenously	76
2.6	Board composition as a function of CEO bias under Timeline III	79
2.7	Shareholders may preempt CEO entrenchment	80
2.8	Board composition as a function of CEO bias under Timeline IV	87

List of Tables

Acknowledgements

A great number of people helped bring this thesis into existence. I would like to express my sincere gratitude to them for their help.

First, I would like to thank my advisor, Professor Tim Baldenius, for his invaluable guidance and consistent encouragement during my Ph.D. study. Tim is the best advisor one can ask. He is very responsive, willing to help and always has the students' best interests in mind. He gave me not only the big-picture guidance, but also many very detailed comments which were right to the point. I believe that his hard working, sharpness, kindness and many other merits will be of great benefit to me for my whole life.

I am grateful to Professor Nahum Melumad for his encouragement and support over the years. It is my great honor to work with him. With his broad knowledge both in economic theory and the real world business practice, he often guided me to look at the problems from a different and inspiring angle. I am also thankful to Professor Wouter Dessein, Navin Kartik and Alexander Nezlobin for serving on my dissertation defense committee. It was my fortune to have taken a course taught by Navin in the economics department. Chapter

1 of this thesis was actually inspired by one of the papers discussed in Navin's class. He gave me a lot of helpful guidance in the project, and I am very grateful to him and Marina Halac for helpful suggestions as to fix a problem in an earlier draft. Their generous support to students is beyond the boundary of departments and schools.

There are others who have taught or helped me in different ways that I would like to thank, and they are Sudhakar Balachandran, Edwige Cheynel, Trevor Harris, Gil Sadka and Catherine Thomas. I am in particular grateful to Yuan Zhang for carefully reading the first part of the work and offering me many helpful comments from the perspective of an empiricist.

I thank Jing Li and Hanna Lee for their peer support and many stimulating discussions we had. I am grateful to Shira Cohen and Beatrice Michaeli for their generous help in proofreading this work and attending many times of my mock presentations.

I give special thanks to my family. I thank my husband, Zongjian Liu, whose love and support helped me overcome the darkest times. My angel daughter, Alethea Liu, has been my biggest inspiration. Her smiling face is the strongest motive for me to work hard. I owe my parents, who, in their senior years, encouraged their only daughter to leave them behind and pursue her dream. I am indebted to my church family, the Bridge, for their great fellowship and constant prayers.

Finally, I gratefully acknowledge financial support from Columbia Business School and the Deloitte Doctoral Fellowship Foundation.

To my family

Chapter 1

Analyst Reputation, Communication and Information Acquisition

1.1 Introduction

Financial analysts add value in the capital market by providing information to investors. One of the stylized facts emerging from the extant literature is that analysts' reports are on average optimistic. Many studies have proposed an incentive-based explanation for analyst optimism. Conventional wisdom and many academic papers (e.g., Fang and Yasuda (2005), Jackson (2005)) suggest that analysts' reputational (future) concerns may be an effective mechanism to curb opportunism. However, recent economic studies (e.g., Holmstrom (1999), Morris (2001), Ely and Valimaki (2003)) have demonstrated that reputational concerns may have *negative* effects in that they sometimes lead agents to make inefficient

decisions or ignore valuable information. For example, using an expert and decision-maker model, Morris (2001) derives that reputational concerns may *reduce* the informativeness of communication, in that no information is conveyed in equilibrium if aligned experts are sufficiently concerned about their future. In this paper, I reexamine the effect of reputational concerns and find that analysts' reputational concerns may indeed be an effective way to encourage truthful reporting, once I endogenize the analysts' information acquisition decisions.

To model the reputation formation process, I build on Morris (2001) and consider a repeated cheap talk game with two communication periods, preceded by an information acquisition stage. In each period, the investor makes an investment decision based on information strategically communicated by the analyst. The investor is uncertain about the analyst's type. An *aligned* analyst always wants the investor to make the correct investment decision. A *misaligned* analyst, in contrast, always prefers a higher investment level due to such factors as underwriting considerations, trading commission incentives, or pressure from clients or covered firms.¹ Each analyst is endowed with some noisy private

¹For underwriting considerations, see Dugar and Nathan (1995), Lin and McNichols (1998), Michaely and Womack (1999), Dechow, Hutton and Sloan (2000), and Hong and Kubik (2003). For trading commission incentives, see Irvine (2004), Jackson (2005), Cowen, Groysberg and Healy (2006). Jackson (2005) empirically documents that optimistic analysts generate higher trading volume. Cowen, Groysberg and Healy (2006) conclude that "analyst optimism is at least partially driven by trading incentives". Although clients are not obligated to deal through the broker whose analyst provided the research triggering the trading decision, they often do allocate more trade through this broker in order to maintain a good relationship with the analyst. Due to the institutional restrictions and costs associated with short-selling (D'Avolio (2002)), positive reports are more effective in generating trading than negative reports. Mola and Guidolin (2009) look at pressure from clients and empirically document that sell-side analysts are likely to assign frequent and favorable ratings to a stock after the analysts' affiliated mutual funds invest in that stock. For pressure from covered firms, see Francis and Philbrick (1993), Das, Levine and Sivaramakrishnan (1998), Lim (2001), Lambert and Sapsford (2001), and Solomon and Frank (2003).

information about the true state of the world. At the outset, the analyst may engage in (unobservable) costly information acquisition to increase the precision of her signal for both periods. At the end of the first period, the investor updates his belief about the analyst's type based on the analyst's first period report and the realized state. This updated belief about the analyst's type is labeled "analyst reputation". The second period then unfolds similarly to the first period, with a new state of the world.

To demonstrate the main result, it is useful to first examine the communication game in the second period. Since this is the last period, analysts do not care about maintaining their reputation. Consequently, the aligned analyst will report truthfully and the misaligned analyst will issue a high report independent of her signal. Hence if the investor receives a low report, he learns with certainty that the analyst is aligned. If the investor receives a high report, analyst reputation (formed in the first period) matters in that the greater the assessed likelihood that the analyst is aligned, the more seriously the investor will take the analyst's report and invest accordingly. As a result, both analysts benefit from a high reputation. In addition, since the misaligned analyst is more likely to exploit her reputation (because she always issues a high report in the second period), she benefits more from a high reputation than does the aligned analyst.

Now consider the analysts' incentives to acquire information. Note that, loosely speaking, analysts benefit from better information through two channels. First, better information increases the analysts' ability to build reputation. Recall that the misaligned analyst benefits from a high reputation even more than does the aligned analyst. Second, better

information enables analysts to guide investors toward more profitable decisions, holding reputation constant. Since the aligned analyst internalizes the investors' preferences, this increases the aligned analyst's payoff. In contrast, precision *per se* does not matter to the misaligned analyst because her payoff is independent of the state. Combining these two arguments, it is not clear, *a priori*, which type of analyst benefits more from greater precision. My main result shows that overall, the aligned analyst has a stronger incentive to acquire information. Hence, if information-gathering costs are moderate, only the aligned analyst will acquire information and as a result, analyst's reputational concerns may become an effective way to encourage truthful reporting.

The above *endogenous* link between analysts' levels of misalignment and their precision reconciles this study with Morris (2001). In the latter, analysts' precision is exogenously given and identical. Therefore there is only one way for the aligned analyst to signal her type; that is, to issue a low report, because the misaligned analyst is known to have an upward bias. As a result, the aligned analyst with high future concerns will tend to issue low reports independent of her signals, which makes communication uninformative. However, if, on the other hand, on the equilibrium path the aligned analyst has higher precision than does the misaligned analyst, then informative (first period) communication may resurface. The reason is that now there are two possible ways to build reputation: (1) by issuing a report as accurately as possible; and (2) by issuing a low report. When the misaligned analyst's reputational concerns are sufficiently high, the first mechanism emerges in equilibrium, so that both types of analysts have reputational incentives to report truthfully. That

is, the detrimental role of reputational concerns, documented in Morris (2001), disappears; instead, reputational concerns serve as an effective disciplining device to curb opportunistic analyst behavior.

If analysts' information-gathering costs are sufficiently small and the analyst's future concerns are sufficiently important, again only the aligned analyst will acquire information, but now first period communication becomes uninformative. The reason is as follows: For sufficiently small information-gathering costs, the aligned analyst is always better off acquiring information (because she wants the investor to make profitable decisions). At the same time, the misaligned analyst has no incentive to acquire information because the benefit of doing so is zero. To see this, note that if both types of analysts acquire information, then in any first period informative communication, the misaligned analyst must have positive probability to issue a high report, no matter what her signal is. Therefore, the misaligned analyst's utility at the information acquisition stage can be calculated by assuming that she always issues a high report. Thus acquiring higher precision brings no benefit to the misaligned analyst. Given that only the aligned analyst acquires information, first period communication has to be uninformative, because otherwise the misaligned analyst would also choose to acquire information for reputation building purposes. This result leads to the somewhat counterintuitive finding that a decrease in analysts' information-gathering costs may locally *reduce* social welfare because of its detrimental effect on information transmission.

Related Literature. Pioneered by Kreps and Wilson (1982) and Milgrom and Roberts

(1982), and extended in Fudenberg and Levine (1989), the “reputation effect” literature confirms the conventional wisdom that reputation effect is good in the sense that it enhances the players’ commitment power and therefore increases the players’ long-run payoffs. However, some recent studies highlight the *negative* effect of reputational concerns. Specifically, Ely and Valimaki (2003) present an example in which a long-run mechanic needs to perform the necessary repairs to get the short-run motorists’ participation; however, performing a big repair whenever it is necessary prompts the risk of being interpreted as a bad mechanic who always performs big repairs. Therefore, if the long-run mechanic sufficiently cares about his reputation, he may choose not to perform the big repair even when it is necessary, which makes the short-run motorists unwilling to participate. Ely, Fudenberg and Levine (2008) extend the above idea and characterize a class of games in which the *bad reputation effect* exists. Through rather different mechanisms, Scharfstein and Stein (1990), Prendergast and Stole (1996), and Holmstrom (1999) also show that a reputational incentive may lead to information loss or inefficient decisions.

The effect of reputation is also studied in the cheap talk literature initiated by Crawford and Sobel (1982). Focusing on reputation dynamics, Sobel (1985), Benabou and Laroque (1992), Kim (1996), Stocken (2000), Morris (2001) and Wang (2009) study repeated cheap talk games and examine how reputational concerns affect communication. Specifically, in Sobel (1985) and Benabou and Laroque (1992), the good expert is non-strategic, therefore reputational concerns encourage the bad expert (with the opposite interest of the decision maker) to truthfully report her information and hence have *positive* effect. Morris (2001)

endogenizes the aligned expert's strategic behavior and examines the possible *negative* effect of reputation. To be specific, he shows that if the misaligned expert has strong tendency to issue certain message, then the aligned expert may have an incentive to avoid sending this particular message in order to build reputation, which he refers as "political correctness". In those papers, the expert's precision is exogenously given. My paper builds on Morris (2001) and further endogenizes the expert's precision. As a result, Morris' "*political correctness*" result may be overturned and the *positive* effect of reputational concerns may be restored.²

Suurmond, Swank and Visser (2004) and Xu (2011) also try to restore the positive effect of reputational concerns by allowing the agents to acquire information. They demonstrate that reputational concerns may motivate the agent to acquire better information, which is socially beneficial. My paper is closely related to Xu (2011). The main difference is that in Xu (2011) the misaligned expert does not have reputational concerns and therefore his reporting decision is straightforward, which is he always issuing the particular message consistent with his current incentives. As a result of that, the aligned expert again has reputational incentives to avoid sending this particular message, and Morris' "political correctness" result still holds. In other words, in Xu (2011), reputational concerns still have a negative effect on communication, although overall, they have a positive effect on social welfare through the influence on information acquisition. In contrast, the focus of my paper

²There is another (lesser) difference in the modeling: the misaligned analyst's payoff takes the form of a quadratic loss function in my paper, while Morris (2001) deals with a linear payoff function for the misaligned analyst. Because I endogenize the analysts' precision choices, the payoff structure of the misaligned analyst should be comparable with that of the aligned analyst.

is the potential positive effect of reputational concerns on communication when both types of experts care about future and hence face non-trivial reporting decisions.

The effect of reputation on analysts' communication behavior in a static model is the focus of Trueman (1994) and Jackson (2005). Trueman (1994) finds that analysts, in order to enhance investors' assessments of their forecasting abilities, tend to release forecasts closer to prior expectations than is warranted given their private information, and analysts with less ability are more likely to herd. Jackson (2005) examines how analysts trade off short-term incentives to generate more trade against long-term gains from building reputation. In both papers, analysts' reputation (type) is with regard to their precision (ability), which is exogenously given, whereas in this paper, I allow analysts to choose the precision of their information, and analysts' reputation (type) is with regard to their levels of misalignment. In addition, the reputation benefit function is exogenous in the static models.

Prior research has studied analysts' communication and information acquisition behavior without reputational concerns. Morgan and Stocken (2003) study information transmission between analysts and investors when investors are uncertain about analysts' incentives and analysts' information sets are exogenous.³ Hayes (1998) examines how incentives to generate commissions affect analysts' information acquisition decisions, but assumes that analysts report truthfully. Fischer and Stocken (2010) endogenize both analysts' infor-

³Beyer and Guttman (2007) study the interaction between the analyst and the investor in a signalling model. In their model, reputational concerns constitute part of analysts' misreporting costs.

mation acquisition and their reporting behavior. They investigate how public information affects analysts' information acquisition decisions and their communication with investors.

The remainder of the paper is organized as follows: Section 1.2 lays out the model. Section 1.3 studies the communication game in each period for exogenous and commonly known analysts' precision. Section 1.4 fully characterizes the equilibrium of the model where analysts' precision choices are endogenous and unobservable. Section 1.5 discusses the welfare consequences of changes in information-gathering costs. Section 1.6 presents the empirical implications and Section 1.7 concludes. All proofs are contained in the Appendix.

1.2 Model Setup

In this section, I describe the basic setup of the model, which follows Morris (2001). I consider an investor ("he") who is uninformed about the state of the world and makes decisions based on the advice provided by an analyst ("she"). With probability λ , the analyst's preference is aligned with the investor (A); that is, she wants the investor to make correct investment decision in each period. With probability $1 - \lambda$, the analyst is misaligned (M) and always wants the investor to make "buy" decisions (independent of the state of the world). The investor is uncertain about the analyst's type $J \in \{A, M\}$ and only knows the prior probability of the analyst being aligned (λ).

The game has one information acquisition stage and two communication periods. At

stage 0, the analyst may choose to exert *unobservable* effort c to increase the precision of her signal for the following two periods. In period 1, the state of the world w_1 can take the value of 0 or 1; each state occurs with equal probability. The analyst observes an informative signal $s_1 \in \{0, 1\}$ about the state of the world, and the default precision is $\gamma_L : Pr(s_1 = w_1 | w_1) = \gamma_L \in (1/2, 1)$. If the analyst acquires information at stage 0, she will increase her precision to $\gamma_H \in (1/2, 1) > \gamma_L$ in the subsequent two periods. After observing the signal, the analyst issues a report $m_1 \in \{0, 1\}$. The investor then makes an investment decision $a_1 \in R$ according to his inference about the state based on the analyst's report m_1 . After the action a_1 is taken, the state of the world w_1 is publicly observed. Then the investor updates his belief about the analyst's type based on the realized state w_1 and the received report m_1 . As a result, the analyst now has reputation $\lambda_2 = \Lambda(m_1, w_1)$ (to be specified below) entering period 2. Period 2 then unfolds similarly to period 1, with a new and independent state w_2 (again equally likely to be 0 or 1), a new signal s_2 , a new report m_2 sent by the analyst, and a new action a_2 taken by the investor.

The sequence of events is as follows:

In each period, the investor aims to adjust his investment decision a_t to the state of the world w_t . His utility in each period, t , is given by a quadratic loss function

$$-(a_t - w_t)^2.$$

The aligned analyst has identical preferences over a_t as the investor. The utility of the

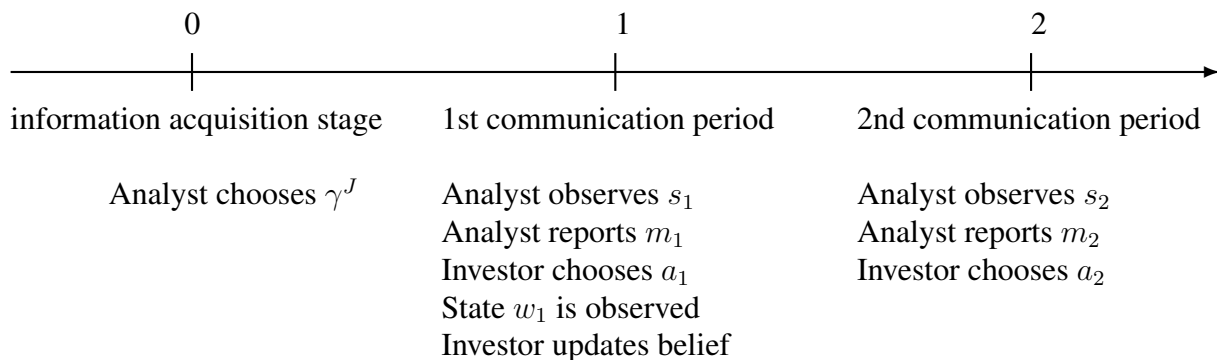


Figure 1.1: Timeline

aligned analyst is given by

$$-x^A(a_1 - w_1)^2 - (1 - x^A)(a_2 - w_2)^2 - C(\gamma^A).$$

The misaligned analyst, in contrast, always wants the high action to be chosen, independent of the state. Her utility is given by

$$-x^M(a_1 - 1)^2 - (1 - x^M)(a_2 - 1)^2 - C(\gamma^M),$$

where $0 < x^J < 1$ captures the weight type J analyst puts on period 1 utility and $1 - x^J$ is the weight on period 2 utility. I sometimes refer to $1/x^J$ as the exogenous *future (reputational) concerns*. $C(\cdot)$ represents each analyst's disutility of acquiring certain precision, with $C(\gamma_L) = 0$ and $C(\gamma_H) = c$. Note that the cost of acquiring information is assumed to be independent of the analyst's type.

An equilibrium in this game is characterized by the analyst's information acquisition strategy at stage 0, the analyst's communication strategy in each period, the decision rule for the investor in each period, and the belief function of the investor. The type J analyst's information acquisition strategy specifies the precision she will choose at stage 0; I denote it by $\gamma^J \in \{\gamma_L, \gamma_H\}$. The type J analyst's communication strategy in period t is a function $\sigma_t^J : \{0, 1\} \times \{\gamma_L, \gamma_H\}^3 \rightarrow [0, 1]$, where $\sigma_t^J(s_t, \gamma^J | \tilde{\gamma})$ is the probability of the type J analyst reporting 1 in period t when her signal is s_t and her precision is γ^J while the investor's conjecture of the analysts' precision is $\tilde{\gamma} \equiv (\tilde{\gamma}^A, \tilde{\gamma}^M)$.⁴ The investor's decision rule in period t is a function $a_t : \{0, 1\} \times [0, 1] \times \{\gamma_L, \gamma_H\}^2 \rightarrow R$, where $a_t(m_t, \lambda_t, \tilde{\gamma})$ is the investor's action in period t when he receives message m_t , his belief of the analyst being aligned is λ_t and his conjecture of the analyst's precision is $\tilde{\gamma}$. As is implied by the notation, I only consider pure strategies for the analyst's information acquisition decision; however I do allow the analyst to play mixed communication strategies.

Let $\phi_t^J(m_t | w_t)$ denote the investor's conjecture about the probability of the type J analyst sending message m_t given state w_t in period t :⁵

$$\phi_t^J(1 | w_t) = \tilde{\gamma}^J \tilde{\sigma}_t^J(w_t, \tilde{\gamma}^J | \tilde{\gamma}) + (1 - \tilde{\gamma}^J) \tilde{\sigma}_t^J(1 - w_t, \tilde{\gamma}^J | \tilde{\gamma}),$$

and $\phi_t^J(0 | w_t) = 1 - \phi_t^J(1 | w_t)$. The belief function $\Gamma_t(m_t, \lambda_t, \tilde{\gamma})$ states the investor's infer-

⁴Strictly speaking, the analyst's communication strategy should depend on her conjecture about the investor's action. However, the analyst will infer that the investor's action depends on his conjecture about the analyst's precision, $\tilde{\gamma} \equiv (\tilde{\gamma}^A, \tilde{\gamma}^M)$. Hence, I write out $\tilde{\gamma}$ instead of $\tilde{a}_t(\cdot)$.

⁵I adopt the standard notation in the literature where $\tilde{\cdot}$ represents the conjecture.

ence of the actual state being 1 in period t . By Bayes rule, it is given by

$$\Gamma_t(m_t, \lambda_t, \tilde{\gamma}) = \frac{\lambda_t \phi_t^A(m_t|1) + (1 - \lambda_t) \phi_t^M(m_t|1)}{\lambda_t \phi_t^A(m_t|1) + (1 - \lambda_t) \phi_t^M(m_t|1) + \lambda_t \phi_t^A(m_t|0) + (1 - \lambda_t) \phi_t^M(m_t|0)} \quad (1.1)$$

$\Gamma_t(m_t, \lambda_t, \tilde{\gamma})$ is well defined when the denominator is nonzero. I adopt the convention that $\Gamma_t(m_t, \lambda_t, \tilde{\gamma}) = 1/2$, the prior, if the denominator is zero. That is, when the posterior belief of the state is undefined according to Bayes rule, the investor keeps his prior belief about the state. At the end of period 1, the investor updates his belief about the analyst's type. In particular, $\lambda_1 = \lambda$ is the prior reputation, and $\lambda_2 = \Lambda(m_1, w_1|\tilde{\gamma})$ denotes the posterior reputation, defined as the investor's belief of the analyst being aligned if report m_1 is received and state w_1 is realized:

$$\Lambda(m_1, w_1|\tilde{\gamma}) = \frac{\lambda \phi_1^A(m_1|w_1)}{\lambda \phi_1^A(m_1|w_1) + (1 - \lambda) \phi_1^M(m_1|w_1)}. \quad (1.2)$$

Again, I adopt the convention that $\Lambda(m_1, w_1|\tilde{\gamma}) = \lambda$, the prior, if the denominator is zero.

At this point, I am in a position to define the equilibrium of the game.⁶

Definition 1 *A Perfect Bayesian Nash equilibrium of the game is a strategy-belief profile*

$(\gamma^{A}, \gamma^{M*}, \sigma_t^A(\cdot), \sigma_t^M(\cdot), a_t(\cdot), \Gamma_t(\cdot), \Lambda(\cdot))$ satisfying the following properties:*

⁶When I define the equilibrium here, I suppress the functional dependence of the players' strategies and belief functions on $\tilde{\gamma} = (\tilde{\gamma}^A, \tilde{\gamma}^M)$. Such conjecture will be borne out in equilibrium. The formal definition of the equilibrium is relegated to the Appendix.

(I) *The communication strategy of the type J analyst in period t , $\sigma_t^J(s_t, \gamma^{J^*})$, maximizes her utility in period t given $a_t(m_t, \lambda_t)$.*

(II) *The investor's action in period t , $a_t(m_t, \lambda_t)$, is optimal given the state inference function $\Gamma_t(m_t, \lambda_t)$.*

(III) *The information acquisition strategy of the type J analyst, γ^{J^*} , maximizes her utility at the information acquisition stage.*

(IV) *The state and type inference functions, $\Gamma_1(m_1, \lambda)$, $\Gamma_2(m_2, \Lambda)$ and $\Lambda(m_1, w_1)$, are derived from the analyst's equilibrium strategy according to inference rules (1.1) and (1.2).*

To facilitate the following arguments, I formally define “informative” communication in period t :

Definition 2 (1) *Communication in period 1 takes the form of “babbling” if $\Gamma_1(0, \lambda) =$*

$$\Gamma_1(1, \lambda) = \frac{1}{2} \text{ and } \Lambda(1, 1) = \Lambda(0, 1) = \Lambda(1, 0) = \Lambda(0, 0) = \lambda.$$

(2) *Communication in period 2 is babbling if $\Gamma_2(0, \lambda_2) = \Gamma_2(1, \lambda_2) = \frac{1}{2}$.*

(3) *Communication in period t is “informative” if and only if it is not babbling.*

Communication in the first period may be informative in terms of either the analyst's type or the underlying state, whereas the only relevant dimension of informativeness in the second period is with regard to the underlying state. In the following, I refer to an equilibrium in which communication in *each* period is babbling as a babbling equilibrium, and

an equilibrium in which communication in *either* period is informative as an informative equilibrium.

Extending standard arguments from the cheap talk literature in which babbling equilibria always exist, in my model there always exists an equilibrium in which neither analyst acquires information and communication in each period is babbling. Suppose that in each period the analyst issues report randomly, independent of her type and signal. Then the investor will rationally make his investment decision solely based on his prior knowledge of the state. Given such response of the investor, the analyst has incentive neither to deviate from her uninformative report, nor to become better informed. Therefore, a babbling equilibrium always exists and neither analyst will acquire information. The interesting question is whether and when there exist informative equilibria and which, if any, type of analyst chooses to acquire information. In the following analysis, without loss of generality, I assume $a_t(1, \lambda_t, \tilde{\gamma}) \geq a_t(0, \lambda_t, \tilde{\gamma})$.

Before proceeding, I discuss some of the key features of the model. First, the model captures the information asymmetry between the investor and the analyst about the latter's preference. The prior literature has agreed that little is known about analysts' preferences. Different analysts may have different preferences due to their respective compensation contracts, different effectiveness of the "Chinese wall" between investment banking and research groups of their respective employers, or their different levels of integrity, etc. However, the investor has little knowledge of those attributes and can only try to infer the analysts' types through their track record. Secondly, in the model, two dimensions in which

analysts can differ: their preferences and their precision. I treat the former as the primitive difference between analysts, and the latter as the endogenously derived difference. The motivation for this specific model choice is that in practice, we do observe that analysts actively engage in information acquisition through various channels such as developing industry knowledge and analyzing financial reports. Hence it appears more descriptive to allow analysts to choose their precision levels.

1.3 The Repeated Communication Game — Exogenous and Commonly Known Precision

For now, to illustrate the key features of the communication game, I take the analyst's precision γ^j as exogenously given and commonly known; I will relax this assumption in Section 1.4. The communication game can be solved by backward induction.

1.3.1 The Second Period Communication Game

At the end of period 1, the investor updates his belief about the analyst's type according to (1.2) and the analyst now has a commonly known reputation, λ_2 , entering period 2. Since period 2 is the last period, each type of analyst has no incentive to protect her reputation and simply seeks to maximize her utility in that period.

In line with the cheap talk literature, I assume that informative communication, if it

can be supported in equilibrium, is played in each period. The following argument demonstrates that pure strategy informative communication always obtains in the second period. Suppose this is the case, then $a_2(1, \lambda_2, \gamma) > a_2(0, \lambda_2, \gamma)$.⁷ Therefore the misaligned analyst has a strict incentive to report 1, and the aligned analyst must have a strict incentive to report her signal truthfully.⁸ If the investor receives message 0, he will be sure that the analyst is aligned and truthfully reporting her signal. Given the aligned analyst's precision, γ^A , the investor will assign probability $1 - \gamma^A$ to state 1 and choose action $a_2(0, \lambda_2, \gamma) = 1 - \gamma^A < 1/2$. If the investor receives message 1, he will be uncertain about the analyst's type and choose his action based on the updated belief:⁹

$$a_2(1, \lambda_2, \gamma) = \frac{\frac{1}{2}[\lambda_2\gamma^A + (1 - \lambda_2)]}{\frac{1}{2}[\lambda_2\gamma^A + (1 - \lambda_2)] + \frac{1}{2}[\lambda_2(1 - \gamma^A) + (1 - \lambda_2)]} = \frac{1 - \lambda_2 + \lambda_2\gamma^A}{2 - \lambda_2}.$$

Clearly, $a_2(1, \lambda_2, \gamma) \in [1/2, \gamma^A] > a_2(0, \lambda_2, \gamma)$. Therefore, the misaligned analyst will indeed always report 1. It is also shown in footnote 10 that the aligned analyst will indeed

⁷In Section 1.3, since the analyst's precision γ^J is exogenous and commonly known, the players' strategies and belief functions depend no longer on their conjectures about the analyst's precision, instead they depend on the commonly known $\gamma = (\gamma^A, \gamma^M)$.

⁸The argument is as follows: given that the misaligned analyst reports 1 all the time, for $a_2(1, \lambda_2, \gamma) > a_2(0, \lambda_2, \gamma)$ to hold, the aligned analyst must report 1 more often when she observes signal 1 than when she observes signal 0. Since I focus here on pure strategies, this means the aligned analyst must report her signal truthfully.

⁹This confirms Morgan and Stocken's (2003) finding that the investor's uncertainty about the analyst's incentive makes it impossible for the aligned analyst to credibly reveal good news.

truthfully report her signal.¹⁰ Hence pure strategy informative communication does obtain in the second period.¹¹

In addition, all else equal, the action induced by a high second-period report, $a_2(1, \lambda_2, \gamma)$, is increasing in analyst reputation λ_2 . The higher the probability with which an analyst is believed to be aligned, the more credible her report is perceived to be, and hence the investor will choose a higher action accordingly.

Given analyst reputation λ_2 , write $V^J(\lambda_2, \gamma)$ for the type J analyst's second period expected utility when the analyst's precision is $\gamma \equiv (\gamma^A, \gamma^M)$. The aligned analyst's second period expected utility (anticipating she will report truthfully) is therefore:

$$\begin{aligned}
 V^A(\lambda_2, \gamma) &= -\frac{1}{2}\gamma^A[a_2(1, \lambda_2, \gamma) - 1]^2 - \frac{1}{2}(1 - \gamma^A)[a_2(0, \lambda_2, \gamma) - 1]^2 \\
 &\quad - \frac{1}{2}(1 - \gamma^A)[a_2(1, \lambda_2, \gamma) - 0]^2 - \frac{1}{2}\gamma^A[a_2(0, \lambda_2, \gamma) - 0]^2 \\
 &= -\frac{(\lambda_2 - 1)^2 + 2\gamma^A(\gamma^A - 1)(\lambda_2^2 - 2)}{2(\lambda_2 - 2)^2}. \tag{1.3}
 \end{aligned}$$

¹⁰If the aligned analyst observes signal 0, she will compare her payoff conditional on sending message 0, $U_2^A(m_2 = 0, s_2 = 0, \lambda_2, \gamma^A) = -\gamma^A(a_2(0, \lambda_2, \gamma) - 0)^2 - (1 - \gamma^A)(a_2(0, \lambda_2, \gamma) - 1)^2$, with her payoff conditional on sending message 1, $U_2^A(m_2 = 1, s_2 = 0, \lambda_2, \gamma^A) = -\gamma^A(a_2(1, \lambda_2, \gamma) - 0)^2 - (1 - \gamma^A)(a_2(1, \lambda_2, \gamma) - 1)^2$. It is straightforward to show that $U_2^A(m_2 = 0, s_2 = 0, \lambda_2, \gamma^A) - U_2^A(m_2 = 1, s_2 = 0, \lambda_2, \gamma^A) = (a_2(1, \lambda_2, \gamma) - a_2(0, \lambda_2, \gamma))[a_2(1, \lambda_2, \gamma) + a_2(0, \lambda_2, \gamma) - 2(1 - \gamma^A)] > 0$. Hence the aligned analyst will indeed report 0 when she observes signal 0. Similarly, it can be shown that the aligned analyst will indeed report 1 when she observes signal 1.

¹¹In fact it can be shown that this pure strategy informative equilibrium is the unique informative equilibrium in the second period.

The misaligned analyst's second period expected utility equals:

$$\begin{aligned} V^M(\lambda_2, \gamma) &= -[a_2(1, \lambda_2, \gamma) - 1]^2 \\ &= -\frac{(\lambda_2\gamma^A - 1)^2}{(\lambda_2 - 2)^2}. \end{aligned} \tag{1.4}$$

It is straightforward to show that

$$\begin{aligned} \frac{\partial V^A(\lambda_2, \gamma)}{\partial \lambda_2} &= \frac{(1 - \lambda_2)(2\gamma^A - 1)^2}{(2 - \lambda_2)^3} \geq 0, \\ \frac{\partial V^M(\lambda_2, \gamma)}{\partial \lambda_2} &= \frac{2(2\gamma^A - 1)(1 - \lambda_2\gamma^A)}{(2 - \lambda_2)^3} > 0. \end{aligned}$$

Also

$$\frac{\partial V^A(\lambda_2, \gamma)}{\partial \lambda_2} - \frac{\partial V^M(\lambda_2, \gamma)}{\partial \lambda_2} = -\frac{(2\gamma^A - 1)(3 - 2\gamma^A - \lambda_2)}{(2 - \lambda_2)^3} < 0.$$

Both types of analysts benefit from a high reputation, *with the misaligned analyst benefiting even more*. To generate intuition for this result, notice that the investor's action upon receiving message 1 is increasing in analyst reputation λ_2 , while the action induced by message 0 is independent of λ_2 . The misaligned analyst always reports 1 and hence her reputation pays off in all scenarios. In contrast, with (ex-ante) probability 1/2, the aligned analyst reports 0, in which case her payoff is independent of her reputation, λ_2 . Only with the remaining probability, the aligned analyst reports 1 and may benefit from her reputa-

tion. Therefore, the misaligned analyst is more likely to exploit her reputation and hence benefits more from a high reputation than her aligned peer.

To conclude, in the second period, the aligned analyst reports truthfully and the misaligned analyst always reports 1. Both types of analysts benefit from a high reputation with the misaligned analyst benefiting from it even more than the aligned one.

1.3.2 The First Period Communication Game

In the first communication period, the analyst needs to anticipate the reputational consequences of the second period and takes these into consideration as she chooses her first period communication strategy. Specifically, the aligned analyst's objective in the first communication period includes both her first period payoff and her second period expected utility, and is given by

$$-x^A(a_1 - w_1)^2 + (1 - x^A)V^A(\Lambda(m_1, w_1|\gamma), \gamma).$$

Analogously, the misaligned analyst's objective in the first communication period is given by

$$-x^M(a_1 - 1)^2 + (1 - x^M)V^M(\Lambda(m_1, w_1|\gamma), \gamma).$$

According to definition 2, first period communication can be informative in terms of either the analyst's type or the underlying state. I argue that in equilibrium, however, it has

to convey information about *both* dimensions. Suppose it only conveyed information about the underlying state, while being uninformative about the analyst's type. In this case, the analyst, having no reputational consideration, would act only on her current reporting incentives in that the aligned analyst would tell the truth, while the misaligned analyst would always report 1. However, these optimal reporting strategies themselves are informative about the analyst's type, indicating a contradiction. On the other hand, suppose communication in the first period were uninformative about the underlying state, yet informative about the analyst's type. In that case, since there would be no current reporting consequences, both types of analysts would end up choosing the same communication strategy that will bring them higher reputation, which makes communication completely uninformative, indicating another contradiction. Thus in equilibrium, informative communication in the first period must convey information about both dimensions.

The striking finding of Morris (2001) is that when the aligned analyst cares a lot about the future, i.e., x^A becomes small, then no information can be conveyed in the first period. I replicate Morris' result in my setting in the following Lemma:

Lemma 1 (Morris 2001) *Suppose both types of analysts have the same precision, i.e., $\gamma^A = \gamma^M$. For any x^M , there exist cutoff values $\underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M) \leq \bar{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$ such that if the aligned analyst's future concerns are sufficiently high, i.e., $x^A < \underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$, communication in the first period is babbling. On the other hand, if and only if the aligned analyst's future concerns are sufficiently low, i.e., $x^A > \bar{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$, there exists*

a truth-telling equilibrium in which the aligned analyst truthfully reports her signal in the first period.

To understand the intuition behind this result, firstly I argue that, when both types of analysts have the same precision, there is only one way to build reputation for being aligned, which is to issue a low report. Therefore, both types of analysts will have reputational incentives to report 0. If the aligned analyst's future concerns are sufficiently high, then her reputational reporting incentive dominates and she will report 0 all the time. One would think that such communication can still be informative since the misaligned analyst's communication strategy may be signal-dependent. However the following argument demonstrates that this cannot be the case.

Suppose in an informative communication equilibrium, the aligned analyst always reports 0, independent of her signal. Furthermore, the misaligned analyst reports 1 more often when she observes signal 1 compared with signal 0, i.e., $\sigma_1^M(1) > \sigma_1^M(0)$, because, as assumed without loss of generality above, $a_1(m_1 = 1) \geq a_1(m_1 = 0)$. This implies that the misaligned analyst reports 1 more often given state 1 than he does given state 0. If indeed the aligned analyst reports 0 all the time, then the investor will rationally anticipate that: (1) the possibility of a report 0 coming from the aligned analyst is higher if the state is 1 than if the state is 0, which implies that $\Lambda(0, 1) > \Lambda(0, 0)$; (2) the analyst must be misaligned if the report is 1, i.e., $\Lambda(1, 1) = \Lambda(1, 0) = 0$. Therefore the reputation enhancement by reporting 0 when the state is 1 is greater than that when the state is 0, i.e.,

$\Lambda(0, 1) - \Lambda(1, 1) > \Lambda(0, 0) - \Lambda(1, 0)$. This in turn implies that the misaligned analyst has stronger reputational incentive to report 0 when she observes signal 1 rather than signal 0. Now recall that the misaligned analyst's current reporting incentive is independent of her signal. Therefore, the misaligned analyst will report 1 more often when her signal is 0, i.e., $\sigma_1^M(1) \leq \sigma_1^M(0)$, which contradicts $\sigma_1^M(1) > \sigma_1^M(0)$. Hence, for first period communication to be informative, the aligned analyst cannot report 0 all the time.

On the other hand, if the aligned analyst cares a lot about her current payoff, then her current reporting incentive dominates and she will report her signal truthfully.

The above result is valid for any x^M . Note that those cutoff values, $\underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$ and $\bar{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$, are functions of x^M . Roughly speaking, here in the proposed first period informative communication equilibrium, the misaligned analyst's reputational incentive is to report 0, while her current incentive is to report 1. Therefore, her future concerns, x^M , play an important role in determining her equilibrium strategy, which, in turn, affects how the investor should update his beliefs and further both types of analysts' reputational and current reporting incentives. To get an idea about the values of those cutoffs $\underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$ and $\bar{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$, the following result considers the special case where $x^M \rightarrow 0$.¹²

Corollary 1 *Suppose both types of analysts have the same precision, i.e., $\gamma^A = \gamma^M$. If the misaligned analyst has (weakly) higher future concerns than the aligned analyst, i.e.,*

¹²I would like to give special thanks to Marina Halac and Navin Kartik for pointing out a mistake in the previous version. The new Lemma 1 and Corollary 1 are inspired by their comments.

$x^A \geq x^M$, then for $x^M \rightarrow 0$, a truth-telling equilibrium exists in which the aligned analyst truthfully reports her signal in the first period. In other words, for $x^M \rightarrow 0$, $\underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M) \leq \bar{x}^A(\lambda, \gamma^A, \gamma^M, x^M) \leq x^M$.

When both types of analysts have the same precision, the misaligned analyst *has the ability* to mimic the aligned analyst. If the misaligned analyst cares (almost) only about the future, she will certainly try to mimic her aligned peer. As a result, the two types of analysts will follow sufficiently similar strategies, which, in turn, makes the reputation enhancement by reporting 0 sufficiently small. Therefore, given $x^A \geq x^M$, the aligned analyst's current reporting incentive dominates (even if $x^A \rightarrow 0$) and she will report her signal truthfully.

If the misaligned analyst has more precise information than the aligned analyst, then the above arguments apply, *a fortiori*. Hence, the same result prevails, in that first period communication takes the form of babbling when the aligned analyst's future concerns are sufficiently high. I now ask the central question for the remainder of this section: if the aligned analyst has greater precision, can informative (first period) communication obtain in equilibrium?

Proposition 1 *Suppose it is common knowledge that the aligned analyst is better informed than the misaligned analyst, i.e., $\gamma^A > \gamma^M$; then there exists value $\underline{x}^M(\lambda, \gamma^A, \gamma^M)$ such that given the misaligned analyst's important future concerns, i.e., $x^M < \underline{x}^M(\lambda, \gamma^A, \gamma^M)$, there exists an equilibrium in which both types of analysts truthfully report their signals in the first period.*

The value of $\underline{x}^M(\lambda, \gamma^A, \gamma^M)$ is derived in the Appendix.¹³ If the aligned analyst is better informed than the misaligned analyst, then informative communication may resurface in the first period regardless of the aligned analyst's time preferences.

For an aligned analyst who has an informational advantage, there are two possible ways to build reputation for being aligned: (1) by issuing a report as *accurately* as possible; and (2) by issuing a *low* report, since it is commonly known that the misaligned analyst is upwardly biased. Suppose the first mechanism dominates; then both types of analysts will have reputational incentives to truthfully report their signals. For the aligned analyst, she now has both current and reputational incentives to tell the truth and hence will report her signal truthfully. For the misaligned analyst, when her future concerns are sufficiently important, her reputational reporting incentive dominates and hence she will also truthfully report her signal. Given that both types of analysts tell the truth and the aligned analyst has a higher precision, it is indeed rational for the investor to update favorably about the analyst's type when the report is consistent with the realized state. On the other hand, suppose the second mechanism dominates and the investor updates favorably when he receives a low report, independent of the realized state. Then, when the misaligned analyst's future concerns are sufficiently important, her reputational reporting incentive dominates and she will always issue a low report, which actually makes the investor's updating rule irrational.

¹³Note that the cutoff $\underline{x}^M(\cdot)$ is independent of x^A . The reason is as follows. If the aligned analyst is better informed than the misaligned analyst, in the postulated informative communication equilibrium, both the aligned analyst's reputational and current reporting incentives are to report truthfully. Therefore the aligned analyst's future concerns, x^A , will not affect her equilibrium strategy, and hence will have no effect on the analyst's current and reputational reporting incentives.

Therefore, when the misaligned analyst values her future a lot, the first mechanism prevails in equilibrium; i.e., the analyst builds reputation for being aligned by issuing reports as accurately as possible.

To conclude, for exogenous and commonly known precision, if the misaligned analyst has equal or higher precision than the aligned analyst, and the aligned analyst's future concerns are sufficiently important, first period communication takes the form of babbling. On the contrary, if the aligned analyst is better informed, informative communication can obtain in the first period if the misaligned analyst values future a lot. In the following section, I endogenize the analyst's precision choice and assume it is unobservable to the investor. To better analyze the implications of reputational concerns, from now on, I focus on the more interesting case where the misaligned analyst has important reputational concerns and as a maintained assumption, assume $x^M < \underline{x}^M(\lambda, \gamma^A, \gamma^M)$.

1.4 Unobservable Choice of Precision

In this section, I refer back to my original model where the analyst's precision choice is endogenous and unobservable. First, it is necessary to note that with unobservable precision choice, the communication game starts from a non-singleton information set and hence is not a proper subgame. However, any informative communication in the full game satisfies the properties characterized in the case where the analyst's precision is exogenous and commonly known. This is because in equilibrium, the investor's conjecture about the

analyst's precision is consistent with the analyst's actual precision choice. Hence, although the analyst's chosen precision is unobservable, in equilibrium it will be commonly known.

If information-gathering costs are very high, then neither type of analyst will acquire information. In this case, by Lemma 1, first period communication is babbling when the aligned analyst cares a lot about the future.

If information-gathering costs are moderate, additional analysis is needed in order to understand which analyst will acquire information. To see this, it is helpful to note that, loosely speaking, the analyst benefits from acquiring information through two channels: First, more precise information increases the analyst's ability to build reputation. Recall that while both types of analysts benefit from a high reputation, the misaligned analyst benefits even more so than does the aligned analyst. Second, more precise information enables the analyst to guide the investor toward more profitable decisions, holding reputation constant. The aligned analyst's internalizing the investor's preference in effect increases her payoff. In contrast, precision, *per se*, does not matter to the misaligned analyst because her payoff is independent of the state. For later reference, I label the first channel the "reputation effect" and the second channel the "precision effect". Based on combining these two arguments, it is not clear, *a priori*, which type of analyst has a stronger incentive to acquire information. The following proposition evaluates this trade-off. To that end, denote by Δ^J the type J analyst's benefit from acquiring information.

Proposition 2 *Suppose $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$. If the misaligned analyst's future con-*

cerns are important, i.e., $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, then for moderate information-gathering costs, i.e., $\underline{\Delta} < c < \bar{\Delta}$, there exists an informative equilibrium in which only the aligned analyst acquires information and communication in each period is informative. Specifically, both types of analysts report truthfully in the first period. In the second period, the aligned analyst reports truthfully and the misaligned analyst reports 1 all the time.

The values of $\underline{\Delta}$ and $\bar{\Delta}$ are derived in the Appendix. Following the rationale established in Section 1.3.1 and Proposition 1, it is straightforward to demonstrate that: If the investor holds the (proposed equilibrium) conjectures that: (a) only the aligned analyst acquires information, (b) in the first period both types of analysts truthfully report their signals, and (c) in the second period the aligned analyst reports truthfully and the misaligned analyst always reports 1, then for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, the optimal communication strategies of the analyst with conjectured precision (as in (a)) are indeed consistent with the investor's conjectures about the analyst's communication behaviors (as in (b) and (c)). Now it remains to show that for $\underline{\Delta} < c < \bar{\Delta}$, the analyst's precision choice is consistent with the investor's conjecture (as in (a)). To that end, we need also examine the analyst's communication behaviors for off-equilibrium precision.

Given the investor's above (proposed equilibrium) conjectures about the analyst's precision choice and communication behaviors, as I show in the Appendix, under the technical conditions that $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$, the aligned analyst will report her signal truthfully in each period even if her true precision is the off-equilibrium precision, γ_L . That is, she

will truthfully report in each period irrespective of her true precision. Write $V^J(\lambda_2, \gamma^J, \tilde{\gamma}^A)$ for the type J analyst's second period expected utility when her reputation is λ_2 , actual precision is γ^J and the investor's conjecture of the aligned analyst's precision is $\tilde{\gamma}^A$.¹⁴ Hence the aligned analyst's utility at the information acquisition stage if her actual precision is γ_k , where $k \in \{H, L\}$, is as follows (the arguments in $a_1(\cdot)$ are $m_1, \tilde{\gamma}^A$, and $\tilde{\gamma}^M$):

$$\begin{aligned}
& U_0^A(\gamma^A = \gamma_k | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
= & \frac{1}{2} \gamma_k [-x^A(a_1(1, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A)V^A(\Lambda(1, 1 | \gamma_H, \gamma_L), \gamma_k, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_k) [-x^A(a_1(0, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A)V^A(\Lambda(0, 1 | \gamma_H, \gamma_L), \gamma_k, \gamma_H)] \\
& + \frac{1}{2} \gamma_k [-x^A(a_1(0, \gamma_H, \gamma_L))^2 + (1 - x^A)V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_k, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_k) [-x^A(a_1(1, \gamma_H, \gamma_L))^2 + (1 - x^A)V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_k, \gamma_H)] - \mathbb{1}_H(\gamma_k)c,
\end{aligned}$$

where the indicator variable $\mathbb{1}_H(\cdot) = 1$ if $\gamma_k = \gamma_H$, and $\mathbb{1}_H(\cdot) = 0$ if $\gamma_k = \gamma_L$.

Therefore, the benefit to the aligned analyst from acquiring information can be calculated as

$$\Delta^A \equiv U_0^A(\gamma^A = \gamma_H | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) + c - U_0^A(\gamma^A = \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L).$$

Given that the aligned analyst will truthfully report in each period irrespective of her true

¹⁴Note that $\tilde{\gamma}^M$ does not affect $V^J(\cdot)$ because the misaligned analyst always reports 1 in the second period, independent of her precision. For the same reason, holding reputation λ_2 constant, $V^M(\cdot)$ is also independent of her actual precision γ^M . Hence, later I suppress the argument γ^M in $V^M(\cdot)$.

precision, better information will not only enable the aligned analyst to provide more accurate guidance to the investor; it will also increase her ability to build reputation. Hence the aligned analyst's total benefit from acquiring information consists of "reputation effect" and two periods' "precision effect", i.e., $\Delta^A = RE^A + PE_1^A + PE_2^A = \bar{\Delta}$, where RE^J and PE_t^J represent, respectively, type J analyst's "reputation effect" and period t "precision effect". (RE^A , PE_1^A , and PE_2^A are formally derived in the Appendix - the proof of Proposition 2.)

For the misaligned analyst, in contrast, the alternate logic ensues: Given the investor's conjectures that only the aligned analyst acquires information and both types of analysts tell the truth in the first period, the investor will favorably update his belief about the analyst's type when the report is consistent with the realized state. Hence, the misaligned analyst will have greater reputational incentive to tell the truth when her precision is higher. At the same time, the misaligned analyst's current incentive to report 1 is independent of her precision. As a result, holding the investor's strategy constant, having a higher precision would make the misaligned analyst more likely to report truthfully. Remember that, for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, by Proposition 1, the misaligned analyst with the default precision will report her signal truthfully. It stands to reason, then, that the misaligned analyst will also report her signal truthfully as she becomes better informed. That is, having a higher precision does not affect the misaligned analyst's communication strategy, which in turn leaves her first period payoff unaffected (since the misaligned analyst's first period payoff

depends only on her report).¹⁵ Thus, by acquiring information, the misaligned analyst stands only to benefit by boosting her reputation. Therefore:

$$\begin{aligned}
\Delta^M &= \frac{1}{2}(\gamma_H - \gamma_L)(1 - x^M)\{V^M(\Lambda(0, 0|\gamma_H, \gamma_L), \tilde{\gamma}^A = \gamma_H) - V^M(\Lambda(1, 0|\gamma_H, \gamma_L), \tilde{\gamma}^A = \gamma_H)\} \\
&+ \frac{1}{2}(\gamma_H - \gamma_L)(1 - x^M)\{V^M(\Lambda(1, 1|\gamma_H, \gamma_L), \tilde{\gamma}^A = \gamma_H) - V^M(\Lambda(0, 1|\gamma_H, \gamma_L), \tilde{\gamma}^A = \gamma_H)\} \\
&\equiv RE^M = \underline{\Delta}.
\end{aligned}$$

Now it remains to show that $\bar{\Delta} > \underline{\Delta}$, i.e., the aligned analyst benefits more from acquiring information than does the misaligned analyst. This distinction may seem trivial at first glance, since the aligned analyst benefits from both “reputation effect” and “precision effect”, while the misaligned analyst benefits only from “reputation effect”. However, recall that the misaligned analyst is the one who reaps the greater benefit from a high reputation, thus further analysis is needed in order to evaluate this trade-off.

As will be demonstrated below, the aligned analyst’s “precision effect” is greater than the misaligned analyst’s “reputation effect”, i.e., $PE_1^A + PE_2^A > RE^M$, hence overall, the aligned analyst benefits more from higher precision than does the misaligned analyst. To illustrate this observation, first note that higher precision enables the analyst to provide more accurate guidance to the investor, even holding reputation constant. In other words, the analyst’s “precision effect” comes from her impact on the investor’s decision in each period by improving her report (from inaccurate to accurate). In contrast, the analyst’s

¹⁵Note that the ex ante probability for the analyst to observe 0 and 1 is always $\frac{1}{2}$ irrespective of her precision.

“reputation effect” comes from her impact on the investor’s decision in the second period by improving her reputation. The key to understanding Proposition 2 now is that the aligned analyst’s impact on the investor’s decision in each period through *differential reporting* is greater than the misaligned analyst’s impact on the investor’s second period decision through *differential reputation*. Figure 1.2 and Figure 1.3 demonstrates this point. In the first period, both types of analysts truthfully report their signals. Hence, the impact of differential reporting on the investor’s decision is

$$a_1(m_1 = 1, \lambda) - a_1(m_1 = 0, \lambda) = 2\lambda\gamma_H + 2(1 - \lambda)\gamma_L - 1 > 2\gamma_L - 1 > \gamma_H - \frac{1}{2}$$

given $\gamma_L \geq 0.75$. In the second period, the aligned analyst truthfully reports her signal while the misaligned analyst always reports 1. Hence, $a_2(m_2 = 0, \lambda_2) = 1 - \gamma_H$, and $a_2(m_2 = 1, \lambda_2)$ depends on the analyst’s reputation λ_2 . Clearly $a_2(m_2 = 1, \lambda_2 = 0) = \frac{1}{2}$ and $a_2(m_2 = 1, \lambda_2 = 1) = \gamma_H$. Therefore, the *maximum* effect of *differential reputation* on the investor’s second period decision is $\gamma_H - \frac{1}{2}$, while the *minimum* effect of *differential reporting* on the investor’s decision (each period) is $\gamma_H - \frac{1}{2}$. In other words, the aligned analyst’s *differential reporting* has a bigger impact on the investor’s decision than does the misaligned analyst’s *differential reputation*. Hence, the aligned analyst’s “precision effect” is greater than the misaligned analyst’s “reputation effect”, and thus overall, the aligned analyst benefits more from higher precision than does the misaligned analyst.

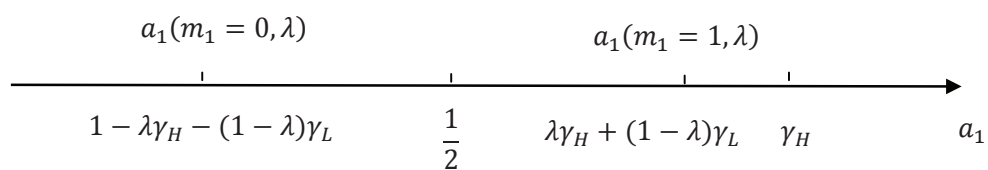


Figure 1.2: The impact of m_1 on $a_1(m_1)$

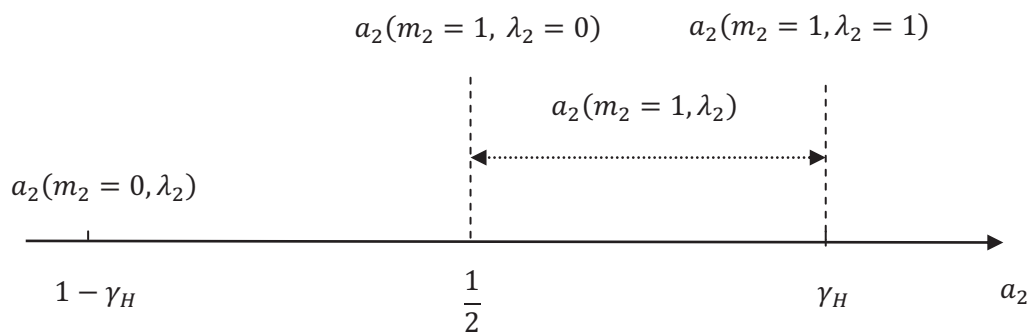


Figure 1.3: The impact of m_2 and λ_2 on $a_2(m_2, \lambda_2)$

I now address the remaining case of small information-gathering costs:

Proposition 3 *Suppose the analyst becomes almost perfectly informed after she acquires information, i.e., $\gamma_H \rightarrow 1$. Then for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$ and $x^A \in (0, \underline{x}^A(\lambda, \gamma_H, \gamma_L, x^M)] \cup [x^M, \infty)$, there exists value c^o such that for small information-gathering costs, i.e., $c < c^o$, there exists a unique informative equilibrium in which only the aligned analyst acquires information, first period communication is babbling and second period communication is informative.*

To facilitate the argument, it is useful to first identify the following property of first period informative communication when both types of analysts have the same precision. That is, in equilibrium the misaligned analyst must have positive probability to report 1 irrespective of her signal. As observed in the identical precision case, in any first period informative communication, it must be true that the misaligned analyst reports 1 more often than does the aligned analyst. Recall, as well, that the aligned analyst truthfully reports 0 if she observes signal 0, because both her reputational and current reporting incentives suggest so. Now suppose that the misaligned analyst also truthfully reports 0 given signal 0. Then, the difference in communication strategy between aligned and misaligned analysts must be based on signal 1. Given that the signal is informative, the reputation enhancement (by reporting 0) when the state is 1 is greater than that when the state is 0, which implies that the misaligned analyst would have stronger reputational incentive to report 0 when she observes signal 1 rather than signal 0. Observing that the misaligned analyst's current

reporting incentive is independent of her signal, the misaligned analyst would thus report 0 more often given signal 1 than given signal 0, which, combined with the supposition that the misaligned analyst reports 0 given signal 0, implies that the misaligned analyst would report 0 all the time. This contradicts the property that the misaligned analyst reports 1 more often than does the aligned analyst. Hence the misaligned analyst will have positive probability to report 1 given signal 0, and also a fortiori, given signal 1.

If analysts' information-gathering costs are small, then the aligned analyst is always better off acquiring information even by the second period "precision effect" alone. Given that the aligned analyst acquires information, I claim that the misaligned analyst will maintain her default precision for the following reasons: as argued in the preceding paragraph, if in equilibrium both types of analysts acquire information, then the misaligned analyst must have positive probability to report 1 irrespective of her signal. It is shown in the Appendix that holding the investor's strategy unchanged, the misaligned analyst with the default precision γ_L also has positive probability to report 1 for either signal. Therefore, by the definition of mixed strategy, the misaligned analyst's utility at the information acquisition stage in either case — with or without acquiring information — can be calculated by assuming that she always reports 1 in the first period (holding the investor's strategy unchanged). As a result, the misaligned analyst has no incentive to acquire information simply because the benefit of doing so is zero. Given that only the aligned analyst acquires information, the first period communication has to be babbling for small information-gathering costs;

otherwise, the misaligned analyst would also choose to acquire information for reputation building purposes (since $c < c^o = \Delta^M$), which contradicts the afore presented argument.¹⁶

At first glance, Proposition 3 seems to conflict with Proposition 1, since the latter states that first period informative communication obtains for $x^M < \underline{x}^M(\lambda, \gamma^A, \gamma^M)$ if the aligned analyst is better informed. However, recall that Proposition 1 speaks to the case where analysts' precision is common knowledge. When analysts' precision choices are unobservable, then one has to take into consideration possible deviations not only regarding analysts' reporting strategies, but also regarding their precision choices when analyzing whether first period informative communication can be sustained in equilibrium.

To conclude, for the most interesting case where both types of analysts' future concerns are important, I find that if information-gathering costs are sufficiently small, only the aligned analyst acquires information and first period communication is uninformative. Further, if information-gathering costs are moderate, again only the aligned analyst acquires information but first period communication becomes informative. If information-gathering costs are high, neither analyst acquires information and first period communication reverts to babbling.

¹⁶The condition $x^A \in (0, \underline{x}^A(\lambda, \gamma_H, \gamma_L, x^M)] \cup [x^M, \infty)$ is required to guarantee that both types of analysts truthfully reporting is the *unique* informative communication for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$. Therefore ruling out such informative communication for $c < c^o$ implies that the first period communication has to be babbling.

1.5 Policy Implications

In general, analysts' information-gathering costs are influenced by many factors: access to management inside information, complexity of industry knowledge, transparency of financial disclosure, etc. All else equal, the more transparent the financial disclosure, the lower the analysts' information-gathering costs. Proponents of increased transparency usually base their arguments on the conventional wisdom that higher transparency would increase the welfare of financially unsophisticated investors. However, it has been argued that regulatory acts requiring increased transparency (e.g., FAS 131 segment reporting) impose significant "implementation costs" on firms and hence, the social welfare consequences of such regulatory acts are up for debate. The following analysis points out that even without such implementation costs, greater transparency may decrease social welfare, by reducing the informativeness of communication between analysts and investors.

Corollary 2 *Suppose $\gamma_H \rightarrow 1$, $\gamma_L \geq 0.75$ and all players (both types of analysts and the investor) have the same future concerns, i.e., $x^A = x^M = x^I = x$. Consider $c_1 = c^o - \delta$ and $c_2 = c^o + \varepsilon$, then if the players' future concerns are sufficiently high, i.e., $x < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, then for $\delta \rightarrow 0$ and $\varepsilon \rightarrow 0$, decreasing analysts' information-gathering costs from c_2 to c_1 reduces social welfare.*

Consider the case where future concerns are important. Then, Proposition 2 has shown that for moderate information-gathering costs, there exists an informative equilibrium in which only the aligned analyst acquires information and both types of analysts tell the

truth in the first period. In contrast, by Proposition 3, given certain regulatory conditions, communication in the first period takes the form of babbling if the analyst's costs of acquiring information are small. That is, if information-gathering costs decrease from c_2 to c_1 , communication in the first period changes from informative to babbling.

First period informative communication not only enables the investor to make better concurrent decisions, but also facilitates learning about the analyst's type, which leads to better decision making in the second period. Similarly, the aligned analyst, who internalizes the investor's payoff, is also better off under informative communication, given that the additional information-gathering costs are negligible as postulated in Corollary 2. In contrast, informative communication leaves the misaligned analyst worse off. To see this, note that the misaligned analyst's payoff in the first period, $-(a_1(\cdot) - 1)^2$, is independent of the state and increasing and concave in the investor's action. If first period communication is informative, then the investor's investment decision will be contingent non-trivially on the analyst's report; given the concavity of the misaligned analyst's first period payoff, this will make her worse off for the first period (by Jensen's inequality). In addition, when first period communication is informative, the misaligned analyst cannot completely pool with the aligned analyst, which reduces her payoff for the second period (recall that babbling amounts to complete pooling of analysts' types).

Given the countervailing effects of first period communication on the various players' payoffs, a tradeoff arises. Corollary 2 evaluates this tradeoff, with social welfare $W(\cdot)$ defined as the investor's ex ante expected utility plus the analyst's ex ante expected utility

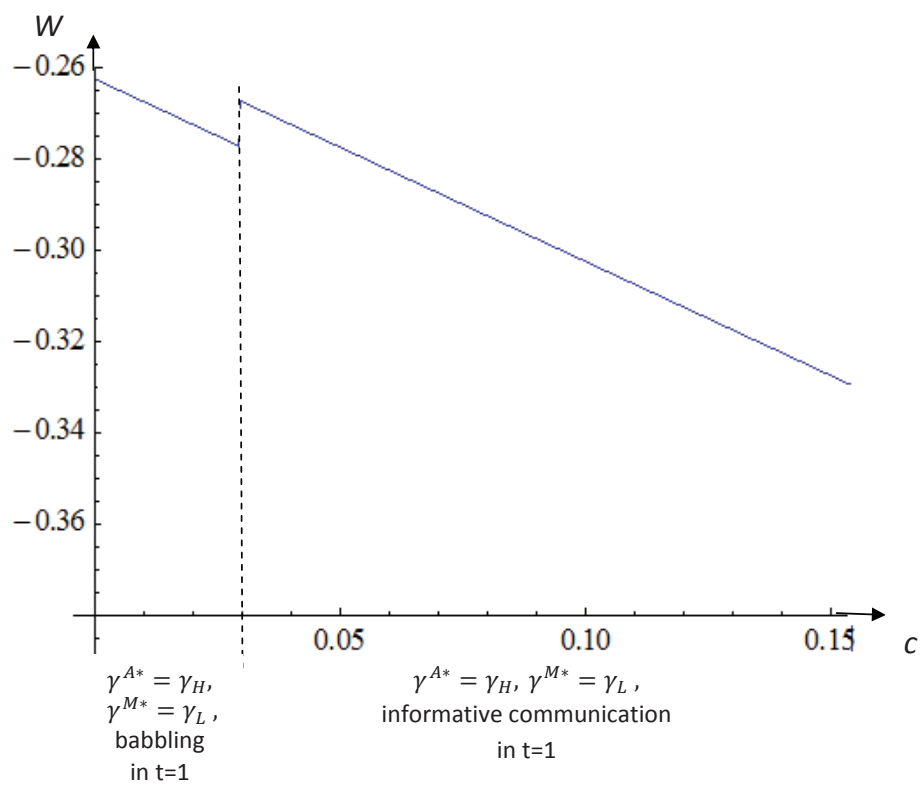


Figure 1.4: Social welfare as a function of c
 For this figure, the parameter values are $x = 0.05$, $\gamma_H \rightarrow 1$, $\gamma_L = 0.8$ and $\lambda = 0.5$.

(before she learns her type). I show in the proof of Corollary 2 that:

$$W(c = c_2) - W(c = c_1) = \frac{1}{2}x\lambda(a_1(1) - a_1(0)) - \lambda(c_2 - c_1).$$

The first term captures the net effect of the investor making informative decisions, whereas the second term represents the difference in analysts' information-gathering costs. Clearly, the first term is positive, and the second term approaches zero when the cost difference is negligible. Hence, social welfare is reduced when information-gathering costs decrease from c_2 to c_1 (see Figure 1.4).

Fischer and Stocken (2010) derive a related result. They find that in some circumstances, more precise public information may “crowd out” analysts' private information, making the investor strictly worse off. My analysis shows that in the presence of reputation formation, increased transparency may sometimes impede communication between analysts and investors, and hence may have negative social welfare consequences.

1.6 Empirical Implications

This paper generates several empirical predictions consistent with extant empirical studies.

First, the model suggests that analysts with better reputation have greater impact on investors' decisions, in line with Stickel (1992), Park and Stice (2000) and Jackson (2005).¹⁷

¹⁷In these studies, analyst reputation is with regard to the precision, not the levels of misalignment. However, these two different dichotomies converge as aligned analysts have higher precision in equilibrium when

Second, higher reputation analysts have better future performance. This, too, is consistent with Stickel (1992) and Desai, Liang and Singh (2000), among others. Third, the model suggests that in general, analysts' reputation is increasing in their forecast accuracy, consistent with Chen, Francis, and Jiang (2005) and Jackson (2005). Fourth, on average, analysts' reports are informative (e.g., Dimson and Marsh (1984), Womack (1996)), but optimistic (e.g., O'Brien (1988), Lys and Sohn (1990), Brown (1993), Dugar and Nathan (1995)). Fifth, if one interprets aligned analysts as unaffiliated and misaligned analysts as affiliated—a reasonable link at least before the Global Settlement—this paper predicts that affiliated analysts issue more optimistic reports than do unaffiliated analysts, which is again confirmed by most empirical findings (e.g., Dugar and Nathan (1995), Lin and McNichols (1998), Michaely and Womack (1999)).

The model also generates new empirical implications. Proposition 2 suggests that in general, aligned analysts tend to have more precise information and are more likely to report truthfully. Hence, from an ex post perspective, my model predicts an endogenous association between analysts' forecast accuracy and bias. More specifically, analysts with higher forecast accuracy are less optimistic. This prediction is confirmed by Conroy and Harris (1995). However, Lim (2001) and several other studies (e.g., Chen and Matsumoto (2006)) hold the contrary view that more optimistic analysts tend to be more accurate because of privileged access to management inside information. Regulation FD aims to

information-gathering costs are moderate. Hence, in equilibrium analyst reputation conveys information both about the levels of misalignment and the precision.

prohibit selective disclosures with the goal of creating a more even playing field among analysts. I therefore expect that my model speaks more to the post-Reg FD regime. Furthermore, my results imply that if misaligned analysts have a significant informational advantage over aligned analysts, then communication may become uninformative. Hence Reg FD has the additional benefit of fostering communication by limiting misaligned analysts' informational advantage.

Somewhat surprisingly, Proposition 3 suggests that if learning more about firms comes at a low cost to analysts (e.g., due to high disclosure standards) and future concerns are weighted heavily, analysts' recommendations will be less informative.

1.7 Conclusion

This paper investigates how the presence of reputation formation affects analysts' strategic communication with investors when analysts can acquire information. If information-gathering costs are moderate, only aligned analysts acquire information; as a result, informative communication can be sustained in equilibrium. That is, analysts' communication in the reputation formation stage conveys information both about their types (the levels of misalignment) and their precision (given the endogenous association between the analysts' types and precision). On the other hand, if information-gathering costs are small and analysts care a lot about the future, communication becomes uninformative. Hence, efforts

aimed at reducing information-gathering costs for analysts may have adverse effects on social welfare.

Certain key features of the model warrant further discussion. For example, I assume that analysts' precision is perfectly correlated over time. In the extreme opposite case of zero serial correlation (i.e., acquiring information only increases analysts' precision in the first period), neither analyst will acquire information regardless of information-gathering costs if both types of analysts care a lot about future transactions.¹⁸ In practice, however, some elements of analysts' information sets are clearly persistent (e.g., industry knowledge, the ability to analyze financial statement). My qualitative findings (assuming perfect correlation) will continue to hold when the correlation is sufficiently large.

Finally, in this paper I only consider analysts' messages about their signals. What if one allows analysts to send multidimensional cheap talk messages? For example, analysts could send a separate message about their precision in addition to the one about their signals. Because I confine my attention to pure strategies for analysts' information acquisition decisions, this possibility does not affect my analysis. The reason is that, when analysts are confined to pure strategies for their information acquisition choices, investors will ignore analysts' separate cheap talk messages about their precision because their precision

¹⁸The key to understanding this result is that misaligned analysts benefit more from a high reputation than do their aligned peers. When both types of analysts have sufficiently important future concerns, they acquire information mainly for reputation building purposes. Then, whenever it is profitable for the aligned analysts to acquire information, it is also optimal for the misaligned analysts to do so. Hence, equilibrium cannot be reached where only the aligned analysts acquire information. At the same time, following the rationale established in Proposition 3, both types of analysts acquiring information cannot constitute equilibrium either, because the misaligned analyst has no incentives to acquire information for any level of information-gathering costs.

choices are deterministic in equilibrium. An interesting model extension is to allow for mixed strategies for analysts' information acquisition decisions. The extent to which this extension would affect my results has yet to be explored.

Chapter 2

Board Composition and CEO power

2.1 Introduction

The board of directors performs the dual function of monitoring the firm's management and advising it on key decisions. In the wake of major corporate governance failures, commentators have questioned the effectiveness of boards in performing these tasks — especially, the monitoring task. This concern is rooted in the inherent information asymmetry between board and management and in the widely held belief that powerful CEOs use their (often significant) influence over the board's composition to ensure a “pliable” board.¹ The Sarbanes-Oxley Act (SOX) of 2002 was designed in part to strengthen the monitoring function of the board. Recent evidence in Faleye et al. (2010) suggests, however, that this has come at the expense of a diminished advisory role. In this paper we develop a model to

¹Tirole (2001, p.5) summarizes this succinctly: “Managers value ... lenient oversight”.

study the link among board composition (holding the board size constant), the assignment of decision-making authority, and the flow of information inside a firm. We then employ this model to assess how CEO power affects these constructs — in particular, the board composition.

We model a firm whose CEO has private information about the NPV-maximizing scale of investment but is an empire builder. Board members, while internalizing the shareholders' objective, have an information disadvantage at the outset. They are either *monitoring* or *advisory* types. The primary task of the “monitors” is to uncover the CEO's information, whereas that of the “advisors” is to discover incremental decision-relevant information that is unbeknownst even to the CEO. In practice, of course, most board members engage in both monitoring and advising. Yet, they allocate their time across the two tasks according to their respective background and expertise.² Our model adopts a reduced form approach by letting the board composition parameterize the monitoring intensity. As a result, the information collectively available to the parties is endogenously determined by the board composition: the more advisors are on the board, the more likely new decision-relevant information will be discovered but the less effective is the board monitoring, holding constant the board size.

To study the interplay of board composition and decision making, we consider two organizational modes. Under *centralization*, the decision is made by the board. Under *del-*

²Board members who are financial experts (e.g., accountants or former CFOs) spend more of their time monitoring and often serve on the audit committee. Former CEOs, consultants or technology and marketing executives, on the other hand, spend more time advising and often serve on the M&A or investment committee (a typical advisory function). See Faleye et al. (2011) for details.

egation, it is made by the CEO unless board monitoring uncovers the CEO's information, in which case the board effectively forces the CEO's hand. Given the information-based notion of monitoring, our model thus allows for *selective intervention*. Consistent with earlier models that have treated the information environment as exogenous (e.g., Dessein 2002, Harris and Raviv 2005, Adams and Ferreira 2007), we find that the less biased is the CEO, the more the board will delegate.³ This holds regardless of the board composition. Exploiting the endogenous nature of information in our model, we show that advisor-heavy boards and centralization of authority are complements. Under centralization the board can perfectly adjust the investment decision to its own information, whereas under delegation part of the information is lost in the process of noisy cheap-talk communication — hence, any new information an advisor-heavy board may discover is more valuable under centralization.

In reality, of course, firms choose the board structure and allocation of authority jointly. In that case, we establish a non-monotonic relationship between CEO bias and board composition. For small biases, as those increase, firms will stick to delegation but add monitors to the board, as the value of selective intervention upon successful monitoring grows larger. As the CEO's bias increases further, the firm will eventually centralize investment decisions. This in turn raises the value of board-generated new information, which calls for more advisors (i.e., fewer monitors) on the board. As CEO bias increases even fur-

³Harris and Raviv (2008) also endogenize the board's information in that acquiring new information is individually costly to board members. They study the associated free-rider problem. However, the board in their model never makes an attempt at uncovering the CEO's information, i.e., there is no board monitoring in the sense of our model.

ther, the firm will again add more monitors to the board, because a highly biased CEO will submit very noisy signals about his information to the board. Empirically, therefore, the relationship between agency problems and monitoring intensity of boards may be *locally decreasing* in the data, if researchers fail to control for the allocation of decision rights within firms.

An often voiced concern is that CEOs who can influence the board composition will nominate fewer monitors in order to have an agreeable board (e.g., Hermalin and Weisbach 1998, Tirole 2001). Recent regulations (e.g., the Sarbanes-Oxley Act, SOX) aim to enhance the monitoring role of boards, among other things. Ironically, our results suggest that for such regulations to increase shareholder value, CEO bias has to be small. Only then does the conventional wisdom prevail that the CEO prefers less board monitoring than do shareholders. In contrast, we show that more biased CEOs may have incentives to nominate a *more* monitor-heavy board, because such a board is more likely to delegate. In other words, the CEO prefers to be in control of the decision — even if subject to greater scrutiny and the threat of selective intervention — rather than having authority outright centralized by an “activist” board that is prone to learn valuable information of its own. Regulations fostering greater monitoring scrutiny may thus exacerbate imbalances in board composition precisely in those circumstances (severe agency problems) for which these regulations were designed.⁴

⁴Numerous studies have considered the welfare effects of SOX, mostly on the direct compliance costs associated with certain SOX provisions, such as Section 404 on internal controls, e.g., Zhang (2007). Our results demonstrate an additional opportunity cost of strengthening the monitoring role of boards, namely that the board becomes less productive as a source of strategic advice.

Short of influencing the board nomination process, CEOs have more subtle means of entrenching themselves and securing authority within the firm, e.g., by choosing investment projects that confer them an information advantage (see also Harris and Raviv 2005).⁵ We refer to such projects as “complex.” To prevent entrenchment, shareholders may be better off committing to an advisor-heavy board at the outset. Given the stickiness of boards, this constitutes a credible commitment to centralize decisions at a later stage even if the CEO, in the interim, were to choose a complex project. Still, even if CEO entrenchment can be preempted in equilibrium, the specter of such behavior creates additional costs to shareholders; again, in the form of an imbalance between the two desired board activities.

Taken together, our results demonstrate that caution is in order when interpreting observed patterns of board composition. Advisor-heavy boards could indicate either negligible agency problems (if CEO entrenchment is not an issue) or severe agency problems (if shareholders need to actively preempt entrenchment by a highly biased CEO). From an empirical perspective, this highlights the need for careful contextual analysis and by controlling for proxies of CEO power such as chairman duality, tenure, founder status, etc.

Related literature. The view of boards serving a dual role as advisors and monitors dates back to Mace (1971). While most of the literature has focused on the relationship between inside and outside board members (e.g., Raheja 2005, Adams and Ferreira 2007, Harris and Raviv 2008, Linck et al. 2008), Linck et al. (2009) and Adams (2009) have

⁵This argument has been made by Shleifer and Vishny (1989) and Edlin and Stiglitz (1995) in connection with CEO replacements.

brought the dual *functional* role (a la Mace) back to the fore by documenting a shift toward board monitoring as a result of SOX (see also Becht et al. 2008).⁶ Faleye et al. (2010) study the tradeoff between monitoring and advising empirically and find that this increase in board scrutiny has come at a significant cost in terms of suboptimal advising, especially as it relates to acquisitions and innovation. Throughout our model we assume the board is fully aligned with the interest of shareholders, which is consistent with it including mostly outsiders.⁷ Our focus is therefore not on the insiders/outside dichotomy but on the question *what kind of outsiders* should be nominated.⁸

Directly examining the different functional roles of boards, Baker and Gompers (2003) study the implications of venture capitalist representation. VCs perform both advisory and monitoring roles, yet Baker and Gompers conclude that VCs mostly substitute for “instrumental directors,” who would be classified as advisors according to our taxonomy. Consistent with our model predictions, the authors find that boards with greater VC representation tend to be more actively involved in decision-making.

⁶Linck et al. (2009, p.3289): “The composition of the boards...changes with relatively more lawyers and financial experts and relatively fewer executives from other firms than before SOX.”

⁷Harris and Raviv (2010) and Chakraborty and Yilmaz (2010) study biased boards in strategic communication settings.

⁸Jensen (1993) observes that insiders make for bad monitors, which suggests equating insiders with advisors, and outsiders with monitors. This approach is taken in numerous empirical studies. For instance, Coles et al. (2008) and Bathala and Rao (1995), respectively, find that firms with higher R&D expenditures or more growth options have more advisory needs and nominate more insiders to the board. More generally, Lehn et al. (2004) show that, over the life cycle of firms, boards become more outsider-heavy, consistent with the view that mature firms are plagued by more severe agency problems. However, this mapping is far from clear. For instance, outsiders may be former entrepreneurs (more likely to be advisors — see Bhagat and Black (1999) on high-growth firms), or bankers, former regulators/public sector officials or CPAs (more likely to be monitors). Caution therefore is in order when linking our results to empirical findings along the insiders/outside dichotomy.

Regarding the issue of CEO power, Shivdasani and Yermack (1999) find that CEOs frequently play a key role in the board nomination process. In a dynamic model, Hermalin and Weisbach (1998) show how CEO power arises endogenously as a result of delivered performance. Shleifer and Vishny (1989) and Edlin and Stiglitz (1995) demonstrate how managers can entrench themselves by choosing complex strategies or products. We will take the extent of CEO power as given in (some parts of) our model and study its effect on board composition.

Technically, our paper is related to Harris and Raviv (2005, 2008, 2010), and Adams and Ferreira (2007). Building on the work of Dessein (2002), these studies analyze strategic communication between the board and the CEO. Our paper adds to this stream of literature by investigating the role of CEO power, monitoring as a source of information for the board, and selective intervention. Adams and Ferreira (2007) also study board monitoring, but they model monitoring as a power struggle between the board and CEO, with independent boards having lower cost of wresting away control from the CEO. Our notion of board monitoring differs in that it centers on information collection. Harris and Raviv (2008) also model the board's acquisition of "new" information (and the attendant free-rider problem), but they do not consider board monitoring.

The remainder of the paper is organized as follows. Section 2.2 lays out the model. Section 2.3 looks at special cases where either the allocation of decision rights or the board composition is exogenously specified. In Section 2.4, shareholders choose both these design variables endogenously. Section 2.5 considers the case where a powerful CEO de-

termines (or influences) the board composition. Section 2.6 analyzes the case of CEO entrenchment through choosing complex projects. Section 2.7 concludes. All proofs are contained in the appendix.

2.2 Model

We model the strategic interaction between a privately informed CEO of a firm and its board of directors. We take the board size as given and focus on the interaction between board composition and the allocation of decision-making authority.⁹ To capture the dual role of boards parsimoniously, let $x \in [0, 1]$ denote the fraction of board members engaged in monitoring (the “monitoring” types) and $1 - x$ the fraction of board members engaged in generating information incremental to that held by the CEO (the “advisory” types). As noted in the introduction, this is a reduced-form representation of a more general model where board members specialize endogenously according to their respective expertise. For instance, former CFOs serving on boards often focus on the monitoring function whereas marketing executives tend to focus on advising.

The firm faces an investment decision, denoted by $y \in \mathbb{R}$. In line with earlier studies by Harris and Raviv (2005, 2008, 2010) and Adams and Ferreira (2007), both the CEO and the board want to adjust the investment decision to some state of the world, θ . The board aims to maximize shareholder value (i.e., the NPV), which is represented by the following

⁹Several empirical studies have shown negative correlation between board size and firm performance beyond a certain board size, e.g., Yermack (1996), Eisenberg et al. (1998).

quadratic loss function (ignoring lump sum terms):

$$-(y - \theta)^2. \quad (2.1)$$

The CEO on the other hand is biased towards larger investments, possibly due to empire benefits, and aims to maximize

$$-(y - \theta - b)^2, \quad b > 0. \quad (2.2)$$

Throughout the paper we assume that the CEO's bias parameter b is commonly known at the outset. We do not model compensation as a means to align the CEO's incentives with that of the shareholders: in practice, the higher the CEO's pay-performance sensitivity, the lower one would expect the bias, b , to be. The state variable is additively separable over two signals, each independently uniformly distributed: $\theta = a + p$, where $a \sim U[0, A]$ and $p \sim U[0, P]$. The CEO observes a with certainty, whereas the board *may* observe a and/or p . We assume that it is common knowledge whether the board is informed about a or p .

The probability that the board learns the state variables depends on the board composition. In particular, for given fraction of monitoring types, x , the board observes a with probability $q_a(x)$, and p with probability $q_p(x)$. The higher is x , the higher is $q_a(x)$ (better monitoring) but the lower is $q_p(x)$ (less gathering of new information). That is, engaging in more monitoring comes at an opportunity cost of reduced board creativity, and vice versa.

Specifically, we assume that $q_a(\cdot) \in [0, 1)$, $q_p(\cdot) \in [0, 1)$, $q'_a(\cdot) > 0$, $q''_a(\cdot) \leq 0$, $q'_p(\cdot) < 0$, $q''_p(\cdot) \leq 0$. Both board activities are subject to decreasing marginal returns, and the respective success probabilities are bounded away from one. It seems plausible that even if the board is entirely devoted to monitoring it will not discover the CEO's information, a , with certainty; the same applies to acquiring new information about p .

Throughout the paper we assume that $A \geq P$, i.e., the CEO's private information is weakly "more important" to the investment decision than the information that may be discovered by the board. This seems descriptive given the CEO's proximity to day-to-day operations. Moreover, in our model, decision-making authority is allocated before the parties observe their respective information. Our base model is therefore broadly in line with the "ex-ante environment" in Harris and Raviv (2005). The main modification is that, if board monitoring uncovers a , the investment decision will be chosen in the interest of the shareholders under both centralization and delegation as a result of selective intervention by the board (see the next section for details).¹⁰

2.2.1 Payoffs for Given Allocation of Decision Rights

If decision-making is centralized, it is the board that chooses y so as to maximize its objective in (2.1) given its available information, which may include p and a . If the board has

¹⁰Harris and Raviv (2005) show that under certain conditions it is preferable to condition the centralization/delegation decision on the realization of p (and possibly a). Harris and Raviv (2008, 2010) further explore this "ex-post" environment. We also ignore conditional delegation, e.g., delegation up to a spending limit (as in Melumad and Shibano 1991, Dessein 2002, Alonso and Matouschek 2007), as this would require a formal contract.

not learned a through monitoring, it will ask for a report, r_a , from the privately informed CEO. In a setting like ours, it is natural to assume that communication takes the form of “cheap-talk” (e.g., Crawford and Sobel 1982). Specifically, the report itself does not affect either player’s payoffs directly; the set of available reports is independent of the realized state; and the parties cannot commit to report-contingent contracts. We denote the board’s information set by Ω_B and the resulting investment choice by $y_B(\Omega_B)$. Then,

$$\begin{aligned}
 y_B(\Omega_B) &\in \arg \max_y E\{-(y - \theta)^2 \mid \Omega_B\} \\
 &= \begin{cases} a + p, & \text{if board learns } a \text{ and } p, \\ a + E[p], & \text{if board only learns } a, \\ E[a \mid r_a] + p, & \text{if board only learns } p, \\ E[a \mid r_a] + E[p], & \text{if board learns neither } a \text{ nor } p. \end{cases} \quad (2.3)
 \end{aligned}$$

We further denote by $\bar{\sigma}_j^2$, $j = A, P$, the prior variance of a or p , which, by (2.1) and (2.2), represents the expected information loss if a decision is made without any knowledge of the realizations of the respective random variables. Similarly, $\sigma_A^2(b)$, to be specified below, denotes the expected information loss regarding a conditional on receiving a message r_a , in case board monitoring has failed to uncover a .¹¹ Since the board learns a with probability $q_a(x)$ and p with probability $q_p(x)$, we can state the board’s (and hence the

¹¹The information losses with and without reporting are also functions of A or P , respectively. To avoid clutter, we suppress this dependency whenever there is no risk of confusion.

shareholders') expected payoff under centralization as follows:

$$\begin{aligned}
\pi^C &= - \left\{ q_a(x)(1 - q_p(x))\bar{\sigma}_P^2 + (1 - q_a(x)) \left[q_p(x)\sigma_A^2(b) + (1 - q_p(x))(\sigma_A^2(b) + \bar{\sigma}_P^2) \right] \right\} \\
&= - \left[(1 - q_p(x))\bar{\sigma}_P^2 + (1 - q_a(x))\sigma_A^2(b) \right] \\
&\equiv -Var^C(x, b),
\end{aligned} \tag{2.4}$$

where Var^C denotes the *residual variance* (with cheap talk) under centralization.

Consider now the case of delegation. If the board learns a , it would be implausible to assume that the CEO can act according to his bias. Instead, the board then intervenes selectively by forcing the CEO's hand toward the investment level which maximizes shareholder value, y_B .¹² Only if the board does not learn a , the CEO has real decision-making authority and chooses y in accordance with (2.2), given his available information set, Ω_{CEO} :

$$y_{CEO}(\Omega_{CEO}) = \begin{cases} a + p, & \text{if board learns } a \text{ and } p, \\ a + E[p], & \text{if board only learns } a, \\ a + E[p | r_p] + b, & \text{if board only learns } p, \\ a + E[p] + b, & \text{if board learns neither } a \text{ nor } p. \end{cases} \tag{2.5}$$

The first two cases of (2.5) effectively replicate those of (2.3) for the case of successful monitoring. If board monitoring was unsuccessful (the board has not learned a), then with probability q_p the CEO can incorporate into the investment decision a noisy cheap-talk

¹²For instance, the board could make use of an implicit (off-equilibrium) threat of dismissal.

report, r_p , sent by the informed board. Denoting by $\sigma_p^2(b)$ the expected information loss regarding p conditional on this report, the board's expected utility under delegation reads:

$$\begin{aligned}
\pi^D &= - \left\{ q_a(x)(1 - q_p(x))\bar{\sigma}_P^2 + (1 - q_a(x)) \left[q_p(x)(\sigma_P^2(b) + b^2) + (1 - q_p(x))(\bar{\sigma}_P^2 + b^2) \right] \right\} \\
&= - \left[(1 - q_p(x))\bar{\sigma}_P^2 + (1 - q_a(x))q_p(x)\sigma_P^2(b) \right] - (1 - q_a(x))b^2 \\
&\equiv -Var^D(x, b) - (1 - q_a(x))b^2,
\end{aligned} \tag{2.6}$$

where Var^D denotes the residual variance under delegation. Delegation imposes an additional bias cost on the shareholders: with probability $(1 - q_a)$ the CEO retains effective control over the investment decision and overinvests.

Now consider the CEO's payoff under the respective regimes. All else equal, the CEO, too, aims to minimize the residual variance. At the same time, he has a preference for being in control (subject to the threat of selective intervention):¹³

$$\pi_{CEO}^C = \pi^C - b^2 \quad \text{and} \quad \pi_{CEO}^D = \pi^D + [1 - 2q_a(x)]b^2. \tag{2.7}$$

Comparing (2.7) with (2.4) and (2.6) yields the following observation (stated without formal proof):

Lemma 2 *Conditional on centralization the CEO prefers the same board composition as the shareholders, whereas conditional on delegation the CEO prefers a lower x than the*

¹³It is easy to show that π^C , π_{CEO}^C and π^D , respectively, each are concave in x . The same however does not necessarily hold for π_{CEO}^D .

shareholders. Furthermore, if the board prefers delegation for some A, P, b , and x , then a fortiori so does the CEO.

All else equal, each of the players prefers to be in control in that $\pi^D \geq \pi^C$ implies $\pi_{CEO}^D > \pi_{CEO}^C$, whereas $\pi_{CEO}^C > \pi_{CEO}^D$ implies $\pi^C > \pi^D$. Moreover, in line with conventional wisdom, we find that the CEO indeed always prefers (weakly, at least) a less monitor-heavy board than the shareholders, holding constant the allocation of decision rights. Under centralization, the board is in control with probability one. Hence, the board composition only affects the residual variance and both parties agree on the variance-minimizing level of x . Under delegation, x not only affects the variance (which, again, both parties prefer to be small) but also the probability of selective intervention. Specifically, the greater is x , the more likely board monitoring will succeed, and thus the more likely the investment decision will ultimately reflect the shareholders' preferences.

2.2.2 Communication Equilibrium

We now briefly describe the communication equilibria and characterize the expected information loss terms $\sigma_A^2(b)$ and $\sigma_P^2(b)$. In a general cheap-talk game, a sender remits a (possibly noisy) signal to a receiver, who then updates his belief according to Bayes' rule and makes a decision. As shown by Crawford and Sobel (1982), all equilibria of this communication game are characterized by a partitioning of the signal space, where the sender only specifies which partition the true realized state of nature belongs to. In line with earlier

studies, we select as our focal equilibrium the most informative one—i.e., the one with the finest feasible partitioning—because it Pareto dominates all other equilibria.¹⁴

We compute the equilibrium using the Crawford and Sobel (1982) approach. Under centralization, if the board does not observe a , it will form its beliefs about a based on the report, r_a , sent by the CEO. The expected information loss is then given by

$$\sigma_A^2(b) = \frac{A^2}{12(N(A, b))^2} + \frac{b^2((N(A, b))^2 - 1)}{3},$$

where $N(A, b)$ is the maximum number of equilibrium partitions under centralization (see “Preliminaries” in the Appendix for details). The expected information loss $\sigma_P^2(b)$ under delegation is computed analogously.¹⁵

2.3 Board Composition and Allocation of Decision Rights

Before addressing the complete model, it is useful to decompose it as follows. We first investigate the optimal allocation of decision rights for a given board composition and then, conversely, the optimal board composition holding constant decision rights. We later

¹⁴The most informative equilibrium is also supported by the NITS (“No Incentive to Separate”) refinement proposed by Chen, Kartik, and Sobel (2008). Note however, as pointed out in Melumad and Shibano (1991), the most informative equilibrium may not be the Pareto-dominant one when the sender’s bias depends on his private information.

¹⁵It is readily verified that both $\sigma_A^2(\cdot)$ and $\sigma_P^2(\cdot)$ are increasing in b everywhere. Moreover, $\sigma_A^2(\cdot)$ and $\sigma_P^2(\cdot)$ are continuous in b , even at those points where $N(A, b)$ and $N(P, b)$ change in a discrete fashion (of course, $\sigma_A^2(\cdot)$ and $\sigma_P^2(\cdot)$ are not differentiable at those thresholds).

synthesize the insights from this preliminary analysis and solve for the optimal endogenous choice of both design variables.

2.3.1 Exogenous Board Composition

We begin by studying the effect of agency problems on the allocation of decision rights, taking as given for now the board composition, x . The profit differential between delegation and centralization as a function of board composition and CEO bias is:

$$\pi^D(x, b) - \pi^C(x, b) = [1 - q_a(x)] \underbrace{\{\sigma_A^2(b) - [b^2 + q_p(x)\sigma_P^2(b)]\}}_{Z(x,b)}. \quad (2.8)$$

Because of selective intervention in case of successful monitoring, only if monitoring fails to uncover a there will be a difference between the two regimes. This differential cost is expressed by the function $Z(\cdot)$ in (2.8). Specifically, under centralization the shareholders incur an information loss due to noisy reporting of a by the CEO; under delegation they incur a bias cost and an information loss due to the noisy communication of p (with probability q_p).

Lemma 3 *When the CEO's bias, b , is sufficiently small, delegation dominates centralization for any $x \in [0, 1]$. When b is sufficiently large, centralization dominates delegation for any $x \in [0, 1]$. For intermediate values of b , there exists a cutoff $\hat{x}(b) \in (0, 1)$ such that the shareholders prefer centralization for $x \leq \hat{x}(b)$, and delegation for $x > \hat{x}(b)$.*

Proof. All proofs are found in the Appendix.

Extending the results in Dessein (2002) and Harris and Raviv (2005), if the CEO's bias, b , is sufficiently small, $Z(x, b) > 0$ for all x ; the reverse holds for b sufficiently large. The novel insight of Lemma 3 revolves around intermediate values of b . For these values we show that monitor-heavy boards call for delegation, and advisor-heavy boards call for centralization. More precisely, the board is indifferent between centralization and delegation for some cutoff level of board composition, $\hat{x}(b) \in (0, 1)$, at which

$$\pi^D(\hat{x}(b), b) \equiv \pi^C(\hat{x}(b), b), \quad (2.9)$$

it prefers to centralize for $x \leq \hat{x}(b)$ and to delegate for $x > \hat{x}(b)$.¹⁶ The reason for this is that *advisor-heavy boards are more likely to discover p , and this information is more valuable under centralization*. To illustrate, suppose the board has discovered p . Under centralization, p will always be perfectly impounded in the board's choice of y_B . Under delegation, when monitoring fails to uncover a , y_{CEO} will only reflect the noisy report about p submitted by the board. Hence, delegation suffers from an information loss regarding p .¹⁷

¹⁶We assume without loss of generality that, whenever the board is indifferent between centralization and delegation, it will choose centralization.

¹⁷Note that Lemma 3 also applies to the special case where $A = P$. This contrasts with the results in Harris and Raviv (2005); in their model centralization is always preferred for $A = P$ because the board *always* observes p . Then for $A = P$, the information loss under centralization due to a noisy report of a is the same as that under delegation given the noisy report of p ; yet the board also incurs the bias cost under delegation. In our model, the board observes p with probability $q_p(x) < 1$, so the information loss due to

Closely related, Baker and Gompers's (2003) study venture capitalist representation on boards. They conclude that VCs mostly substitute for advisory board members, when mapped into our taxonomy. Consistent with Lemma 3, they find that boards with greater VC representation are more actively involved in decision-making.

2.3.2 Exogenous Allocation of Decision Rights

We now analyze the optimal board composition for given CEO bias, taking decision rights as given. Specifically, we aim to compare $x^C(b)$ and $x^D(b)$, where $x^S(b) \in \arg \max_x \pi^S(x, b)$, $S = C, D$. To that end, it is helpful to assess the values to the shareholders of learning a and p , respectively, under centralization and delegation. As argued above, the value of learning p is greater under centralization, which calls for more monitoring under delegation. Now consider the value of learning a under the respective modes. The benefit from learning a is the avoided information loss $\sigma_A^2(b)$ under centralization, whereas under delegation it comprises both the avoided bias cost b^2 and the avoided information loss $\sigma_P^2(b)$ (with probability q_p). (Recall that delegation is subject to selective intervention contingent on the board learning a .) For high values of b the benefit from learning a is greater under delegation (i.e., $Z(\cdot) < 0$ in (2.8)). This reinforces the above value-of-learning- p effect in that it also calls for more monitoring under delegation. Formally:

Lemma 4 *For b sufficiently high, $x^D(b) > x^C(b)$.*

noisy communication is lower under delegation than under centralization. Thus, in our setting, delegation is preferred for b sufficiently small even when $A = P$.

Now consider the case where b and P are small relative to A . In that case $Z(\cdot) > 0$ in (2.8) and the benefit to the shareholders from learning a is higher under centralization. Since now the values of learning a and p each are greater under centralization, the net effect on the optimal board composition is not immediately obvious. To evaluate this tradeoff, additional structure on the technology is needed:¹⁸

Lemma 5 *Suppose $A = P$, $q_a(x) = \eta x^z$, $q_p(x) = \eta(1 - x)^z$, where $\eta \in (0, 1)$ and $z \in (0, 1)$. Then, for η sufficiently large, $x^D(b) > x^C(b)$ holds for any b .*

Given the additional structure, the value-of-learning- p effect is the dominant force, implying that less centralized firms nominate more monitors to the board for any level of CEO bias.

The primary benefit of board monitoring is to mitigate the inefficiencies arising from biased decision-making (under delegation) or noisy reporting (under centralization). Consequently, one would expect the optimal monitoring intensity to go up as b increases. The next result confirms this intuition, holding fixed again the allocation of decision rights.

Lemma 6 *$x^S(b)$ is non-decreasing in b , for given $S = C, D$.*

This result follows directly from the fact that π^S , $S = C, D$, has increasing differences in x and b , because the respective information losses with regard to a and p each are non-decreasing in b . Therefore, the more biased is the CEO, the more important it is for the

¹⁸The exponential technology in Lemma 5 is also used for the running example in some of the figures below.

board to uncover what the CEO knows, regardless of the allocation of decision rights. Conversely, if b is small, there is little value to monitoring; instead the board should devote all of its resources to advising.

In sum, our preliminary analysis has shown that greater CEO bias directly calls for more centralization and more monitoring. However, there is a confounding effect in that more centralization calls for more advising. To derive empirical predictions on the link between CEO bias and board composition, we now turn to the full-fledged model where all design choices are endogenous.

2.4 Equilibrium Choice of Board Composition and Decision Rights

We now consider the board's simultaneous choice of x and S as a function of the economic environment of the firm. While in practice the board nomination process often precedes the assigning of decision rights, it is without loss of generality in our base model to treat those two choices as simultaneous. With an eye on the extant empirical literature, which has focused on the severity of agency problems, we let the bias parameter b vary while keeping the uncertainty parameters (A, P) fixed. The sequence of events is summarized in Figure 2.1:

In this base scenario, the shareholders are in complete control of the firm in that they

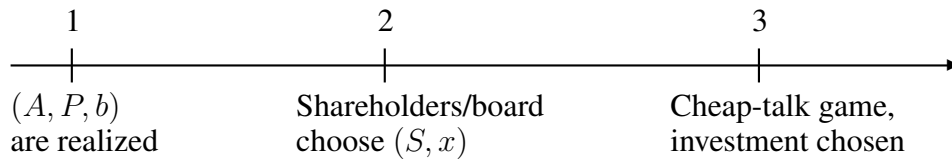


Figure 2.1: Timeline I—Shareholders/board choose x and S

choose both S and x . Since S is a discrete choice variable, the shareholders' problem can be decomposed as follows. First, for any allocation of decision rights, S , find the optimal board composition, $x^S(b)$, as characterized in Section 2.3.2. Then, pick $S = C$, if $\pi^C(x^C(b), b) > \pi^D(x^D(b), b)$, and $S = D$ otherwise. For notational convenience we introduce the indicator variable $\mathbb{1}_D \in \{0, 1\}$, where $\mathbb{1}_D = 1$ indicates delegation, and $\mathbb{1}_D = 0$ indicates centralization. The shareholders' optimization problem at Date 2 can then be stated as follows:

$$\mathcal{P}_0 : \max_{\mathbb{1}_D \in \{0, 1\}} (1 - \mathbb{1}_D)\pi^C(x^C(b), b) + \mathbb{1}_D\pi^D(x^D(b), b), \quad \text{for any } b.$$

Denote the solution to this program by $(S^*(b), x^*(b))$ where $x^*(b) = x^{S^*(b)}(b)$.

Our first main result characterizes the solution to the shareholders' optimization problem, \mathcal{P}_0 .

Proposition 4 (Timeline I)

(i) For sufficiently small levels of CEO bias, b , the shareholders choose delegation, i.e.,

$$S^*(b) = D, \text{ and board composition } x^*(b) = x^D(b).$$

(ii) For b sufficiently large, the shareholders choose centralization, i.e., $S^*(b) = C$, and

$$x^*(b) = x^C(b).$$

(iii) There exists at least one cutoff \hat{b} at which the optimal allocation of decision rights switches from delegation to centralization.

The fraction of monitoring types on the board, $x^*(b)$, is non-decreasing in b almost everywhere, except at such cutoffs, \hat{b} , where $\lim_{\varepsilon \rightarrow 0} x^*(\hat{b} - \varepsilon) = \lim_{\varepsilon \rightarrow 0} x^D(\hat{b} - \varepsilon) > \hat{x}(\hat{b}) > \lim_{\varepsilon \rightarrow 0} x^*(\hat{b} + \varepsilon) = \lim_{\varepsilon \rightarrow 0} x^C(\hat{b} + \varepsilon)$.

Consistent with earlier studies, the more biased is the CEO, the more likely the board will centralize the decision-making. The last part of the proposition is more surprising as it shows that the optimal board composition is non-monotonic in the CEO's bias, b . Starting from small levels of b , the firm optimally delegates authority to the CEO and mostly engages in advising the latter (low x). As b increases, by Lemma 6, x^D , too, will increase. When the CEO's bias reaches the threshold \hat{b} , defined by

$$\pi^D(x^D(\hat{b}), \hat{b}) \equiv \pi^C(x^C(\hat{b}), \hat{b}), \quad (2.10)$$

the board is indifferent as to who makes the decision and hence switches to centralization.¹⁹

¹⁹As suggested by Proposition 4, there may be more than one such threshold \hat{b} at which the optimal

??

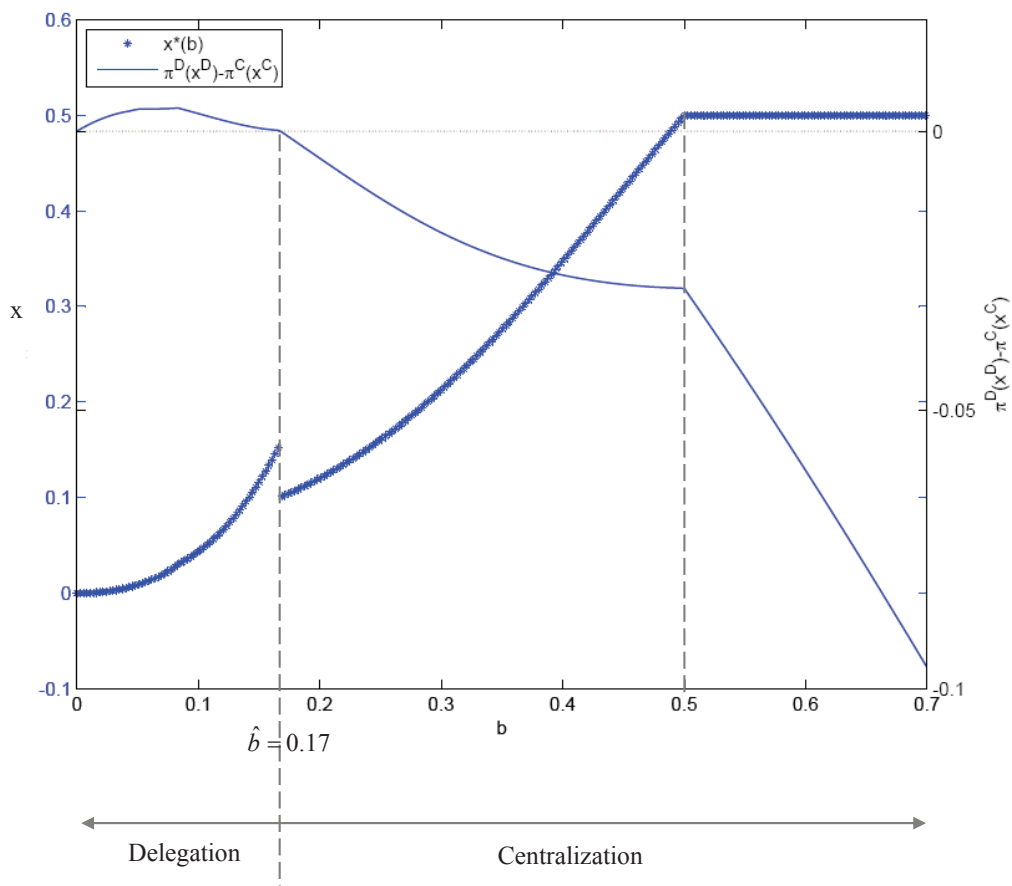


Figure 2.2: Board composition as a function of CEO bias under Timeline I

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Lemma 5. The parameter values are $A = P = 2$, $\eta = 0.8$, $z = 0.5$.

Centralization in turn creates greater demand for advisors because, as argued above, the value of learning p is higher than under delegation — thus the discrete downward jump in $x^*(b)$ at \hat{b} (see Figure 2.2).²⁰

Our result has empirical implications regarding the life cycle of firms. As firms mature, agency problems (as measured by b) tend to increase, and therefore one would expect to see more monitor-heavy boards (e.g., Lehn et al. 2004). However, if tests do not control for the allocation of decision rights, this general trend may be obscured in the data due to an omitted correlated variable problem. Consider two firms that operate in similar industries and use comparable technologies (i.e., have roughly similar A s and P s), but one is more mature and hence plagued by more severe agency problems than the other. The more mature firm may well have fewer monitors on the board as a result of it having centralized more decision-making authority at the board level.

2.5 CEO Power I: Influence Over Board Composition

Critics of corporate governance practice have argued that a key reason why boards exert insufficient control over management is that CEOs often have significant say over the board selection process. This issue was studied analytically by Hermalin and Weisbach (1998)

allocation of decision rights changes. This is due to the fact that $\pi^C(x^C(b), b)$ and $\pi^D(x^D(b), b)$ may intersect more than once, see also Harris and Raviv (2005). To abstract from such technicalities and focus on the main economic tensions, we will henceforth only consider scenarios where such a threshold is unique. The threshold can be shown to be unique for plausible restrictions on the primitives of the model.

²⁰Using the terminology developed in connection with Lemmas 4 and 5, the threshold value \hat{b} is large enough such that, when switching from delegation to centralization, the value-of-learning- p effect dominates the value-of-learning- a effect.

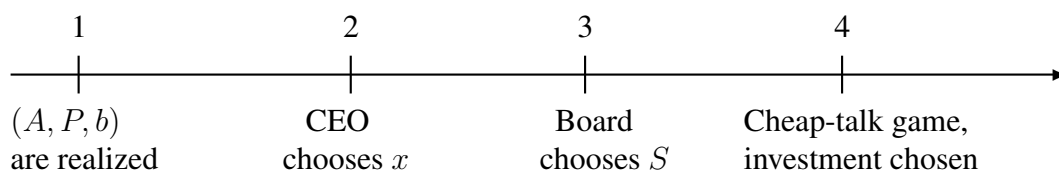


Figure 2.3: Timeline II—CEO influences board composition

(see Shivdasani and Yermack 1999 for empirical evidence). In their model the CEO's bargaining power over board composition evolves dynamically as a result of delivered performance. To address the issue of CEO power—and how it affects the benchmark model of the preceding section—we now consider a modified scenario where the CEO himself chooses the board composition, x , followed by the board allocating the decision rights, S . (Legally, of course, it is up to the shareholders to choose the board; however, a recurring theme in the practitioners' literature is that CEOs can *influence* the board composition—we simply push this notion to the extreme.) We maintain the earlier assumption that the board, even if selected by the CEO, acts in the shareholders' interest when choosing S .²¹ In the Concluding Remarks section, below, we speculate on the possible consequences of biased boards.

The question of interest is how the CEO's preferred board composition differs from that

²¹An implicit assumption is that there is a binding limit on the number of insiders on the board. In that case, a powerful CEO can only pick from outsiders whose objective is to maximize shareholder value. But the CEO may prefer a monitor/advisor mix that diverges from that preferred by the shareholders.

the shareholders would choose if they were in control. A key concern in the practitioners' literature is that a powerful CEO prefers less board monitoring. Holding constant the allocation of decision rights, Lemma 0 has confirmed this intuition. As we will show now, however, this intuition may not always hold if decision rights are assigned in a sequentially rational fashion. Instead, the CEO sometimes prefers *more* monitors on the board than the shareholders do.

The modified optimization program as in Timeline II, with the CEO choosing x at Date 2, reads:

$$\mathcal{P}_1 : \max_{x \in [0,1]} [1 - \mathbb{1}_D(x, b)]\pi_{CEO}^C(x, b) + \mathbb{1}_D(x, b)\pi_{CEO}^D(x, b), \quad \text{for any } b,$$

subject to:

$$\mathbb{1}_D(x, b) \in \arg \max_{\mathbb{1}_D \in \{0,1\}} (1 - \mathbb{1}_D)\pi^C(x, b) + \mathbb{1}_D\pi^D(x, b).$$

The CEO maximizes his expected payoff subject to the constraint that the board will subsequently allocate decision rights in a sequentially rational fashion. Denote the solution to this program by $x_{CEO}(b)$. Our next result compares $x_{CEO}(b)$ with the board composition, $x^*(b)$, selected by the shareholders in the benchmark model as of Proposition 4. (Recall that \hat{b} denotes the threshold bias value at which a shareholder-nominated board under Timeline I is indifferent between centralizing and delegating, as of (2.10).)

Proposition 5 (Timeline II) *Suppose the CEO chooses the board composition, x , before the board allocates the decision rights, S . Then:*

(i) *For $b < \hat{b}$, $x_{CEO}(b) < x^*(b)$ and $S(x_{CEO}(b), b) = S^*(b) = D$.*

(ii) *For b sufficiently large, $x_{CEO}(b) = x^*(b)$ and $S(x_{CEO}(b), b) = S^*(b) = C$.*

(iii) *For $b = \hat{b} + \varepsilon$ with $\varepsilon \rightarrow 0$, $x_{CEO}(b) > x^*(b)$ and $S(x_{CEO}(b), b) = D$ whereas $S^*(b) = C$.*

The key to understanding the first part of the result is Lemma 0. On the one hand, that earlier result has shown that whenever the board prefers delegation in the benchmark model of Section 2.4, then *a fortiori* so does the CEO. (The CEO could simply set $x_{CEO}(b) = x^D(b)$ to induce delegation and act according to his bias.²²) Secondly, Lemma 0 has shown that, under delegation, the CEO prefers a *lower* level of monitoring than the shareholders, as that would reduce $q_a(x)$, the probability of selective intervention. For $b < \hat{b}$, delegation is preferred even by the board in the benchmark model; thus the CEO has no incentive to influence the board's decision and simply sets $x_{CEO}(b) < x^*(b) = x^D(b)$. On the other hand, for very severe agency problems (part (ii)) the board chooses centralization even if composed entirely of monitors.²³ Therefore, the CEO sets $x_{CEO}(b) = x^*(b) = x^C(b)$

²²As we show in the Appendix, for $b < \hat{b}$, $x^D(b) > \hat{x}(b)$. Hence, if the CEO were to (suboptimally) choose $x^D(b)$, the board would delegate the decision, as desired by the CEO.

²³Technically, the critical value $\hat{x}(b)$ at which the board is indifferent to delegate the investment decision would exceed one.

because, conditional on centralization, Lemma 0 has shown that the players' preferences over board composition coincide.

The main insight of Proposition 5 is that, in certain cases, the CEO prefers *excessive* board monitoring (part (iii) of Proposition 5; see Figure 2.4). When the optimal solution in the benchmark model of Program \mathcal{P}_0 is such that the shareholders marginally prefer centralization over delegation (i.e., b exceeds but is close to \hat{b}), then the CEO has incentives to influence the board toward delegating the decision by nominating more monitors to the board. Put succinctly, the CEO prefers delegation — even if subject to the threat of selective intervention by a vigilant board — rather than having control taken away by an advisor-heavy (activist) board.

The past decade has seen a number of regulatory attempts at strengthening the monitoring role of boards, e.g., the Sarbanes-Oxley Act of 2002 (SOX) with its focus on the audit committee (see SOX, Section 301). In the context of our model, we interpret such a regulation as a rule stipulating a lower bound on x , i.e., $x \geq \underline{x}$.²⁴ Clearly, as long as shareholders determine the board composition, any regulatory intervention of this sort will necessarily (weakly) lower the shareholders' payoff.²⁵ However, with a powerful CEO, a welfare-increasing role for regulation may arise because, as Proposition 5 has shown, the CEO's choice of x tends to diverge from that preferred by the shareholders.

²⁴One of the main objectives of SOX is to increase board independence. As explained in the Introduction, there is no clear mapping between independence and our advisor/monitor dichotomy. However, Linck et al. (2009) and Faleye et al. (2010) document a significant shift in favor of monitoring types post-SOX. Section 407 of SOX specifically requires that the audit committee includes at least one qualified financial expert.

²⁵This may not be true if the shareholders are misinformed as in Harris and Raviv (2010).

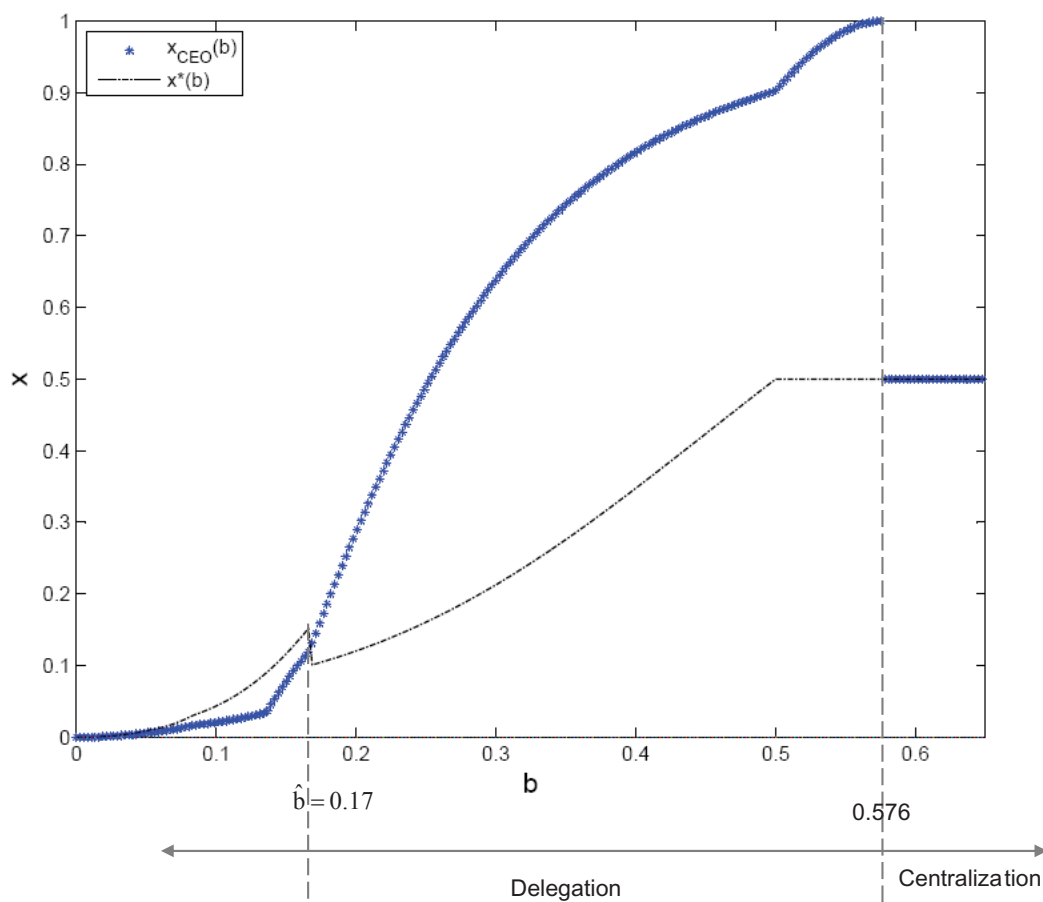


Figure 2.4: Board composition as a function of CEO bias under Timeline II

The dashed line is the optimal board composition under the benchmark model of Timeline I. For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Lemma 5. The parameter values are the same as in Figure 2.2, i.e., $A = P = 2$, $\eta = 0.8$, $z = 0.5$.

Corollary 3 *Suppose the CEO chooses the board composition as in Timeline II. Imposing a constraint that $x \geq \underline{x}$ (weakly) lowers shareholder value if and only if $b > \hat{b}$, and strictly so if this constraint is binding in that $x_{CEO}(b) < \underline{x}$.*

A necessary condition for such regulations to increase social welfare is that agency problems must be sufficiently small, i.e., $b < \hat{b}$. Only in that case the conventional wisdom applies that the CEO prefers a “monitor-light” boards, so that constraining his choice of x from below may indeed improve shareholder value. However, for greater biases, imposing such a constraint can only result in a board composition that departs even further from the benchmark solution. Hence, somewhat ironically, regulations intended to mitigate incentive costs arising from CEO bias and power may be harmful in precisely those circumstances where agency problems are severe.

Stepping outside the scope of our model, Corollary 3 is consistent with the recent trend of increasing board sizes after SOX (Linck et al 2008, 2009). If firms have composed their boards optimally pre-SOX, and then face an exogenous shock forcing them to add more monitors to the board, one would in fact expect to see board size to grow as firms aim to re-balance the board composition.

2.6 CEO Power II: Entrenchment

A somewhat more subtle way in which managers can exert power and obtain *de facto* authority is by choosing investment projects that confer them a greater degree of private

information (Shleifer and Vishny 1989, Edlin and Stiglitz 1995). In this section, we build on the discussion in Harris and Raviv (2005) on “strategic choice of information advantage” and embed it in our model with endogenous board composition. We return to our initial assumption that it is the shareholders, not the CEO, who control the board nomination process.

Suppose the CEO can choose the level of “complexity” of the project endogenously. Specifically, a remains uniformly distributed over $[0, A]$, but we now assume the CEO can freely choose the upper bound A within some range $[\underline{A}, \bar{A}]$. This choice is assumed to be observable but not contractible.²⁶ In the following, for given “default complexity” \underline{A} , we interpret \bar{A} as a measure of the CEO’s *discretion*. While we assume that increasing project complexity comes at no direct cost to the CEO, our model can easily be extended to include such costs.

We first address the case where the CEO chooses the level of complexity *before* the shareholders assemble the board and allocate decision rights (Section 2.6.1). In this setting, A corresponds to a long-term strategy choice whose consequences outlast the tenure of a typical director. Alternatively, it may describe a situation where the CEO of a company opts for a certain technology before a new (non-staggered) board is assembled. In Section 2.6.2, we consider a more short-term notion of project complexity. In that case, the share-

²⁶To illustrate the observability assumption, consider the following example (*Business Week*, January 25, 2006, “Daimler Shakeup: Realignment in the Auto Industry”). Jurgen Schrempp, former CEO of DaimlerChrysler, pursued a strategy of growth and diversification. His successor, Dieter Zetsche, reverted to the traditional strategy of a more focused car maker. In the context of our model, Schrempp’s strategy can be interpreted as a “high- A ” strategy, and that of Zetsche as a “low- A ” strategy.

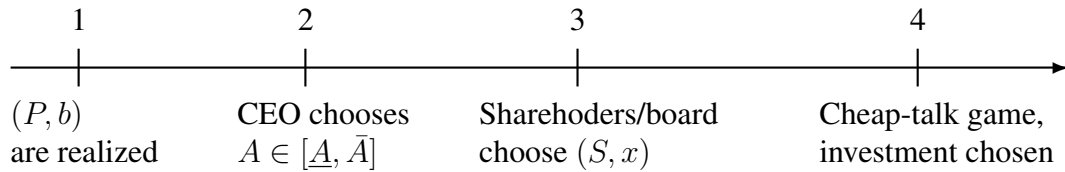


Figure 2.5: Timeline III—CEO chooses A endogenously

holders can assemble the board *before* the CEO’s choice of A . The board composition then potentially plays a strategic role in deterring CEO entrenchment.

2.6.1 Endogenous Project Complexity

Consider the revised sequence of events as depicted in Timeline III.

When allocating decision rights at Date 3, the board trades off the information loss regarding a under centralization with the information loss regarding p *plus* the bias cost under delegation, as in (2.8). As the information advantage of the CEO becomes larger (i.e., as A increases), the board’s willingness to delegate to the CEO increases, as noted by Harris and Raviv (2005). Clearly, the CEO will want to “noise up” the project only if absent such action the board would choose to centralize, i.e., if $b \geq \hat{b}(\underline{A})$. In this case, there exists a threshold $\hat{A}(b)$ given by

$$\pi^C(x^C(\hat{A}(b), b), \hat{A}(b), b) \equiv \pi^D(x^D(b), b), \quad (2.11)$$

such that the board prefers to delegate the investment decision for any $A > \hat{A}(b)$, and to centralize otherwise. Under delegation, neither party's payoff depends on the project complexity A , because the realization of a will always be impounded without noise in the investment decision. In other words, a more complex project (a higher A) adversely affects the payoff of either party only under centralization.

If by choosing $A > \hat{A}(b)$ the CEO succeeds in influencing the board's decision towards delegation, no additional information loss regarding a will be incurred by either party. On the other hand, if the equilibrium path is such that the investment will be made by the board at Date 3, then the CEO's and the board's objectives regarding project complexity are aligned in that both prefer the least complex project, \underline{A} , at Date 2. We assume without loss of generality that, whenever the CEO is indifferent among various levels of A , he will choose the lowest one.²⁷

The following observation describes the circumstances under which CEO entrenchment will occur:

Observation 1 *[Timeline III] Suppose the CEO chooses project complexity A before the shareholders/board choose S and x . Then:*

- (i) *For any $b < \hat{b}(\underline{A})$, the CEO chooses $A = \underline{A}$, followed by the board delegating the investment decision and setting $x = x^D(b)$.*

²⁷This would be the equilibrium choice in a more general model where (i) the CEO observes A with probability less than one or (ii) the CEO incurs personal costs of increasing the project complexity.

- (ii) For any $b \in [\hat{b}(\underline{A}), b^*)$, where $b^* \leq \hat{b}(\bar{A})$, the CEO chooses $A = \hat{A}(b) + \varepsilon$, $\varepsilon \rightarrow 0$, followed by the board delegating the investment decision and setting $x = x^D(b)$.
- (iii) For any $b \geq \hat{b}(\bar{A})$, the CEO chooses $A = \underline{A}$, followed by the board centralizing the investment decision and setting $x = x^C(\underline{A}, b)$.

For negligible agency problems (part (i)), there is no need for the CEO to entrench himself, as the board is willing to delegate even for the default complexity, \underline{A} . For intermediate levels of agency conflicts (part (ii)), it is both feasible and profitable for the CEO to entrench himself.²⁸ However, for sufficiently severe agency problems (part (iii)), the board is so strongly inclined to centralize that the level of complexity required to trigger delegation exceeds the maximum complexity. Entrenchment then is not feasible. Instead, the CEO will choose the least complex project, \underline{A} , as that will maximize his own payoff conditional on centralization (see Figure 2.6 for illustration).

As argued above, if delegation obtains along the equilibrium path, then increased project complexity is without informational costs to all players. However, if an investment decision is delegated which absent entrenchment would have been centralized, the shareholders incur additional bias costs. We close our analysis by asking, if and how shareholders can

²⁸To determine whether entrenchment is optimal for the CEO, provided it is feasible, define $g(b) \equiv \pi_{CEO}^D(x^D(b), b) - \pi_{CEO}^C(x^C(\underline{A}, b), \underline{A}, b)$, and b^{**} such that $g(b^{**}) \equiv 0$. It is easy to show that if b^{**} exists, then $b^{**} > \hat{b}(\underline{A})$. The reason is that for any $b \leq \hat{b}(\underline{A})$, the board (weakly) prefers delegation, which by (2.7) is then also the preferred choice for the CEO. Since $g(b)$ is continuous and positive for $b \leq \hat{b}(\underline{A})$ and for $b \rightarrow \infty$, it follows that either no b^{**} exists or an even number of b^{**} exist. The cutoff in part (ii) of Observation 1 follows as $b^* = \min\{\hat{b}(\bar{A}), \min b^{**}\}$.

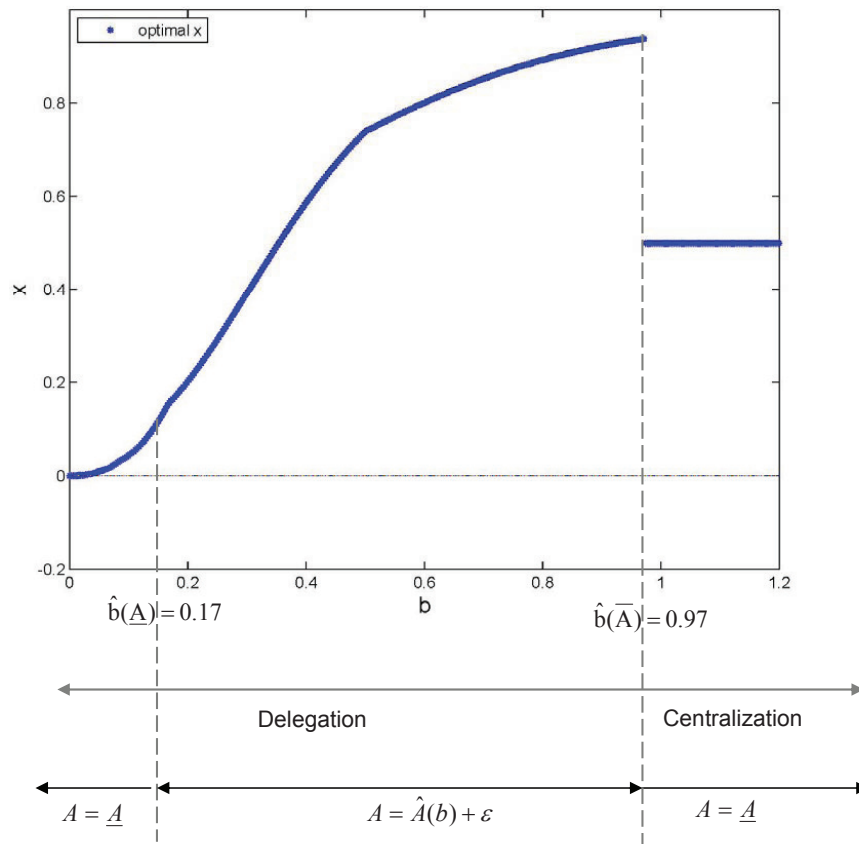


Figure 2.6: Board composition as a function of CEO bias under Timeline III

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Lemma 5. The parameter values are $\underline{A} = P = 2$, $\bar{A} = 3.5$, $\eta = 0.8$, $z = 0.5$.

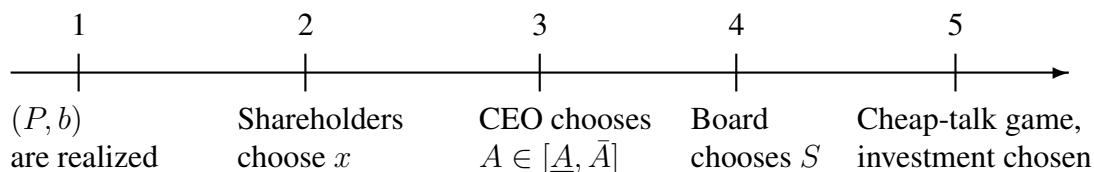


Figure 2.7: Shareholders may preempt CEO entrenchment

protect themselves against entrenchment by assembling the board in a strategic fashion at the outset.

2.6.2 Preempting CEO Entrenchment

Board composition in practice is fairly sticky in that board members serve for multiple years. Hence, when considering an entrenchment strategy, a CEO often has to take the existing board as given. Shareholders, in turn, when nominating the board, need to be cognizant of the threat of entrenchment. We now consider the sequence of events depicted in Timeline IV where the shareholders first assemble the board, followed by the CEO choosing the project complexity, and finally, the board deciding whether to delegate.

The board composition now takes on a strategic role in that it may influence the CEO's subsequent choice of project complexity. Put differently, the shareholders may want to deviate from the board composition that optimally balances the advisory and monitoring roles (as described in Proposition 4) so as to preempt CEO entrenchment.

To ensure that entrenchment is a meaningful threat, the following assumption is maintained throughout this section:

Assumption 1 $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$.

This assumption postulates that absent CEO entrenchment the board would centralize the decision (because $b \geq \hat{b}(\underline{A})$). At the same time, it is feasible for the CEO to entrench himself (because $b \leq \hat{b}(\bar{A})$ is equivalent to $\hat{A}(b) \leq \bar{A}$).²⁹

The shareholders may now be able to preempt CEO entrenchment by nominating an advisor-heavy board (i.e., a low level of x) at Date 2 as a way to commit to centralization at Date 4 in Timeline IV. Recall from Lemma 3 that advisor-heavy boards and centralization are complements. Denote by $\tilde{A}(x, b)$ the level of project complexity that leaves the board indifferent between centralizing and delegating for given levels of bias and board composition (Timeline IV):

$$\pi^C(x, \tilde{A}(x, b), b) \equiv \pi^D(x, b). \quad (2.12)$$

Using the same logic as in Section 2.6.1, $\tilde{A}(x, b)$ is decreasing in x : the more advisor-heavy the board, the less it is willing to delegate — a greater level of project complexity is thus required for the CEO to gain control. By setting x below $\hat{x}(\bar{A}, b)$ — generalizing

²⁹Assumption 1 postulates that, relative to the admissible range of bias parameters b , the CEO's discretion as measured by \bar{A} must be sufficiently large. Lemma 7, below, invokes this assumption, yet considers the case of “small” CEO discretion. For this case to be compatible with Assumption 1, the range of b has to be correspondingly small.

the definition in (2.9) to allow for A to vary³⁰ — the shareholders can credibly commit the board to centralize *for any* feasible level of project complexity. Even if the CEO were to choose the most complex project, \bar{A} , a sufficiently advisor-heavy board would value new information enough so as to ensure centralization. The CEO, anticipating centralization along the equilibrium path, then will refrain from entrenching himself.

The shareholders' decision problem at Date 2 essentially boils down to a discrete profit comparison between two alternative courses of action:

- (i) to *preempt* CEO entrenchment by setting $x \leq \hat{x}(\bar{A}, b)$, or
- (ii) to *acquiesce* to CEO entrenchment and adapt x optimally to the fact that, along the equilibrium path, the investment decision will be delegated.

More formally, this discrete choice for the shareholders can be expressed by comparing the values of the following two optimization programs:

\mathcal{P}_C : (*Preempt CEO entrenchment*)

$$\begin{aligned} & \max_{x \in [0,1]} \pi^C(x, \underline{A}, b), \\ & \text{subject to: } \tilde{A}(x, b) \geq \bar{A} \quad (\text{equivalently, } x \leq \hat{x}(\bar{A}, b)). \end{aligned} \tag{C}$$

³⁰In Section 2.3.1 we suppressed the functional dependence of $\hat{x}(\cdot)$ on A and P . Since we now allow for A to vary, we explicitly write $\hat{x}(\cdot)$ as a function of A and b .

\mathcal{P}_D : (Acquiesce to CEO entrenchment)

$$\begin{aligned} & \max_{x \in [0,1]} \pi^D(x, b), \\ & \text{subject to: } \tilde{A}(x, b) < \bar{A} \quad (\text{equivalently, } x > \hat{x}(\bar{A}, b)). \end{aligned} \quad (\text{D})$$

If entrenchment is preempted in equilibrium (Program \mathcal{P}_C), the CEO chooses the default complexity level \underline{A} at Date 3, because greater complexity would merely add to the informational loss without yielding any offsetting benefits.

By definition of $\hat{b}(\underline{A})$, the board is indifferent between centralization and delegation for $b = \hat{b}(\underline{A})$ and strictly prefers centralization for any $b > \hat{b}(\underline{A})$, given $A = \underline{A}$. Thus, if both constraints (C) and (D) in the respective optimization programs were slack, the analysis would revert to the benchmark setting of Timeline I, above. In that case, for any $b \geq \hat{b}(\underline{A})$, the value of \mathcal{P}_C would exceed that of \mathcal{P}_D ; hence the board would prefer centralization given Assumption 1. As a result, a necessary condition for delegation to obtain in equilibrium is that constraint (C) has to be binding. More generally, the shadow prices of the constraints (C) and (D) indicate whether there is a cost to the shareholders associated with the threat of CEO entrenchment under the respective organizational modes.

Lemma 7 *Suppose Assumption 1 holds. Then:*

(i) *Constraint (D) in program \mathcal{P}_D is always slack.*

(ii) *For \bar{A} sufficiently small, constraint (C) in program \mathcal{P}_C is always slack. For \bar{A} suf-*

ficiently large, constraint (C) is binding at $b = \hat{b}(\underline{A})$ and there exists at least one $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$ such that (C) holds with equality at b^C , i.e., $x^C(\underline{A}, b^C) = \hat{x}(\bar{A}, b^C)$.

To illustrate part (i), suppose the shareholders, anticipating delegation along the equilibrium path, choose the benchmark board composition $x^D(b)$, as of Timeline I. Then, by Assumption 1, the CEO can always choose a level of project complexity that induces the board at Date 4 to delegate conditional on $x^D(b)$. Hence, constraint (D) in the “Acquiesce” program \mathcal{P}_D is always slack and the shareholders’ Date-1 choice of $x^D(b)$ then would indeed be optimal. This follows naturally from the fact that the CEO values being in control.

Now consider the “Preempt” program \mathcal{P}_C (part (ii) of Lemma 7). Suppose the shareholders anticipate that the decision will be centralized at Date 4. Absent the threat of entrenchment, they would then prefer the board composition $x^C(\underline{A}, b)$. The question is whether this board composition is sufficiently advisor-heavy to preempt entrenchment. If the CEO has limited discretion in that $\bar{A} - \underline{A}$ is sufficiently small, the answer is yes for any admissible level of b satisfying Assumption 1. However, if the CEO has significant discretion, then the answer will depend on his bias parameter. If b is small, the CEO can implement a project sufficiently complex so as to violate constraint (C) if the shareholders were to choose $x^C(\underline{A}, b)$. This follows from the fact that high discretion and low bias both predispose the board towards delegating. For high b , on the other hand, the board prefers to centralize at Date 4, even if the CEO were to choose \bar{A} . In that case, constraint (C) will be slack at $x^C(\underline{A}, b)$.

In summary, CEO entrenchment is never a credible threat if the CEO has little discretion, i.e., if \bar{A} is small. This implies the following:

Observation 2 (Timeline IV) [Preempt at no cost] *Suppose Assumption 1 holds and the shareholders choose x before the CEO chooses $A \in [\underline{A}, \bar{A}]$. For \bar{A} sufficiently small, the shareholders choose $x = x^C(\underline{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.*

On the other hand, if the CEO enjoys significant discretion, by Lemma 7, the shareholders face the following tradeoff: Absent CEO entrenchment they would prefer centralization; yet centralization may be subject to entrenchment—i.e., a positive shadow cost of constraint (C)—whereas delegation is immune to entrenchment. We now evaluate this tradeoff by asking whether it is optimal (or even feasible) for the shareholders to preempt entrenchment.³¹ (Recall that constraint (C) in \mathcal{P}_C is binding for $b < b^C$.)

Proposition 6 (Timeline IV) *Suppose Assumption 1 holds and the shareholders choose x before the CEO chooses $A \in [\underline{A}, \bar{A}]$. Then, for \bar{A} sufficiently large:*

(i) **[Acquiesce]** *If $b = \hat{b}(\underline{A}) + \delta$, with δ “small”, then the shareholders choose $x = x^D(b)$ at Date 2, the CEO chooses $A = \tilde{A}(x^D(b), b) + \varepsilon > \underline{A}$ at Date 3, and the board delegates the decision at Date 4.*

(ii) **[Preempt at a cost]** *If $b = b^C - \delta$, with δ “small”, then the shareholders choose*

³¹If there exist multiple thresholds $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$, as defined in Lemma 7, we will focus on the smallest one.

$x = \hat{x}(\bar{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.

(iii) **[Preempt at no cost]** If $b = b^C + \delta$, with δ “small”, then the shareholders choose $x = x^C(\underline{A}, b)$ at Date 2, the CEO chooses $A = \underline{A}$ at Date 3, and the board centralizes the decision at Date 4.

Suppose the CEO has significant discretion in his project choice. Then, to commit to centralization (if feasible), the shareholders would have to deviate from the benchmark board composition, $x^C(\underline{A}, b)$. If the shareholders are only marginally predisposed to centralizing absent CEO entrenchment (i.e., for $b = \hat{b}(\underline{A}) + \delta$), it is not worthwhile for them to do so and bear the attendant shadow price of constraint (C). Instead, they acquiesce to the CEO’s entrenchment threat by choosing $x = x^D(b)$.

However, as the CEO’s bias increases further, then in the benchmark setting of Timeline I (without the threat of CEO entrenchment) the shareholders’ preference for centralization becomes more pronounced. At the same time, as the bias approaches b^C , the shadow price of constraint (C) under centralization vanishes. The shareholders then are better off preempting entrenchment by deviating from $x^C(\cdot)$. By strategically adding more advisors to the board, the shareholders effectively commit to centralization. This deviation from the benchmark case of Timeline I is less costly than the additional bias cost that would result from acquiescing to CEO entrenchment (as under Timeline III).

When the CEO’s bias becomes sufficiently high (exceeds b^C), then the board is so pre-

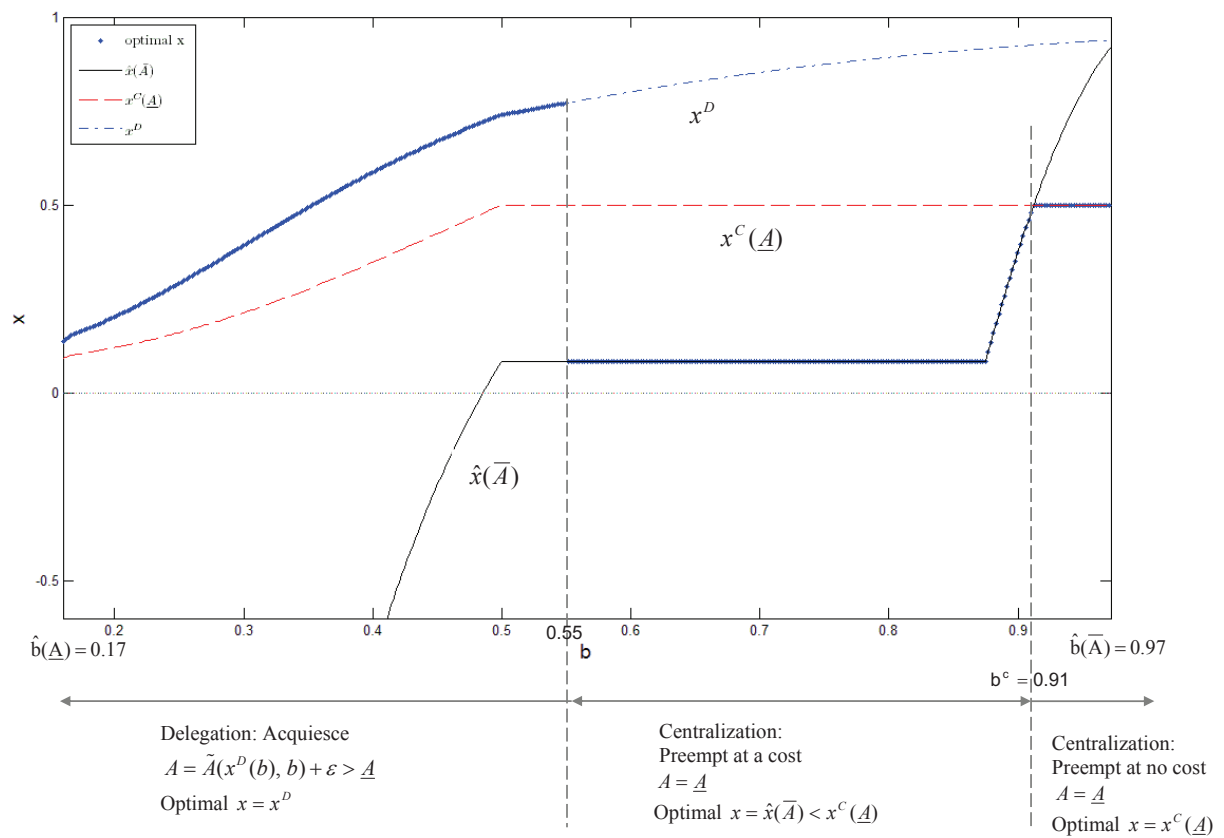


Figure 2.8: Board composition as a function of CEO bias under Timeline IV

For this figure, $q_a(x)$ and $q_p(x)$ take the exponential function form as in Lemma 5. The parameter values are the same as in Figure 2.6, i.e., $\underline{A} = P = 2$, $\bar{A} = 3.5$, $\eta = 0.8$, $z = 0.5$.

disposed to centralize that constraint (C) becomes slack. In other words, even if the CEO were to choose the maximum project complexity \bar{A} , the board would still find centralization the sequentially optimal choice at Date 4 given its prior choice of $x^C(\underline{A}, b)$. The unconstrained optimal board composition, $x^C(\cdot)$, then constitutes a credible commitment to centralize. As a result, the shareholders can preempt entrenchment at no cost (see Figure 2.8 for illustration).

As argued in connection with Corollary 3, regulatory interventions bounding x from below can in fact make shareholders worse off by exacerbating distortions in board composition if the board is effectively nominated by the CEO. Proposition 6 reveals another potential downside of such regulations in settings where project complexity is endogenous: they may jeopardize shareholders' ability to preempt CEO entrenchment.

Taken together, the preceding results show that caution is in order when interpreting observed patterns of board composition. A given board composition can be consistent with different severity levels of agency problems, depending on the sequence of events — specifically, whether there are entrenchment options. For instance, an advisor-heavy board can indicate negligible agency problems if CEO entrenchment is not an issue (Propositions 4 and 5), or severe agency problems if shareholders need to actively preempt entrenchment by a highly biased CEO (Proposition 6).

2.7 Concluding Remarks

Shareholder activists, regulators, and researchers have been increasingly interested in, and concerned about, CEO power. This paper examines two aspects of CEO power. First, if a powerful CEO controls the board nomination process, we find, somewhat surprisingly, that the CEO may in fact appoint a board that is excessively focused on monitoring. Such a board is less likely to generate incrementally new information, which in turn implies more delegation. Second, CEOs may entrench themselves by strategically exacerbating project complexity. Here, we show that when shareholders appoint the board, they sometimes choose a more advisor-heavy board as compared to that chosen in the absence of entrenchment. In both cases, the board composition serves a strategic purpose, over and above trading off the monitoring and advising functions, in that it commits other players to certain actions desired by the party in control of the nomination process.

Our model has explored various sequences of actions chosen by shareholders/boards and by the CEO, as well as different assumptions regarding the extent to which the parties can precommit to their actions. There appears to be anecdotal support for all of those timelines. It would be desirable if future empirical work could shed more light on the balance of power among the various actors involved in corporate governance.

Throughout the paper we assume that board members internalize the shareholders' preferences. Despite strict fiduciary duty laws, board members' preferences may diverge from those of shareholders, or even differ amongst themselves. An interesting extension

would be to incorporate divergent preferences into our model. Harris and Raviv (2010) and Chakraborty and Yilmaz (2009) show that shareholders may sometimes be better off choosing board members whose preferences over the scale of investment are intermediate between their own and those of the CEO. Doing so improves communication between CEO and the board. It is an interesting avenue for future research to investigate how the desirability of biased boards interacts with the board composition along the monitor/advisor dimension that lies at the heart of our paper.

Appendix A

Proofs for Chapter 1

A.1 Formal definition of the equilibrium

I begin by introducing the analyst's utility at each decision node. To that end, denote by $U_t^J(m_t, s_t, \lambda_t, \gamma^J)$ the type J analyst's utility at period $t = 1, 2$ when she reports m_t and her signal is s_t , reputation is λ_t and precision is γ^J . Specifically, $\lambda_1 = \lambda$ and $\lambda_2 = \Lambda(m_1, w_1)$. Analyst J 's utility at the information acquisition stage is denoted by $U_0^J(\gamma^J)$.¹

Given the investor's second period optimal decision rule $a_2(m_2, \Lambda)$, type J analyst's

¹To save on notation, I suppress the functional dependence of the players' utilities and strategies on their conjectures about their counterparties' actions.

second period utility is:

$$U_2^A(m_2, s_2, \Lambda, \gamma^A) = -\gamma^A(a_2(m_2, \Lambda) - s_2)^2 - (1 - \gamma^A)[a_2(m_2, \Lambda) - (1 - s_2)]^2,$$

$$U_2^M(m_2, s_2, \Lambda, \gamma^M) = -(a_2(m_2, \Lambda) - 1)^2.$$

Denote by $\sigma_t^J(s_t, \gamma^J, \lambda_t)$ the type J analyst's optimal communication strategy at period t upon observing signal s_t when her precision is γ^J and reputation is λ_t . Then,

$$\sigma_2^J(s_2, \gamma^J, \Lambda) \in \operatorname{argmax}_{\sigma_2^J \in [0,1]} \sigma_2^J U_2^J(1, s_2, \Lambda, \gamma^J) + (1 - \sigma_2^J) U_2^J(0, s_2, \Lambda, \gamma^J).$$

Hence, type J analyst's second period expected utility (anticipating that she will communicate optimally) is²

$$V^J(\Lambda, \gamma^J) = \frac{1}{2} \sum_{s_2=0,1} \{ \sigma_2^J(s_2, \gamma^J, \Lambda) U_2^J(1, s_2, \Lambda, \gamma^J) + (1 - \sigma_2^J(s_2, \gamma^J, \Lambda)) U_2^J(0, s_2, \Lambda, \gamma^J) \}.$$

Given the investor's first period optimal decision rule $a_1(m_1, \lambda)$, type J analyst's 1st

²Again, here I suppress the dependence of $V^J(\cdot)$ on $\tilde{\gamma}$.

period utility is then given by:

$$\begin{aligned}
U_1^A(m_1, s_1, \lambda, \gamma^A) &= x^A \{-\gamma^A (a_1(m_1, \lambda) - s_1)^2 - (1 - \gamma^A) [a_1(m_1, \lambda) - (1 - s_1)]^2\} \\
&\quad + (1 - x^A) \{\gamma^A V^A(\Lambda(m_1, s_1), \gamma^A) + (1 - \gamma^A) V^A(\Lambda(m_1, 1 - s_1), \gamma^A)\}, \\
U_1^M(m_1, s_1, \lambda, \gamma^M) &= -x^M (a_1(m_1, \lambda) - 1)^2 \\
&\quad + (1 - x^M) \{\gamma^M V^M(\Lambda(m_1, s_1), \gamma^M) + (1 - \gamma^M) V^M(\Lambda(m_1, 1 - s_1), \gamma^M)\}.
\end{aligned}$$

Then,

$$\sigma_1^J(s_1, \gamma^J, \lambda) \in \operatorname{argmax}_{\sigma_1^J \in [0,1]} \sigma_1^J U_1^J(1, s_1, \lambda, \gamma^J) + (1 - \sigma_1^J) U_1^J(0, s_1, \lambda, \gamma^J).$$

Thus, type J analyst's utility at the information acquisition stage is given by

$$U_0^J(\gamma^J) = \frac{1}{2} \sum_{s_1=0,1} \{\sigma_1^J(s_1, \gamma^J, \lambda) U_1^J(1, s_1, \lambda, \gamma^J) + (1 - \sigma_1^J(s_1, \gamma^J, \lambda)) U_1^J(0, s_1, \lambda, \gamma^J)\} - C(\gamma^J).$$

Now I am in a position to define the equilibrium of the game formally.

Definition 3 *A Perfect Bayesian Nash equilibrium of the game is a strategy-belief profile*

$(\gamma^{A*}, \gamma^{M*}, \sigma_t^A(\cdot), \sigma_t^M(\cdot), a_t(\cdot), \Gamma_t(\cdot), \Lambda(\cdot))$ *satisfying the following properties:*

(1)

$$\sigma_t^J(s_t, \gamma^{J*}, \lambda_t) \in \operatorname{argmax}_{\sigma_t^J \in [0,1]} \sigma_t^J U_t^J(1, s_t, \lambda_t, \gamma^{J*}) + (1 - \sigma_t^J) U_t^J(0, s_t, \lambda_t, \gamma^{J*}).$$

(2)

$$a_t(m_t, \lambda_t) \in \operatorname{argmax}_{a_t \in R} -\Gamma_t(m_t, \lambda_t)(a_t - 1)^2 - (1 - \Gamma_t(m_t, \lambda_t))a_t^2.$$

(3)

$$\gamma^{J^*} \in \operatorname{argmax}_{\gamma^J \in \{\gamma_L, \gamma_H\}} U_0^J(\gamma^J).$$

(4) *The state and type inference functions, $\Gamma_1(m_1, \lambda)$, $\Gamma_2(m_2, \Lambda)$ and $\Lambda(\cdot)$, are derived from the analyst's equilibrium strategy according to inference rules (1.1) and (1.2).*

Specifically,

$$\phi_t^J(1|w_t) = \gamma^{J^*} \sigma_t^J(w_t, \gamma^{J^*}, \lambda_t) + (1 - \gamma^{J^*}) \sigma_t^J(1 - w_t, \gamma^{J^*}, \lambda_t).$$

When the posteriors are undefined according to Bayes rule, I adopt the convention that the investor sticks to his priors.

The analysis in Section 1.3.1 shows that the analyst's second period optimal reporting strategy is independent of her reputation λ_2 . Furthermore, for the first period, $\lambda_1 (= \lambda)$ is a constant, hence I drop the argument λ_t in $\sigma_t^J(\cdot)$ in the main text.

A.2 Properties of informative communication in the first period

For later use, I first extend the insights from Morris (2001) and characterize the properties of first period informative communication for exogenous analyst's precision $\gamma = (\gamma^A, \gamma^M)$.

Let me introduce the notations under exogenous and commonly known $\gamma = (\gamma^A, \gamma^M)$: $\sigma_t^J(s_t|\gamma)$ is the probability of type J analyst reporting 1 in period t when her signal is s_t ; $a_t(m_t, \lambda_t, \gamma)$ is the investor's action in period t when he receives message m_t and his belief of analyst being aligned is λ_t ; $\Gamma_t(m_t, \lambda_t, \gamma)$ states the investor's inference of the actual state being 1 in period t ; $\Lambda(m_1, w_1|\gamma)$ is the investor's belief of the analyst being aligned if message m_1 is received and state w_1 is realized. $V^J(\Lambda(\cdot), \gamma^A)$ is the type J analyst's second period expected utility when she has reputation $\Lambda(\cdot)$. Note that $V^J(\cdot)$ is independent of γ^M . The reason is that the misaligned analyst always reports 1 in the second period and hence her precision has no effect on the investor's action.

Write $\Pi_C^J(s_1|\gamma)$ for the net current expected gain to the type J analyst of reporting 1 rather than 0, when she observes signal s_1 and the analyst's precision $\gamma = (\gamma^A, \gamma^M)$:³

$$\begin{aligned}\Pi_C^A(s_1|\gamma) &= -x^A\{\gamma^A(a_1(1, \gamma) - s_1)^2 + (1 - \gamma^A)(a_1(1, \gamma) - (1 - s_1))^2\} \\ &\quad + x^A\{\gamma^A(a_1(0, \gamma) - s_1)^2 + (1 - \gamma^A)(a_1(0, \gamma) - (1 - s_1))^2\}, \\ \Pi_C^M(1|\gamma) &= \Pi_C^M(0|\gamma) = -x^M(a_1(1, \gamma) - 1)^2 + x^M(a_1(0, \gamma) - 1)^2.\end{aligned}\tag{A.1}$$

³To save on notation, from now on, I suppress the functional dependence of $a_1(\cdot)$ and $\Gamma_1(\cdot)$ on λ .

Write $\Pi_R^J(s_1|\gamma)$ for the net expected reputational gain to the type J analyst of reporting 0 rather than 1, when she observes signal s_1 and the analyst's precision is γ :

$$\begin{aligned} \Pi_R^J(s_1|\gamma) = & (1 - x^J)\gamma^J[V^J(\Lambda(0, s_1|\gamma), \gamma) - V^J(\Lambda(1, s_1|\gamma), \gamma)] \\ & + (1 - x^J)(1 - \gamma^J)[V^J(\Lambda(0, 1 - s_1|\gamma), \gamma) - V^J(\Lambda(1, 1 - s_1|\gamma), \gamma)] \end{aligned} \quad (\text{A.2})$$

The type J analyst has a strict incentive to report 1 when she observes s_1 if and only if $\Pi_C^J(s_1|\gamma) > \Pi_R^J(s_1|\gamma)$. I refer to $\Pi_C^J(\cdot)$ and $\Pi_R^J(\cdot)$ as the type J analyst's *current reporting incentive* and *reputational reporting incentive*, respectively.

To save on notation, I suppress the functional dependence of the above functions on $\gamma = (\gamma^A, \gamma^M)$ when there is no confusion. I also suppress the argument λ in $\Gamma_1(\cdot)$ and $a_1(\cdot)$.

Following Morris (2001)'s notation, let

$$f(\lambda, \delta) = (1 - \gamma^A) \min\left\{V^A(\lambda) - V^A\left(\frac{1}{1 + \frac{1-\lambda}{\lambda}(1 + \delta)}\right), V^A\left(\frac{1}{1 + \frac{1-\lambda}{\lambda}\frac{1}{1+\delta}}\right) - V^A(\lambda)\right\}$$

and let $h(\lambda, k)$ be the unique value of δ solving

$$k = V^M\left(\frac{1}{1 + \frac{1-\lambda}{\lambda}\frac{1}{1+\delta}}\right) - V^M\left(\frac{1}{1 + \frac{1-\lambda}{\lambda}(1 + \delta)}\right)$$

if $k < V^M(1) - V^M(0)$; if $k \geq V^M(1) - V^M(0)$, let $h(\lambda, k) = \infty$.

Define

$$\underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M) \equiv 1 - \frac{1}{x_{oo}^A(\lambda, \gamma^A, \gamma^M, x^M) + 1}, \quad (\text{A.3})$$

and

$$x_{oo}^A(\lambda, \gamma^A, \gamma^M, x^M) = \frac{f\left(\lambda, \min\left\{\frac{1-\gamma^A}{2\gamma^A}, \frac{2\gamma^A-1}{2(1-\gamma^A)}, h\left(\lambda, \frac{1}{2} \frac{x^M}{1-x^M} \left[-\left(\frac{\gamma^A}{\gamma^A+\frac{1}{2}} - 1\right)^2 + \left(\frac{1}{2} - 1\right)^2\right]\right)\right\}\right)}{\hat{u}^A(\gamma^A, 1) - \hat{u}^A(1 - \gamma^A, 1)},$$

Where $\hat{u}^A(a_1, s_1) = -\gamma^A(a_1 - s_1)^2 - (1 - \gamma^A)(a_1 - (1 - s_1))^2$ denotes the aligned analyst's expected first period payoff if she observes signal s_1 and the investor's action is a_1 .

Claim 1 *Suppose it is common knowledge that the aligned analyst is almost perfectly informed and has higher precision than the misaligned analyst, i.e., $\gamma^A = \gamma_H \rightarrow 1 > \gamma^M$ and $x^A \in (0, \underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)] \cup [x^M, \infty)$. Then any informative communication in the first period has to satisfy the following properties:*

- (1) $\Lambda(0, 0) \geq \Lambda(1, 0)$;
- (2) *The aligned analyst always reports truthfully, i.e., $\sigma_1^A(s_1) = s_1$;*
- (3) *The misaligned analyst always reports 1 when she observes signal 1, i.e., $\sigma_1^M(1) = 1$.*

I prove the result in eight steps. Each step identifies a property that must hold for first period informative communication when $\gamma^A = \gamma_H \rightarrow 1 > \gamma^M$ and $x^A \in (0, \underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)] \cup [x^M, \infty)$. The first four steps hold for generic $\gamma = (\gamma^A, \gamma^M)$ and x^A .

Property 1. $\Lambda(0, 0) \geq \Lambda(1, 0)$.

Prove by contradiction. Suppose $\Lambda(0, 0) < \Lambda(1, 0)$, there are 2 subcases:

(i) $\Lambda(0, 0) < \Lambda(1, 0)$ and $\Lambda(1, 1) > \Lambda(0, 1)$.

Now $\Pi_R^M(s_1) < 0$ and $\Pi_C^M(s_1) \geq 0$ for $s_1 = 0, 1$. Therefore $\sigma_1^M(0) = \sigma_1^M(1) = 1$. Then if $\sigma_1^A(0) = \sigma_1^A(1) = 1$, the communication in the first period is uninformative, which is a contradiction. If $\sigma_1^A(0) \neq 1$ or $\sigma_1^A(1) \neq 1$, then $\Lambda(0, 1) = 1$ or $\Lambda(1, 0) = 1$, contradicting “ $\Lambda(0, 0) < \Lambda(1, 0)$ and $\Lambda(1, 1) > \Lambda(0, 1)$ ”.

(ii) $\Lambda(0, 0) < \Lambda(1, 0)$ and $\Lambda(1, 1) \leq \Lambda(0, 1)$.

In this case, $\Pi_R^M(1) > \Pi_R^M(0)$, which, together with the fact that $\Pi_C^M(1) = \Pi_C^M(0)$, implies that either $\sigma_1^M(1) = 0$, or $\sigma_1^M(0) = 1$. Then, $\phi_1^M(1|1) \leq \phi_1^M(1|0)$. On the other hand, by the definition of Λ in (1.2), $\Lambda(0, 0) < \Lambda(1, 0) \Rightarrow \phi_1^A(1|0) > \phi_1^M(1|0)$, and $\Lambda(1, 1) \leq \Lambda(0, 1) \Rightarrow \phi_1^A(1|1) \leq \phi_1^M(1|1)$. Therefore, $\phi_1^A(1|1) \leq \phi_1^M(1|1) \leq \phi_1^M(1|0) < \phi_1^A(1|0)$. Thus, by (1.1), $a_1(1) = \Gamma_1(1) < \Gamma_1(0) = a_1(0)$. A contradiction (recall that $a_1(1) \geq a_1(0)$ is assumed without loss of generality).

To conclude, $\Lambda(0, 0) < \Lambda(1, 0)$ is impossible for informative communication. Therefore $\Lambda(0, 0) \geq \Lambda(1, 0)$.

Property 2. $a_1(1) = \Gamma_1(1) > \Gamma_1(0) = a_1(0)$.

Given property 1, $\Lambda(0, 0) \geq \Lambda(1, 0)$, there are two subcases. The idea is to show that $\Gamma_1(1) > \Gamma_1(0)$ holds for both subcases.

(i) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) \geq \Lambda(1, 1)$.

First, I claim that at least one of the inequalities is strict. Suppose not, that is, both

inequalities hold with equality. If $\Gamma_1(1) = \Gamma_1(0)$, the communication is noninformative, which is a contradiction. If $\Gamma_1(1) > \Gamma_1(0)$, then the misaligned analyst will always report 1 since there are no reputational consequences. Thus, the investor will know for sure that the analyst is aligned if he receives message 0, i.e., $\Lambda(0, 0) = 1$, which implies that $\Lambda(1, 0) = 1$ given that $\Lambda(0, 0) = \Lambda(1, 0)$. Obviously, $\Lambda(1, 0) = 1$ is contradicting $\sigma_1^M(0) = \sigma_1^M(1) = 1$. Therefore, at least one of the inequalities is strict. As a result, $\Pi_R^J(s_1) > 0$ for $J = M, A$ and $s_1 = 0, 1$. Suppose $\Gamma_1(1) = \Gamma_1(0)$, then it follows that $\Pi_C^J(s_1) = 0$ for $J = M, A$ and $s_1 = 0, 1$. Therefore, $\sigma_1^J(s_1) = 0$, the communication is noninformative. A contradiction.

(ii) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) < \Lambda(1, 1)$.

In this case, $\Pi_R^M(0) > \Pi_R^M(1)$ and $\Pi_R^A(0) > \Pi_R^A(1)$. Suppose $\Gamma_1(1) = \Gamma_1(0)$. Then $\Pi_C^J(s_1) = 0$ for each $J = M, A$ and $s_1 = 0, 1$. Therefore, for $J = M, A$, either $\sigma_1^J(0) = 0$, or $\sigma_1^J(1) = 1$. I.e., $\sigma_1^J(1) \geq \sigma_1^J(0)$.

In addition, by (1.1), $\Gamma_1(1) = \Gamma_1(0)$ implies that:

$$\begin{aligned}
& \frac{1}{\Gamma_1(1)} - 1 = \frac{1}{\Gamma_1(0)} - 1 \\
\iff & \frac{\lambda\phi_1^A(1|0) + (1-\lambda)\phi_1^M(1|0)}{\lambda\phi_1^A(1|1) + (1-\lambda)\phi_1^M(1|1)} = \frac{\lambda\phi_1^A(0|0) + (1-\lambda)\phi_1^M(0|0)}{\lambda\phi_1^A(0|1) + (1-\lambda)\phi_1^M(0|1)} \\
\iff & \frac{\lambda\phi_1^A(1|0) + (1-\lambda)\phi_1^M(1|0)}{\lambda\phi_1^A(1|1) + (1-\lambda)\phi_1^M(1|1)} = 1 \\
\iff & \lambda(2\gamma^A - 1)[\sigma_1^A(1) - \sigma_1^A(0)] + (1-\lambda)(2\gamma^M - 1)[\sigma_1^M(1) - \sigma_1^M(0)] = 0.
\end{aligned}$$

(A.4)

As a result, $\Gamma_1(1) = \Gamma_1(0)$ holds only when $\sigma_1^J(1) = \sigma_1^J(0)$ for $J = M, A$. There are four scenarios then: (a) if $\sigma_1^A(1) = \sigma_1^A(0) = \sigma_1^M(1) = \sigma_1^M(0) = 1$, then the communication is noninformative. A contradiction; (b) if $\sigma_1^A(1) = \sigma_1^A(0) = \sigma_1^M(1) = \sigma_1^M(0) = 0$, again the communication is noninformative. A contradiction; (c) if $\sigma_1^A(1) = \sigma_1^A(0) = 1$ and $\sigma_1^M(1) = \sigma_1^M(0) = 0$, then $\Lambda(0, 0) = 0$ and $\Lambda(1, 0) = 1$, contradicting $\Lambda(0, 0) \geq \Lambda(1, 0)$; (d) if $\sigma_1^A(1) = \sigma_1^A(0) = 0$ and $\sigma_1^M(1) = \sigma_1^M(0) = 1$, then $\Lambda(1, 1) = 0$ and $\Lambda(0, 1) = 1$, contradicting $\Lambda(0, 1) < \Lambda(1, 1)$.

To summarize, in any first period informative communication, $\Gamma_1(1) > \Gamma_1(0)$ has to hold.

Property 3. $\sigma_1^A(0) = 0 \Rightarrow \sigma_1^A(1) > 0$.

Prove by contradiction. Suppose $\sigma_1^A(0) = \sigma_1^A(1) = 0$, then for $\Gamma_1(1) > \Gamma_1(0)$ to hold, it must be true that $\sigma_1^M(1) > \sigma_1^M(0)$. Thus, $\phi_1^M(1|1) > \phi_1^M(1|0)$, which directly leads to $\phi_1^M(0|1) < \phi_1^M(0|0)$. On the other hand, $\sigma_1^A(0) = \sigma_1^A(1) = 0$ implies that $\phi_1^A(0|1) = \phi_1^A(0|0) = 1$. Hence, $\Lambda(0, 1) > \Lambda(0, 0)$. In addition, $\sigma_1^A(1) = \sigma_1^A(0) = 0$ implies $\Lambda(1, 1) = \Lambda(1, 0) = 0$. Therefore, $\Pi_R^M(1) > \Pi_R^M(0)$. Recall that $\Pi_C^M(1) = \Pi_C^M(0)$, hence $\sigma_1^M(1) \leq \sigma_1^M(0)$ has to hold, contradicting $\sigma_1^M(1) > \sigma_1^M(0)$. Therefore, $\sigma_1^A(0) = 0$ and $\sigma_1^A(1) = 0$ cannot hold at the same time. Hence, $\sigma_1^A(0) = 0 \Rightarrow \sigma_1^A(1) > 0$.

Property 4. $\sigma_1^A(0) = 0 \Rightarrow \sigma_1^M(0) \leq \sigma_1^M(1)$.

Suppose $\sigma_1^A(0) = 0$ and $\sigma_1^M(0) > \sigma_1^M(1)$. Then, by Property 3, $\sigma_1^A(1) > 0$. Thus, $\phi_1^A(1|1) > \phi_1^A(1|0)$ and $\phi_1^A(0|1) < \phi_1^A(0|0)$. On the other hand, $\sigma_1^M(0) > \sigma_1^M(1)$ implies that $\phi_1^M(1|1) < \phi_1^M(1|0)$ and $\phi_1^M(0|1) > \phi_1^M(0|0)$. Therefore, $\Lambda(0, 0) > \Lambda(0, 1)$ and

$\Lambda(1, 1) > \Lambda(1, 0)$. Then $\Pi_R^M(1) < \Pi_R^M(0)$. Recall that $\Pi_C^M(1) = \Pi_C^M(0)$, so $\sigma_1^M(0) \leq \sigma_1^M(1)$, contradicting $\sigma_1^M(0) > \sigma_1^M(1)$. Hence, $\sigma_1^A(0) = 0$ and $\sigma_1^M(0) > \sigma_1^M(1)$ cannot hold at the same time. Therefore, $\sigma_1^A(0) = 0 \Rightarrow \sigma_1^M(0) \leq \sigma_1^M(1)$.

Property 5. $\sigma_1^A(0) = 0$.

Recall that

$$\Pi_R^A(0) = (1 - x^A)[\gamma^A(V^A(\Lambda(0, 0)) - V^A(\Lambda(1, 0))) + (1 - \gamma^A)(V^A(\Lambda(0, 1)) - V^A(\Lambda(1, 1)))].$$

Hence, $\lim_{\gamma^A \rightarrow 1} \Pi_R^A(0) = (1 - x^A)[V^A(\Lambda(0, 0)) - V^A(\Lambda(1, 0))]$. By Property 1, $\Lambda(0, 0) \geq \Lambda(1, 0)$, it follows that $\lim_{\gamma^A \rightarrow 1} \Pi_R^A(0) \geq 0$. At the same time, Property 2, $\Gamma_1(1) > \Gamma_1(0)$, implies that $\Pi_C^A(0) < 0$. Therefore $\sigma_1^A(0) = 0$.

Property 6. $\sigma_1^A(1) > 0$ and $\sigma_1^M(0) \leq \sigma_1^M(1)$.

These directly follow from Property 3, Property 4, and Property 5.

Property 7. $\sigma_1^A(1) = 1$.

Given property 1, there are two subcases:

(1) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) \leq \Lambda(1, 1)$.

Then $\lim_{\gamma^A \rightarrow 1} \Pi_R^A(1) = (1 - x^A)[V^A(\Lambda(0, 1)) - V^A(\Lambda(1, 1))] \leq 0$. At the same time,

$\Pi_C^A(1) > 0$. Therefore $\sigma_1^A(1) = 1$.

(2) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) > \Lambda(1, 1)$.

If $x^A \leq \underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)$, then the similar proof as in Proposition 2 of Morris (2001)

shows that there is no informative equilibrium satisfying $\Lambda(0,0) \geq \Lambda(1,0)$ and $\Lambda(0,1) > \Lambda(1,1)$.

If $x^A \geq x^M$, then $\sigma_1^A(1) = 1$. The proof includes the following two steps:

(i)

$$\Lambda(0,1) > \Lambda(1,1) \Rightarrow \begin{cases} \sigma_1^M(1) > 0 \\ V^M(\Lambda(0,0)) - V^M(\Lambda(1,0)) \geq V^M(\Lambda(0,1)) - V^M(\Lambda(1,1)). \end{cases}$$

By property 6, $\sigma_1^M(0) \leq \sigma_1^M(1)$. This includes the following 4 cases:

(a) $\sigma_1^M(0) = \sigma_1^M(1) = 0$.

In this case, $\Lambda(1,1) = 1$, contradicting $\Lambda(0,1) > \Lambda(1,1)$.

(b) $0 \leq \sigma_1^M(0) < \sigma_1^M(1) \leq 1$.

$$0 < \sigma_1^M(1) \leq 1 \Rightarrow \Pi_R^M(1) \leq \Pi_C^M(1); \text{ and } 0 \leq \sigma_1^M(0) < 1 \Rightarrow \Pi_R^M(0) \geq \Pi_C^M(0).$$

Therefore, $\Pi_R^M(0) \geq \Pi_C^M(0) = \Pi_C^M(1) \geq \Pi_R^M(1)$. By the definition of $\Pi_R^M(\cdot)$, $\Pi_R^M(0) \geq \Pi_R^M(1) \Rightarrow V^M(\Lambda(0,0)) - V^M(\Lambda(1,0)) \geq V^M(\Lambda(0,1)) - V^M(\Lambda(1,1))$.

(c) $0 < \sigma_1^M(0) = \sigma_1^M(1) < 1$.

$$0 < \sigma_1^M(1) < 1 \Rightarrow \Pi_R^M(1) = \Pi_C^M(1); \text{ and } 0 < \sigma_1^M(0) < 1 \Rightarrow \Pi_R^M(0) = \Pi_C^M(0).$$

Therefore, $\Pi_R^M(0) = \Pi_C^M(0) = \Pi_C^M(1) = \Pi_R^M(1)$. By the definition of $\Pi_R^M(\cdot)$, $\Pi_R^M(0) = \Pi_R^M(1) \Rightarrow V^M(\Lambda(0,0)) - V^M(\Lambda(1,0)) = V^M(\Lambda(0,1)) - V^M(\Lambda(1,1))$.

(d) $\sigma_1^M(0) = \sigma_1^M(1) = 1$.

In this case, it must be true that $\Lambda(0,1) = \Lambda(0,0) = 1$. At the same time, $\sigma_1^A(0) <$

$\sigma_1^A(1)$ and $\sigma_1^M(0) = \sigma_1^M(1) = 1$ imply $\Lambda(1, 0) < \Lambda(1, 1)$. Thus, $V^M(\Lambda(0, 0)) - V^M(\Lambda(1, 0)) > V^M(\Lambda(0, 1)) - V^M(\Lambda(1, 1))$.

(ii)

$$\left. \begin{aligned} V^M(\Lambda(0, 0)) - V^M(\Lambda(1, 0)) &\geq V^M(\Lambda(0, 1)) - V^M(\Lambda(1, 1)) \\ \Lambda(0, 1) &> \Lambda(1, 1) \\ \sigma_1^M(1) &> 0 \end{aligned} \right\} \Rightarrow \sigma_1^A(1) = 1.$$

Now, we have

$$\begin{aligned} \Pi_R^M(1) &= (1 - x^M)[\gamma^M[V^M(\Lambda(0, 1)) - V^M(\Lambda(1, 1))] + (1 - \gamma^M)[V^M(\Lambda(0, 0)) - V^M(\Lambda(1, 0))] \\ &\geq (1 - x^M)[V^M(\Lambda(0, 1)) - V^M(\Lambda(1, 1))] \\ &> (1 - x^M)[V^A(\Lambda(0, 1)) - V^A(\Lambda(1, 1))] \\ &= \frac{1 - x^M}{1 - x^A} \lim_{\gamma^A \rightarrow 1} \Pi_R^A(1). \end{aligned}$$

The last inequality comes from $\Lambda(0, 1) > \Lambda(1, 1)$ and $\frac{\partial V^A(\lambda_2)}{\partial \lambda_2} < \frac{\partial V^M(\lambda_2)}{\partial \lambda_2}$. At the same

time, $\lim_{\gamma^A \rightarrow 1} \Pi_C^A(1) = \frac{x^A}{x^M} \Pi_C^M(1)$. In addition, $\sigma_1^M(1) > 0 \Rightarrow \Pi_R^M(1) \leq \Pi_C^M(1)$.

Therefore, $\lim_{\gamma^A \rightarrow 1} \Pi_C^A(1) = \frac{x^A}{x^M} \Pi_C^M(1) \geq \frac{x^A}{x^M} \Pi_R^M(1) > \frac{x^A(1-x^M)}{x^M(1-x^A)} \lim_{\gamma^A \rightarrow 1} \Pi_R^A(1) \geq$

$\lim_{\gamma^A \rightarrow 1} \Pi_R^A(1)$ for $x^A \geq x^M$. As a result, $\sigma_1^A(1) = 1$ when $\gamma^A \rightarrow 1$.

Property 8. $\sigma_1^M(1) = 1$.

Again, there are two subcases:

(1) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) \leq \Lambda(1, 1)$.

In this case, $\Pi_R^M(0) > \Pi_R^M(1)$, which together with $\Pi_C^M(0) = \Pi_C^M(1)$, implies $\sigma_1^M(1) = 1$ or $\sigma_1^M(0) = 0$. Suppose $\sigma_1^M(1) \neq 1$, then $\sigma_1^M(0) = 0$. Recall that $\sigma_1^A(0) = 0$ and $\sigma_1^A(1) = 1$. Therefore,

$$\begin{aligned} \frac{\phi_1^M(1|0)}{\phi_1^A(1|0)} &= \frac{(1 - \gamma^M)\sigma_1^M(1)}{1 - \gamma^A} < \frac{1 - \gamma^M\sigma_1^M(1)}{1 - \gamma^A} = \frac{\phi_1^M(0|1)}{\phi_1^A(0|1)}, \\ \frac{\phi_1^M(1|1)}{\phi_1^A(1|1)} &= \frac{\gamma^M\sigma_1^M(1)}{\gamma^A} < \frac{1 - (1 - \gamma^M)\sigma_1^M(1)}{\gamma^A} = \frac{\phi_1^M(0|0)}{\phi_1^A(0|0)}. \end{aligned}$$

As a result, $\Lambda(1, 0) > \Lambda(0, 1)$ and $\Lambda(1, 1) > \Lambda(0, 0)$. Hence, $\Lambda(1, 1) > \Lambda(0, 0) \geq \Lambda(1, 0) > \Lambda(0, 1)$. Therefore, $\Pi_R^M(1) < 0$. At the same time, $\Pi_C^M(1) > 0$. Then $\sigma_1^M(1) = 1$ follows immediately.

(2) $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(0, 1) > \Lambda(1, 1)$.

$\Lambda(0, 1) > \Lambda(1, 1)$ implies that $\phi_1^M(1|1) > \phi_1^A(1|1) = \gamma^A$. If $\gamma^A \rightarrow 1$, then $\phi_1^M(1|1) \rightarrow 1$, which implies that $\Lambda(0, 1) \rightarrow \Lambda(1, 1)$ and $\sigma_1^M(s_1) \rightarrow 1$ for $s_1 = 0, 1$. At the same time, $\phi_1^A(1|0) = 1 - \gamma^A \rightarrow 0$ and $\phi_1^M(1|0) \rightarrow 1$, which implies $\Lambda(0, 0) \gg \Lambda(1, 0)$. As a result, $\Pi_R^M(0) > \Pi_R^M(1)$. Recall that $\Pi_C^M(0) = \Pi_C^M(1)$, therefore either $\sigma_1^M(1) = 1$ or $\sigma_1^M(0) = 0$. Suppose $\sigma_1^M(1) \neq 1$, then $\sigma_1^M(0) = 0$. Then $\phi_1^M(1|1) = \gamma^M\sigma_1^M(1) < \gamma^A = \phi_1^A(1|1)$, which contradicts $\Lambda(0, 1) > \Lambda(1, 1)$. Therefore, $\sigma_1^M(1) = 1$. ■

Claim 2 *If the misaligned analyst is (weakly) better informed than the aligned analyst,*

i.e., $\gamma^A \leq \gamma^M$, then any informative communication in the first period has to satisfy the following properties:

- (1) The investor updates favorably his belief about the analyst's type when receiving report 0. More specifically, $\Lambda(0, 1) \geq \Lambda(1, 1)$ and $\Lambda(0, 0) \geq \Lambda(1, 0)$, and at least one of the inequalities is strict;
- (2) The aligned analyst reports 0 when she observes signal 0, i.e. $\sigma_1^A(0) = 0$. If she observes signal 1, the aligned analyst reports 1 with positive probability, i.e. $\sigma_1^A(1) > 0$;
- (3) The misaligned analyst reports 1 with positive probability for both signals. In addition, she reports 1 more often when she observes signal 1 than when she observes signal 0, i.e. $\sigma_1^M(1) \geq \sigma_1^M(0) > 0$.

I prove the result in nine steps. Each step identifies a property that must hold for first period informative communication when $\gamma^A \leq \gamma^M$.

Property 1-4. The first four steps are the same as the proof of Claim 1.

Property 5. $\Lambda(1, 1) \leq \Lambda(0, 1)$.

Prove by contradiction. Suppose instead $\Lambda(1, 1) > \Lambda(0, 1)$. By Property 1, $\Lambda(0, 0) \geq \Lambda(1, 0)$. Thus, $\Pi_R^J(1) < \Pi_R^J(0)$ for $J = M, A$. On the other hand, $\Pi_C^J(1) \geq \Pi_C^J(0)$.

Therefore, either $\sigma_1^J(0) = 0$, or $\sigma_1^J(1) = 1$. At the same time,

$$\begin{aligned} \Lambda(1,1) > \Lambda(0,1) &\implies \phi_1^A(1|1) > \phi_1^M(1|1) \\ &\implies \gamma^A \sigma_1^A(1) + (1 - \gamma^A) \sigma_1^A(0) > \gamma^M \sigma_1^M(1) + (1 - \gamma^M) \sigma_1^M(0) \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} \Lambda(0,0) \geq \Lambda(1,0) &\implies \phi_1^A(1|0) \leq \phi_1^M(1|0) \\ &\implies \gamma^A \sigma_1^A(0) + (1 - \gamma^A) \sigma_1^A(1) \leq \gamma^M \sigma_1^M(0) + (1 - \gamma^M) \sigma_1^M(1) \end{aligned} \quad (\text{A.6})$$

There are four subcases:

(a) $\sigma_1^A(0) = \sigma_1^M(0) = 0$.

In this case, (A.5) can be reduced to $\gamma^A \sigma_1^A(1) > \gamma^M \sigma_1^M(1)$, which implies that $\sigma_1^A(1) > \sigma_1^M(1)$ because $\gamma^M \geq \gamma^A$. On the other hand, by (A.6), $(1 - \gamma^A) \sigma_1^A(1) \leq (1 - \gamma^M) \sigma_1^M(1)$, which leads to $\sigma_1^A(1) \leq \sigma_1^M(1)$. A contradiction.

(b) $\sigma_1^A(1) = \sigma_1^M(1) = 1$.

In this case, by (A.6),

$$\begin{aligned} \gamma^A \sigma_1^A(0) + 1 - \gamma^A &\leq \gamma^M \sigma_1^M(0) + 1 - \gamma^M \implies \gamma^A(1 - \sigma_1^A(0)) \geq \gamma^M(1 - \sigma_1^M(0)) \\ &\implies \sigma_1^A(0) \leq \sigma_1^M(0). \end{aligned}$$

On the other hand, (A.5) can be reduced to:

$$\begin{aligned} \gamma^A + (1 - \gamma^A)\sigma_1^A(0) &> \gamma^M + (1 - \gamma^M)\sigma_1^M(0) \\ \Rightarrow (\gamma^M - \gamma^A)(\sigma_1^A(0) - 1) + (1 - \gamma^M)[\sigma_1^A(0) - \sigma_1^M(0)] &> 0 \\ \Rightarrow \sigma_1^A(0) &> \sigma_1^M(0). \end{aligned}$$

Again a contradiction.

$$(c) \sigma_1^A(0) = 0 \text{ and } \sigma_1^M(1) = 1.$$

In this case, (A.5) can be reduced to $\gamma^A\sigma_1^A(1) > \gamma^M + (1 - \gamma^M)\sigma_1^M(0)$, which leads to $\sigma_1^A(1) > \frac{\gamma^M}{\gamma^A} \geq 1$. A contradiction.

$$(d) \sigma_1^A(1) = 1 \text{ and } \sigma_1^M(0) = 0.$$

By (A.6),

$$\gamma^A\sigma_1^A(0) + 1 - \gamma^A \leq (1 - \gamma^M)\sigma_1^M(1) \Rightarrow \begin{cases} \sigma_1^M(1) \geq \frac{1 - \gamma^A}{1 - \gamma^M} > 1, & \text{if } \gamma^M > \gamma^A \\ \sigma_1^M(1) = 1 \text{ and } \sigma_1^A(0) = 0. & \text{if } \gamma^M = \gamma^A \end{cases}$$

If $\gamma^M > \gamma^A$, $\sigma_1^M(1) > 1$. A contradiction. If $\gamma^M = \gamma^A$, then $\sigma_1^M(1) = 1$ and $\sigma_1^A(0) = 0$, which, combining with $\sigma_1^A(1) = 1$ and $\sigma_1^M(0) = 0$, implies that $\phi_1^M(1|1) = \phi_1^A(1|1)$, contradicting (A.5).

To summarize, $\Lambda(1, 1) > \Lambda(0, 1)$ leads to a contradiction in every possible scenario.

Hence, $\Lambda(1, 1) \leq \Lambda(0, 1)$.

Property 6. $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(1, 1) \leq \Lambda(0, 1)$, and at least one of these inequalities is strict.

By Property 1 and Property 5, $\Lambda(0, 0) \geq \Lambda(1, 0)$ and $\Lambda(1, 1) \leq \Lambda(0, 1)$. Suppose both hold with equality. Then, there is no reputational consequence of reporting 1, and $\Gamma_1(1) > \Gamma_1(0)$ implies $\sigma_1^M(1) = \sigma_1^M(0) = 1$. Therefore, $\Lambda(0, 0) > \Lambda(1, 0)$, contradicting $\Lambda(0, 0) = \Lambda(1, 0)$.

Property 7. $\sigma_1^A(0) = 0$.

By Property 6, $\Pi_R^A(0) > 0$; by Property 2, $\Pi_C^A(0) < 0$. Hence, $\sigma_1^A(0) = 0$.

Property 8. $\sigma_1^A(1) > 0$ and $\sigma_1^M(0) \leq \sigma_1^M(1)$.

These directly follow from Property 3, Property 4, and Property 7.

Property 9. $\sigma_1^M(0) > 0$.

Prove by contradiction. Suppose that $\sigma_1^M(0) = 0$. Then, $\Lambda(0, 0) \geq \Lambda(1, 0)$ (Property 1) implies that $\sigma_1^A(1) \leq \sigma_1^M(1)$ for $\gamma^A \leq \gamma^M$. At the same time,

$$\frac{\phi_1^M(1|1)}{\phi_1^A(1|1)} = \frac{\gamma^M \sigma_1^M(1)}{\gamma^A \sigma_1^A(1)} \geq \frac{(1 - \gamma^M) \sigma_1^M(1)}{(1 - \gamma^A) \sigma_1^A(1)} = \frac{\phi_1^M(1|0)}{\phi_1^A(1|0)},$$

and

$$\frac{\phi_1^M(0|1)}{\phi_1^A(0|1)} = \frac{1 - \gamma^M \sigma_1^M(1)}{1 - \gamma^A \sigma_1^A(1)} \leq \frac{1 - \gamma^A \sigma_1^M(1)}{1 - \gamma^A \sigma_1^A(1)} \leq \frac{1 - (1 - \gamma^A) \sigma_1^M(1)}{1 - (1 - \gamma^A) \sigma_1^A(1)} \leq \frac{1 - (1 - \gamma^M) \sigma_1^M(1)}{1 - (1 - \gamma^A) \sigma_1^A(1)} = \frac{\phi_1^M(0|0)}{\phi_1^A(0|0)}.$$

The inequality comes from $\gamma^M \geq \gamma^A$ and $\sigma_1^A(1) \leq \sigma_1^M(1)$. Therefore, $\Lambda(1, 1) \leq \Lambda(1, 0)$

and $\Lambda(0, 1) \geq \Lambda(0, 0)$. Then $\Lambda(0, 1) \geq \Lambda(0, 0) \geq \Lambda(1, 0) \geq \Lambda(1, 1)$. As a result, $\Pi_R^M(1) \geq \Pi_R^M(0)$. Hence, $\sigma_1^M(1) \leq \sigma_1^M(0)$ (because $\Pi_C^M(1) = \Pi_C^M(0)$). Since $\sigma_1^M(0) = 0$, it must be true that $\sigma_1^M(1) = 0$. Then, $\Lambda(1, 1) = 1 > \Lambda(0, 1)$, contradicting Property 5, $\Lambda(1, 1) \leq \Lambda(0, 1)$. ■

Proof of Lemma 1 is the same as the proof of Proposition 2 in Morris (2001) and thus is omitted.

A.3 Proof of Corollary 1

Suppose the investor holds conjectures that:

$$\tilde{\sigma}_1^A(0|\gamma^A = \gamma^M) = 0, \quad \tilde{\sigma}_1^A(1|\gamma^A = \gamma^M) = 1, \quad \tilde{\sigma}_1^M(0|\gamma^A = \gamma^M) = \mu \in (0, 1], \quad \tilde{\sigma}_1^M(1|\gamma^A = \gamma^M) = 1,$$

then

$$\begin{aligned} \Lambda(1, 1|\gamma^A = \gamma^M) &= \frac{1}{1 + \frac{1-\lambda}{\lambda}(1 + \frac{1-\gamma^A}{\gamma^A}\mu)}; & \Lambda(0, 1|\gamma^A = \gamma^M) &= \frac{1}{1 + \frac{1-\lambda}{\lambda}(1 - \mu)}; \\ \Lambda(1, 0|\gamma^A = \gamma^M) &= \frac{1}{1 + \frac{1-\lambda}{\lambda}(1 + \frac{\gamma^A}{1-\gamma^A}\mu)}; & \Lambda(0, 0|\gamma^A = \gamma^M) &= \frac{1}{1 + \frac{1-\lambda}{\lambda}(1 - \mu)}; \\ a_1(1, \gamma^A = \gamma^M) &= 1 - \gamma^A + \frac{2\gamma^A - 1}{1 + (1 - \lambda)\mu}; & a_1(0, \gamma^A = \gamma^M) &= 1 - \gamma^A. \end{aligned} \quad (\text{A.7})$$

For this to be an equilibrium, it must be true that the investor's above conjectures are consistent with the analysts' best responses. That is, the following inequalities have to

hold:

$$\Pi_R^A(0|\gamma^A = \gamma^M) \geq \Pi_C^A(0|\gamma^A = \gamma^M) \quad (\text{A.8})$$

$$\Pi_R^A(1|\gamma^A = \gamma^M) \leq \Pi_C^A(1|\gamma^A = \gamma^M) \quad (\text{A.9})$$

$$\Pi_R^M(0|\gamma^A = \gamma^M) \leq \Pi_C^M(0|\gamma^A = \gamma^M) \quad (\text{A.10})$$

$$\Pi_R^M(1|\gamma^A = \gamma^M) \leq \Pi_C^M(1|\gamma^A = \gamma^M) \quad (\text{A.11})$$

The following four steps show that when $x^A \geq x^M \rightarrow 0$, we could indeed find μ such that the inequalities (A.8)-(A.11) hold.

(I) Condition (A.8) always holds.

Given the type inference function and the state inference function derived in (A.7), it is readily to show that $\Pi_R^A(0|\gamma^A = \gamma^M) > 0$ and $\Pi_C^A(0|\gamma^A = \gamma^M) < 0$. Therefore, (A.8) always holds.

(II) (A.10) \Rightarrow (A.11).

According to (A.7), $\Lambda(0, 0|\gamma^A = \gamma^M) = \Lambda(0, 1|\gamma^A = \gamma^M)$ and $\Lambda(1, 0|\gamma^A = \gamma^M) < \Lambda(1, 1|\gamma^A = \gamma^M)$, which imply that $\Pi_R^M(1|\gamma^A = \gamma^M) < \Pi_R^M(0|\gamma^A = \gamma^M)$. Recall that $\Pi_C^M(1|\gamma^A = \gamma^M) = \Pi_C^M(0|\gamma^A = \gamma^M)$, hence if (A.10) holds, (A.11) must hold.

(III) When $x^A \geq x^M \rightarrow 0$, (A.10) \Rightarrow (A.9).

We first write out $\Pi_C^M(0|\gamma^A = \gamma^M)$, $\Pi_C^A(1|\gamma^A = \gamma^M)$, $\Pi_R^M(0|\gamma^A = \gamma^M)$ and $\Pi_R^A(1|\gamma^A =$

γ^M).

$$\begin{aligned}\Pi_C^M(0|\gamma^A = \gamma^M) &= -x^M[a_1(1, \gamma^A = \gamma^M) - 1]^2 + x^M[a_1(0, \gamma^A = \gamma^M) - 1]^2 \\ &= x^M \frac{(2\gamma^A - 1)[1 + 2\gamma^A\mu(1 - \lambda)]}{[1 + \mu(1 - \lambda)]^2},\end{aligned}$$

$$\begin{aligned}\Pi_C^A(1|\gamma^A = \gamma^M) &= -x^A[\gamma^A(a_1(1, \gamma^A = \gamma^M) - 1)^2 + (1 - \gamma^A)(a_1(1, \gamma^A = \gamma^M))^2] \\ &\quad + x^A[\gamma^A(a_1(0, \gamma^A = \gamma^M) - 1)^2 + (1 - \gamma^A)(a_1(0, \gamma^A = \gamma^M))^2] \\ &= x^A \frac{(2\gamma^A - 1)^2[1 + 2\mu(1 - \lambda)]}{[1 + \mu(1 - \lambda)]^2},\end{aligned}$$

$$\begin{aligned}\Pi_R^M(0|\gamma^A = \gamma^M) &= (1 - x^M)\gamma^A[V^M(\Lambda(0, 0|\gamma^A = \gamma^M), \gamma^A) - V^M(\Lambda(1, 0|\gamma^A = \gamma^M), \gamma^A)] \\ &\quad + (1 - x^M)(1 - \gamma^A)[V^M(\Lambda(0, 1|\gamma^A = \gamma^M), \gamma^A) - V^M(\Lambda(1, 1|\gamma^A = \gamma^M), \gamma^A)] \\ &= (1 - x^M)\gamma^A \frac{\partial V^M(\lambda_2, \gamma^A)}{\partial \lambda_2} \Big|_{\lambda_2 = \lambda_2^*} [\Lambda(0, 0|\gamma^A = \gamma^M) - \Lambda(1, 0|\gamma^A = \gamma^M)] \\ &\quad + (1 - x^M)(1 - \gamma^A) \frac{\partial V^M(\lambda_2, \gamma^A)}{\partial \lambda_2} \Big|_{\lambda_2 = \lambda_2^{**}} [\Lambda(0, 1|\gamma^A = \gamma^M) - \Lambda(1, 1|\gamma^A = \gamma^M)] \\ &= (1 - x^M)\gamma^A \frac{2(2\gamma^A - 1)(1 - \lambda_2^*\gamma^A)}{(2 - \lambda_2^*)^3} \frac{\frac{1-\lambda}{\lambda} \frac{1}{1-\gamma^A}\mu}{[1 + \frac{1-\lambda}{\lambda}(1 - \mu)][1 + \frac{1-\lambda}{\lambda}(1 + \frac{\gamma^A}{1-\gamma^A}\mu)]} \\ &\quad + (1 - x^M)(1 - \gamma^A) \frac{2(2\gamma^A - 1)(1 - \lambda_2^{**}\gamma^A)}{(2 - \lambda_2^{**})^3} \frac{\frac{1-\lambda}{\lambda} \frac{1}{\gamma^A}\mu}{[1 + \frac{1-\lambda}{\lambda}(1 - \mu)][1 + \frac{1-\lambda}{\lambda}(1 + \frac{1-\gamma^A}{\gamma^A}\mu)]},\end{aligned}$$

where $\lambda_2^* \in [\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{\gamma^A}{1-\gamma^A}\mu)}, \frac{1}{1+\frac{1-\lambda}{\lambda}(1-\mu)}]$ and $\lambda_2^{**} \in [\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{1-\gamma^A}{\gamma^A}\mu)}, \frac{1}{1+\frac{1-\lambda}{\lambda}(1-\mu)}]$.

$$\begin{aligned}
\Pi_R^A(1|\gamma^A = \gamma^M) &= (1-x^A)(1-\gamma^A)[V^A(\Lambda(0,0|\gamma^A = \gamma^M), \gamma^A) - V^A(\Lambda(1,0|\gamma^A = \gamma^M), \gamma^A)] \\
&\quad + (1-x^A)\gamma^A[V^A(\Lambda(0,1|\gamma^A = \gamma^M), \gamma^A) - V^A(\Lambda(1,1|\gamma^A = \gamma^M), \gamma^A)] \\
&= (1-x^A)(1-\gamma^A)\frac{\partial V^A(\lambda_2, \gamma^A)}{\partial \lambda_2}\Big|_{\lambda_2=\lambda_2^+}[\Lambda(0,0|\gamma^A = \gamma^M) - \Lambda(1,0|\gamma^A = \gamma^M)] \\
&\quad + (1-x^A)\gamma^A\frac{\partial V^A(\lambda_2, \gamma^A)}{\partial \lambda_2}\Big|_{\lambda_2=\lambda_2^{++}}[\Lambda(0,1|\gamma^A = \gamma^M) - \Lambda(1,1|\gamma^A = \gamma^M)] \\
&= (1-x^A)(1-\gamma^A)\frac{(2\gamma^A-1)^2(1-\lambda_2^+)}{(2-\lambda_2^+)^3}\frac{\frac{1-\lambda}{\lambda}\frac{1}{1-\gamma^A}\mu}{[1+\frac{1-\lambda}{\lambda}(1-\mu)][1+\frac{1-\lambda}{\lambda}(1+\frac{\gamma^A}{1-\gamma^A}\mu)]} \\
&\quad + (1-x^A)\gamma^A\frac{(2\gamma^A-1)^2(1-\lambda_2^{++})}{(2-\lambda_2^{++})^3}\frac{\frac{1-\lambda}{\lambda}\frac{1}{\gamma^A}\mu}{[1+\frac{1-\lambda}{\lambda}(1-\mu)][1+\frac{1-\lambda}{\lambda}(1+\frac{1-\gamma^A}{\gamma^A}\mu)]},
\end{aligned}$$

where $\lambda_2^+ \in [\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{\gamma^A}{1-\gamma^A}\mu)}, \frac{1}{1+\frac{1-\lambda}{\lambda}(1-\mu)}]$ and $\lambda_2^{++} \in [\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{1-\gamma^A}{\gamma^A}\mu)}, \frac{1}{1+\frac{1-\lambda}{\lambda}(1-\mu)}]$.

Suppose (A.10) holds. If $x^M \rightarrow 0$, by (A.10), then $\mu \rightarrow 0$. This implies that

$\Pi_R^M(0|\gamma^A = \gamma^M) = \Pi_C^M(0|\gamma^A = \gamma^M)$. At the same time, $\mu \rightarrow 0$ implies that $\frac{1}{1+\frac{1-\lambda}{\lambda}(1-\mu)} \rightarrow$

λ , $\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{\gamma^A}{1-\gamma^A}\mu)} \rightarrow \lambda$, and $\frac{1}{1+\frac{1-\lambda}{\lambda}(1+\frac{1-\gamma^A}{\gamma^A}\mu)} \rightarrow \lambda$. Therefore, $\lambda_2^* \rightarrow \lambda$, $\lambda_2^{**} \rightarrow \lambda$, $\lambda_2^+ \rightarrow \lambda$,

and $\lambda_2^{++} \rightarrow \lambda$. Hence,

$$\begin{aligned}
\lim_{x^M \rightarrow 0} \frac{\Pi_C^M(0|\gamma^A = \gamma^M)}{x^M} &= (2\gamma^A - 1), \\
\lim_{x^M \rightarrow 0} \frac{\Pi_C^A(1|\gamma^A = \gamma^M)}{x^A} &= (2\gamma^A - 1)^2, \\
\lim_{x^M \rightarrow 0} \frac{\Pi_R^M(0|\gamma^A = \gamma^M)}{\mu} &= (1-x^M)\frac{2(2\gamma^A-1)(1-\lambda\gamma^A)}{(2-\lambda)^3}\lambda(1-\lambda)\left(\frac{\gamma^A}{1-\gamma^A} + \frac{1-\gamma^A}{\gamma^A}\right), \\
\lim_{x^M \rightarrow 0} \frac{\Pi_R^A(1|\gamma^A = \gamma^M)}{\mu} &= 2(1-x^A)\frac{(2\gamma^A-1)^2(1-\lambda)}{(2-\lambda)^3}\lambda(1-\lambda).
\end{aligned}$$

If $x^A \geq x^M$, then

$$\begin{aligned}
\lim_{x^M \rightarrow 0} \frac{\Pi_R^A(1|\gamma^A = \gamma^M)}{\Pi_R^M(0|\gamma^A = \gamma^M)} &= \frac{(1 - x^A)(2\gamma^A - 1)(1 - \lambda)}{(1 - x^M)(1 - \lambda\gamma^A)(\frac{\gamma^A}{1-\gamma^A} + \frac{1-\gamma^A}{\gamma^A})} \\
&\leq \frac{(2\gamma^A - 1)(1 - \lambda)}{(1 - \lambda\gamma^A)(\frac{\gamma^A}{1-\gamma^A} + \frac{1-\gamma^A}{\gamma^A})}, \\
\lim_{x^M \rightarrow 0} \frac{\Pi_C^A(1|\gamma^A = \gamma^M)}{\Pi_C^M(0|\gamma^A = \gamma^M)} &= (2\gamma^A - 1)\frac{x^A}{x^M} \\
&\geq 2\gamma^A - 1.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\lim_{x^M \rightarrow 0} \frac{\Pi_R^A(1|\gamma^A = \gamma^M)}{\Pi_R^M(0|\gamma^A = \gamma^M)} - \lim_{x^M \rightarrow 0} \frac{\Pi_C^A(1|\gamma^A = \gamma^M)}{\Pi_C^M(0|\gamma^A = \gamma^M)} \\
&\leq (2\gamma^A - 1)\left(\frac{(1 - \lambda)}{(1 - \lambda\gamma^A)(\frac{\gamma^A}{1-\gamma^A} + \frac{1-\gamma^A}{\gamma^A})} - 1\right) \\
&< 0.
\end{aligned} \tag{A.12}$$

Recall that for $x^M \rightarrow 0$, $\Pi_R^M(0|\gamma^A = \gamma^M) = \Pi_C^M(0|\gamma^A = \gamma^M)$. Therefore, by (A.12), $\lim_{x^M \rightarrow 0} \Pi_R^A(1|\gamma^A = \gamma^M) < \lim_{x^M \rightarrow 0} \Pi_C^A(1|\gamma^A = \gamma^M)$. Hence, for $x^A \geq x^M \rightarrow 0$, (A.10) implies (A.9).

(IV) Now the problem is reduced to see whether we could find μ such that (A.10) holds.

Given that there is only one condition (A.10) to satisfy, it immediately follows that we can always find μ satisfying the condition.

Therefore, the truth-telling equilibrium in which the aligned analyst tells the truth can always obtain in the first period when $x^A \geq x^M \rightarrow 0$. ■

A.4 Proof of Proposition 1

Suppose the investor holds conjectures that:

$$\tilde{\sigma}_1^A(0|\gamma^A > \gamma^M) = 0, \quad \tilde{\sigma}_1^A(1|\gamma^A > \gamma^M) = 1, \quad \tilde{\sigma}_1^M(0|\gamma^A > \gamma^M) = 0, \quad \tilde{\sigma}_1^M(1|\gamma^A > \gamma^M) = 1,$$

then

$$\begin{aligned} \Lambda(1, 1|\gamma^A > \gamma^M) &= \Lambda(0, 0|\gamma^A > \gamma^M) = \frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma^M}{\gamma^A}}, \\ \Lambda(1, 0|\gamma^A > \gamma^M) &= \Lambda(0, 1|\gamma^A > \gamma^M) = \frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma^M}{1-\gamma^A}}, \\ a_1(1, \gamma^A > \gamma^M) &= \gamma^A \lambda + \gamma^M (1 - \lambda), \\ a_1(0, \gamma^A > \gamma^M) &= 1 - \gamma^A \lambda - \gamma^M (1 - \lambda). \end{aligned} \tag{A.13}$$

Clearly $\Lambda(1, 1|\gamma^A > \gamma^M) = \Lambda(0, 0|\gamma^A > \gamma^M) > \Lambda(1, 0|\gamma^A > \gamma^M) = \Lambda(0, 1|\gamma^A > \gamma^M)$. Now I examine whether indeed the analysts' best responses are consistent with the investor's conjectures. It is readily to show that both types of analysts have *reputational incentives* to tell the truth, i.e., $\Pi_R^J(0|\gamma^A > \gamma^M) > 0$ and $\Pi_R^J(1|\gamma^A > \gamma^M) < 0$, $J \in \{A, M\}$. The aligned analyst also has *current incentive* to tell the truth, while the misaligned an-

analyst always has *current incentive* to report 1. Therefore, the aligned analyst will indeed truthfully report her signal and the misaligned analyst will truthfully report 1 if she observes signal 1. On the other hand, if the misaligned analyst observes signal 0, she has *reputational incentive* to report 0, while her *current incentive* is to report 1.

$$\begin{aligned}\Pi_C^M(0|\gamma^A > \gamma^M) &= -x^M(a_1(1, \gamma^A > \gamma^M) - 1)^2 + x^M(a_1(0, \gamma^A > \gamma^M) - 1)^2 \\ &> 0,\end{aligned}$$

$$\begin{aligned}\Pi_R^M(0|\gamma^A > \gamma^M) &= (1 - x^M)(2\gamma^M - 1)[V^M(\Lambda(0, 0|\gamma^A > \gamma^M), \gamma^A) \\ &\quad - V^M(\Lambda(1, 0|\gamma^A > \gamma^M), \gamma^A)] \\ &> 0\end{aligned}$$

Therefore, if $x^M < \underline{x}^M(\lambda, \gamma^A, \gamma^M) \equiv 1 - \frac{1}{x_o^M(\lambda, \gamma^A, \gamma^M) + 1}$, where

$$\begin{aligned}x_o^M(\lambda, \gamma^A, \gamma^M) &\equiv \frac{(2\gamma^M - 1)[V^M(\Lambda(0, 0|\gamma^A > \gamma^M), \gamma^A) - V^M(\Lambda(1, 0|\gamma^A > \gamma^M), \gamma^A)]}{-(a_1(1, \gamma^A > \gamma^M) - 1)^2 + (a_1(0, \gamma^A > \gamma^M) - 1)^2} \\ &= \frac{(2\gamma^M - 1)[V^M(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma^M}{\gamma^A}}, \gamma^A) - V^M(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma^M}{1-\gamma^A}}, \gamma^A)]}{2(1 - \lambda)\gamma^M + 2\lambda\gamma^A - 1},\end{aligned}$$

then $\Pi_C^M(0|\gamma^A > \gamma^M) < \Pi_R^M(0|\gamma^A > \gamma^M)$. Hence, $\sigma_1^M(0|\gamma^A > \gamma^M) = 0$, consistent with the investor's conjecture.

To summarize, if $\gamma^A > \gamma^M$, the truth-telling equilibrium in which both types of analysts truthfully report their signals obtains in the first period for $x^M < \underline{x}^M(\lambda, \gamma^A, \gamma^M)$. ■

A.5 Proof of Proposition 2

As Section 1.3.1 and Proposition 1 have shown, if the investor conjectures that only the aligned analyst acquires information, both types of analysts truthfully report in the first period, and in the second period the aligned analyst reports truthfully and the misaligned analyst always reports 1, then the best responses of the analysts who have conjectured precision will indeed be consistent with the investor's conjectures. Now it remains to show that given the investor's conjectures, for moderate information-gathering costs c , it is indeed the aligned analyst's best response to become better informed, while the misaligned analyst's optimal choice is to keep the default precision.

First, I introduce the analyst's current reporting incentive and reputational reporting incentive when the analyst's precision is unobservable. Write $\Pi_C^J(s_1, \gamma^J | \tilde{\gamma})$ for the current expected gain to the type J analyst of reporting 1, rather than reporting 0, given signal s_1 , precision γ^J , and the investor's conjecture of analysts' precision $\tilde{\gamma} = (\tilde{\gamma}^A, \tilde{\gamma}^M)$. Therefore,

$$\begin{aligned} \Pi_C^A(s_1, \gamma^A | \tilde{\gamma}) &= -x^A \{ \gamma^A (a_1(1, \tilde{\gamma}) - s_1)^2 + (1 - \gamma^A) (a_1(1, \tilde{\gamma}) - (1 - s_1))^2 \} \\ &\quad + x^A \{ \gamma^A (a_1(0, \tilde{\gamma}) - s_1)^2 + (1 - \gamma^A) (a_1(0, \tilde{\gamma}) - (1 - s_1))^2 \}, \\ \Pi_C^M(1, \gamma^M | \tilde{\gamma}) &= \Pi_C^M(0, \gamma^M | \tilde{\gamma}) = -x^M (a_1(1, \tilde{\gamma}) - 1)^2 + x^M (a_1(0, \tilde{\gamma}) - 1)^2. \end{aligned}$$

Write $\Pi_R^J(s_1, \gamma^J | \tilde{\gamma})$ for the expected reputational gain to the type J analyst of reporting 0 rather than reporting 1, when she observes signal s_1 and has precision γ^J and the investor's

conjecture of analysts' precision is $\tilde{\gamma}$. Hence,

$$\begin{aligned}\Pi_R^A(s_1, \gamma^A | \tilde{\gamma}) &= (1 - x^A)\gamma^A[V^A(\Lambda(0, s_1 | \tilde{\gamma}), \gamma^A, \tilde{\gamma}) - V^A(\Lambda(1, s_1 | \tilde{\gamma}), \gamma^A, \tilde{\gamma})] \\ &+ (1 - x^A)(1 - \gamma^A)[V^A(\Lambda(0, 1 - s_1 | \tilde{\gamma}), \gamma^A, \tilde{\gamma}) - V^A(\Lambda(1, 1 - s_1 | \tilde{\gamma}), \gamma^A, \tilde{\gamma})], \\ \Pi_R^M(s_1, \gamma^M | \tilde{\gamma}) &= (1 - x^M)\gamma^M[V^M(\Lambda(0, s_1 | \tilde{\gamma}), \tilde{\gamma}) - V^M(\Lambda(1, s_1 | \tilde{\gamma}), \tilde{\gamma})] \\ &+ (1 - x^M)(1 - \gamma^M)[V^M(\Lambda(0, 1 - s_1 | \tilde{\gamma}), \tilde{\gamma}) - V^M(\Lambda(1, 1 - s_1 | \tilde{\gamma}), \tilde{\gamma})].\end{aligned}$$

Thus a type J analyst has a strict incentive to report 1 when she observes s_1 and has precision γ^J if and only if $\Pi_C^J(s_1, \gamma^J | \tilde{\gamma}) > \Pi_R^J(s_1, \gamma^J | \tilde{\gamma})$.

To examine analysts' best responses with regard to information acquisition decisions, we also need to check their best communication decisions off-the-equilibrium path given the investor's conjecture that $\tilde{\gamma}^A = \gamma_H$ and $\tilde{\gamma}^M = \gamma_L$.

(1) The aligned analyst:

(i) Given the investor's conjectures, for $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$, the aligned analyst's best response in the *second* period is to tell the truth no matter what her true precision is.

If the investor conjectures that $\tilde{\gamma}^A = \gamma_H$, then the aligned analyst with the conjectured precision $\gamma^A = \gamma_H$ will tell the truth. For the aligned analyst with precision γ_L , when she observes signal 1, she will compare her payoff conditional on sending message 0, $-\gamma_L(a_2(0, \lambda_2, \gamma_H) - 1)^2 - (1 - \gamma_L)(a_2(0, \lambda_2, \gamma_H) - 0)^2$, with her payoff conditional on sending message 1, $-\gamma_L(a_2(1, \lambda_2, \gamma_H) - 1)^2 - (1 - \gamma_L)(a_2(1, \lambda_2, \gamma_H) - 0)^2$. The dif-

ference (of sending message 0 over 1) is $[a_2(1, \lambda_2, \gamma_H) - a_2(0, \lambda_2, \gamma_H)][a_2(1, \lambda_2, \gamma_H) + a_2(0, \lambda_2, \gamma_H) - 2\gamma_L] < 0$. Hence the aligned analyst with precision γ_L will report 1 when her signal is 1. On the other hand, if the aligned analyst has precision γ_L and observes signal 0, she will again compare her payoff conditional on sending message 0, $-\gamma_L(a_2(0, \lambda_2, \gamma_H) - 0)^2 - (1 - \gamma_L)(a_2(0, \lambda_2, \gamma_H) - 1)^2$, with her payoff conditional on sending message 1, $-\gamma_L(a_2(1, \lambda_2, \gamma_H) - 0)^2 - (1 - \gamma_L)(a_2(1, \lambda_2, \gamma_H) - 1)^2$. The difference (of sending message 0 over 1) is

$$[a_2(1, \lambda_2, \gamma_H) - a_2(0, \lambda_2, \gamma_H)][a_2(1, \lambda_2, \gamma_H) + a_2(0, \lambda_2, \gamma_H) - 2(1 - \gamma_L)] > 0 \quad \text{for } \gamma_L \geq 0.75.$$

Therefore, the aligned analyst with precision γ_L will report 0 when her signal is 0. Then

$$\begin{aligned} & V^A(\lambda_2, \gamma^A = \gamma_H, \tilde{\gamma}^A = \gamma_H) - V^A(\lambda_2, \gamma^A = \gamma_L, \tilde{\gamma}^A = \gamma_H) \\ &= -\frac{1}{2}(\gamma_H - \gamma_L)[(a_2(1, \lambda_2, \gamma_H) - 1)^2 - (a_2(0, \lambda_2, \gamma_H) - 1)^2 + (a_2(0, \lambda_2, \gamma_H))^2 - (a_2(1, \lambda_2, \gamma_H))^2] \\ &= (\gamma_H - \gamma_L)[a_2(1, \lambda_2, \gamma_H) - a_2(0, \lambda_2, \gamma_H)] \\ &= (\gamma_H - \gamma_L)\frac{2\gamma_H - 1}{2 - \lambda_2}. \end{aligned}$$

(ii) Given the investor's conjectures, for $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$, the aligned analyst's best response in the *first* period is to tell the truth no matter what her true precision is.

Given the investor's conjectures, the type inference function and the state inference function are derived as in (A.13). By Proposition 1, the aligned analyst with the conjectured

precision γ_H will indeed tell the truth. For the aligned analyst with precision γ_L ,

$$\begin{aligned}
& \Pi_C^A(1, \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
&= x^A [a_1(0, \gamma_H, \gamma_L) - a_1(1, \gamma_H, \gamma_L)] [a_1(0, \gamma_H, \gamma_L) + a_1(1, \gamma_H, \gamma_L) - 2\gamma_L] \\
&= x^A [a_1(0, \gamma_H, \gamma_L) - a_1(1, \gamma_H, \gamma_L)] (1 - 2\gamma_L) \\
&> 0, \\
& \Pi_R^A(1, \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
&= (1 - x^A) (1 - 2\gamma_L) [V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_L, \tilde{\gamma}^A = \gamma_H) - V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_L, \tilde{\gamma}^A = \gamma_H)]
\end{aligned}$$

Given that the aligned analyst tells the truth off-the-equilibrium path in the second period,

$$\begin{aligned}
V^A(\lambda_2, \gamma_L, \tilde{\gamma}^A = \gamma_H) &= -\frac{1}{2}\gamma_L [a_2(1, \lambda_2, \gamma_H) - 1]^2 - \frac{1}{2}(1 - \gamma_L) [a_2(0, \lambda_2, \gamma_H) - 1]^2 \\
&\quad - \frac{1}{2}(1 - \gamma_L) [a_2(1, \lambda_2, \gamma_H) - 0]^2 - \frac{1}{2}\gamma_L [a_2(0, \lambda_2, \gamma_H) - 0]^2
\end{aligned}$$

For $\gamma_H \rightarrow 1$, $\Lambda(0, 0 | \gamma_H, \gamma_L) \rightarrow \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}$ and $\Lambda(1, 0 | \gamma_H, \gamma_L) \rightarrow 0$. Therefore,

$$\begin{aligned}
& \lim_{\gamma_H \rightarrow 1} \{V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_L, \tilde{\gamma}^A = \gamma_H) - V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_L, \tilde{\gamma}^A = \gamma_H)\} \\
&= -\frac{1}{2}\gamma_L [a_2(1, \lambda_2 = \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}, \gamma_H) - 1]^2 - \frac{1}{2}(1 - \gamma_L) [a_2(1, \lambda_2 = \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}, \gamma_H)]^2 \\
&\quad + \frac{1}{2}\gamma_L [a_2(1, \lambda_2 = 0, \gamma_H) - 1]^2 + \frac{1}{2}(1 - \gamma_L) [a_2(1, \lambda_2 = 0, \gamma_H)]^2 \\
&= \frac{1}{2} \underbrace{[a_2(1, 0, \gamma_H) - a_2(1, \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}, \gamma_H)]}_{<0} [a_2(1, 0, \gamma_H) + a_2(1, \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}, \gamma_H) - 2\gamma_L]
\end{aligned}$$

Notice that

$$\begin{aligned}
\lim_{\gamma_H \rightarrow 1} \left\{ a_2(1, 0, \gamma_H) + a_2\left(1, \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}, \gamma_H\right) - 2\gamma_L \right\} &= \frac{1}{2} + \frac{1}{2 - \frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}} - 2\gamma_L \\
&= \frac{\lambda + (\lambda + 2(1 - \lambda)\gamma_L)(2 - 4\gamma_L)}{4\gamma_L(1 - \lambda) + 2\lambda} \\
&< 0 \quad \text{for } \gamma_L \geq 0.75
\end{aligned}$$

Therefore,

$$\lim_{\gamma_H \rightarrow 1} \Pi_R^A(1, \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) < 0 \quad \text{for } \gamma_L \geq 0.75.$$

Hence, the aligned analyst with precision $\gamma^A = \gamma_L$ will truthfully report 1 if she observes 1, i.e., $\sigma_1^A(1, \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) = 1$. Analogously, it is easy to show that $\sigma_1^A(0, \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) = 0$ for $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$. To summarize, given the investor's conjecture that $\tilde{\gamma}^A = \gamma_H$, $\tilde{\gamma}^M = \gamma_L$ and both types of analysts tell the truth in the first period, for $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$, the aligned analyst will tell the truth on and off the equilibrium path.

(iii) The aligned analyst's choice of precision:

Given that the aligned analyst tells the truth in the first period on and off the equilibrium path, her utility at the information acquisition stage if she becomes better informed is as

follows (the arguments in $a_1(\cdot)$ are m_1 , $\tilde{\gamma}^A$, and $\tilde{\gamma}^M$):

$$\begin{aligned}
& U_0^A(\gamma^A = \gamma_H | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
= & \frac{1}{2} \gamma_H [-x^A (a_1(1, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A) V^A(\Lambda(1, 1 | \gamma_H, \gamma_L), \gamma_H, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_H) [-x^A (a_1(0, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A) V^A(\Lambda(0, 1 | \gamma_H, \gamma_L), \gamma_H, \gamma_H)] \\
& + \frac{1}{2} \gamma_H [-x^A (a_1(0, \gamma_H, \gamma_L))^2 + (1 - x^A) V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_H) [-x^A (a_1(1, \gamma_H, \gamma_L))^2 + (1 - x^A) V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H)] - c.
\end{aligned}$$

If the aligned analyst keeps her default precision, it's straightforward that

$$\begin{aligned}
& U_0^A(\gamma^A = \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
= & \frac{1}{2} \gamma_L [-x^A (a_1(1, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A) V^A(\Lambda(1, 1 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_L) [-x^A (a_1(0, \gamma_H, \gamma_L) - 1)^2 + (1 - x^A) V^A(\Lambda(0, 1 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
& + \frac{1}{2} \gamma_L [-x^A (a_1(0, \gamma_H, \gamma_L))^2 + (1 - x^A) V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
& + \frac{1}{2} (1 - \gamma_L) [-x^A (a_1(1, \gamma_H, \gamma_L))^2 + (1 - x^A) V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)].
\end{aligned}$$

Define Δ^A as the benefit to the aligned analyst from acquiring information. Then,

$$\begin{aligned}
\Delta^A &\equiv U_0^A(\gamma^A = \gamma_H | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) + c - U_0^A(\gamma^A = \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
&= \underbrace{(\gamma_H - \gamma_L)x^A[a_1(1, \gamma_H, \gamma_L) - a_1(0, \gamma_H, \gamma_L)]}_{PE_1^A} \\
&+ \underbrace{(\gamma_H - \gamma_L)(1 - x^A)[V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H)]}_{RE^A} \\
&+ \underbrace{\gamma_L(1 - x^A)[V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)]}_{\text{Part i of } PE_2^A} \\
&+ \underbrace{(1 - \gamma_L)(1 - x^A)[V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_L, \gamma_H)]}_{\text{Part ii of } PE_2^A} \\
&= (\gamma_H - \gamma_L)x^A[2\gamma_L(1 - \lambda) + 2\gamma_H\lambda - 1] \\
&+ (\gamma_H - \gamma_L)(1 - x^A)[V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma_L}{\gamma_H}}, \gamma_H, \gamma_H) - V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma_L}{1-\gamma_H}}, \gamma_H, \gamma_H)] \\
&+ \gamma_L(1 - x^A)[V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma_L}{\gamma_H}}, \gamma_H, \gamma_H) - V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma_L}{\gamma_H}}, \gamma_L, \gamma_H)] \\
&+ (1 - \gamma_L)(1 - x^A)[V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma_L}{1-\gamma_H}}, \gamma_H, \gamma_H) - V^A(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma_L}{1-\gamma_H}}, \gamma_L, \gamma_H)] \\
&\equiv \bar{\Delta}
\end{aligned}$$

where PE_t^A stands for Precision Effect in period t for the aligned analyst and RE^A stands for Reputation Effect for the aligned analyst.

Clearly, for $c < \bar{\Delta}$, the aligned analyst will have incentive to acquire information,

consistent with the investor's conjecture. Also notice that

$$\begin{aligned}
PE_1^A &= (\gamma_H - \gamma_L)x^A[a_1(1, \gamma_H, \gamma_L) - a_1(0, \gamma_H, \gamma_L)] \\
&= (\gamma_H - \gamma_L)x^A[2\gamma_L(1 - \lambda) + 2\gamma_H\lambda - 1] \\
&> (\gamma_H - \gamma_L)x^A(2\gamma_L - 1) \\
&> (\gamma_H - \gamma_L)x^A(\gamma_H - \frac{1}{2}) \text{ for } \gamma_L \geq 0.75 \\
PE_2^A &= \gamma_L(1 - x^A)[V^A(\Lambda(0, 0|\gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(0, 0|\gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
&+ (1 - \gamma_L)(1 - x^A)[V^A(\Lambda(1, 0|\gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(1, 0|\gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
&> (1 - x^A)[V^A(\Lambda(1, 0|\gamma_H, \gamma_L), \gamma_H, \gamma_H) - V^A(\Lambda(1, 0|\gamma_H, \gamma_L), \gamma_L, \gamma_H)] \\
&\geq (1 - x^A)(\gamma_H - \gamma_L)(\gamma_H - \frac{1}{2})
\end{aligned}$$

Hence,

$$\begin{aligned}
\bar{\Delta} &= PE_1^A + RE^A + PE_2^A \\
&> PE_1^A + PE_2^A \\
&> (\gamma_H - \gamma_L)(\gamma_H - \frac{1}{2})
\end{aligned}$$

(2) The misaligned analyst:

(i) Given the investor's conjectures, the misaligned analyst's best response in the *second* period is to always report 1 no matter what her true precision is.

(ii) Given the investor's conjectures, for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, the misaligned analyst's best response in the *first* period is to report truthfully no matter what her true precision is.

Note that the misaligned analyst would have more reputational incentive to tell the truth when the precision is higher (because $\Lambda(0, 0|\gamma_H, \gamma_L) > \Lambda(1, 0|\gamma_H, \gamma_L)$ and $\Lambda(0, 1|\gamma_H, \gamma_L) < \Lambda(1, 1|\gamma_H, \gamma_L)$), while her current incentive to report 1 is independent of her precision. Therefore, having a higher precision would make the misaligned analyst more likely to tell the truth. According to the argument in Proposition 1, for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, the misaligned analyst with the equilibrium precision $\gamma^M = \gamma_L$ will tell the truth, hence the misaligned analyst with precision $\gamma^M = \gamma_H$ will also tell the truth.

(iii) The misaligned analyst's choice of precision:

The misaligned analyst's utility at the information acquisition stage conditional on keeping the default precision is:

$$\begin{aligned}
& U_0^M(\gamma^M = \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
= & \frac{1}{2}[-x^M(a_1(1, \gamma_H, \gamma_L) - 1)^2] + \frac{1}{2}[-x^M(a_1(0, \gamma_H, \gamma_L) - 1)^2] \\
& + \frac{1}{2}(1 - x^M)[\gamma_L V^M(\Lambda(1, 1|\gamma_H, \gamma_L), \gamma_H) + (1 - \gamma_L)V^M(\Lambda(0, 1|\gamma_H, \gamma_L), \gamma_H)] \\
& + \frac{1}{2}(1 - x^M)[\gamma_L V^M(\Lambda(0, 0|\gamma_H, \gamma_L), \gamma_H) + (1 - \gamma_L)V^M(\Lambda(1, 0|\gamma_H, \gamma_L), \gamma_H)].
\end{aligned}$$

The misaligned analyst's utility at the information acquisition stage conditional on acquir-

ing information is:

$$\begin{aligned}
& U_0^M(\gamma^M = \gamma_H | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
= & \frac{1}{2}[-x^M(a_1(1, \gamma_H, \gamma_L) - 1)^2] + \frac{1}{2}[-x^M(a_1(0, \gamma_H, \gamma_L) - 1)^2] \\
& + \frac{1}{2}(1 - x^M)[\gamma_H V^M(\Lambda(1, 1 | \gamma_H, \gamma_L), \gamma_H) + (1 - \gamma_H)V^M(\Lambda(0, 1 | \gamma_H, \gamma_L), \gamma_H)] \\
& + \frac{1}{2}(1 - x^M)[\gamma_H V^M(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H) + (1 - \gamma_H)V^M(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H)].
\end{aligned}$$

Then the benefit of becoming better informed for the misaligned analyst is

$$\begin{aligned}
\Delta^M & \equiv U_0^M(\gamma^M = \gamma_H | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) + c - U_0^M(\gamma^M = \gamma_L | \tilde{\gamma}^A = \gamma_H, \tilde{\gamma}^M = \gamma_L) \\
& = \frac{1}{2}(1 - x^M)(\gamma_H - \gamma_L)[V^M(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H) - V^M(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H)] \\
& + \frac{1}{2}(1 - x^M)(\gamma_H - \gamma_L)[V^M(\Lambda(1, 1 | \gamma_H, \gamma_L), \gamma_H) - V^M(\Lambda(0, 1 | \gamma_H, \gamma_L), \gamma_H)] \\
& = \underbrace{(1 - x^M)(\gamma_H - \gamma_L)[V^M(\Lambda(0, 0 | \gamma_H, \gamma_L), \gamma_H) - V^M(\Lambda(1, 0 | \gamma_H, \gamma_L), \gamma_H)]}_{RE^M} \\
& = (1 - x^M)(\gamma_H - \gamma_L)\left[V^M\left(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{\gamma_L}{\gamma_H}}, \gamma_H\right) - V^M\left(\frac{1}{1 + \frac{1-\lambda}{\lambda} \frac{1-\gamma_L}{1-\gamma_H}}, \gamma_H\right)\right] \\
& \equiv \underline{\Delta}
\end{aligned}$$

For $c > \underline{\Delta}$, the misaligned analyst will not acquire information, again consistent with the investor's conjecture.

Last, we need to check whether $\underline{\Delta} < c < \bar{\Delta}$ is an empty set. Notice that

$$\begin{aligned}
\underline{\Delta} &< (1 - x^M)(\gamma_H - \gamma_L)[V^M(\lambda_2 = 1, \gamma_H) - V^M(\lambda_2 = 0, \gamma_H)] \\
&= (1 - x^M)(\gamma_H - \gamma_L)(\gamma_H - \frac{1}{2})(2 - \gamma_H - \frac{1}{2}) \\
&< (1 - x^M)(\gamma_H - \gamma_L)(\gamma_H - \frac{1}{2}) \\
&< \bar{\Delta}.
\end{aligned}$$

Therefore, for $\underline{\Delta} < c < \bar{\Delta}$, the aligned analyst will acquire information and the misaligned analyst will keep her default precision, consistent with the investor' conjectures. ■

A.6 Proof of Proposition 3

Define $c^o \equiv \min\{(1 - x^A)(\gamma_H - \gamma_L)(\gamma_L - \frac{1}{2}), \underline{\Delta}\}$. The idea of the proof is to show that (1) If $c < c^o$, then there exists an equilibrium in which $\gamma^{A*} = \gamma_H$, $\gamma^{M*} = \gamma_L$, first period communication is babbling and second period communication is informative. (2) For $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, $x^A \in (0, \underline{x}^A(\lambda, \gamma^A, \gamma^M, x^M)] \cup [x^M, \infty)$, $\gamma_H \rightarrow 1$ and $c < c^o$, there is no other informative equilibria.

(1) Suppose the investor conjectures that $\tilde{\gamma}^A = \gamma_H$, $\tilde{\gamma}^M = \gamma_L$, first period communication is babbling, and second period communication is informative. Then the analysts' best responses in each communication period are consistent with these conjectures. For the mis-

aligned analyst, both her first period communication (babbling) and second period communication (always reports 1) are independent of her actual precision, hence she doesn't have incentive to acquire information. However, the aligned analyst's benefit from acquiring information is $(1 - x^A)[V^A(\lambda, \gamma_H, \gamma_H) - V^A(\lambda, \gamma_L, \gamma_H)] = (1 - x^A)(\gamma_H - \gamma_L)\frac{2\gamma_H - 1}{2 - \lambda} > c^o$, hence the aligned analyst will acquire information. Therefore, for $c < c^o$, there exists an equilibrium in which $\gamma^{A*} = \gamma_H$, $\gamma^{M*} = \gamma_L$, first period communication is babbling and second period communication is informative.

(2) I prove this result by ruling out the impossible equilibria one by one. As shown in the main text, informative communication always obtains in the second period.

(i) Suppose in the postulated equilibrium, the aligned analyst keeps the default precision γ_L . Then, in the second period, the aligned analyst with the conjectured precision γ_L will always tell the truth (similar arguments as in Section 1.3.1). In addition, the aligned analyst with $\gamma^A = \gamma_H$ will also tell the truth. Hence,

$$\begin{aligned}
& V^A(\lambda_2, \gamma^A = \gamma_H, \tilde{\gamma}^A = \gamma_L) - V^A(\lambda_2, \gamma^A = \gamma_L, \tilde{\gamma}^A = \gamma_L) \\
&= -\frac{1}{2}(\gamma_H - \gamma_L)[(a_2(1, \lambda_2, \gamma_L) - 1)^2 - (a_2(0, \lambda_2, \gamma_L) - 1)^2 + (a_2(0, \lambda_2, \gamma_L))^2 - (a_2(1, \lambda_2, \gamma_L))^2] \\
&= (\gamma_H - \gamma_L)[a_2(1, \lambda_2, \gamma_L) - a_2(0, \lambda_2, \gamma_L)] \\
&= \frac{(\gamma_H - \gamma_L)(2\gamma_L - 1)}{2 - \lambda_2} \geq (\gamma_H - \gamma_L)\frac{(2\gamma_L - 1)}{2}.
\end{aligned} \tag{A.14}$$

If first period communication is babbling, then the aligned analyst's benefit from acquiring information is $(1 - x^A)[V^A(\lambda, \gamma_H, \gamma_L) - V^A(\lambda, \gamma_L, \gamma_L)] > (1 - x^A)(\gamma_H - \gamma_L)(\gamma_L - \frac{1}{2}) \geq$

c^o , therefore the aligned analyst will have incentive to acquire information, upsetting the postulated equilibria.

If first period communication is informative, then according to Claim 2, $\sigma_1^A(0, \gamma_L | \gamma_L, \tilde{\gamma}^M) = 0$ and $\sigma_1^A(1, \gamma_L | \gamma_L, \tilde{\gamma}^M) > 0$. Therefore, by the definition of mixed strategy, the aligned analyst's utility if she keeps the default precision (given the investor's equilibrium conjectures) can be calculated by treating $\sigma_1^A(1, \gamma_L | \gamma_L, \tilde{\gamma}^M) = 1$. On the other hand, by reveal preference, the aligned analyst's utility when having precision γ_H (given the investor's conjecture that $\tilde{\gamma}^A = \gamma_L$) under the optimal communication strategy is (weakly) greater than her utility under the truth-telling strategy. Therefore, the aligned analyst's benefit from acquiring information is given by:

$$\begin{aligned}
& U_0^A(\gamma^A = \gamma_H | \tilde{\gamma}^A = \gamma_L, \tilde{\gamma}^M) + c - U_0^A(\gamma^A = \gamma_L | \tilde{\gamma}^A = \gamma_L, \tilde{\gamma}^M) \\
\geq & \underbrace{(\gamma_H - \gamma_L)x^A[a_1(1, \gamma_L, \tilde{\gamma}^M) - a_1(0, \gamma_L, \tilde{\gamma}^M)]}_I \\
& + \underbrace{\frac{1}{2}(\gamma_H - \gamma_L)(1 - x^A)[V^A(\Lambda(0, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L) - V^A(\Lambda(1, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)]}_{II} \\
& + \underbrace{\frac{1}{2}(\gamma_H - \gamma_L)(1 - x^A)[V^A(\Lambda(1, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L) - V^A(\Lambda(0, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)]}_{III} \\
& + \frac{1}{2}\gamma_H(1 - x^A)[V^A(\Lambda(0, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_H, \gamma_L) - V^A(\Lambda(0, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)] \\
& + \frac{1}{2}(1 - \gamma_H)(1 - x^A)[V^A(\Lambda(1, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_H, \gamma_L) - V^A(\Lambda(1, 0 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)] \\
& + \frac{1}{2}\gamma_H(1 - x^A)[V^A(\Lambda(1, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_H, \gamma_L) - V^A(\Lambda(1, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)] \\
& + \frac{1}{2}(1 - \gamma_H)(1 - x^A)[V^A(\Lambda(0, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_H, \gamma_L) - V^A(\Lambda(0, 1 | \gamma_L, \tilde{\gamma}^M), \gamma_L, \gamma_L)]
\end{aligned}$$

Notice that $\sigma_1^A(1, \gamma_L | \gamma_L, \tilde{\gamma}^M) > 0$ implies that $\Pi_C^A(1, \gamma_L | \gamma_L, \tilde{\gamma}^M) \geq \Pi_R^A(1, \gamma_L | \gamma_L, \tilde{\gamma}^M)$,

which implies that $I + II + III > 0$. Hence,

$$U_0^A(\gamma^A = \gamma_H | \tilde{\gamma}^A = \gamma_L, \tilde{\gamma}^M) + c - U_0^A(\gamma^A = \gamma_L | \tilde{\gamma}^A = \gamma_L, \tilde{\gamma}^M) > (1 - x^A)(\gamma_H - \gamma_L)(\gamma_L - \frac{1}{2}) \geq c^o.$$

Therefore, the aligned analyst will have incentive to acquire information, again upsetting the postulated equilibria where the aligned analyst keeps the default precision.

(ii) Suppose in the postulated equilibria, both analysts acquire information. Then by Claim 2, $\sigma_1^M(1, \gamma_H | \gamma_H, \gamma_H) \geq \sigma_1^M(0, \gamma_H | \gamma_H, \gamma_H) > 0$, which implies that the misaligned analyst's utility if she acquires information (fixing the investor's postulated equilibrium strategy) can be calculated by treating $\sigma_1^M(1, \gamma_H | \gamma_H, \gamma_H) = \sigma_1^M(0, \gamma_H | \gamma_H, \gamma_H) = 1$. On the other hand, $\sigma_1^M(1, \gamma_H | \gamma_H, \gamma_H) \geq \sigma_1^M(0, \gamma_H | \gamma_H, \gamma_H)$ implies that $\Pi_R^M(1, \gamma_H | \gamma_H, \gamma_H) \leq \Pi_R^M(0, \gamma_H | \gamma_H, \gamma_H)$, which in turn implies that

$$V^M(\Lambda(0, 1 | \gamma_H, \gamma_H), \gamma_H) - V^M(\Lambda(1, 1 | \gamma_H, \gamma_H), \gamma_H) \leq V^M(\Lambda(0, 0 | \gamma_H, \gamma_H), \gamma_H) - V^M(\Lambda(1, 0 | \gamma_H, \gamma_H), \gamma_H).$$

Therefore, $\Pi_R^M(1, \gamma_L | \gamma_H, \gamma_H) \leq \Pi_R^M(0, \gamma_L | \gamma_H, \gamma_H) \leq \Pi_R^M(0, \gamma_H | \gamma_H, \gamma_H)$. Recall that $\Pi_C^M(\cdot)$ does not depend on the misaligned analyst's signal nor her true precision. As a result, $\sigma_1^M(1, \gamma_L | \gamma_H, \gamma_H) \geq \sigma_1^M(0, \gamma_L | \gamma_H, \gamma_H) \geq \sigma_1^M(0, \gamma_H | \gamma_H, \gamma_H) > 0$, which implies that the misaligned analyst's utility if she keeps precision γ_L (fixing the investor's postulated equilibrium strategy) can again be calculated by treating $\sigma_1^M(1, \gamma_L | \gamma_H, \gamma_H) =$

$\sigma_1^M(0, \gamma_L | \gamma_H, \gamma_H) = 1$. As a result, the misaligned analyst has no incentive to increase her precision because the benefit of doing so is zero. Thus there do not exist equilibria in which both analysts acquire information.

(iii) Suppose in the postulated equilibria, only the aligned analyst acquires information and communication in each period is informative. Then by Claim 1, for $\gamma_H \rightarrow 1$ and $x^A \in (0, \underline{x}^A(\lambda, \gamma_H, \gamma_L, x^M)] \cup [x^M, \infty)$, in any first period informative communication, it has to be true that $\sigma_1^A(s_1, \gamma_H | \gamma_H, \gamma_L) = s_1$ and $\sigma_1^M(1, \gamma_H | \gamma_H, \gamma_L) = 1$. Therefore, for $x^M < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$, the *unique* first period informative communication is such that both analysts report their signals truthfully. Then the misaligned analyst's benefit from acquiring information (by the proof of Proposition 2) is

$$\Delta^M = \underline{\Delta} \geq c^o.$$

Hence for $c < c^o$, the misaligned analyst has incentive to increase her precision, which upsets the postulated equilibria. ■

A.7 Proof of Corollary 2

According to Proposition 3, if $x^M = x^A < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$ and $\gamma_H \rightarrow 1$, then for $c < c^o$, the first period communication is babbling. In addition, given $x^M = x^A = x$, $\gamma_H \rightarrow 1$ and $\gamma_L \geq 0.75$, it is readily to verify that $c^o = \underline{\Delta} \rightarrow \frac{1}{2}(1-x)(1-\gamma_L)[\frac{1}{2} - \frac{2\gamma_L^2(1-\lambda)^2}{(\lambda+2\gamma_L(1-\lambda))^2}]$. At

the same time, according to the proof of Proposition 2, for $c = \underline{\Delta} + \varepsilon > c^o$, both types of analysts tell the truth in the first period.

(1) Social welfare of having $c = c_2 = \underline{\Delta} + \varepsilon > c^o$ for $x < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$.

In this case, the equilibrium is such that only the aligned analyst acquires information, both types of analysts tell the truth in the first period, and in the second period, the aligned analyst tells the truth and misaligned analyst always reports 1.

Given this equilibrium, social welfare is calculated. To that end, I first calculate the investor's expected utility for the second period (with posterior analyst reputation λ_2):

$$\begin{aligned}
V^I(\lambda_2) &= -Pr(w_2 = 1, m_2 = 1)[a_2(1, \lambda_2, \gamma_H) - 1]^2 - Pr(w_2 = 1, m_2 = 0)[a_2(0, \lambda_2, \gamma_H) - 1]^2 \\
&\quad - Pr(w_2 = 0, m_2 = 1)[a_2(1, \lambda_2, \gamma_H) - 0]^2 - Pr(w_2 = 0, m_2 = 0)[a_2(0, \lambda_2, \gamma_H) - 0]^2 \\
&= -\frac{1}{2}(\gamma_H \lambda_2 + 1 - \lambda_2)[a_2(1, \lambda_2, \gamma_H) - 1]^2 - \frac{1}{2}(1 - \gamma_H)\lambda_2[a_2(0, \lambda_2, \gamma_H) - 1]^2 \\
&\quad - \frac{1}{2}((1 - \gamma_H)\lambda_2 + 1 - \lambda_2)[a_2(1, \lambda_2, \gamma_H) - 0]^2 - \frac{1}{2}\gamma_H \lambda_2[a_2(0, \lambda_2, \gamma_H) - 0]^2 \\
&= -\frac{1 - [1 + 2(\gamma_H - 1)\gamma_H]\lambda_2}{2(2 - \lambda_2)}.
\end{aligned}$$

It's easy to show that $V^I(\lambda_2)$ is increasing and convex in λ_2 .

Note that in this equilibrium, both types of analysts truthfully report their signals in the first period. This amounts to that the investor observes the signal himself in the first period with (average) precision $\lambda\gamma_H + (1 - \lambda)\gamma_L$. Hence $a_1(1) = \lambda\gamma_H + (1 - \lambda)\gamma_L$ and $a_1(0) = 1 - a_1(1) = 1 - \lambda\gamma_H - (1 - \lambda)\gamma_L$, and the investor's ex ante expected utility is as

follows:

$$\begin{aligned}
\lim_{\gamma_H \rightarrow 1} EU^I(c = c_2) &= \lim_{\gamma_H \rightarrow 1} \left\{ \frac{1}{2}[\lambda\gamma_H + (1 - \lambda)\gamma_L][-x(a_1(1) - 1)^2 + (1 - x)V^I(\Lambda(1, 1))] \right. \\
&\quad + \frac{1}{2}(1 - \lambda\gamma_H - (1 - \lambda)\gamma_L)[-x(a_1(0) - 1)^2 + (1 - x)V^I(\Lambda(0, 1))] \\
&\quad + \frac{1}{2}(1 - \lambda\gamma_H - (1 - \lambda)\gamma_L)[-x(a_1(1) - 0)^2 + (1 - x)V^I(\Lambda(1, 0))] \\
&\quad \left. + \frac{1}{2}[\lambda\gamma_H + (1 - \lambda)\gamma_L][-x(a_1(0) - 0)^2 + (1 - x)V^I(\Lambda(0, 0))] \right\} \\
&= -xa_1(1)a_1(0) + (1 - x)\left\{(1 - \lambda)(1 - \gamma_L)V^I(0) + (\lambda + (1 - \lambda)\gamma_L)V^I\left(\frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}\right)\right\}.
\end{aligned}$$

The aligned analyst's expected utility is ⁴:

$$\begin{aligned}
\lim_{\gamma_H \rightarrow 1} EU^A(c = c_2) &= \frac{1}{2}[-x(a_1(1) - 1)^2 + (1 - x)V^A(\Lambda(1, 1))] \\
&\quad + \frac{1}{2}[-x(a_1(0) - 0)^2 + (1 - x)V^A(\Lambda(0, 0))] - c_2 \\
&= -\frac{x}{2}[(a_1(1) - 1)^2 + a_1(0)^2] + (1 - x)V^A\left(\frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}\right) - c_2 \\
&= -xa_1(0)^2 + (1 - x)V^A\left(\frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}\right) - c_2.
\end{aligned}$$

The misaligned analyst's expected utility is:

$$\lim_{\gamma_H \rightarrow 1} EU^M(c = c_2) = -\frac{x}{2}[a_1(1)^2 + a_1(0)^2] + (1 - x)\left\{\gamma_L V^M\left(\frac{\lambda}{\lambda + (1 - \lambda)\gamma_L}\right) + (1 - \gamma_L)V^M(0)\right\}.$$

⁴Here, I suppress dependence of $V^J(\cdot)$ on the aligned analyst's equilibrium precision.

Adding the players' utilities, social welfare of having $c = c_2$ equals:

$$W(c = c_2) = EU^I(c = c_2) + \lambda EU^A(c = c_2) + (1 - \lambda)EU^M(c = c_2).$$

(2) Social welfare of having $c = c_1 < c^o$ for $x < \underline{x}^M(\lambda, \gamma_H, \gamma_L)$.

According to Proposition 3, in this case, the first period communication is babbling.

Based on this equilibrium, the investor's ex ante expected utility is calculated analogously:

$$EU^I(c = c_1) = -\frac{x}{4} + (1 - x)V^I(\lambda).$$

The aligned analyst's expected utility is:

$$EU^A(c = c_1) = -\frac{x}{4} + (1 - x)V^A(\lambda) - c_1.$$

The misaligned analyst's expected utility is:

$$EU^M(c = c_1) = -\frac{x}{4} + (1 - x)V^M(\lambda).$$

Adding up all players' utilities, social welfare of having $c = c_1$ is

$$W(c = c_1) = EU^I(c = c_1) + \lambda EU^A(c = c_1) + (1 - \lambda)EU^M(c = c_1).$$

(3) Now I compare the the social welfare of having $c = c_1$ and $c = c_2$.

For the investor,

$$\begin{aligned} & \lim_{\gamma_H \rightarrow 1} [EU^I(c = c_2) - EU^I(c = c_1)] \\ = & \underbrace{x\left[\frac{1}{4} - a_1(1)a_1(0)\right]}_{\text{first period, } > 0} + \underbrace{(1-x)[(1-\lambda)(1-\gamma_L)V^I(0) + (\lambda + (1-\lambda)\gamma_L)V^I\left(\frac{\lambda}{\lambda + (1-\lambda)\gamma_L}\right) - V^I(\lambda)]}_{\text{second period, } > 0}. \end{aligned}$$

The fact that the first period difference is greater than 0 comes directly from $a_1(1) + a_1(0) =$

1; and the second period difference being greater than 0 is due to the convexity of V^I .

For the aligned analyst,

$$\begin{aligned} & \lim_{\gamma_H \rightarrow 1} [EU^A(c = c_2) - EU^A(c = c_1)] \\ = & \underbrace{x\left(\frac{1}{4} - a_1(0)^2\right)}_{\text{first period, } > 0} + \underbrace{(1-x)\left[V^A\left(\frac{\lambda}{\lambda + (1-\lambda)\gamma_L}\right) - V^A(\lambda)\right]}_{\text{second period, } > 0} - \underbrace{(c_2 - c_1)}_{\text{cost difference, } \rightarrow 0}. \end{aligned}$$

For the misaligned analyst,

$$\begin{aligned} & \lim_{\gamma_H \rightarrow 1} [EU^M(c = c_2) - EU^M(c = c_1)] \\ = & \underbrace{-\frac{x}{2}[a_1(1)^2 + a_1(0)^2 - \frac{1}{2}]}_{\text{first period, } < 0} - \underbrace{\frac{(1-x)(1-\gamma_L)\lambda^2[(4-3\lambda)\lambda + 4\gamma_L(1-\lambda)(3-2\lambda)]}{4(2-\lambda)^2(2\gamma_L(1-\lambda) + \lambda)^2}}_{\text{second period, } < 0}. \end{aligned}$$

Social welfare:

$$\lim_{\gamma_H \rightarrow 1} [W(c = c_2) - W(c = c_1)] = \underbrace{\frac{1}{2}x\lambda(a_1(1) - a_1(0))}_{\text{first period, } > 0} - \underbrace{\lambda(c_2 - c_1)}_{\text{cost difference, } \rightarrow 0} .$$

Thus, social welfare increases when information-gathering costs increase from c_1 to c_2 .

■

Appendix B

Proofs for Chapter 2

B.1 Preliminaries.

We follow the procedure in Crawford and Sobel (1982). The prior variance of p given that p is uniformly distributed over $[0, P]$ is $\bar{\sigma}_P^2 = P^2/12$. As mentioned in the text, we focus on the most informative equilibrium, i.e., the one with the finest partitioning. The expected information loss regarding the random variable $a \sim U[0, A]$, conditional on a cheap talk report made by a privately informed party equals (we suppress the argument b whenever there is no scope for confusion):

$$\sigma_A^2 = \frac{A^2}{12(N(A, b))^2} + \frac{b^2((N(A, b))^2 - 1)}{3}. \quad (\text{B.1})$$

Here, $N(A, b)$ is the maximum size of the equilibrium partition under centralization, as given by

$$N(A, b) = \left\langle -\frac{1}{2} + \frac{1}{2} \left(1 + \frac{2A}{b}\right)^{\frac{1}{2}} \right\rangle,$$

where $\langle t \rangle$ denotes the smallest integer greater or equal to t . The number of partitions, $N(A, b)$, jumps from $n + 1$ to n at $b = b_n(A) \equiv \frac{2A}{(2n+1)^2 - 1}$. In particular, there will be no information conveyed ($N(\cdot) = 1$) if and only if $b > b_1(A) = A/4$. Lastly, σ_P^2 is computed analogously.

B.2 Proof of Lemma 3.

Simple algebra establishes the profit differential between delegation and centralization, $\pi^D(x, b) - \pi^C(x, b) = (1 - q_a(x))Z(x, b)$, as stated in (2.8), where¹

$$Z(x, b) \equiv \sigma_A^2 - [b^2 + q_p(x)\sigma_P^2].$$

Consider first the case of small CEO biases. Following the arguments in Dessein (2002), in the limit, as $b \rightarrow 0$, $N(A, b) \simeq \sqrt{\frac{A}{2b}}$. Therefore,

$$\lim_{b \rightarrow 0} \sigma_A^2 = \frac{Ab}{3} - \frac{b^2}{3}, \quad \text{and} \quad \lim_{b \rightarrow 0} \frac{\sigma_A^2}{b} = \frac{A}{3}.$$

¹As in the main text we suppress the functional dependence of key functions on A and P , if there is no risk of confusion.

Analogously,

$$\lim_{b \rightarrow 0} \frac{\sigma_P^2}{b} = \frac{P}{3}.$$

Recall that $A \geq P$ and $q_p(x) < 1$, therefore,

$$\lim_{b \rightarrow 0} \frac{Z}{b} = \frac{A}{3} - q_p(x) \frac{P}{3} > 0.$$

That is, when b tends to 0, delegation converges to first-best faster than centralization. So when b is sufficiently small, delegation dominates centralization for all $x \in [0, 1]$.

Now consider the case of high CEO bias. When $b \geq \frac{A}{2\sqrt{3}}$, $b^2 \geq \frac{A^2}{12} \geq \sigma_A^2$. Then $Z(\cdot) = \sigma_A^2 - b^2 - q_p(x)\sigma_P^2 < 0$ for any x , i.e., centralization dominates delegation for all $x \in [0, 1]$.

Lastly, if b is in the intermediate region, define $\hat{x}(b)$ by $Z(\hat{x}(b), b) = 0$. Because $q_p(\cdot)$ is decreasing in x , $Z(\cdot)$ is increasing in x . As a result, when $x > \hat{x}(b)$, $Z(x, b) > Z(\hat{x}(b), b) = 0$, which implies that shareholders prefer delegation, and vice versa. ■

B.3 Proof of Lemma 4.

Differentiate equation (2.8) with respect to x at $x^C(b)$, we get²

$$\frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x=x^C} = \frac{\partial \pi^C(x, b)}{\partial x} \Big|_{x=x^C} - q'_a(x^C)Z(x^C, b) - (1 - q_a(x^C))q'_p(x^C)\sigma_P^2$$

²To avoid clutter, in the proof, we suppress the argument b in $x^S(\cdot)$, where $S = C, D$.

Since $\frac{\partial \pi^C(x, b)}{\partial x} \Big|_{x=x^C} = 0$,

$$\frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x=x^C} = -q'_a(x^C)Z(x^C, b) - (1 - q_a(x^C))q'_p(x^C)\sigma_P^2$$

By Lemma 3, $Z(\cdot) < 0$ for b sufficiently high, in which case $\frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x=x^C} > 0$. Recall that $\pi^D(x, b)$ is concave in x , then $x^D(b) > x^C(b)$ is immediate. ■

B.4 Proof of Lemma 5.

The optimal $x^C(b)$ and $x^D(b)$ are obtained, respectively, by inspecting the following first-order derivatives:

$$\frac{\partial \pi^C(x, b)}{\partial x} = q'_p(x)\bar{\sigma}_P^2 + q'_a(x)\sigma_A^2, \quad (\text{B.2})$$

$$\frac{\partial \pi^D(x, b)}{\partial x} = q'_p(x)\bar{\sigma}_P^2 + q'_a(x)b^2 + B(x)\sigma_P^2, \quad (\text{B.3})$$

where

$$B(x) \equiv q'_a(x)q_p(x) - (1 - q_a(x))q'_p(x) > 0.$$

The assumptions stated in the lemma ensure interior solutions for both regimes for any $b > 0$. The optimal interior solutions $x^C(b)$ and $x^D(b)$, respectively, are obtained by setting

(B.2) and (B.3) each equal to zero. Therefore,

$$\begin{aligned}\frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x=x^C} &= q'_p(x^C) \bar{\sigma}_P^2 + q'_a(x^C) b^2 + B(x^C) \sigma_P^2 \\ &= -q'_a(x^C) \sigma_A^2 + q'_a(x^C) b^2 + B(x^C) \sigma_P^2\end{aligned}$$

If $A = P$, then $\sigma_A^2 = \sigma_P^2 = \sigma^2$, and therefore

$$\begin{aligned}\frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x=x^C} &= -\sigma^2 \{q'_a(x^C) - B(x^C)\} + q'_a(x^C) b^2 \\ &= -\sigma^2 \{q'_a(x^C)(1 - q_p(x^C)) + (1 - q_a(x^C))q'_p(x^C)\} + q'_a(x^C) b^2.\end{aligned}$$

Let

$$h(x) \equiv q'_a(x)(1 - q_p(x)) + (1 - q_a(x))q'_p(x).$$

Given the exponential technology assumed in the lemma:

$$h(x) = \eta z(1 - x)^{z-1} x^{z-1} [(1 - x)^{1-z} - x^{1-z} - \eta(1 - 2x)].$$

Let $w(x) \equiv (1 - x)^{1-z} - x^{1-z} - \eta(1 - 2x)$. Then, $\lim_{\eta \rightarrow 1} w(0) = \lim_{\eta \rightarrow 1} (1 - \eta) = 0$ and $w(\frac{1}{2}) = 0$. Moreover, $w''(x) = (1 - z)z(x^{-z-1} - (1 - x)^{-z-1}) \geq 0$ for $x \in [0, 1/2]$, i.e., $w(x)$ is a convex function for $x \in [0, 1/2]$. Therefore, by Jensen's inequality, $w(x) \leq 0$ for $x \in [0, 1/2]$. Hence $h(x) \leq 0$ for $x \in [0, 1/2]$.

Next, we show that $x^C(b) \in (0, 1/2]$, where (the interior) $x^C(b)$ is again obtained by

setting (B.2) equal to zero. If $b > A/4$ and $A = P$, $\sigma_A^2 = \bar{\sigma}_P^2$. In this case, x^C is given by $q'_p(x^C) + q'_a(x^C) = 0$, which, given the assumed exponential technology, implies $x^C(b) = 1/2$. Since $x^C(b)$ is non-decreasing in b (Lemma 6), $x^C(b) \in (0, 1/2]$ for all b .

Therefore,

$$\begin{aligned} \frac{\partial \pi^D}{\partial x} \Big|_{x=x^C} &= -\sigma^2 h(x^C) + q'_a(x^C) b^2 \\ &> 0 = \frac{\partial \pi^D}{\partial x} \Big|_{x=x^D}. \end{aligned}$$

Then, by global concavity of $\pi^D(x, b)$, we have $x^D(b) > x^C(b)$. ■

B.5 Proof of Proposition 4.

This proof provides the solution to the shareholders's optimization program \mathcal{P}_0 .

Part (i). For b sufficiently small, Lemma 3 shows that $\pi^D(x, b) > \pi^C(x, b)$ for any x . Then, $\pi^D(x^D(b), b) \geq \pi^D(x^C(b), b) > \pi^C(x^C(b), b)$. The first inequality is by revealed preference and the second inequality follows from $\pi^D(x, b) > \pi^C(x, b)$ for any x . Hence, for b sufficiently small, the shareholders chooses delegation.

Part (ii). The proof follows analogously as in part (i).

Part (iii). The preceding parts show that $\pi^D(x^D(b), b) - \pi^C(x^C(b), b) > 0$ for b small

and $\pi^D(x^D(b), b) - \pi^C(x^C(b), b) < 0$ for b large. Then by continuity, there must exist at least one \hat{b} such that $\pi^D(x^D(\hat{b}), \hat{b}) - \pi^C(x^C(\hat{b}), \hat{b}) = 0$.

We now show that $x^C(\hat{b}) < \hat{x}(\hat{b}) < x^D(\hat{b})$. We proceed in two steps: we first show that this chain of inequalities holds in weak form, and then rule out equality. First, suppose that $x^C(\hat{b}) > \hat{x}(\hat{b})$. By Lemma 3, then

$$\pi^D(x^C(\hat{b}), \hat{b}) > \pi^C(x^C(\hat{b}), \hat{b}).$$

Yet, by definition of \hat{b} , $\pi^D(x^D(\hat{b}), \hat{b}) \equiv \pi^C(x^C(\hat{b}), \hat{b})$. Thus, $\pi^D(x^C(\hat{b}), \hat{b}) > \pi^D(x^D(\hat{b}), \hat{b})$, which contradicts the optimality of $x^D(\cdot)$. As a result, $x^C(\hat{b}) \leq \hat{x}(\hat{b})$. A similar argument shows that $\hat{x}(\hat{b}) \leq x^D(\hat{b})$. Therefore, $x^C(\hat{b}) \leq \hat{x}(\hat{b}) \leq x^D(\hat{b})$.

Now, we show that the preceding chain of inequalities holds in a strict sense. Suppose that $x^D(\hat{b}) = x^C(\hat{b}) \equiv \bar{x}$. Then it must also be true that $\hat{x}(\hat{b}) = \bar{x}$. Then, by virtue of $x^S(b) \in \arg \max_x \pi^S(x, b)$, the following must hold:

$$\frac{\partial}{\partial x} \pi^C(\bar{x}, \hat{b}) = \frac{\partial}{\partial x} \pi^D(\bar{x}, \hat{b}) = 0$$

Using (B.2) and (B.3), the preceding condition implies that

$$\begin{aligned}
& q'_a(\bar{x}) \left[\sigma_A^2 - \left(\hat{b}^2 + \frac{B(\bar{x})}{q'_a(\bar{x})} \sigma_P^2 \right) \right] = 0 \\
\iff & \underbrace{Z(\bar{x}, \hat{b})}_{=0} + \left[q_p(\bar{x}) - \frac{q'_a(\bar{x})q_p(\bar{x}) - (1 - q_a(\bar{x}))q'_p(\bar{x})}{q'_a(\bar{x})} \right] \sigma_P^2 = 0 \\
& \iff (1 - q_a(\bar{x}))q'_p(\bar{x}) = 0,
\end{aligned}$$

which contradicts our maintained assumption that $q_j(x) < 1$ and $q'_j(x) \neq 0$ for any x and $j = a, p$. Thus, $x^D(\hat{b}) > x^C(\hat{b})$. Using a simple geometrical argument shows that $\hat{x}(\hat{b}) \in (x^C(\hat{b}), x^D(\hat{b}))$ must then hold.

Hence, at the cutoff \hat{b} where the optimal organization structure switches from delegation to centralization, $x^*(b)$ jumps down from $x^D(\hat{b})$ to $x^C(\hat{b})$. Other than at such discontinuities, $x^*(b)$ is increasing in b , by Lemma 6. ■

B.6 Proof of Proposition 5.

We now solve the CEO's program \mathcal{P}_1 and compare its solution, denoted by $x_{CEO}(b)$, with that of the benchmark program \mathcal{P}_0 , denoted by $x^*(b)$.

The CEO's program \mathcal{P}_1 can be decomposed as follows. First, find the CEO's optimal

payoff $\pi_{CEO}^{S^*}(b)$ for given mode $S = C, D$:

$$\max_x \pi_{CEO}^D(x, b), \quad \text{s.t. } x > \hat{x}(b), \quad (\text{B.4})$$

$$\max_x \pi_{CEO}^C(x, b), \quad \text{s.t. } x \leq \hat{x}(b). \quad (\text{B.5})$$

In a second step, compare $\pi_{CEO}^{D^*}(b)$ and $\pi_{CEO}^{C^*}(b)$ and select the larger one.

Denote by $x_{CEO}^{D^*}(b)$ and $x_{CEO}^{C^*}(b)$, respectively, the solutions to the constrained maximization problems (B.4) and (B.5). Analogously, denote by $x_{CEO}^D(b)$ and $x_{CEO}^C(b)$, respectively, the solutions to the corresponding *unconstrained* maximization problems. It is immediate that $x_{CEO}^C(b) = x^C(b)$, because under centralization the CEO's preference over the board composition coincides with that of the shareholders (Lemma 0).

Part (i): $b < \hat{b}$. First, we show that $\hat{x}(b) < x^D(b)$ for $b < \hat{b}$. Suppose on the contrary that $\hat{x}(b) \geq x^D(b)$. Then,

$$\pi^C(x^C(b), b) \geq \pi^C(x^D(b), b) \geq \pi^D(x^D(b), b).$$

The first inequality follows by revealed preference and the second inequality follows from $\hat{x}(b) \geq x^D(b)$ together with Lemma 3. However, by the definition of \hat{b} and the regularity condition that \hat{b} is unique (footnote 19), $\pi^D(x^D(b), b) > \pi^C(x^C(b), b)$ for $b < \hat{b}$ — a contradiction. Thus, $\hat{x}(b) < x^D(b)$ for $b < \hat{b}$.

Secondly, since $x^D(b) > \hat{x}(b)$ for $b < \hat{b}$, a revealed preference argument shows that

$\pi_{CEO}^{D^*}(b) \geq \pi_{CEO}^D(x^D(b), b)$. At the same time, $\pi_{CEO}^{C^*}(b) \leq \pi_{CEO}^C(x_{CEO}^C(b), b)$ by revealed preference. Therefore:

$$\begin{aligned}
\pi_{CEO}^{D^*}(b) - \pi_{CEO}^{C^*}(b) &\geq \pi_{CEO}^D(x^D(b), b) - \pi_{CEO}^C(x_{CEO}^C(b), b) \\
&= \pi_{CEO}^D(x^D(b), b) - \pi_{CEO}^C(x^C(b), b) \quad (\text{by } x_{CEO}^C(b) = x^C(b)) \\
&= \pi^D(x^D(b), b) + [1 - 2q_a(x^D(b))]b^2 - \pi^C(x^C(b), b) + b^2 \quad (\text{by (2.7)}) \\
&= \pi^D(x^D(b), b) - \pi^C(x^C(b), b) + 2[1 - q_a(x^D(b))]b^2 \\
&> 0 \quad (\text{since } \pi^D(x^D(b), b) > \pi^C(x^C(b), b) \text{ when } b < \hat{b})
\end{aligned}$$

Hence, when $b < \hat{b}$, the CEO will choose $x_{CEO}(b) = x_{CEO}^{D^*}(b)$, followed by the board choosing $S^*(x_{CEO}(b)) = D$.

Next, we show that $x_{CEO}^{D^*}(b) < x^D(b)$. To see this, notice that

$$\frac{\partial \pi_{CEO}^D(x, b)}{\partial x} < \frac{\partial \pi^D(x, b)}{\partial x},$$

for any x . Then

$$\frac{\partial \pi_{CEO}^D(x, b)}{\partial x} \Big|_{x \geq x^D(b)} < \frac{\partial \pi^D(x, b)}{\partial x} \Big|_{x \geq x^D(b)} \leq 0$$

The last inequality holds by global concavity of $\pi^D(x, b)$ and optimality of $x^D(b)$. Therefore, $\hat{x}(b) \leq x_{CEO}^{D^*}(b) < x^D(b)$. (Recall that $\hat{x}(b) < x^D(b)$ when $b < \hat{b}$.)

Part (ii): b is sufficiently large. In this case, centralization dominates delegation for all

x (by Lemma 3). Then, manipulating x to obtain delegation is not feasible for CEO. By Lemma 0, then, the CEO will choose $x_{CEO}(b) = x^C(b)$.

Part (iii): $\hat{b} \leq b \leq \hat{b} + \varepsilon$, where $\varepsilon \rightarrow 0$. Proposition 4 shows that $x^C(b) < \hat{x}(b) < x^D(b)$ at $b = \hat{b}$. By continuity, $x^C(b) < \hat{x}(b) < x^D(b)$ holds also for $\hat{b} \leq b \leq \hat{b} + \varepsilon$. Then the same revealed preference arguments as in part (i) show that $\pi_{CEO}^{D*}(b) \geq \pi_{CEO}^D(x^D(b), b)$ and $\pi_{CEO}^{C*}(b) \leq \pi_{CEO}^C(x_{CEO}^C(b), b)$. Therefore:

$$\begin{aligned} \pi_{CEO}^{D*}(b) - \pi_{CEO}^{C*}(b) &\geq \pi_{CEO}^D(x^D(b), b) - \pi_{CEO}^C(x_{CEO}^C(b), b) \\ &= \pi_{CEO}^D(x^D(b), b) - \pi_{CEO}^C(x^C(b), b) \\ &= \pi^D(x^D(b), b) - \pi^C(x^C(b), b) + 2[1 - q_a(x^D(b))]b^2 \end{aligned}$$

Clearly, when $\varepsilon \rightarrow 0$, $[\pi^D(x^D(b), b) - \pi^C(x^C(b), b)] \rightarrow 0$, and therefore

$$\pi_{CEO}^{D*}(b) - \pi_{CEO}^{C*}(b) > 0 \text{ because } q_a(x^D(b)) < 1.$$

So, if b is “slightly above” \hat{b} , CEO will choose $x_{CEO}(b) = x_{CEO}^{D*}(b) \geq \hat{x}(b) > x^C(b)$ followed by the board choosing $S^*(x_{CEO}(b)) = D$. ■

B.7 Proof of Lemma 7

Part (i). In the proof of Proposition 5 part (i), we show that $\hat{x}(b) < x^D(b)$ for $b < \hat{b}$. Adding back the suppressed argument A , it is immediate that for $b < \hat{b}(\bar{A})$, $\hat{x}(\bar{A}, b) < x^D(b)$. (Note that $x^D(\cdot)$ is independent of A .) That is, Constraint (D) is always slack for $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$.

Part (ii). To prove that for \bar{A} sufficiently small, constraint (C) in program \mathcal{P}_C will be slack for $b \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$, we first show that $\hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$. Suppose not and instead $\hat{x}(\underline{A}, b) < x^C(\underline{A}, b)$. Then, by Lemma 3, $\pi^D(x^C(\underline{A}, b), b) > \pi^C(x^C(\underline{A}, b), \underline{A}, b)$. By Proposition 4, $\pi^C(x^C(\underline{A}, b), \underline{A}, b) \geq \pi^D(x^D(b), b)$ for any $b \geq \hat{b}(\underline{A})$. Therefore, $\pi^D(x^C(\underline{A}, b), b) > \pi^D(x^D(b), b)$, which contradicts the optimality of $x^D(b)$. As a result, $\hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$. Therefore $\lim_{\bar{A} \rightarrow \underline{A}} \hat{x}(\bar{A}, b) = \hat{x}(\underline{A}, b) \geq x^C(\underline{A}, b)$ for any $b \geq \hat{b}(\underline{A})$, i.e., for \bar{A} sufficiently small, constraint (C) in program \mathcal{P}_C is slack for any $b \geq \hat{b}(\underline{A})$.

On the other hand, if $\bar{A} > \tilde{A}(x^C(\underline{A}, \hat{b}(\underline{A})), \hat{b}(\underline{A}))$, which is equivalent to $x^C(\underline{A}, \hat{b}(\underline{A})) > \hat{x}(\bar{A}, \hat{b}(\underline{A}))$, constraint (C) will be binding at $\hat{b}(\underline{A})$. At the same time,

$$\begin{aligned} x^C(\underline{A}, \hat{b}(\bar{A})) &< x^C(\bar{A}, \hat{b}(\bar{A})), && \text{since } \partial x^C / \partial A > 0 \\ &< \hat{x}(\bar{A}, \hat{b}(\bar{A})), && \text{by Proposition 4} \end{aligned}$$

Then by continuity of $x^C(\cdot)$ and $\hat{x}(\cdot)$, there must exist at least one $b^C \in (\hat{b}(\underline{A}), \hat{b}(\bar{A}))$ such that $x^C(\underline{A}, b^C) = \hat{x}(\bar{A}, b^C)$. ■

B.8 Proof of Proposition 6

As shown in Lemma 7, part (ii), when $\bar{A} > \tilde{A}(x^C(\underline{A}, \hat{b}(\underline{A})), \hat{b}(\underline{A}))$, constraint (C) will be binding at $\hat{b}(\underline{A})$ and there must exist at least one $b^C \in [\hat{b}(\underline{A}), \hat{b}(\bar{A})]$ such that $x^C(\underline{A}, b^C) =$

$\hat{x}(\bar{A}, b^C)$. If there exist multiple thresholds b^C , we focus on the smallest one. Denote by $V_C(\cdot)$ as the value of program \mathcal{P}_C and by $V_D(\cdot)$ the value of program \mathcal{P}_D .

Part (i). [Acquiesce] Since constraint (C) is binding at $\hat{b}(\underline{A})$, then by continuity, constraint (C) is also binding at $\hat{b}(\underline{A}) + \delta$, where $\delta \rightarrow 0$. Hence $V_C(\hat{b}(\underline{A}) + \delta) < \pi^C(x^C(\underline{A}, \hat{b}(\underline{A}) + \delta), \underline{A}, \hat{b}(\underline{A}) + \delta)$ due to the shadow cost of (C). On the other hand, $V_D(\cdot) = \pi^D(x^D(b), b)$ since constraint (D) is always slack, by Lemma 7. Therefore:

$$\begin{aligned} \lim_{\delta \rightarrow 0} V_C(\hat{b}(\underline{A}) + \delta) &< \lim_{\delta \rightarrow 0} \pi^C(x^C(\underline{A}, \hat{b}(\underline{A}) + \delta), \underline{A}, \hat{b}(\underline{A}) + \delta) \\ &= \lim_{\delta \rightarrow 0} \pi^D(x^D(\hat{b}(\underline{A}) + \delta), \hat{b}(\underline{A}) + \delta) \\ &= \lim_{\delta \rightarrow 0} V_D(\hat{b}(\underline{A}) + \delta). \end{aligned}$$

Hence for b “slightly above” $\hat{b}(\underline{A})$, shareholders prefer to acquiesce to CEO entrenchment and choose $x = x^D(b)$. In response, the CEO will choose $A = \tilde{A}(x^D(b), b) + \varepsilon$ to noise up the project complexity by just enough so as to ensure delegation.

Part (ii). [Preempt at a cost] At the threshold b^C , constraint (C) is just binding, there-

fore $V_C(b^C) = \pi^C(x^C(\underline{A}, b^C), \underline{A}, b^C)$. For $b = b^C - \delta$,

$$\begin{aligned}
 \lim_{\delta \rightarrow 0} V_C(b^C - \delta) = V_C(b^C) &= \pi^C(x^C(\underline{A}, b^C), \underline{A}, b^C) \\
 &> \pi^D(x^D(b^C), b^C), \quad \text{since } b^C > \hat{b}(\underline{A}) \\
 &= \lim_{\delta \rightarrow 0} \pi^D(x^D(b^C - \delta), b^C - \delta) \\
 &= \lim_{\delta \rightarrow 0} V_D(b^C - \delta)
 \end{aligned}$$

Hence for $b = b^C - \delta$, the constraint (C) is still binding, yet the shareholders are better off choosing $x = \hat{x}(\bar{A}, b)$ to preempt CEO entrenchment. In response, the CEO will choose $A = \underline{A}$.

Part (iii). [Preempt at no cost] For $b = b^C + \delta$, similarly as part (ii), we can show that $\lim_{\delta \rightarrow 0} V_C(b^C + \delta) > \lim_{\delta \rightarrow 0} V_D(b^C + \delta)$. Hence for b slightly above b^C , constraint (C) becomes slack, the shareholders will choose $x = x^C(\underline{A}, b)$ to preempt CEO entrenchment. In response, the CEO will choose $A = \underline{A}$. ■

Bibliography

Adams, R. (2009), Asking directors about their dual roles, working paper, University of Queensland.

Adams, R., and D. Ferreira (2007), A theory of friendly boards, *Journal of Finance* 62 (1), 217-250.

Adams, R., B. E. Hermalin, and M. S. Weisbach, (2009), The role of boards of directors in corporate governance: a conceptual framework & survey, *Journal of Economic Literature*, forthcoming.

Aghion, P., and J. Tirole (1997), Formal and real authority in organizations, *Journal of Political Economy* 105(1), 1-29.

Alonso, R., and N. Matouschek (2007), Relational delegation, *Rand Journal of Economics* 38, 1070-89.

Baker, M., and P. Gompers (2003), The determinants of board structure at the initial public offering, *Journal of Law and Economics*, October 46(2), 569-597.

Barber, B., Lehavy, R., McNichols, M., and Trueman, B. (2006), Buys, Holds, and Sells: The Distribution of Investment Banks. Stock Ratings and the Implications for the Profitability of Analysts Recommendations, *Journal of Accounting and Economics* 41, 87-117.

Bathala, C. and R. Rao (1995), The determinants of board composition: an agency theory perspective, *Managerial and Decision Economics* 16, 59-69.

Becht, M., J. Franks, C. Mayer, and S. Rossi (2008), Returns to shareholder activism: evidence from a clinical study of the Hermes U.K. Focus Fund, *Review of Financial Studies*, forthcoming.

Benabou, R., and Laroque, G. (1992), Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility, *Quarterly Journal of Economics* 107, 921-958.

Beyer, A., and Guttman, I. (2007), The Effect of Trading Commissions on Analysts' Forecast Bias, working paper, Stanford University.

Bhagat, S. and B. Black (1999), The uncertain relationship between board composition and firm performance, *Business Lawyer* 54, 921-963.

Brown, L. D. (1993), Earnings forecasting research: Its implications for capital markets research, *International Journal of Forecasting* 9, 295-320.

Chang, C. , Daouk, H., and Wang, A., (2008), Do Investors Learn About Analyst Accuracy? Working paper, Cornell University.

Chakraborty and Yilmaz (2009), Authority, consensus and governance, working paper, York University.

Chen, Y., N. Kartik, and J. Sobel (2008), Selecting cheap-talk equilibria, *Econometrica* 76(1), 117-136.

Chen, Q., Francis, J., and Jiang, W. (2005), Investor learning about analyst predictive ability, *Journal of Accounting and Economics* 39, 3-24.

Chen, S., and Matsumoto, D. (2006), Favorable versus unfavorable recommendations: the impact on analyst access to management-provided information, *Journal of Accounting Research* 44, 657-689.

Coles, J., N. Daniel, N., L. Naveen (2008), Boards: does one size fit all? *Journal of Financial Economics* 87, 329-356.

Conroy, R., and Harris, R. (1995), Analysts' earnings forecasts in Japan: Accuracy and sell-side optimism, *Pacific-Basin Finance Journal* 3(4), 393-408.

Cowen, A., Groyberg, B., and Healy, P. (2006), Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41, 119-146.

Crawford, V.P., and Sobel, J. (1982), Strategic Information Transmission, *Econometrica* 50, 1431-1451.

D'Avolio, G. (2002), The market for borrowing stock, *Journal of Financial Economics* 66, 271-306.

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

Dechow, P., Hutton, A., and Sloan, R. (2000), The Relation between Analysts' Fore-

casts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings, *Contemporary Accounting Research* 1-32.

Desai, H., Liang, B., and Singh, A. K. (2000), Do All-Stars Shine? Evaluation of Analyst Recommendations. *Financial Analysts Journal* 56, 20-30.

Dessein, W. (2002), Authority and communication in organizations, *Review of Economic Studies* 69, 811-838.

Dimson, E., and Marsh, P.R. (1984), An analysis of brokers' and analysts' unpublished forecasts of U.K. stock returns, *Journal of Finance* 39, 1257-1292.

Dugar, A. and Nathan, S. (1995), The Effect of Investment Banking Relationships on Financial Analyst's Earnings Forecasts and Investment Recommendations, *Contemporary Accounting Research* 12, 131-160.

Edlin, A., and J. Stiglitz (1995), Discouraging rivals: managerial rent-seeking and economic inefficiencies, *American Economic Review*, 85(5), 1301-1312.

Eisenberg, T., S. Sundgren, and M. Wells (1998), Larger board size and decreasing firm value in small firms, *Journal of Financial Economics* 48, 35-54.

Ely, J. and Valimaki, J. (2003), Bad reputation, *the Quarterly Journal of Economics* CXVIII, 785-814.

Ely, J., Fudenberg, D., and Levine, D. (2008), When is reputation bad? *Games and Economic Behavior* 63, 498-526.

Faleye, O., R. Hoitash, U. Hoitash (2010), The costs of intense board monitoring, *Journal of Financial Economics*, forthcoming.

Faleye, O., R. Hoitash, U. Hoitash (2011), Advisory directors, working paper, Northeastern University.

Fang, L. and Yasuda, A.(2005), Analyst reputation, underwriting pressure, and forecast accuracy, working paper, University of Pennsylvania.

Fischer, P.E. and Stocken, P.C. (2010), Analyst Information Acquisition and Communication, *The Accounting Review* 85(6), 1985-2010.

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

Fudenberg, D. and Levine D.K. (1989), Reputation and Equilibrium Selection in Games with a Single Patient Player, *Econometrica* LVII, 251-268.

Harris, M., and A. Raviv (2005), Allocation of decision-making authority, *Review of Finance* 9, 353-383.

Harris, M., and A. Raviv (2008), A theory of board control and size, *Review of Financial Studies* 21(4), 1797-1832.

Harris, M., and A. Raviv (2010), Control of corporate decisions: shareholders vs. management, *Review of Financial Studies*, forthcoming.

Hayes, R.M. (1998), The Impact of Trading Commission Incentives on Analysts' Stock Coverage Decisions and Earnings Forecasts, *Journal of Accounting Research* 36, 299-320.

Holmstrom, B. (1999), Managerial incentive problems: a dynamic perspective. *Review of Economic Studies* 66 (1), 169- 182.

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

Irvine, P. (2004), Analysts' forecasts and brokerage-firm trading, *The Accounting Review* 79, 125-149.

Jackson, A. (2005), Trade generation, reputation and sell-side analysts, *Journal of Finance* 60, 673-717.

Jensen, M., (1993), The modern industrial revolution, exit, and the failure of internal control systems, *Journal of Finance*, 48(3), 831-880.

Lehn, K., S. Patro, and M. Zhao (2004), Determinants of the size and structure of corporate boards, working paper, University of Pittsburgh.

Linck, J., J. Netter, and T. Yang (2008), The determinants of board structure, *Journal of Financial Economics* 87, 308-387.

Linck, J., J. Netter, and T. Yang (2009), The effects and unintended consequences of the Sarbanes-Oxley act on the supply and demand for directors, *Review of Financial Studies* 22(8), 3287-3328.

Kadan, O., Madureira, L., Wang, R., and Zach, T. (2009), Conflicts of Interest and Stock Recommendations: The Effects of the Global Settlement and Related Regulations, *Review of Financial Studies* 22, (10), 4189-4217.

Kim, J.-Y.(1996), Cheap Talk and Reputation in Repeated Pretrial Negotiation, *RAND Journal of Economics* 27, 787-802.

Kreps, D. M. and Wilson, R.(1982), Reputation and Imperfect Information, *Journal of Economic Theory* XXVII, 253-279.

Lambert, W. and Sapsford, J. (2001), J.P. Morgan Memo to Analysts Raises Eyebrows, *Wall Street Journal*.

Lin, H.W. and McNichols, M.F. (1998), Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations, *Journal of Accounting and Economics* 25, 101-127.

Lim, T. (2001), Rationality and analysts' forecast bias. *Journal of Finance* 56, 369-385.

Lys, T. and Sohn, S. (1990), The association between revisions of analysts' earnings forecasts and security-price changes, *Journal of Accounting and Economics* 13, 341-363.

Malmendier, U. and Shanthikumar, D. (2007), Are small investors naive about incentives? *Journal of Financial Economics* 85, 457-489.

McNichols, M. and O'Brien, P.C. (1997), Self-Selection and Analyst Coverage. *Journal of Accounting Research* 35, 167-199.

Melumad, N., and Shibano, T. (1991), Communication in settings with no transfers, *The RAND Journal of Economics* 22(2), 173-198.

Michaely, R. and Womack, K.L. (1999), Conflict of Interest and the Credibility of Underwriter Analyst Recommendations, *Review of Financial Studies* 12, 653-686.

Milgrom, P. and Roberts, J. (1982), Predation, Reputation, and Entry Deterrence, *Journal of Economic Theory* XXVII, 280-294.

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

Morgan, J. and Stocken, P.C. (2003), An analysis of Stock Recommendations, *The RAND Journal of Economics* 34(1), 183-203.

Morris, S. (2001), Political Correctness, *Journal of Political Economy* 109, 231-265.

O'Brien, Patricia C. (1988), Analysts forecasts as earnings expectations, *Journal of Accounting and Economics* 10, 53-83.

Park, C. W., and Stice, E. K. (2000), Analyst forecasting ability and the stock price reaction to forecast revisions, *Review of Accounting Studies* 5, 259-272.

Prendergast, C., and Stole, L. (1996), Impetuous youngsters and jaded old-timers: acquiring a reputation for learning, *Journal of Political Economy* 104 (6), 1105-1134.

Raheja, C. (2005), Determinants of board size and composition: a theory of corporate boards, *Journal of Financial and Quantitative Analysis*, 40(2), 283-306.

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

Shivdasani, A., and D. Yermack (1999), CEO involvement in the selection of new board members: an empirical analysis, *Journal of Finance* 54, 1829-1853.

Shleifer, A., and R. Vishny (1989), Management entrenchment: the case of manager-specific investments, *Journal of Financial Economics* 25, 123-139.

Sobel, J. (1985), A Theory of Credibility, *Review of Economic Studies* 52, 557-573.

Solomon, D., and Frank, R. (2003), You Don't Like Our Stock? You Are Off the List

– SEC Sets New Front On Conflicts By Taking Aim at Companies That Retaliate Against Analysts, *Wall Street Journal*.

Stickel, S.E. (1992), Reputation and Performance among Security Analysts, *Journal of Finance* 47, 1811-1836.

Stocken, P.C. (2000), Credibility of Voluntary Disclosure, *The RAND Journal of Economics* 31(2), 359-374.

Suurmond, G., Swank, O.H., and Visser, B.(2004), On the bad reputation of reputational concerns, *Journal of Public Economics* 88, 2817-2838.

Tirole, J. (2001), Corporate governance, *Econometrica* 69(1), 1-35.

Trueman, B. (1994), Analyst Forecasts and Herding Behavior, *Review of Financial Studies* 7, 97-124.

Wang, H. (2009), Reputation acquisition of underwriter analysts — Theory and evidence, *Journal of Applied Economics* 12(2), 331-363.

Womack, K.L. (1996), Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance* 51, 137-167.

Xu, H.(2011), Reputational Concern with Endogenous Information Acquisition, working paper, Washington University in St. Louis.

Yermack D. (1996) , Higher market valuation of companies with a small board of directors, *Journal of Financial Economics* 40, 185-211.

Zhang, I. (2007), Economic consequences of the Sarbanes-Oxley Act of 2002, *Journal of Accounting and Economics* 44 (1-2), 74-115.