STRATEGIC INFORMATION TRANSMISSION IN THE STOCK MARKET AND THE FIRM

A thesis presented

by

Jordi Blanes i Vidal

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

London School of Economics

University of London

2004

Abstract

In this thesis I apply the tools of information economics to analyse the way that information is transmitted in the stock market and the firm. In particular, I use models of asymmetric information to explain a number of empirical regularities that affect the modern corporation and the modern capital market.

In the first chapter I show that delegation of decision-making rights can stimulate the career concerns of subordinates in organisations. I show that, when an employer takes a decision following the proposal of her subordinate, a winner's curse reduces the subordinate's prospects in the labour market, muting incentives. This can be solved by delegating decision-making rights to the worker.

The second chapter is a joint work with Marc Moller. We use tools of information economics to propose a model of leadership in order to understand why many leaders are unduly confident in their own judgment, a fact that frequently afflicts modern organisations. We show that overconfidence can improve a leader's use of private information although it harms the aggregation of external advice. Overconfidence can therefore improve overall efficiency if the cost of consulting externally is sufficiently high.

In the last chapter I apply similar tools to study how investors in the stock market react to public messages that may be optimistically biased. I first construct a communication game between an investor and a (possibly) biased securities analyst. I find an equilibrium characterized by the following properties: first, the investor reacts more to bad news than to good news, and second, the difference in this reaction is higher when the investor has a greater prior suspicion that the analyst is a biased type. I then use parametric and nonparametric techniques and a large database of earnings and forecasts to test these predictions, and find that the evidence supports them.

Contents

Introduction

The rapid progress of information economics has offered academics a valuable framework to analyse the way that information is transmitted in human organisations. In this thesis, I apply tools developed in the last two decades to advance the study of two of these organisations: the modern company and the capital market.

I start by analysing the way that information is accumulated and used in hierarchical organisations. In my first chapter, I offer a new argument why delegation of decision-making rights may be beneficial to the efficiency of an organisation. In particular, I suggest that such delegation may stimulate the career concerns of subordinates. I show this by building a theoretical model where an employer/superior and a subordinate have to seek information and take a decision about the strategic direction for the company. I show that, when the employer retains authority over the decision and takes the decision following the proposal of her subordinate, she has better information about the performance (ability) of the worker than is available to the labour market. When this happens, the labour market wage offers to the subordinate are likely to suffer from "the winner's curse", so rewards for good performance are reduced. With such muted incentives the subordinate has no interest in seeking information. I conclude this chapter by showing how this problem can be solved by delegating decision-making rights to the worker.

The second chapter of my thesis is joint work with Marc Moller. We study a hierarchical scenario where the main function of the leader is to gather information and to take decisions. An important ingredient of our model is that often the leader worries about the impact on his subordinate's motivation of his decisions. We show that inefficiency arises when the leader's private information does not point in the direction that is most effective for motivational purposes. Since the leader does not fully internalise the cost of the effort incurred by his subordinate, he is willing to adopt a suboptimal decision that induces higher motivation in her. In this context we compare alternative leadership styles, using the confidence of the leader in his ability to make good decisions as a differentiating variable. "Overconfident" leaders overestimate such ability and we show that they gather too little external information, as compared to "realistic" leaders. Overconfident leaders are, however, less willing to distort their decisions to boost their subordinates' motivation, which makes their decision-making more efficient. We show that appointing an overconfident leader may improve overall efficiency when the cost of gathering external information is high and his private information is very precise.

In the last chapter I apply the tools of information economics to the study of the transmission of information in the stock market. In particular, I study how investors react to public messages that may be optimistically biased. I first construct a communication game between an investor and a (possibly) biased securities analyst. My model is in the tradition of the strategic information transmission literature, and I find an equilibrium characterized by the following properties: first, the investor reacts more to bad news than to good news, and second, the difference in this reaction is higher when the investor has a greater prior suspicion that the analyst is a biased type. I then use parametric and nonparametric techniques and a large database of earnings and forecasts to test these predictions, and find that the evidence supports them. Lastly, I also provide some evidence in favour of the claim that analysts bias their forecasts in order to avoid losing access to the company management and the information deriving from it.

Chapter 1
Authority, Delegation and the Winner's Curse

"Mr. Bloomblerg seems happier than his predecessor to let officials run their departments and take credit where it is due, without micro-managing their every decision" [13]

1.1 Introduction

Every good manager understands that it is impossible to give responsibility without giving power. In recent years, numerous organisations have engaged in a process of formal empowerement of their subordinates, allowing them the right to take decisions without interferences from upper management. According to a survey by Osterman [10] around 45% of workers have large discretion over the mode of doing their job. Similarly, casual observation often points towards decentralisation of decision-making powers and responsibilities, with flatter hierarchies and greater autonomy of teams. Rajan and Zingales [11] assert, perhaps excessively, that "the biggest challenges for the owners or top management today is to manage in an atmosphere of diminished authority".

Economic theory has started in the last years to address the fact that authority within organisations matters. Following Aghion and Tirole [1], the main emphasis has been on the congruence of objectives between a firm and its employees. They stress that as interests become more aligned, delegation of power motivates employees without causing severe disruption of the decision-making process.

This chapter follows a somewhat different path, based on the observation that an important part of the incentives in agency relationships are career-concerns based. As the above quote suggests, the chapter's aim is to highlight the fact that delegation often improves the allocation of responsibility by allowing employees to take credit for the decisions they have had an influence on. The reason for this is two-fold. First of all, delegation makes the performance of a worker more transparent to third parties, by removing the possible noise of a superior reversing the worker's decisions. Secondly, it also makes information about performance more symmetric, since the superior becomes an observer of the decision's result, rather than a participant of the decision process. Consequently she has as much information as outside parties.

The second reason forms the central of the analysis in this chapter. I show how when the current employer of a worker holds better information about his performance (and therefore his ability) than prospective future employers, a winner's curse effect appears in the labour market. To elaborate, following an outside offer to the worker, the current employer will decide to match it only when the worker is of relatively high ability. Outside employers, being often left with the worst of the pick, will react by reducing offers to all workers, regardless of their observed performance. The low responsiveness of future wages to performance -that is, the weakness of the career concerns incentives- will reduce the information acquisition effort of the worker. It follows therefore from the analysis that the interplay between career concerns and the winner's curse effect creates a trade-off between increasing the incentives of the worker ex ante and extracting a rent from him ex post. By delegating (i.e. disappearing from the decision process) a principal is committing to not know more than third parties about the performance of a worker. That is costly ex post since it decreases the quality of the decisions, and it also denies the possibility of keeping an able worker at depressed wages. The advantage, however, stems from the improved ex ante incentives that the worker has to appear as able.

Model and Results. The interaction between these different forces is studied in an organisation consisting of an employer (she) and a worker of uncertain ability (he) engaged in a principal-agent relationship with moral hazard in information acquisition. A decision must be taken with respect to the choice between two alternative projects that, if successful, deliver private benefits to the employer. A distinct feature of the model is that the knowledge of the two agents is complementary in a particular way. The worker is informed about the technical side of the choice, that is, the prospects of success of the two projects. The employer, on the other hand, observes the private value that she attaches to the success of the different projects. To simplify, it is assumed that both of them remain completely uninformed about the other side's knowledge. This specific information structure is proposed to stress the possibility of intervention ex post, where, if she retains the right to do so, the employer may decide to reverse the proposal of the worker after observing that the project with less chances of success is nevertheless worth undertaking. I believe that this kind of intervention is very common in reality. Think, for instance, of a football club. Technically the coach of the team is obviously better informed about what players should be hired for the following season. On the other hand, the president of the club may not resist intervening in the hiring process in order to, for instance, sign a popular scorer who does not fit into the coach's scheme. The marketing benefits of doing so may outweigh the risk of jeopardising the chances of winning the league in the following season. Similarly, a mutual fund manager may be obliged to invest in a certain underperforming company in which his superior holds a personal stake. Another example is the fact that securities analysts working for investment banks have often found themselves pressured to recommend companies to investors in order to win more investment banking business for their employers. It is worth stressing that the employers in the three examples above do not hold better information than their employees about the technical side of the choice. However, they still find it worth reversing their employees proposals.

The accuracy of the proposal by the worker depends both on his ability and on his information acquisition. This acquisition is private and costly, which introduces a moral hazard problem. I adopt an incomplete contracts approach by assuming than neither the choice of project nor the private value to the employer can be contracted upon. However, incentives for information acquisition can still be provided through career concerns. There are two reasons for that. On the one hand, more able employees make more accurate proposals. On the other hand, outside prospective employers observe the success or failure of the decision, and adjust their expectation about the ability of the worker accordingly. Those reputation incentives are rather obvious in the examples above, but incentives in many other professions are, to some extent, career concerns based.

An unfortunate fact that weakens the strength of the career concerns is that, if the employer keeps ultimate authority, outside firms can only observe the success or failure of the project choice, and not whether the decision followed the worker's proposal or was reversed by the employer. I believe that this is quite realistic, capturing the fact that two members of an organisation working closely together can establish a communication channel without the knowledge of third parties. An immediate effect of the inability by the outside market to spot when the superior reverses the worker's proposal, is that the result of the decision becomes a worse estimate of the ability of the worker. The reason is that following the observation of a successful project, the outside market will be unsure as to whom they should allocate the credit for the decision. More importantly, I show below how it also creates a winner's curse. Because the current employer is able to allocate the credit better, she will seek to retain a worker when the credit is due to him, failing to match outside offers when he is of relative low ability. Being left with the worse workers, the optimal response of outside employers will be to reduce their wage offers to any worker, independently of their observed quality. The market's response is ex post beneficial for the employer, since this way she will usually manage to retain good workers at low wages, therefore extracting rents from them. The down side is the incentives of the worker. The decrease in the responsiveness of future wages to the accuracy of the proposals of the worker will have a demotivating effect, leading to lower information acquisition.

When the employer commits not to interfere in the decision-making process, the information about the worker's performance becomes more transparent and more symmetric. Consequently, the incentives for the worker to signal high ability -by acquiring information about the best project- and be rewarded for it -in the form of future wages- become stronger. The employer gains ex ante from a better informed worker, but fails to be able to block projects that are obviously not in his interest. Furthermore, the employer ceases to hold superior information about the ability of the worker, and therefore cannot extract a rent on him.

Outline. This chapter is organised as follows. Section 1.2 contains a discussion of related literature. Section 1.3 describes the assumptions of the model and comments on the alternative decision rights arrangements. Section 1.4 shows why career concerns incentives are stronger under delegation than under authority and studies under what circumstances delegation is more efficient. Section 1.5 concludes.

1.2 Related work

The most direct link of this chapter is to the delegation literature started by Aghion and Tirole [1]. They argue that workers are more motivated to become informed and learn about different projects when they get to decide. If the interests of firm and worker are sufficiently congruent, the firm may benefit from this increase in effort when it delegates authority over a decision. Carbonara [3] extends their basic model by showing that delegation is optimal when parties' interests are neither too divergent nor too close. Alternatively, Baker, Gibbons and Murphy [2] claim that decision rights can be delegated through self-enforcing incentive contracts. While this chapter shares, with the above papers, the view that decision making rights shape the acquisition of information by the worker, the focus is on the effect that such rights have on the career concerns incentives of the worker. In this respect this chapter endogenises the congruence of interests by underlining the potential conflict between the technical success of a project and the value the principal appropriates from it.

The career concerns literature was started by Fama [6] and Holmstrom [8]. They show how, in the absence of contracts, the market will provide implicit, though insufficient, incentives for an agent of uncertain ability. The impact of different information structures on incentives is analysed by Dewatripont, Jewitt and Tirole [5], who study a generalised version of Holmstrom's model. All of them take the information structure as given, without focusing on the impact of the organisation design. Ortega [9] is probably the most closely related to the present chapter. Like me, he studies the effect of the redistributing power on the visibility -the career concerns incentives- of the workers. However, he concentrates on the distribution of power among agents, and assumes that information about worker's performance is always symmetric. This chapter, however, involves a principal and an agent and focuses on the implications of asymmetric information and the role of the winner's curse.

Lastly, two papers whose focus is also close are Greenwald [7] and Cremer [4]. Greenwald is the first study on the winner's curse in the labour market. He shows how the existence of asymmetric information damages a worker's freedom to change jobs, as well as reducing his wage below his productivity level. While I use some of his intuitions, I show how the employer may gain by altering the information structure, therefore reducing the effect of the winner's curse. Cremer shows that a principal can gain by committing not to learn the ability of an under-performing agent, since it makes it more difficult to threaten him by terminating his contract. This chapter also praises the virtues of strategic ignorance, although it emphasises different issues.

1.3 The model

Consider a world consisting of a risk neutral agent (the worker) hired for a single period by a risk neutral principal (the employer). A large number of potential future employers of the worker have production technologies identical to his current employer and are hereinafter referred to as the market.

Production Technology. Employer and worker have to choose between two projects, 0 and 1, which are mutually exclusive (i.e., only one of the two projects can be adopted). The projects' success depend upon the state of the world, $x \in \{0, 1\}$, so that project 0 is successful if and only if it is adopted when $x = 0$ and similarly for project 1. The state of the world is unknown ex ante, but it is common knowledge that both states are equally likely to occur.

The projects' payoffs are received by the employer, and are conditional upon success in the following manner: a successful project pays either B or 0; a failed project always brings a revenue of 0.

This payoff structure is intended to capture, in a very simplified way, the fact that a project that has chances to prosper, from the technical point of view, may not always be worth undertaking for a firm (i.e. when it is known to pay 0 instead of B). In the absence of other considerations, an employer will prefer to have the final say over the decision, and be able to block projects that, even if successful, would deliver meagerly.

Information Structure. Prior to the selection of the project, the worker receives imperfect information about the state of the world. If state x has occurred the worker receives signal x with probability $p(e, a)$, depending on the innate ability of the worker, a, and his choice of effort $e \in [0, \infty)$. The function $p(.,.)$ is assumed to be bounded ($p \in (1/2, 1)$), concave in both arguments, and with positive cross partial derivative ($p_{ea}(e, a) > 0$), implying that the marginal return of an extra unit of effort is higher for more able workers. The effort of the worker brings a disutility of e and is not observable by the employer, which creates a moral hazard problem.

The employer receives information about the (private) payoffs of the projects, that is, observes whether each of the successful projects would pay B or 0. Again, if the employer keeps authority and observes that one of the projects pays 0 under success, she will impose the adoption of the other project. The chances of this happening are assumed to be λ, where $\lambda \in (0, 1/2)$. The fact that λ is a low number indicates that the need for intervention under authority is not very high (i.e., in most occasions the project adopted coincides with the proposal of the worker).

The market only observes the success or failure of the project after it has been undertaken. Equivalently, it observes the project choice and the state of the world. It does not observe the private payoff to the employer, and therefore cannot infer who is to be given credit for the project choice.

Ability. The ability of the worker a is symmetrically uncertain at the initial stage. It is common knowledge, however, that it can be only of two types, high and low, $a \in \{h, l\}, h > l$, with both types being equally likely.

Contracting and wages. Following both the delegation and the career concerns literature, I assume that wages and projects cannot be conditioned. This further implies that both the result of the project and the payoffs to the employer are not contractible variables.

However, due to the fact that the performance of the worker is known to the outside market, career concerns incentives can be provided. Following the observation of the result of the project and the allocation of decision rights, the market makes a wage offer to the worker, with the current employer being allowed a counteroffer. The worker then decides for whom to work. This decision is modelled in the following way: with probability $1 - u$ the worker accepts the highest offer. With probability u, however, the worker leaves his current employer irrespective of the wage offer received. This simplified behaviour is consistent with a rational agent that in a proportion u of cases receives a large and negative utility shock.

Decision rights. Two alternative decision-making arrangements are considered:

-Authority: Here the employer has the ultimate right to choose her preferred project, after consultation with the worker and observing her own information

-Delegation: The decision-making is delegated to the worker, with no interference from the employer.

Timing: To sum up, consider the timing of the game. Period 1; *(i)* The employer offers the worker an unconditional wage and an allocation of decision rights; *(ii)* Following the selection of x by nature, the worker puts effort e and receives a signal about the state of the world, which indicates the project with more chances of being successful. Similarly, the employer observes the payoffs of both projects conditional on success; *(iii)* If decision rights are delegated to the worker, he chooses the project. Otherwise he communicates his signal to the employer, who in light of this and his own information selects a project; *(iv)* The result of the project is observed by everybody and the employer receives her corresponding payoff. Period 2; *(i)* The market makes a wage offer to the worker, after which the employer decides whether to match it or not. Then the worker decides for whom to work. *(ii)-(iv)* are as in Period 1.

1.4 Analysis

1.4.1 The first best choice of effort

Start by defining the precision $\widehat{p}$ of a worker undertaking information acquisition effort e who is believed to belong to the high ability group with probability q. The precision of the signal of such a worker is $\widehat{p}(e,q) = qp(e,h) + (1-q)p(e,l)$. Note that the precision function inherits all the properties of $p(.,.)$.

I characterise now the first best solution. With contractible effort and a prior $q = 1/2$, the problem for the employer in each period becomes:

$$\max_e \widehat{p}(e, 1/2)(1-\lambda)B - e$$

with a rearranged first order condition:

$$\widehat{p}_e(e^{FB}, 1/2) = \frac{1}{(1-\lambda)B} \tag{1.1}$$

The first-best effort depends positively on the chances of the worker affecting the final decision, $(1-\lambda)$, as well as on payoffs from a profitable and successful effort, B. Note that since the agent is risk-neutral and with unlimited liability, the first best could be achieved by a well-designed contract which conditioned wages on payoffs. The concern of this chapter, however, is on economic situations in which the payoffs received by the principal are private, but the result of the project is observed by outsiders.

1.4.2 The value of a worker in the second period

Proceeding by backward induction, I study the expected value V that a worker of ability q has for a principal in the second period. $V(q) = (1-\lambda)[\widehat{p}(0,q) - 1/2]B$. Since there is no future period, the worker incurs no effort and his precision is $\widehat{p}(0,q)$, still higher, by assumption, that the prior $1/2$, held by the worker.

I consider now the posterior ability q of a worker following his performance in period 1. If he is believed to have put effort $\widehat{e}$ and is known to have received the correct signal about the world, Bayesian updating gives his posterior ability as $q_S(\widehat{e}) = \Pr[a = h \mid success, \widehat{e}] = \frac{p(\widehat{e},h)}{p(\widehat{e},h)+p(\widehat{e},l)}$. Similarly for a worker who has received the wrong signal with certainty, the revised probability is $q_F(\widehat{e}) = \Pr[a = h \mid failure, \widehat{e}] = \frac{1-p(\widehat{e},h)}{1-p(\widehat{e},h)+1-p(\widehat{e},l)}$. Since the prior ability of a worker is $1/2$, it is obvious that $q_S(\widehat{e}) > 1/2 > q_F(\widehat{e})$.

Lastly, I calculate the difference in value of a worker following success and failure in period 1:

$$V_S(q(\widehat{e})) - V_F(q(\widehat{e})) = (1-\lambda)[\widehat{p}(0, q_S(\widehat{e})) - \widehat{p}(0, q_F(\widehat{e}))]B > 0$$

In a competitive labour market the worker is paid his expected value and workers with a record of success in period 1 are able to command higher compensation. That, in turn, gives them incentives to be informed in the first period. I show below how the current employer of a worker can affect the difference in compensation, and therefore the effort of a worker, by adjusting the decision rights allocation.

1.4.3 Delegation

I turn now to consider the problem in the first period by comparing the two alternative decision rights arrangements. Start with delegation. When the worker is entirely responsible for the decision, his success or failure is free of the noise introduced by the employer. Similarly, the information held by the current employer and the market is identical: just the success or failure of the project chosen. With symmetric information a competitive market appears, and the worker is paid its expected value. Formally, $w_S^D(q(\widehat{e})) = V_S(q(\widehat{e}))$ and $w_F^D(q(\widehat{e})) = V_F(q(\widehat{e}))$. In period 1, the problem for the worker is:

$$\max_e \widehat{p}(e, 1/2)[V_S(q(\widehat{e})) - V_F(q(\widehat{e}))] - e$$

By rearranging the first order condition and evaluating at $\widehat{e} = e^D$:

$$\widehat{p}_e(e^D, 1/2) = \frac{1}{[V_S(q(e^D)) - V_F(q(e^D))]}$$

Comparing this expression with (1.1), it can be seen that the level of incentives is suboptimal. The profits earned by the employer under delegation are:

$$\Pi^D = (1 - \lambda)\widehat{p}(e^D, 1/2)B \tag{1.2}$$

Expression (1.2) shows the two drawbacks of delegation. On one side the precision of a motivated worker is useful to the employer only in the $(1 - \lambda)$ percentage of cases in which there is coincidence of interests ex post. On the other, the employer earns profits only during the first period. By failing to have superior information over the ability of the worker, the employer is forced to pay competitive wages in period 2, and therefore cannot earn a rent on him.

1.4.4 Authority

The analysis is a bit more complicated in this subsection, since it deals with the study of the winner's curse effect. Since the employer has the final say, the decision-making process is ex post efficient, using all the information available at a particular time. With a probability λ the employer will not rubber-stamp the proposal of the worker, instead choosing an alternative project, so that the choice of project is not one-to-one related to the signal of the worker. Because the market cannot observe directly the signal but the employer can, asymmetric information about the ability of the worker appears. The corresponding winner's curse is a well-known phenomenon studied first in the auction literature. In the labour market Greenwald [7] was the first to show that the market offers will lie strictly above the lower bound of the worker's value only if there is exogenous turnover with positive probability.

To illustrate the mechanics of the winner's curse and show why it reduces the future wage and the sensitivity of the wage to the quality of the signal, imagine that the project selected has been a success. The market is unsure as to whom allocate the credit, but knows that with probability $t_S(\widehat{e}) = \frac{(1-\lambda)\widehat{p}(\widehat{e},1/2)}{(1-\lambda)\widehat{p}(\widehat{e},1/2)+\lambda[1-\widehat{p}(\widehat{e},1/2)]}$ the project has been chosen by the worker (i.e. the value of the worker is $V_S(q(\widehat{e}))$ and with complementary probability it was chosen by the employer, and the value of the worker is $V_F(q(\widehat{e})$. The employer knows the value with certainty.

Imagine that the market naively offers the worker his expected value, that is an offer of $t_S(\widehat{e})V_S(q(\widehat{e}) + [1 - t_S(\widehat{e})]V_F(q(\widehat{e})$. The employer will match the market's offer when he observes the worker's value to be $V_S(q(\widehat{e})$. It follows that the market will employ the worker only when he is of lower ability, paying him a wage above his value, and therefore making a loss on him.

Taking into account that with probability u the worker will leave his current employer regardless of the wage offer, the equilibrium wage offer by the market must solve:

$$u\{t_S(\widehat{e})V_S(q(\widehat{e}) + [1 - t_S(\widehat{e})]V_F(q(\widehat{e}) - w_S^A(q(\widehat{e}))\} + (1-u)[1 - t_S(\widehat{e})]\left[V_F(q(\widehat{e}) - w_S^A(q(\widehat{e}))\right] = 0$$

The expression above illustrates that the market makes profits on the unconditional quitters, and losses on the workers on whom the employer ceases to match the market's offer. The losses made on these last workers has received in the literature the name of the winner's curse effect. Note that this effect is high, reducing the equilibrium wages, when exogenous turnover (u) is low and the importance of the employer's private information (λ) is high. Solving $w_S^A(q(\widehat{e}))$ from the expression above,

$$w_S^A(q(\widehat{e})) = \frac{ut_S(\widehat{e})V_S(q(\widehat{e}) + [1 - t_S(\widehat{e})]V_F(q(\widehat{e})}{u + (1-u)[1 - t_S(\widehat{e})]}$$

Note that $w_S^A(q(\widehat{e}))$ varies between $V_F(q(\widehat{e})$ and $t_S(\widehat{e})V_S(q(\widehat{e}) + [1 - t_S(\widehat{e})]V_F(q(\widehat{e})$ as u moves from 0 to 1. Without exogenous turnover the market employs the worker only when he is of low ability and pays him accordingly. With complete exogenous turnover there is no winner's curse and the worker is paid his expected value.

Similarly,

$$w_F^A(q(\widehat{e})) = \frac{ut_F(\widehat{e})V_S(q(\widehat{e}) + [1 - t_F(\widehat{e})]V_F(q(\widehat{e})}{u + (1-u)[1 - t_F(\widehat{e})]}$$

where $t_F(\widehat{e}) = \frac{(1-\lambda)(1-\widehat{p}(\widehat{e},1/2))}{(1-\lambda)(1-\widehat{p}(\widehat{e},1/2))+\lambda\widehat{p}(\widehat{e},1/2)}$. What a worker is paid following success exceeds what he is paid following failure by an amount $w_S^A(q(\widehat{e})) - w_F^A(q(\widehat{e})) < V_S(q(\widehat{e})) - V_F(q(\widehat{e}))$.

The problem of the worker under authority becomes:

$$\max_e \widehat{p}(e, 1/2)\left[w_S^A(q(\widehat{e})) - w_F^A(q(\widehat{e}))\right] - e$$

Rearranging the first order condition and evaluating at $e^A = \widehat{e}$, the effort provided by the worker under authority solves:

$$\widehat{p}_e(e^A, 1/2) = \frac{1}{w_S^A(q(\widehat{e})) - w_F^A(q(\widehat{e}))}$$

The profits earned by the employer under authority are:

$$\Pi^A = (1-\lambda)\widehat{p}(e^A, 1/2)B + \lambda\left(1 - \widehat{p}(e^A, 1/2)\right)B + (1-u)R$$

The employer is able now to extract profits from three different sources. The first one is due to the worker's effort to gather information in order to appear more informed. Further, the decision-making process is more efficient ex post, since it uses also the information of the employer. Lastly, the employer is able to extract a rent R from the worker in the second period, due to her informational advantage over the market.

1.4.5 Authority versus Delegation

I conclude this section by presenting the theoretical results

Proposition 1 *The effort provided by the worker is always higher under delegation than under authority, ($e^D > e^A$)*

Proof Note that $w_S^A(q) - w_F^A(q) < V_S(q) - V_F(q)$, as $\lambda \in [0, 1/2]$ and $u \in [0, 1]$. The rest follows from the concavity of $\widehat{p}$. ■

The intuition of Proposition 1 is very straightforward. When the employer holds superior information about the worker's ability, the market offers a flatter wage for fear of suffering the winner's curse. Career concern incentives are weak and the worker takes too little effort.

The following Proposition compares profits under the two arrangements:

Proposition 2 *Under some parameter values it is profitable for the employer to delegate decision-making rights to the worker*

Proof Make $u = 0$ and λ arbitrarily close to 0. It is straightforward that $\Pi^D - \Pi^A = \left[\widehat{p}(e^D, 1/2) - \widehat{p}(e^A, 1/2)\right] B$, which is always positive by Proposition 1. ■

1.5 Conclusion

This chapter identifies a new advantage of delegation over authority. When career concerns effects are important and the effect of the winner's curse effect is strong in the labour market, small advantages in information held by the current employer can be highly detrimental to the incentives of a worker. Delegating decision-making rights can therefore be used as a way to establish an "arm's length relationship", credibly committing not to exploit privileged information on the worker's ability.

1.6 References

Aghion, P. and J. Tirole, 1997 "Formal and Real Authority in Organisations", *Journal of Political Economy,* 105, 1-29.

Baker, G., R. Gibbons and K. Murphy, 1999, "Informal Authority in Organisations" *Journal of Law, Economics and Organisation,* 15, 56-73

Carbonara, E., 1999, "The Optimal Allocation of Power in Organisations", Unpublished Manuscript

Cremer, J. 1995, "Arm's Length Relationships" *Quarterly Journal of Economics,* 110, 275-295.

Dewatripont, M., Jewitt, I. and J. Tirole 1999, "The Economics of Career Concerns, Part I: Comparing Economic Structures", *Review of Economic Studies,* 66, 183-198.

Fama, E., 1980, "Agency Problems and the Theory of the Firm", *Journal of Political Economy,* 87, 288-307.

Greenwald, B. 1986, "Adverse Selection in the Labour Market", *Review of Economic Studies,* 53, 325-347.

Holmstrom, B., 1982, "Managerial Incentive Problems: A Dynamic Perspective" in *Essays in Economics and Management in Honour of Lars Wahlbeck,* Helsinki: Swedish School of Economics.

Ortega, J. 2003, "Power in the Firm and Managerial Career Concerns", *Journal of Economics and Management Strategy,* 13: 212-244

Osterman, P., 1994, "Supervision, Discretion and Work Organisation", *American Economic Review Papers and Proceedings,* 84, 380-384.

Rajan, R. and L. Zingales, 2004, "Saving Capitalism from Capitalists", MIT Press

Stein, J., 1997, "Internal Capital Markets and the Competition for Corporate Resources", *Journal of Finance,* 52, 111-133.

The Economist, 27th of April, 2002

Chapter 2
Strategic Decisions, Motivation and Leadership Styles

Joint work with Marc Moller

2.1 Introduction

It is often argued that leaders should be good at listening and responsive to opinions other than their own. Morris, Willkers and Knasel [11], for instance, assert that "the act of listening is all the time helping to mark you out as a leader and to build your colleagues' commitment to you". Since the early 70s, the management literature on leadership has mostly shared this view. Democratic decision–makers, who seek ample advice and build consensus around their decisions, have been generally regarded as more effective than those who adopt a command–and–control, top–down style of leadership. Implicit in this view is the notion that a better decision is reached when a multitude of perspectives and opinions are taken into account. For example, Heifetz [6] states that "autocratic decision-making assumes that authorities have little to learn, and it limits their ability to test basic substantive, political and moral assumptions". But the presumed superiority of democratic leaders also stems from a new emphasis on the effects of decisions on the motivation of subordinates. Korsgaard, Shweiger and Sapienza [8] state that "a more complete view of effective decision processes should consider not only the quality of decisions but also the impact of such processes on team members' affective responses, such as commitment to the decision". In this respect, an inclusive approach to decision–making is widely believed to be better at enhancing the motivation of subordinates.

Yet, many leaders have adopted autocratic, decisive styles and thrived as a result of it; think of Jack Welch at General Electric or Margaret Thatcher in British politics. Many successful leaders pay little regard to the reaction of their subordinates to their decisions, even at the risk of alienating them. Instead the ability to take risks and stand by one's beliefs is often deemed to be one of the greatest assets of these leaders.

Furthermore, empirical evidence supporting the superiority of a democratic style of leadership remains inconclusive. Locke and Sweiger [10], for instance, review 46 studies testing the effects of participative decision–making on productivity. While 22% of these studies find a democratic style to induce higher team productivity, another 22% report the superiority of an autocratic style, and 56% find no significant difference[1]. In a similar exercise, Locke and Sweiger find that 60% of studies conclude that democratic styles lead to higher group satisfaction than autocratic styles, whereas only 9% of studies report the opposite finding. In short, democratic leaders may produce happier, perhaps more motivated, subordinates, but not necessarily more productive ones.

Our aim in this chapter is to study theoretically the trade-offs between different leadership styles. We posit that leadership is a combination of the ability to make good decisions and to motivate subordinates in charge of implementing those decisions[2]. We view

1 See also Amason [1], Filley, House and Kerry [3], Latham, Winters and Locke [9], Wagner [19], Wagner and Gooding [20] and Wagner, Leana, Locke and Schweiger [21].

2 This view is common in the management literature on leadership. See, for instance, Vroom and Yetton [18]. For other papers in the economics literature, see Rotemberg and Saloner [13], Hermalin [7], Goel and Thakor [4] and Van den Steen [17].

these two aspects as complementary. The success of a (good) decision depends partly on the effort spent by the workforce to implement it. We propose a model in which the leader receives private information about the best choice of project for the organisation. He has to make two choices: whether or not to seek further evidence, and how to decide among the alternative projects.

The main trade-off of the model arises because the leader's private information might not point in the direction that is most effective for motivational purposes. This occurs, for instance, when the leader privately believes that some project is the most likely to be successful, but the worker would be more motivated with respect to another project, either because public evidence supports it or because its productivity, if it were the right choice, would be higher. In such cases, the leader might be willing to distort his decision and choose a suboptimal project which induces higher effort from the subordinate. We show that in an incomplete contracts world this is inefficient, as the leader should be transparent in his decision–making and let the worker adjust his effort according to the organisation's prospects. It is a tempting response because the leader does not fully internalise the cost of the worker's effort. A trade-off therefore arises in our model between the quality of a decision and the motivation of subordinates induced by it. There is evidence that real decision–makers explicitly recognise the existence of this trade-off. For instance, Thomas [16] states that "leaders struggle with the following problem all the time. From a leadership point of view, you always want to move toward telling the hard truths and helping people cope with the realities of change. But as a manager, you might be more inclined to minimize the complexity of a situation so things can run smoothly for as long as possible."

In this framework we compare democratic and autocratic leadership styles, using the self–confidence of the leader as a differentiating variable. We assume that the quality of the private information received by different types of leaders is identical. However, while some leaders are "realistic" about the quality of their private information, other leaders overestimate it, and are therefore "overconfident"[3]. We first show that an overconfident leader is less likely to compromise in his decisions. Since he (wrongly) believes that he is very well informed, he is unwilling to distort his project choice in order to boost the motivation of the workers. Hence, his behavioural bias makes him more willing to "tell the hard truths" and make the right decisions, even at the cost of reducing the workers' motivation. A realistic leader, on the other hand, understands the limits of his private information and is willing to compromise in his project choice thereby adopting an inefficient decision.

While more efficient at taking the right decisions given their personal opinion, we show that overconfident leaders fail to gather enough external evidence which might complement and improve their private information. Since such leaders overestimate the quality of their private information, they undervalue the benefits that consulting with external sources might generate, thus devoting too few resources to such consultation. In our model a trade-off arises endogenously between the acquisition of information and its use. Realistic leaders are more democratic. They frequently seek external advice and they take into account their subordinates' reaction to their decisions. Overconfident leaders, on the other hand, fail to consult enough and disregard their subordinates' motivation, thereby adopt-

[3] In reality, leaders also differ in their ability, i.e. in the quality of their private information, and it is beneficial to appoint a more able leader. We prefer to abstract from these differences and focus on the alternative styles induced by different levels of self–confidence.

ing a more autocratic approach to decision–making. In this chapter we evaluate the relative benefits from appointing an overconfident leader. We find that appointing an overconfident leader improves an organizations efficiency if external consultation is sufficiently costly. If external consultation is expensive the relative disadvantage of inefficiently low external consultation is outweighed by the leader's improved decision–making[4].

The plan of this chapter is as follows. In Section 2.2 we briefly discuss the related literature. Section 2.3 contains the model. In Section 2.4 we deal with motivational issues whereas Section 2.5 discusses a leader's incentive for external consultation. In Section 2.6 we consider the relationship between efficiency and a leader's level of overconfidence. Section 2.7 discusses the robustness of our results and Section 2.8 concludes. Proofs which do not serve the intuition for our results are relegated to the Appendix.

2.2 Related work

The economic literature on leadership so far has failed to consider how decision–making and motivational aspects interrelate and sometimes conflict with each other. Hermalin [7] proposes a model of team production where the leader does essentially the same task as any of his subordinates: put non–contractible effort into a common project. The leader differs from his subordinates in that he is privately informed about the productivity of the project. As he benefits from any extra effort incurred, the leader cannot credibly transmit to the

---

[4] In this paper we assume that a leadership style is uniquely associated with each leader. In reality, however, there are examples of leaders who have changed their styles according to the circumstances. The Economist [15], for instance, comments that "Mr Blair used to be liked because he said things that nearly everyone could agree with and because he seemed to be a decent and reasonably normal chap. But now Mr Blair is widely seen as a resolute leader, prepared to take risks for what he believes in."

team that the productivity of the project is high. However, Hermalin shows that the leader can signal this fact by exerting high effort himself; a mechanism that he calls "leading by example". While we share with Hermalin the notion that the leader has some privileged information about the company, in our model the leader's task is to take decisions, rather than to engage in the implementation process. Both views are, however, complementary. Hermalin's "leading by example" mechanism is probably important in small teams where subordinates can observe a leader's effort. On the other hand our model applies to larger organisations where subordinates and leaders cannot monitor each other and often do not even know each other personally. Rotemberg and Saloner [13] examine a framework where subordinates can create innovative ideas and the company can choose whether or not to test them. Because both activities are costly, the company will test only the most promising ideas and subordinates devote inefficiently little effort to innovation. Rotemberg and Saloner show how the company can gain by appointing a leader who empathises with the subordinates and is therefore not completely profit–maximising. The reason is that such a leader will test more ideas than optimal ex post, thereby creating an atmosphere where innovation is encouraged. Rotemberg and Saloner's setup applies to organisations where information/ideas reside at the bottom, whereas in our model information resides at the top.

A large part of the literature has focused on "irrational" leaders. Goel and Thakor [4] argue that managers who underestimate project risks have a higher likelihood of winning in tournaments and thus of getting elected as a leader. We share with them the study of overconfidence as a leader's attribute. Our focus, however, is on the relative benefits of different levels of overconfidence, rather than the forces which make leaders gain their position. Lastly, a few studies, such as Prendergast [12] or Branderburger and Ben-Polak [2] find that a decision–maker may avoid the project that he considers to be the best in favor of a different project that another agent who is less informed regards highly. Prendergast proposes a very simple reason; the use of subjective compensation by firms rewards workers if they choose to do those projects that their superiors prefer. Branderburger and Ben-Polak's focus is on the stock market, where they show how the fixation of managers with the share price can lead them to adopt the strategies that are preferred by investors. Unlike those papers, however, we show that appointing an overconfident decision–maker is a potential solution to the problem.

2.3 The model

Our environment consists of a leader and a single worker; the leader is in charge of making decisions and the worker has to implement these decisions. More specifically, the leader has to choose between two alternative and mutually exclusive projects, $d \in \{0,1\}$ and the worker has to choose whether to exert effort ($e = 1$) or not ($e = 0$) in implementing the project.

There are two states of the world $X \in \{0,1\}$. We assume that even if the worker exerts effort, the success of a project is not guaranteed. Success depends also on the decision matching the state of the world. If the leader chooses $d = X$, and the worker exerts effort, then the project creates a revenue of $V_d > 0$, but otherwise revenue is zero. Projects are heterogeneous in the sense that one of them leads to higher potential revenue than the other. Without loss of generality we normalize $V_1 = 1$ and $V_0 = v \in (0,1)$.

Information. The state of the world, X, is initially unknown, but it is common knowledge that both states are equally likely to occur. The decision–maker can gather information about X from two different sources: a costless private signal received as a result of being in a leadership position, and a costly external consultation or gathering of evidence that might improve the information provided by his private signal.

The leader has the opportunity to consult external experts or conduct research in order to determine the state of the world. This consultation may be successful, in which case it provides perfect knowledge of the state of the world, or it may fail and produce no evidence at all. We assume that consultation is more likely to be successful when more resources are spent on it. More specifically, the leader chooses a probability $p \in [0, 1]$ with which consultation reveals the state of the world and pays a cost γp^2 where $\gamma > 0$.

If consultation fails the leader relies solely on his own opinion about X. We assume that he receives a private signal, $S \in \{0, 1\}$, of imperfect precision $\pi = \Pr(S = X) \in (\frac{1}{2}, 1)$.

Overconfidence. Note that π represents the leader's ability to make the right decision when he has not obtained external evidence. The objective of this model is to investigate the implications of appointing an overconfident leader. We therefore allow for the possibility that the leader believes the precision of his private signal to be $\pi^0 \geq \pi$. Then $(\pi^0 - \pi)$ is the leader's level of overconfidence, and a leader with $\pi^0 = \pi$ is a "realistic" leader. We assume that π and π^0 are common knowledge[5].

[5] Common knowledge here implies that the leader knows that the worker believes that the signal's precision is π and that the worker knows that the leader believes that the signal's precision is π^0.

Effort decision. The worker pays a cost C of exerting effort which is privately observed and uniformly distributed in $[0, 1]$.

Payoffs and contracts. Both agents are assumed to be risk neutral. We adopt an incomplete contracting approach by positing that restrictions on the form of contracts make it unfeasible to fully compensate the worker for his effort. In other words, it is not possible to "sell the firm to the worker". Instead we assume that the worker receives a fixed share $\beta \in (0, 1)$ of the project's revenue.

Let $D(X, d)$ denote a dummy variable which takes the value 1 when $X = d$ and 0 otherwise and note that we can write the project's revenue as $R = D(X, d)eV_d$. Letting $R_L = (1 - \beta)R$ and $R_W = \beta R$ denote the leader's and the worker's share of the project's revenue, the leader's payoff becomes $P_L = R_L - \gamma p^2$ and the worker's payoff is $P_W = R_W - Ce$. Total surplus is $P = P_L + P_W = D(X, d)eV_d - \gamma p^2 - Ce$.

Expectations of a leader with overconfidence π^0 will be denoted by $E\pi_0$. E_X and $E^S_{\pi_0}$ are the leader's expectations conditional on the consultation process being successful or not but before he learns X or S respectively[6].

Timing. The timing of events is as follows. *(i)* Nature determines the state of the world X and the worker's cost C of exerting effort. *(ii)* The leader chooses the probability p of finding evidence about X. If evidence is found the state of the world X is revealed to everybody. If no evidence is found the leader receives the private signal S about the state of the world. *(iii)* The leader chooses the project *d*. *(iv)* The worker chooses whether to exert effort or not and the project's revenue is shared.

---

[6] Note that we have used lower case letters to denote choice variables and greek letters to denote the exogeneous parameters of the model.

2.4 Motivating subordinates

The aim of this section is to show that, when there is uncertainty about the state of the world, the decision–maker may face a trade-off between making the right decision and motivating his subordinate to implement it. To show this we solve the game by backward induction. Note first that in period *(iv)* it is optimal for the worker to exert effort if and only if his expected gains from doing so are larger than his costs C. In turn, the expected gains from exerting effort depend on the decision d taken by the leader in *(iii)* and on whether the consultation process in period *(ii)* has been successful or not. We consider the different outcomes of the consultation process separately, and study how the optimal effort choice of the worker affects the leader's project choice.

2.4.1 Motivation in the absence of uncertainty

Consider first the case where the consultation process has been successful and the state of the world has been revealed to everybody. In such case the leader's optimal choice of project is straightforward. If he chooses $d \neq X$ then the worker will not exert effort as the project has no possibility of delivering any revenue. If instead he chooses $d = X$ then the worker will exert effort if his cost C of doing so is smaller than his share of the revenue βV_X. As C is uniformly distributed in $[0, 1]$ the worker will exert effort with probability βV_X. The leader's expected revenue from choosing $d = X$ is therefore $(1 - \beta)\beta V_X^2 > 0$ so that $d = X$ is the optimal strategy for the leader.

In Section 2.5 we will discuss the leader's choice of the probability p for finding evidence. For this purpose we need to determine the revenue that the leader can expect if he finds unambiguous evidence about X. As both states of the world are equally likely ex ante we have $E^X R_L = \frac{1}{2}(1-\beta)\beta(1+v^2)$.

The following lemma summarizes these findings.

Lemma 1 *Consider the subgame after evidence about the state of the world has been found. In the unique equilibrium of this subgame the leader chooses* $d = X$ *and the worker exerts effort if and only if* $d = X$ *and* $C < \beta V_X$. *The leader's ex ante expected revenue from finding evidence is* $E^X R_L = \frac{1}{2}(1-\beta)(1+v^2)$.

Note that the leader would prefer to observe $X = 1$ rather than $X = 0$, since project 1 delivers higher revenue and makes the worker more likely to exert effort. However, once $X = 0$ has been observed, the leader has no incentive to choose $d = 1$, since this would reduce the probability of adopting the productive project *and* reduce the motivation of the worker. In the absence of uncertainty about X there is no trade-off between the quality of a decision and the motivation induced by it.

2.4.2 Motivation in the presence of uncertainty

Suppose that the consultation process has been unsuccessful and that the leader has to rely solely on his private signal when making his project choice. Note that from the worker's perspective both projects are equally likely to be the right choice. However, contingent on being the right choice, project 1 leads to higher gains than project 0. For motivational purposes the leader is therefore inclined to favour project 1. If a realistic leader with $\pi_0 = \pi$ observes $S = 0$ then assuming that the worker exerts effort the leader expects a revenue of $(1-\beta)\pi v$ if he chooses project 0 and $(1-\beta)(1-\pi)$ if he chooses project 1. If the signal's precision is small compared to the heterogeneity of the projects' revenues, that is if $\pi \leq \frac{1}{1+v}$, then the leader's expected revenue from project 1 would be larger than his expected revenue from project 0 irrespective of his signal. In this case the leader's favoured project would always be identical to the one which is optimal for motivational purposes and the leader's optimal project choice would trivially be $d = 1$. In order to restrict attention to the more interesting case we make the following assumption.

Assumption 1 *The signal S is valuable, that is $\pi > \frac{1}{1+v}$.*

Assumption 1 states that conditional on the worker exerting effort regardless of the project choice, the leader's expected revenue is higher when he makes his project choice contingent on his signal, $d = S$.

However, the worker's compensation depends on the outcome of the production process and his effort might be influenced by the leader's project choice. Hence, in the presence of uncertainty about X, the leader may be inclined to choose project 1 after observing signal 0 even when Assumption 1 holds.

To show this point suppose that the leader is realistic ($\pi_0 = \pi$) and that he follows the efficient strategy $d = S$. The worker will exert effort on project 0 if $C < \beta\pi v$ and on project 1 if $C < \beta\pi$. After observing $S = 0$ the leader thus gains by deviating to $d = 1$ if and only if $(1-\beta)\beta\pi^2 v^2 < (1-\beta)\beta\pi(1-\pi)$, that is, if $\pi < \frac{1}{1+v^2}$. We assume that this is the case, so that a realistic leader cannot credibly commit to choose $d = S$, and overconfidence may improve the efficiency of the decision–making. We therefore make the following assumption.

Assumption 2 *The signal S is sufficiently inaccurate to prevent a realistic leader from an efficient project choice, that is $\pi < \frac{1}{1+v^2}$.*

The reason why a realistic leader does not find it optimal to follow $d = S$ is that, while both agents benefit from a higher level of revenue, only the worker pays the cost of implementing the project. Because the leader benefits from an increase in the effort of the worker without paying its extra cost, he has an incentive to motivate the worker with his project choice. This increase in motivation is not costless for the leader, as it requires adopting a project that he considers less likely to be the right choice. Under Assumption 2 the benefits of this increase in motivation are higher than its costs. The gains from distorting the project choice depend positively on the heterogeneity in the projects revenues and negatively on the precision of the leader's private signal. Although under Assumption 2 a realistic leader cannot adopt the efficient project choice, an overconfident leader may do so. The following Lemma characterizes equilibrium behaviour for different levels of overconfidence π^0. The Intuitive Criterion is used as a selection device. Strictly speaking, such criterion is defined for signaling games but its extension to our case is straightforward and its definition is omitted.

Lemma 2 *Consider the subgame after no evidence about the state of the world has been found. Depending on the leader's level of overconfidence there exist different types of equilibria.*

- *If* $\pi_0 \leq (1 + 2\pi v^2)^{-1}$ *then in every equilibrium the leader chooses* $d = 1$ *irrespective of his signal and the worker exerts effort on project* 1 *if and only if* $C < \frac{1}{2}\beta$. *The leader's ex ante expected revenue from using his private signal is* $E^S_{\pi_0} R_L = \frac{1}{4}(1-\beta)\beta$.

- *If* $(1 + 2\pi v^2)^{-1} < \pi_0 < (1 + v^2)^{-1}$ *then there exists a unique equilibrium which survives the Intuitive Criterion. The leader chooses* $d = 1$ *if* $S = 1$. *When* $S = 0$ *he chooses* $d = 1$ *with probability* $q = \frac{\pi(1-\pi_0(1+v^2))}{\pi\pi_0 v^2 - (1-\pi)(1-\pi_0)}$ *and* $d = 0$ *with probability* $1 - q$ *where* $0 < q < 1$. *The worker exerts effort on project* 1 *if and only if* $C < \beta\frac{\pi\pi_0 v^2}{(1-\pi_0)}$. *He exerts effort on project* 0 *if and only if* $C < \beta\pi v$. *The leader's ex ante expected revenue from using his private signal is* $E^S_{\pi_0} R_L = \frac{1}{2}(1-\beta)\beta v^2 \pi \frac{\pi_0}{1-\pi_0}$

- *If* $(1 + v^2)^{-1} \leq \pi_0$ *then there exists a unique equilibrium. The leader chooses* $d = S$. *The worker exerts effort if and only if* $C < \beta\pi V_d$. *The leader's ex ante expected revenue from using his private signal is* $E^S_{\pi_0} R_L = \frac{1}{2}(1-\beta)\beta\pi\pi_0(1 + v^2)$.

Proof We first show that there cannot be an equilibrium in which the leader chooses $d = 0$ with positive probability after $S = 1$. By contradiction suppose that there is such an equilibrium. For any project choice d of the leader let w_d denote the worker's belief about the probability with which the leader has chosen the right project. The worker will optimally exert effort if and only if $C < \beta V_d w_d$. Now suppose that $S = 1$. As in equilibrium the leader chooses $d = 0$ with positive probability it has to hold that $(1-\beta)\beta(1-\pi_0)w_0 v^2 \geq (1-\beta)\beta\pi_0 w_1$. However, as $w_d \in [1-\pi, \pi]$ it holds that

$(1-\pi_0)w_0v^2 \leq (1-\pi_0)\pi v^2 < (1-\pi_0)\pi \leq (1-\pi)\pi_0 \leq \pi_0 w_1$ which is a contradiction. It follows that in every equilibrium the leader chooses $d = 1$ if $S = 1$. We will now discuss the remaining possibilities for an equilibrium. Suppose first that the leader chooses $d = 1$ irrespective of his signal S. As the leader does not condition the project choice on his signal, the worker knows that the leader chooses the right project with probability $\frac{1}{2}$. It is therefore optimal for the worker to exert effort on project 1 if and only if $C < \frac{1}{2}$. If the worker observes the off–equilibrium choice $d = 0$ we suppose that he believes that $S = 0$ with probability $w \in [0,1]$. It is then optimal to exert effort on project 0 if and only if $C < \beta v(w\pi + (1-w)(1-\pi))$. Given the worker's equilibrium behaviour the leader's expected revenue from following the equilibrium strategy is $(1-\beta)\beta\frac{1}{2}\pi_0$ if $S = 1$ and $(1-\beta)\beta\frac{1}{2}(1-\pi_0)$ if $S = 0$. The leader's expected revenue from deviating to $d = 0$ is $(1-\beta)\beta(1-\pi_0)(w\pi + (1-w)(1-\pi))v^2$ if $S = 1$ and $(1-\beta)\beta\pi_0(w\pi + (1-w)(1-\pi))v^2$ if $S = 0$. Possible gains from deviating are larger if $S = 0$ and deviations are not profitable if $\frac{1}{2}(1-\pi_0) > \pi_0(w\pi + (1-w)(1-\pi))v^2$ which is equivalent to $\pi_0 < (1 + 2(w\pi + (1-w)(1-\pi))v^2)^{-1}$. Note that this threshold is minimal for $w = 1$. For $\pi_0 < (1 + 2\pi v^2)^{-1}$ choosing $d = 1$ for all S is an equilibrium no matter the off–equilibrium belief of the worker.

The leader's ex ante expected revenue from using his private signal is $E^S_{\pi_0}R_L = \frac{1}{2}(1-\beta)\beta(\frac{1}{2}\pi_0 + \frac{1}{2}(1-\pi_0)) = \frac{1}{4}(1-\beta)\beta$.

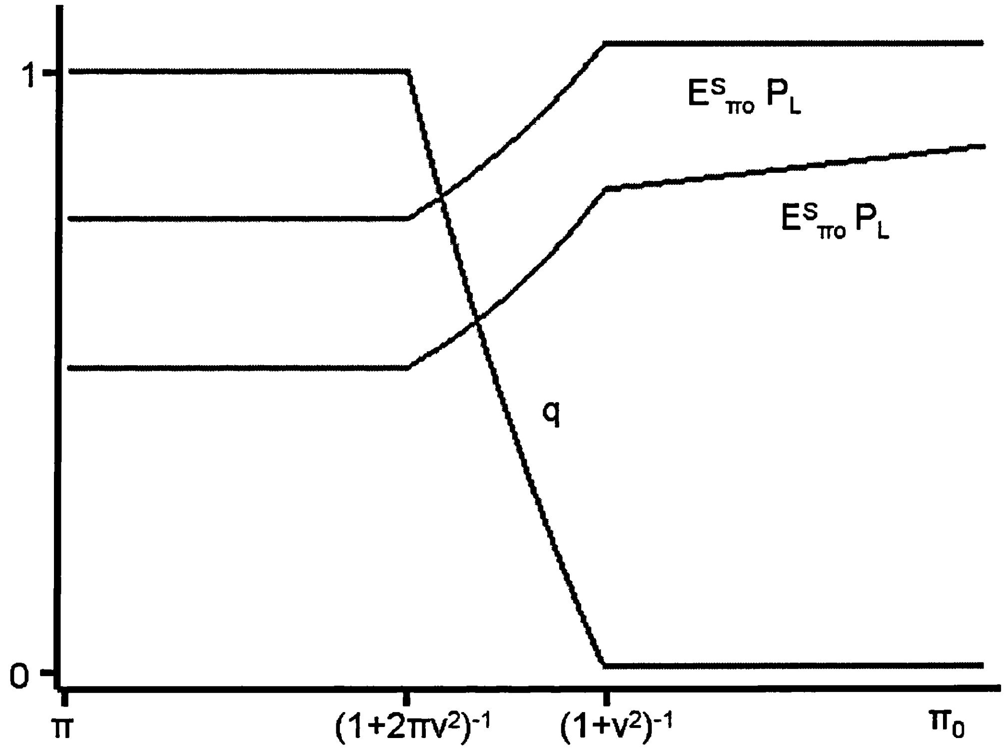
For $(1 + 2\pi v^2)^{-1} < \pi_0 < (1 + 2(1-\pi)v^2)^{-1}$ however, choosing $d = 1$ for all S is an equilibrium only if the worker believes that $S = 1$ with positive probability after $d = 0$, that is if $w < 1$. We will now show that for $(1 + 2\pi v^2)^{-1} < \pi_0 < (1 + 2(1-\pi)v^2)^{-1}$ such beliefs are ruled out by the Intuitive Criterion. To see this consider a leader of type $S = 1$. For any possible off-equilibrium belief of the worker the leader's payoff from deviating to $d = 0$ is smaller than $(1 - \beta)\beta(1 - \pi_0)\pi v^2$. His payoff from following the equilibrium path is $(1 - \beta)\beta\pi_0\frac{1}{2}$. It holds that $(1 - \pi_0)\pi v^2 < \pi_0(1 - \pi)v^2 < \frac{1}{2}(1 - \pi_0) < \frac{1}{2}\pi_0$ where the second inequality follows from $\pi_0 < (1 + 2(1 - \pi)v^2)^{-1}$. This implies that a leader of type $S = 1$ cannot profit from a deviation to $d = 0$ for any belief of the worker. For $(1 + 2\pi v^2)^{-1} < \pi_0 < (1 + 2(1 - \pi)v^2)^{-1}$ the equilibrium in which $d = 1$ for all S is therefore ruled out by the Intuitive Criterion.

Now suppose that the leader chooses $d = 1$ if $S = 1$ and $d = 1$ with probability q if $S = 0$. The worker can calculate the probability that the right project was chosen using Bayesian updating. For $d = 0$ the worker knows that $S = 0$ such that it is optimal for him to exert effort if and only if $C < \beta\pi v$. If $d = 1$ exerting effort is optimal if and only if $C < \beta\frac{\pi+q(1-\pi)}{1+q}$. If $S = 0$ the leader has to be indifferent between choosing $d = 0$ or $d = 1$. If he chooses $d = 0$ his expected revenue is $(1 - \beta)\beta\pi_0 v^2$ and for $d = 1$ he expects $(1 - \beta)\beta(1 - \pi_0)\frac{\pi+q(1-\pi)}{1+q}$. Setting these two revenues equal and solving for q leads to $q = \frac{\pi(1-\pi_0(1+v^2))}{\pi\pi_0 v^2-(1-\pi)(1-\pi_0)}$. Note that Assumption 1 implies that the denominator of q is positive. The nominator is positive if and only if $\pi_0 < (1 + v^2)^{-1}$and $q < 1$ if and only if $(1 + 2\pi v^2)^{-1} < \pi_0$. It remains to check that it is not profitable for the leader to choose $d = 0$ if $S = 1$. This deviation gives the leader an expected revenue of $(1-\beta)\beta(1-\pi_0)\pi v^2$ whereas by following the equilibrium strategy he gets $(1 - \beta)\beta\pi_0\frac{v^2\pi\pi_0}{1-\pi_0}$. As $1 - \pi_0 < \pi_0$ this deviation is not profitable. For $(1+2\pi v^2)^{-1} < \pi_0 < (1+v^2)^{-1}$ the proposed strategies therefore constitute an equilibrium and the leader's ex ante expected revenue from using his private signal is $E^S_{\pi_0} R_L = \frac{1}{2}(1-\beta)\beta\left((1-q)\pi_0\pi v^2 + (\pi_0 + q(1-\pi_0)\frac{\pi+q(1-\pi)}{1+q})\right) = \frac{1}{2}(1-\beta)\beta v^2\pi\frac{\pi_0}{1-\pi_0}$.

Finally suppose that the leader chooses $d = S$. The worker optimally exerts effort if and only if $C < \beta\pi V_d$. If the leader's signal is S his expected revenue from following the equilibrium is $(1-\beta)\beta\pi_0 V_S^2$. By deviating he expects $(1-\beta)\beta(1-\pi_0)\pi V_{1-S}^2$. As $\pi_0 > 1-\pi_0$ a deviation can only be profitable if $S = 0$. Deviating after $S = 0$ is not profitable if and only if $1-\pi_0 < \pi_0 v^2$ which is equivalent to $(1+v^2)^{-1} < \pi_0$. If this equilibrium is played the leader's ex ante expected revenue from using his private signal is $E^S_{\pi_0} R_L = \frac{1}{2}(1-\beta)\beta\pi_0(1+v^2)$. ■

Figure 1 shows that the equilibrium probability q of choosing project 1 when the leader's signal is 0 is a decreasing function of the leader's level of overconfidence π_0. It also shows that the payoff which the leader expects to obtain from using his private signal, $E^S_{\pi_0} P_L$, is an increasing function of π_0. To understand the intuition of this equilibrium note first that when π_0 is below $(1+2\pi v^2)^{-1}$, the leader is not very confident about his ability to predict the state of the world. As a result he regards the costs of distorting the project choice as low. Since in equilibrium the leader chooses $d = 1$ independently of his private signal, his expected revenue is independent of his level of confidence about the precision of such signal. However, as π_0 increases above $(1+2\pi v^2)^{-1}$, q decreases. The leader starts using his private signal, and his expected revenue increases in π_0 for two reasons. The first reason is that as π_0 increases, q decreases and the decision-making becomes more efficient, thus increasing both the likelihood of choosing the right project and the expected effort of the worker. The second reason is that, as π_0 increases, the leader becomes more and more confident in his ability to choose the right project. Lastly when $\pi_0 > (1+v^2)^{-1}$ the project choice of the leader is completely efficient in equilibrium and the increase in the leader's expected revenue is solely due to the second reason.

Figure 1



2.5 Consulting advisors

In this section we move to period *(ii)* to analyse the leader's decision to consult externally and seek further evidence about the state of the world X. The leader takes into account the payoffs that he expects to obtain in the subgames where evidence about X has been found or not. He chooses p, the probability that evidence will be found, to maximize his expected payoff $E_{\pi_0} P_L$, that is

$$p^* = \underset{p\in[0,1]}{arg\ max}\ pE^X R_L + (1-p)E^S_{\pi_0} R_L - \gamma p^2$$

Lemma 1 has shown that the leader's ex ante expected revenue from finding evidence is $E^X R_L = \frac{1}{2}(1-\beta)\beta(1+v^2)$. According to Lemma 2 the leader's ex ante expected revenue from using his private signal $E^S_{\pi_0} R_L$ depends on his level of overconfidence. The leader's optimal choice, p^*, will therefore depend on π_0, as stated in the following lemma.

Lemma 3 *The leader's equilibrium consultation effort p^* is strictly positive and depends on his level of overconfidence.*

- *For* $\pi_0 < (1+2\pi v^2)^{-1} : p^* = min\left(\frac{1}{4\gamma}(1-\beta)\beta(\frac{1}{2}+v^2), 1\right)$
- *For* $(1+2\pi v^2)^{-1} < \pi_0 < (1+v^2)^{-1} : p^* = min\left(\frac{1}{4\gamma}(1-\beta)\beta(1+v^2-\frac{\pi\pi_0 v^2}{1-\pi_0}), 1\right)$
- *For* $(1+v^2)^{-1} < \pi_0 : p^* = min\left(\frac{1}{4\gamma}(1-\beta)\beta(1+v^2)(1-\pi\pi_0), 1\right)$

2.6 The implications of overconfidence

Having derived the equilibrium consultation decision and the equilibrium project choice, we now turn to the analysis of the relative advantages of appointing an overconfident leader.

2.6.1 Improved decision-making

Suppose that the leader's consultation effort is fixed exogenously, say at $\overline{p} < 1$. This would, for example, be the case if consultation of certain advisors is obligatory. Consider the leader's use of his private information. Efficiency would require that the leader chooses $d = S$. This is because by doing so the leader maximizes the probability of choosing the right project such that the worker's effort is least likely to be wasted. If the leader could commit to choose $d = S$ it would be optimal for the worker to exert effort if and only if $C < \beta\pi V_d$. Then expected total surplus would be

$$E_\pi P^{eff} = \frac{1}{2}\beta(1 - \frac{1}{2}\beta)(\overline{p} + (1 - \overline{p})\pi^2)(1 - v^2) - \gamma\overline{p}^2$$

Lemma 2 has shown that a sufficiently realistic leader distorts his project choice with strictly positive probability in order to motivate the worker to exert effort. If the leader is sufficiently overoptimistic, however, he overestimates the benefits of adopting the project that his signal points at, therefore choosing it efficiently. Proposition 3 formalises the advantages of appointing an overconfident rather than a realistic leader for the organisation's total surplus.

Proposition 3 *For a constant level of consultation, $\overline{p} < 1$, consider the dependence of total surplus on the leader's level of overconfidence. Expected total surplus is*

$$E_\pi P = \frac{1}{2}\beta(1 - \frac{1}{2}\beta)(\overline{p}(1 + v^2) + (1 - \overline{p})f(\pi_0)) - \gamma\overline{p}^2$$

- *If $\pi_0 \leq (1 + 2\pi v^2)^{-1}$ then $f(\pi_0) = \frac{1}{2}$ and total surplus is constant and strictly smaller than its efficient level.*

- *If $\pi_0 \in [(1 + 2\pi v^2)^{-1}, (1 + v^2)^{-1}]$ then $f(\pi_0) = \frac{\pi^2 v^2((1-\pi_0)^2 + v^2\pi_0(\pi_0 - 2\pi))}{(\pi_0 - 1)(\pi\pi_0(v^2 - 1) + \pi_0 + \pi - 1)}$ and total surplus is strictly increasing and smaller than its efficient level.*

- *If* $\pi_0 \geq (1+v^2)^{-1}$ *then* $f(\pi_0) = \pi^2(1+v^2)$ *and total surplus is equal to its efficient level.*

In Figure 1 in Section 2.4 we have plotted the expected total surplus from using the leader's private signal, $E^S_\pi P$, against the leader's level of overconfidence. It can be seen that the increase in total surplus is a result of a decrease in the leader's likelihood of distorting his project choice in order to boost motivation. Higher decision–making efficiency is beneficial to the team, as it increases the probability of adopting the productive project. It is also beneficial because it improves the worker's ability to regulate his effort level according to the productivity of the team, thereby minimising the possibility that his effort will be wasted. Note that when π_0 is higher than $(1+v^2)^{-1}$, q drops to zero, decision-making is completely efficient, and total surplus ceases to increase in π_0.

2.6.2 Insufficient consultation

In this section we show that overconfident leaders devote very limited resources to seek external advice. To see this, note that while total surplus can be written as

$$E_\pi P = pE^X(R - Ce) + (1-p)E^S_\pi(R - Ce) - \gamma p^2$$

the leader chooses p to maximize

$$E_{\pi_0} P_L = pE^X(R_L) + (1-p)E^S_{\pi_0}(R_L) - \gamma p^2$$

Let $p^{eff} = arg\ max_{p\in[0,1]}\ E_\pi P$ denote the efficient effort level and let $p^*(\pi_0) = arg\ max_{p\in[0,1]} E_{\pi_0} P_L$ denote the leader's optimal choice. There are two reasons why the leader's choice is inefficiently low. The first reason is the familiar team production effect by which the leader fails to internalise the fact that the worker benefits when he obtains better information about X. By comparing $E_\pi P$ with $E_{\pi_0} P_L$ we can see that the leader focuses on R_L rather than on $R - Ce$. The leader therefore neglects the worker's share of revenue as well as his cost of implementing the project.

The second reason for $p^*(\pi_0) \leq p^{eff}$ is that an overconfident leader overestimates the expected revenue from using his private signal. Note that p^* depends negatively on the relative advantage of finding clear evidence rather than having to rely on the imprecise private signal, $E^X R_L - E^S_{\pi_0} R_L$. It follows from Lemma 2 and Figure 1 that $E^S_{\pi_0}(R_L)$ is constant for $\pi_0 < (1+2\pi v^2)^{-1}$ and strictly increasing in the leader's level of overconfidence for $\pi_0 > (1+2\pi v^2)^{-1}$. Clearly, when the leader is more overconfident, he expects a higher revenue from using his private signal, and he chooses a lower p^*. The following proposition summarizes these findings.

Proposition 4 *The leader's consultation effort is inefficiently low and weakly decreasing in his level of overconfidence.*

2.6.3 Can overconfidence improve efficiency?

In the previous sections, we have seen that overconfidence improves the use of a leader's information but hinders its acquisition. In this section, we study the elements that determine the relative advantages of appointing an overconfident rather than a realistic leader. The following proposition characterizes the leader's level of self–confidence which maximizes total expected surplus.

Proposition 5 *Whether overconfidence can improve efficiency depends on the leader's cost of consultation γ. If $\gamma \leq \gamma^*$ then total surplus is maximized for a realistic leader ($\pi_0 = \pi$). If $\gamma > \gamma^*$ then overconfidence improves efficiency. In this case total surplus is maximized for an overconfident leader with $\pi_0 = (1 + v^2)^{-1} > \pi$.*

The optimal level of confidence depends positively on γ. When γ increases, external consultation becomes more and more wasteful compared to the costless use of the leader's private signal. An overconfident leader with $\pi_0 \geq (1 + v^2)^{-1}$ can use his private signal efficiently whereas a realistic leader makes inefficient project choices due to his strong motivational concerns. For sufficiently large γ, total surplus is thus higher if the leader is sufficiently overconfident.

For $\gamma > \gamma^*$ the optimal level of confidence $\pi_0 = (1 + v^2)^{-1}$ is decreasing in v. When v is low, the worker's motivation associated with different project choices differs strongly. It thus requires a high level of overconfidence to prevent the inefficiencies associated with the leader's motivational concerns.

2.7 Robustness

The model outlined in Section 2.3 contains some assumptions that are necessary for our results, and others that only serve to simplify the argument. In this section we discuss these assumptions separately.

2.7.1 Simplifying assumptions

In our model, we assume that the leader incurs the entire cost of consulting externally. This is not necessarily the case in an organisation, and a more realistic scenario would assume that such cost is divided among the two agents. This would be the case, for example, if the consultation process took the form of a delay of the project until better evidence is available and future payoffs are discounted. It is clear, however, that the argument that an overconfident leader underestimates the benefits of consulting does not hinge on the exact share of the costs he incurs, as long as this share is strictly positive.

The assumption that consultation produces either perfect knowledge of X or no evidence at all is also a simplifying assumption. The same holds for the assumption that the worker can perfectly observe whether consultation was successful or not. Relaxing these assumptions would considerably complicate the algebra. However, the conclusion that consultation increases the expected revenue of the organisation would remain unchanged. Similarly, in our model the leader observes the value of the private signal only after he has chosen the level of consultation. Alternatively, it could be assumed that he receives his signal before consultation takes place. In this case the level of consultation would generally depend on his private information. His consultation effort could therefore serve as a signal of his type, that is whether he observed $S = 0$ or $S = 1$. However, both types of leaders, if they are overconfident, would engage in too little consultation effort. Similarly, as long as such signalling process does not reveal the leader's type perfectly, he would still have the option to distort his project choice.

The assumption that the implementation cost of the worker is distributed uniformly greatly simplifies our calculations, as it causes the probability with which he exerts effort to increase linearly in his expectation of the project's productivity. Alternative assumptions about such relation would undoubtedly affect the incentives of the leader to distort the project choice. Such incentive will be present, however, as long as the worker's motivation is not independent of his view on the team prospects.

Lastly, a comment on our interpretation of the concept of "common knowledge". We assume throughout the chapter that the only deviation from full rationality is the fact that the leader overestimates the precision of his private signal. In particular, the worker knows the true precision, and the leader is aware of this fact. Alternatively, we could assume that the leader also overestimates the worker's perception of his ability, or even that he is "charismatic", and he manages to convince the worker that he is smarter than he actually is. These alternatives would clearly reinforce our result that overconfidence can serve as a counterweight to the temptation to improve the motivation of subordinates.

2.7.2 Necessary assumptions

Our results depend critically on the assumption that different project choices induce different levels of motivation amongst subordinates. We have chosen to model such asymmetry as arising from the difference between the projects' payoffs conditional on success, $1 - v$. It is important to emphasise that such asymmetry could be caused by other reasons. One could assume, for instance, that $v = 1$, but that the information technology is not symmetric, with different precisions associated with different signals, $\Pr(X = 1 \mid S = 1) =$

$\pi_1 > \pi_0 = \Pr(X = 0 \mid S = 0)$. In this case, the worker would be more motivated after observing $d = 1$ because he would expect that the leader is "more certain" about his choice than if he observed $d = 0$. Assumptions 1 and 2, could be rewritten as $\pi_0 > 1/2$ and $\pi_1(1 - \pi_0) > \pi_0^2$ respectively. Similarly, every result could be rewritten in terms of this new asymmetry.

Another important assumption is the fact that the production process cannot be sold to the worker completely. If the leader's payoff is independent of the team's revenue, then the leader would lose any incentive to motivate the worker with his project choice. There are different justifications for this assumption. The leader might receive a private, uncontractible, benefit when the project is successful. Alternatively, the worker may be subject to limited liability. In our model effort and the quality of the project choice are strategic complements, a frequent assumption in the economic literature. While the fact that the leader might distort his project choice in order to increase the motivation of the worker does not depend critically on this assumption, the direction of such distortion does. If effort and project choice were strategic substitutes, the leader would be tempted to choose the least productive project, hence trying to portray a bleaker picture to the worker in the hope of extracting higher effort from him.

2.8 Concluding comments

In the last two decades, the management literature on leadership has debated widely about how much impact CEOs have on their companies' economic performance. On one side, traditional management theorists have argued that CEOs have significant impact, as they are able to shape their companies' strategic direction. On the other side, organizational ecology researchers[7] have claimed that CEOs are so constrained by their environments that they have little ability to affect company performance. In this chapter we have adopted the mixed view that leaders have the formal authority to choose the direction of an organisation, but that the success of their choices depends partly on the motivation of the workforce. As a result, leaders are implicitly constrained in their choices, and may prove irrelevant if they are not confident enough about their own ability to guide the company in the right direction. Using this framework we have shown that the choice of leaders is influenced by a trade-off between the collection of information and its use. The question of whether efficiency could be achieved simultaneously in both fields, however, remains. In practice, we observe that, while organisations often appoint leaders with strong self-confidence, they also strive to explicitly formalise the consultation process. Organisational teams, for instance, often have weekly meetings where their leader gets to hear, whether he likes it or not, about all the different opinions within the team. Similarly, political leaders are often obliged to consult parliamentary committees before reaching any decision. Because the decision–maker is himself the most interested in reaching the right decision, it is perhaps a bit puzzling why he is not left to decide in what circumstances alternative opinions would enrich the quality of his decisions. Using the framework of this chapter, however, this can be interpreted as an attempt to counterweight the drawbacks of appointing an overconfident leader.

2.9 Appendix

[7] See for example Hannan and Freeman [5]

Proof of Lemma 3 The first order condition of the leader's maximization problem reads $E^X R_L - E^S_{\pi_0} R_L - 2\gamma p^* \geq 0$. The leader's expected revenue from finding evidence is strictly larger than his expected revenue from using his private signal no matter his level of overconfidence, $E^X R_L > E^S_{\pi_0} R_L$. The leader's optimal choice p^* will therefore be strictly positive and the values of p^* are found by substituting $E^X R_L$ and $E^S_{\pi_0} R_L$ taken from Lemma 1 and Lemma 2 respectively.

Proof of Proposition 3 Consider the leader's use of his private signal. Suppose that $\pi_0 \leq (1 + 2\pi v^2)^{-1}$. Lemma 2 has shown that in this case the leader will choose $d = 1$ no matter his signal. He will select the right project with probability $\frac{1}{2}$. Expected total surplus becomes

$$\frac{1}{2}\beta(1 - \frac{1}{2}\beta)(\overline{p}(1 + v^2) + (1 - \overline{p})\frac{1}{2}) - \gamma\overline{p}^2 < E_\pi P^{eff}$$

If $\pi_0 \geq (1+v^2)^{-1}$Lemma 2 has shown that the leader will choose $d = S$. He will therefore select the right project with probability π and expected total surplus becomes

$$\frac{1}{2}\beta(1 - \frac{1}{2}\beta)(\overline{p} + (1 - \overline{p})\pi^2)(1 + v^2) - \gamma\overline{p}^2 = E_\pi P^{eff}$$

Finally consider the intermediate case where $(1 + 2\pi v^2)^{-1} < \pi_0 < (1 + v^2)^{-1}$. In this case Lemma 2 predicts that the leader chooses project 1 if $S = 1$. However if $S = 0$ he randomizes, choosing project 1 with probability q and project 0 with probability $1 - q$. It follows that if $d = 0$ the leader's choice will be correct with probability π. If $d = 1$ project 1 is the right choice with probability $\frac{\pi+q(1-\pi)}{1+q}$. Expected total surplus in this case is

$$\frac{1}{2}\beta(1 - \frac{1}{2}\beta)\left(\overline{p}(1 + v^2) + (1 - \overline{p})f(\pi_0)\right) - \gamma p^2$$

where $f(\pi_0) = (1-q)\pi^2 v^2 + \frac{(\pi+q(1-\pi))^2}{1+q}$. Note that $f((1+2pv^2)^{-1}) = \frac{1}{2}$ and $f((1+v^2)^{-1}) = \pi^2(1+v^2)$ which implies that expected total surplus is continuous in π_0. Substituting for q from Lemma 1 gives

$$f(\pi_0) = \frac{\pi^2 v^2((1-\pi_0)^2 + v^2\pi_0(\pi_0 - 2\pi))}{(\pi_0 - 1)(\pi\pi_0(v^2-1) + \pi_0 + \pi - 1)^2}$$

For the first derivative one gets

$$\frac{\partial f}{\partial \pi_0} = \frac{\pi^2 v^4(2\pi - 1)(\pi(1-(1-v^2)\pi_0^2) + 2\pi_0(\pi_0 - 1))}{(\pi_0 - 1)^2(\pi\pi_0(v^2-1) + \pi_0 + \pi - 1)^2}$$

The derivative is strictly positive if $\pi(1-(1-v^2)\pi_0^2) + 2\pi_0(\pi_0 - 1) > 0$. This last expression is minimized at $\pi_0 = (2-(1-v^2)\pi)^{-1}$ where it takes the value $\frac{2\pi-(1-v^2)\pi^2-1}{2-(1-v^2)\pi}$. It follows from Assumption 1 that this value is strictly positive as it is strictly increasing in π and zero for $\pi = (1+v^2)^{-1}$. Thus expected total surplus is strictly increasing in π_0 for $\pi_0 \in [(1+2\pi v^2)^{-1}, (1+v^2)^{-1}]$.

Proof of Proposition 4 As the worker exerts effort if and only if his share of the project's expected revenue is larger than C it holds that $E^X(Ce) = \frac{1}{2}\beta E^X(R)$ and $E^S_\pi(Ce) = \frac{1}{2}\beta E^S_\pi(R)$. It follows that

$$\begin{aligned} E^X(R-Ce) - E^S_\pi(R-Ce) &= (1-\frac{1}{2}\beta)(E^X(R) - E^S_\pi(R)) \\ &> (1-\beta)(E^X(R) - E^S_\pi(R)) = E^X(R_L) - E^S_\pi(R_L) > 0 \end{aligned}$$

The first order conditions thus imply that $0 < p^*(\pi) \leq p^{eff}$ and that $p^*(\pi) < p^{eff}$ if $p^{eff} < 1$. It follows from Lemma 1 that $E^X(R) - E^S_{\pi_0}(R)$ is strictly positive and weakly decreasing in π_0. This implies that $p^*(\pi)$ is weakly decreasing in π_0.

Proof of Proposition 5 Using the leader's equilibrium effort of seeking advice p^* from Lemma 3 and $f(\pi_0)$ as defined in Proposition 3 the equilibrium level of total surplus for a leader with overconfidence π_0 is

$$E_\pi P(\pi_0) = \frac{1}{2}\beta(1-\frac{1}{2}\beta)(p^*(1+v^2)+(1-p^*)f(\pi_0)) - \gamma(p^*)^2$$

Let $\overline{\pi} = max(\pi, (1+2pv^2)^{-1})$. If $\pi_0 \leq \overline{\pi}$, the leader will choose $d = 1$ no matter his signal and $E_\pi P(\pi_0) = E_\pi P(\pi)$ as p^* and $f(.)$ are constant. If $\pi_0 \geq (1+v^2)^{-1}$ the leader chooses $d = S$ and it holds that $E_\pi P(\pi_0) \leq E_\pi P(\frac{1}{1+v^2})$ as p^* is non–increasing and $f(\pi_0) = \pi^2(1+v^2) < (1+v^2)$. We will now show that $E_\pi P(\pi_0) \leq max\left(E_\pi P(\overline{\pi}), E_\pi P(\frac{1}{1+v^2})\right)$ for any π_0 such that $\overline{\pi} < \pi_0 < (1+v^2)^{-1}$. In this range the leader chooses $d = 1$ with probability $q = \frac{\pi(1-\pi_0(1+v^2))}{\pi\pi_0 v^2-(1-\pi)(1-\pi_0)} \in (0,1)$ after observing the signal $S = 0$. By inverting the relationship between q and π_0 we get

$$\pi_0 = \frac{\pi + q(1-\pi)}{\pi(1+v^2)+q(1-\pi+\pi v^2)}$$

Substituting this expression into the formula for p^* from Lemma 3 one finds

$$p^* = min\left(\frac{1}{4\gamma}(1-\beta)\beta\frac{1-\pi+v^2+q(\pi+v^2)}{1+q}, 1\right)$$

Moreover

$$f(\pi_0) = (1-q)\pi^2 v^2 + \frac{(\pi+q(1-\pi))^2}{1+q}$$

so that we can write total surplus as a function of q rather than π_0. As π_0 increases from $\overline{\pi}$ to $(1+v^2)^{-1}$, q decreases from some $\overline{q} > 0$ to 0. Let q^* be the smallest $q \in [0, \overline{q}]$ for which $p^* = 1$. In the case that such a q^* does not exist set $q^* = \overline{q}$. As p^* is non-increasing in π_0 it follows that $p^* = 1$ for all $q \in [q^*, \overline{q}]$. As evidence is found for sure it follows that total surplus is equal to $E_\pi P(\overline{\pi})$ in $[q^*, \overline{q}]$. We will now show that total surplus is a convex function of q for $q \in [0, q^*]$ which implies that $E_\pi P(\pi_0) \leq max\left(E_\pi P(\overline{\pi}), E_\pi P(\frac{1}{1+v^2})\right)$ for all π_0 such that $\overline{\pi} < \pi_0 < (1+v^2)^{-1}$. For the second derivative of total surplus with respect to q we find

$$\begin{aligned}\frac{\partial^2 E_\pi P}{\partial q^2} &= \frac{1}{2}\beta(1-\frac{1}{2}\beta)\left((1-p^*)\frac{\partial^2 f}{\partial q^2} + (1+v^2-f(q))\frac{\partial^2 p^*}{\partial q^2} - \frac{\partial f}{\partial q}\frac{\partial p^*}{\partial q}\right)\\ &\quad -2\gamma\left(\left(\frac{\partial p^*}{\partial q}\right)^2 - p^*\frac{\partial^2 p^*}{\partial q^2}\right)\end{aligned}$$

Substituting the derivatives and simplifying leads to

$$\frac{\partial^2 E_\pi P}{\partial q^2} = \frac{1}{8}\beta^2(1-\beta)\frac{1-2\pi}{\gamma(1+q)^4}(g_0(q) + \beta g_1(q))$$

where

$$g_0(q) = (4(1-v^2)\pi^2 - 8\pi + 4)q^2 + (2 - 6\pi - 4v^2(1+\pi^2))q - 4\pi^2 + 8\pi - 5 - 4v^2$$

and

$$g_1(q) = 2(v^2\pi^2 - (1-\pi)^2)q^2 + (v^2(2\pi^2+3) + 4\pi - 1)q + (4-2\pi)(1-\pi) + 3v^2$$

Note that $g_1(q) > 0$ for all $q \in [0,1]$ as $v^2\pi^2 > (1-\pi)^2$ by Assumption 1. It follows that $g_0(q) + \beta g_1(q) < g_0(q) + g_1(q) = -2(v^2\pi^2 - (1-\pi)^2)q^2 - (v^2(1+2\pi^2) + 2\pi - 1)q - v^2 - (1 - 2\pi(1-\pi)) < 0$ which implies that $\frac{\partial^2 E_\pi P}{\partial q^2} > 0$.

Our argument so far implies that $E_\pi P(\pi_0) \leq max\left(E_\pi P(\pi), E_\pi P(\frac{1}{1+v^2})\right)$ for all $\pi_0 \geq \pi$. It remains to compare the values of total surplus at these two points. Define $\underline{\gamma} = \frac{1}{4}(1-\beta)(1+v^2-\pi)$ and $\overline{\gamma} = \frac{1}{4}(1-\beta)\beta(\frac{1}{2}+v^2)$. If the cost of consultation is sufficiently low, that is $\gamma \leq \underline{\gamma}$ then a leader with $\pi_0 = (1+v^2)^{-1}$ will choose $p^* = 1$ such that

$$E_\pi P(\frac{1}{1+v^2}) = \frac{1}{2}\beta(1-\frac{1}{2}\beta)(1+v^2) - \gamma$$

If $\gamma > \underline{\gamma}$ then he will choose $p^* < 1$ and

$$\begin{aligned} E_\pi P(\frac{1}{1+v^2}) &= \frac{1}{2}\pi^2\beta(1-\frac{1}{2}\beta)(1+v^2) \\ &+\frac{\beta^2}{16\gamma}(1-\beta)(1+v^2-\pi)\left((1-(2-\beta)\pi^2)(1+v^2)+(1-\beta)\pi\right) \end{aligned}$$

Note that $(1-(2-\beta)\pi^2)(1+v^2)+(1-\beta)\pi \geq (1-\pi^2)(1+v^2) > 0$. $E_\pi P(\frac{1}{1+v^2})$ is continuous and decreasing in γ. Consider the case where $\pi < (1+2\pi v^2)^{-1}$. In this case a realistic leader chooses $d=1$ for all S. If the cost of consultation is such that $\gamma \leq \underline{\gamma}$ then a realistic leader will choose $p^* = 1$ such that

$$E_\pi P(\pi) = \frac{1}{2}\beta(1-\frac{1}{2}\beta)(1+v^2) - \gamma$$

For $\gamma > \overline{\gamma}$ it holds that $p^* < 1$ and one gets

$$E_\pi P(\pi) = \frac{1}{4}\beta(1-\frac{1}{2}\beta) + \frac{1}{16\gamma}\beta^2(1-\beta)(\frac{1}{2}+v^2)^2$$

$E_\pi P(\pi)$ is continuous and decreasing in γ. For $\gamma \leq \underline{\gamma}$ it holds that $E_\pi P(\frac{1}{1+v^2}) = E_\pi P(\pi)$. For $\underline{\gamma} < \gamma \leq \overline{\gamma}$ one finds that $E_\pi P(\pi) - E_\pi P(\frac{1}{1+v^2}) = \frac{1}{\gamma}h_1(\gamma)$ where

$$\begin{aligned} h_1(\gamma) &= -\gamma^2 + \frac{\beta}{2}(1-\frac{\beta}{2})(1-\pi^2)(1+v^2)\gamma \\ &-\frac{\beta^2}{16}(1-\beta)(1+v^2-\pi)\left((1-(2-\beta)\pi^2)(1+v^2)+(1-\beta)v\right) \end{aligned}$$

Note that $h_1(.)$ is strictly concave and $h_1(\underline{\gamma}) = 0$. Also note that

$$\frac{\partial h_1}{\partial \gamma}(\underline{\gamma}) = \frac{1}{2}\beta\left((1-\frac{\beta}{2})(1+v^2-\pi^2(1+v^2)) - (1-\beta)(1+v^2-\pi)\right) > 0$$

For $\gamma \geq \overline{\gamma}$ one gets $E_\pi P(\pi) - E_\pi P(\frac{1}{1+v^2}) = \frac{1}{\gamma}h_2(\gamma)$ where

$$\begin{aligned} h_2(\gamma) &= \frac{1}{2}\beta(1-\frac{1}{2}\beta)\left(\frac{1}{2}-(1+v^2)\pi^2\right)\gamma - \\ &\frac{\beta^2}{16}(1-\beta)\left((1+v^2-\pi)\left((1-(2-\beta)\pi^2)(1+v^2)+(1-\beta)\pi\right)-(\frac{1}{2}+v^2)^2\right) \end{aligned}$$

Note that $\frac{\partial h_2}{\partial \gamma} = \frac{1}{2}\beta(1 - \frac{1}{2}\beta)\left(\frac{1}{2} - (1+v^2)\pi^2\right) \leq \frac{1}{2}\beta(1 - \frac{1}{2}\beta)\left(\frac{1}{2} - \frac{1+v^2}{(1+v)^2}\right) < 0$ and $h_2(\overline{\gamma}) = h_1(\overline{\gamma})$. If $h_2(\overline{\gamma}) > 0$ then let γ^* be defined by $h_2(\gamma^*) = 0$, that is

$$\gamma^* = \frac{\beta(1-\beta)}{2(2-\beta)} \frac{(1+v^2-\pi)((1-(2-\beta)\pi^2)(1+v^2) + (1-\beta)\pi) - (\frac{1}{2}+v^2)^2}{1-2(1+v^2)\pi^2}$$

Otherwise let γ^* be the largest solution to $h_1(\gamma^*) = 0$. It then follows that $E_\pi P(\pi) < E_\pi P(\frac{1}{1+v^2})$ if and only if $\gamma > \gamma^*$. The case where $\pi \geq (1+2\pi v^2)^{-1}$ is similar, the only difference being that a realistic leader randomizes his project choice when $S = 0$ such that the expressions for $E_\pi P(\pi)$ differ.

2.10 References

Amason, A., 1996, "Distinguishing the Effects of Functional and Dysfunctional Conflict on Strategic Decision Making: Resolving a Paradox for Top Management Teams", *The Academy of Management Journal* 39(1), 123-148

Branderburger, A., and B. Polak, 1996, "When Managers Cover their Posteriors: Making the Decisions the Market Wants to See", *RAND Journal of Economics*, 27(3), 527-541.

Filley, A. C., House, R. J., and S. Kerr, 1976, "Managerial Process and Organizational Behaviour", Scott, Foresman.

Goel, A. M.,and A.V. Thakor, 2000, "Rationality, Overconfidence and Leadership", Unpublished Manuscript, University of Michigan Business School.

Hannan, M. T., and J. Freeman, 1989, "Organizational Ecology", Harvard University Press.

Heifetz, R. 1998, "Leadership without Easy Answers", Belknap Press.

Hermalin, B. 1998, "Toward an Economic Theory of Leadership: Leading by Example", *American Economic Review* 88(5), 1188-1206.

Korsgaard, M., Schweiger, D., and H. Sapienza, 1995, "Building Commitment, Attachment and Trust in Strategic Decision-Making Teams: The Role of Prodecural Justice" *The Academy of Management Journal* 38(1), 60-84.

Latham, G., Winters, D., and E. Locke, 1994, "Cognitive and Motivational Effects of Participation: A Mediator Study", *Journal of Organizational Behavior*, 15(1), 49-63.

Locke, E. A., and D.M. Sweiger, 1978, "Participation in Decision-Making: One More Look", in Staw, B.M. (Ed.) Research in Organizational Behaviour, Volume 1, JAI Press, 265-339.

Morris, S., Willkers, G., and E. Knasel, 1999, "How to Lead a Winning Team". Prentice.

Prendergast, C. 1993, "A Theory of Yes-Men", *American Economic Review*, 83(4) 757-770.

Rotemberg, J., and G. Saloner, 1994, "Leadership Style and Incentives", *Management Science*, 39(11), 1299-1318.

Rotemberg, J., and G. Saloner, 2000, "Visionaries, Managers and Strategic Direction", *RAND Journal of Economics*, 31(4), 693-717.

The Economist, 17th of April 2003.

Thomas, D., 2001 quoted in Harvard Business School Bulletin 2001, Harvard University Press.

Van den Steen, E. 2002, "Organizational Beliefs and Managerial Vision", Unpublished Manuscript, MIT.

Vroom, V. H., and P.W. Yetton, 1973, "Leadership and Decision-Making" University of Pittsburgh Press.

Wagner, J. 1994, "Participation's Effects on Performance and Satisfaction: A Reconsideration of Research Evidence", *The Academy of Management Review*, 19(2), 312-330.

Wagner, J., and R. Gooding, 1987, "Shared Influence and Organizational Behavior: A Meta-Analysis of Situational Variables Expected to Moderate Participation-Outcome Relationships", *The Academy of Management Journal*, 30(3), 524-541.

Wagner, J., Leana, C., Locke, E., and D. Schweiger 1997, "Cognitive and Motivational Frameworks in U.S. Research on Participation: A Meta-Analysis of Primary Effects", *Journal* 18(1), 49-65.

Chapter 3
Credibility and Cheap Talk of Securities Analysts: Theory and Evidence

"THERE are two theories about Wall Street's role in the bubble years of the new economy. Either investment analysts were swept up, like everybody else.... Or Wall Street saw a golden chance to peddle dirt.... New York's attorney-general, Eliot Spitzer, is a promoter of the second theory. (He painted) Merrill Lynch's share-buying recommendations ... as little more than a pretext to stuff gullible buyers with the shares of rotten businesses... (Merrill) argues that the state-attorney misunderstands markets." The Economist [42].

3.1 Introduction

It is well known that securities analysts' advice is optimistically biased[8]. However, the manner in which such bias affects the transmission of information from analysts to investors in the stock market is yet unclear. Two issues have been particularly disputed among regulators and practitioners in the last years: whether analysts are deliberate in their optimism, and whether investors are able to understand this bias and anticipate it in their trading decisions. These two issues have been at the core of a passionate public debate, an illustration of which is the quote at the beginning of this chapter.

On one side of this debate the New York attorney-general, Eliot Spitzer, among others, has maintained that analysts are well aware of the real state of the companies they cover

[8] See, for instance, Boni and Womack [9].

but that, being subjected to distorted incentives, they benefit from propagating enthusiasm about these companies. Further, Spitzer's argument continues, investors' gullibility makes them take analysts' advice at face value and be systematically misled by it. Consistently with his view, Spitzer has introduced regulatory reforms to adjust the incentives of analysts and align them with the investors who they are supposed to serve.

On the other side, analysts' employers, usually investment banks, have claimed that analysts report their private beliefs truthfully, although it is possible that some of them may suffer from "honest" (i.e. irrational) optimism about the companies they cover. Crucially, banks have stated that such "honest" bias is perfectly understood and anticipated by the average investor, hence making regulatory intervention unnecessary.

This chapter examines earnings forecasts, an important part of analysts' advice, and argues that none of the above views accurately explain how forecasts are made by analysts and reacted to by investors. Alternatively, it claims that *both analysts and investors are strategic in their behaviour* (i.e. that analysts deliberately try to deceive investors and investors are aware of this fact and heavily discount analysts' advice).

To show this, I first study a model of information transmission between an (almost fully) strategic analyst who is privately informed about the earnings of a company, and a strategic investor. In my model the analyst issues a forecast (i.e. a "message") that the investor must interpret and react to with an appropriate action. The analyst may be either "honest", in which case he truthfully reports his information, or "biased", in which case he strategically raises his message to induce the investor into believing that the situation is better than he knows it to be. Crucially, the investor is uninformed about the type of the analyst, but she can use the message to update the likelihood that the analyst is biased rather than honest. Optimistic forecasts are assigned a lower probability of mirroring the true information (i.e. they are not "credible") and are strongly discounted.

In the most informative continuous equilibrium of this model the action of the investor is an increasing and concave function of the analyst's message. Further, the slope of the action function is lower if the investor has a higher *prior suspicion* that the analyst is a biased type. I show that these predictions do not arise under either of the frameworks endorsed by Eliot Spitzer or the investment banks.

To test these predictions I use a large database of company earnings, analysts forecasts and the stock market's reaction to them. I first find that analysts are indeed biased, especially when they announce good news. For instance, when an average analyst announces an optimistic forecast of 10 cents above the consensus, actual earnings are, on average, only 2 cents higher. However if the analyst announces bad news of 10 cents below the consensus, actual earnings are lower, on average, by 13 cents. Using parametric and nonparametric techniques I find that the function relating actual earnings and forecasts is concave, as predicted by the model.

I next show that the average investor understands this and reacts more to pessimistic forecasts than to optimistic ones. When an analyst announces good news of 10 cents, the price on the shares of the company increases, on average, by 0.15% on the day on which the announcement is made. However, if he reports a bad news of 10 cents the price decreases by 0.7%, more than 5 times higher in magnitude. In my sample the standard deviation of the daily return is 3.6%, so 0.7% is an economically significant magnitude. I show that the function relating the stock market reaction to analysts forecasts is also a concave curve.

I then study the effect on investors' reaction of their prior suspicion that the analyst is a biased type. I conjecture that if an analyst has predicted above actual earnings in most of his forecasts *in the past*, investors will consider him very likely to be an optimistically biased type. Indeed I find that when such analyst announces an optimistic forecast of 10 cents, the share price does not increase. However if the same forecast is announced by an analyst who has not issued highly optimistic forecasts in the past, the share price reacts by a magnitude of 0.27%. Again this is consistent with my model and economically significant.

It is important to stress that these are the first empirical findings to provide support for a model of strategic information transmission (i.e. cheap talk). Following Crawford and Sobel [14], an important theoretical literature has developed in the last two decades that studies the strategic communication between parties with conflicting interests. Unfortunately, this interest has not been followed by a similar surge in studies testing its empirical implications. As a result very little is yet known about its predictive power. One major contribution of this chapter is to combine a theoretical model in the Crawford and Sobel tradition with an empirical analysis that supports the intuitive predictions of the model.

My findings also have direct policy implications. Although the average investor understands the bias in earnings forecasts, I show that information is communicated in a very inefficient way. In particular, optimistic news about the earnings of a company are completely disregarded by investors and are not credibly communicated by analysts. Thus I conclude that analysts' misalignment of incentives is preventing them from serving investors appropriately, and regulatory reforms aimed at reducing it would have beneficial results for the financial industry.

Outline

The rest of the chapter is organised as follows: Section 3.2 discusses related work. Section 3.3 outlines the model, and Section 3.4 derives the theoretical results, which are restated in the form of empirical predictions in Section 3.5. The data is presented in Section 3.6 and the main empirical results in Section 3.7. Section 3.8 discriminates among alternative causes for analysts' misaligned incentives. Section 3.9 briefly concludes. All the figures, tables and proofs are in the Appendix.

3.2 Related work

The concept of credibility was first introduced in a cheap talk framework by Sobel [41], who analyses the incentives to build a reputation in a dynamic binary-state setting (see also Benabou and Laroque [7] and Fisher and Heinkel [17] for similar but richer models). More closely related to this chapter is Morgan and Stocken [36]. They study a continuous-state framework with two types of strategic senders, biased and unbiased. In one of their equilibria the reaction function is concave, with the action of the receiver increasing one-to-one in the message if this is pessimistic enough, and being completely flat if the message is too optimistic. The change in credibility is discrete: at a certain point the message shifts from being taken at face value to being (locally) ignored. Their equilibrium has, however, the undesirable feature that information can only be credibly transmitted in the set of messages that the advisor would *never* use if he was a biased type. Such set is empty if the state space is unbounded, an assumption that is realistic in many real-life scenarios, including the study of earnings per share that is the focus of my chapter. Contrary to Morgan and Stocken, I assume the unbiased sender to be non-strategic. In this setting, I find that information can be transmitted even if the state, message and action spaces are unbounded. In my model the strict concavity of the equilibrium reaction function is caused by the credibility of a message decreasing continuously with the optimism embedded in it. Furthermore, I show that the slope of the reaction function increases with the prior of sender's honesty.

Ottaviani and Squintani [39] introduce a non-strategic receiver in a Crawford and Sobel framework. They find an equilibrium where communication is inflated, information is not lost due to strategic reasons and the action is *ex ante* biased. This chapter complements their work by studying the case of a sender who may be non-strategic with some probability. However, my interest is in characterising the shape of the reaction function in equilibrium, an issue that Ottaviani and Squintani do not explore.

A large empirical literature has studied the recommendations and earnings' forecasts of securities analysts[9]. Jackson [26] finds that analysts have an incentive to generate trade via optimistic forecasts, but are partly restrained by reputation concerns if investors can learn about their records. Michaely and Womack [35] show that, following an IPO, stocks recommended by affiliated analysts perform poorly in comparison to the ones recommended by unaffiliated analysts. They interpret this as evidence that underwriter analysts are positively biased in their recommendations. Malmedier and Shanthikumar [33]

[9] See, for instance, Chen and Jiang [12], Francis and Soffer [20], Keane and Runkle [28], Lim [31], Lin and McNichols [32], and Zitzewitz [45].

build on Michaely and Womack by showing that small investors take recommendations at face value and do not discriminate among analysts. Large investors, on the other hand, are much less naive about analysts' incentives. I study the behaviour of the *average* investor (Badrinath and Wahal [6] show that large institutional investors account for the great majority of volume traded) and I improve on the above mentioned studies on several grounds. By focusing on earnings forecasts (a continuous variable) rather than on stock recommendations (a discrete variable), I am able to study the curvature of the reaction function, and provide much sharper evidence of the way in which investors discount analysts' messages. Secondly, the use of a theoretical model allows me to enrich the set of testable predictions and to consider alternative hypotheses about the rationality of the agents. Lastly, my predicted variable (earnings) is not affected by the recommendation of the analyst, and therefore my test does not suffer from the potential endogeneity problems present when the message (recommendation), if believed, can alter the state (share price) of the world.

3.3 Credibility in a Cheap Talk Model

A representative investor (she) wants to learn about future earnings of a company. At a given moment in time, the investor (and thus the market) holds an underlying consensus, c, about those future earnings and such consensus is incorporated in the price of the shares. The intention of the investor is to react to news about the company earnings, buying when the news is favourable in relation to the consensus and selling otherwise.

Information Structure

The investor has no independent means of gathering information about the company, either due to limited attention (there may be many alternative investment opportunities to consider) or lack of the expertise required to understand its accounts or strategy. A securities analyst (he) has those means. He uses them to collect information and, in conjunction with c, compute an optimal (private) guess, g, about future earnings. Define $x = g - c$ as the news above or below consensus arising from the private information of the analyst. Assume that x is distributed normally with zero mean, $x \sim N(0, \sigma^2)$, where $x \in X \equiv \Re$, and $\phi(.)$ and $\Phi(.)$ are the normal p.d.f and c.d.f respectively.

Preferences

The investor's gain from learning x arises from the opportunity to rebalance her investment portfolio due to the new information. Define a as the action of buying (if $a > 0$) or selling (if $a < 0$) shares, where $a \in A \equiv \Re$. With each piece of news and action is associated a benefit, $U^I = U^I(a, x)$. I simplify the portfolio decision by assuming that the utility of the investor reaches a unique maximum for $a^I = x$ and is:

$$U^I(a, x) = -(a - x)^2$$

This function can be interpreted as the utility loss suffered by the investor when she fails to adjust her portfolio efficiently in the light of new existing information.

The central conflict of the model arises because the preferences of the investor and the analyst *may not* be perfectly aligned. In particular, I assume that the analyst may prefer the investor to take a higher action than is optimal for her, $a^A = x + b > x = a^I$, where a^A is the preferred action of this type of analyst and $b > 0$ is the bias, or degree of dissonance between the preferences of both agents. Assume that q is the prior probability that the analyst is a biased type. There are different ways of deriving such conflict of interest from the preferences of the analyst. I follow Morgan and Stocken [36] and assume that the objective function of the biased analyst consists of two elements: a benefit associated with inflating the investor's action and a cost associated with poor performance caused by the distortion of the information revealed to the investor.

$$U^A(a, x, b) = 2ba - (a - x)^2$$

I also assume that, with probability $1 - q$, the analyst is honest and non-strategic and reports his information truthfully. This assumption is also used by Ottaviani and Squintani [39] and by Benabou and Laroque [7]. The justifications for this assumption can be behavioural (an analyst adhering to a code of ethics by which distorting information is considered immoral) or in terms of payoff uncertainty. In reality, an analyst who misrepresents his private information runs some risk of being discovered and fined huge penalties. For instance, when Henry Blodget, a Merrill Lynch star analyst, was discovered to hold different views in his private e-mails from the ones that he maintained in public, he was fined \$4m and barred from the securities industry for life[10]. The assumption of non-strategic honesty could therefore apply to analysts who are very risk averse or who estimate the chances of being caught lying as very high. I believe, however, that assuming that the analyst is strategic but cares about being perceived as unbiased would not qualitatively alter the theoretical results.

[10] Henry Blodget "described the shares of excite@home, which he publicly rated as a short-term accumulate and long-term buy, as “such a piece of crap”. Of InfoSpace, both a short- and a long-term buy, he claimed in private emails that the “stock is a powder keg”" The Economist [42]. In a similar scandal, another star analyst employed by Citigroup was fined \$15m and also barred from the industry.

Information Transmission and Beliefs

The investor has no means of inducing truthful revelation through a contract. The analyst can, however, issue a message m, with $m \in M \equiv \Re$, to influence the investor's expectation of x. The message takes the form of a forecast f about earnings, where $m = f - c$.

If the analyst is honest, he truthfully announces $m = x$. For the biased analyst, a strategy consists of a family of functions $(v(\cdot \mid x))_{x\in\Re}$, where, for each x, $v(\cdot \mid x)$ is a p.d.f. on the message space M, and $\int_{\Re} v(m \mid x)dm = 1$. If the analyst follows a pure strategy over some range of the state space, then $v(\cdot \mid x)$ is degenerate in this range and is represented by the function $\mu : x \rightarrow m$.

After observing m, the investor updates her expectation about x. Define $(\beta(\cdot \mid m))_{m\in\Re}$ as the family of functions that compose the investor's beliefs of x. For each m, $\beta(\cdot \mid m)$ is a p.d.f on the state space X and $\int_{\Re} \beta(x \mid m)dx = 1$.

Timing

The timing is as follows: *(i)* The analyst learns his type, honest or biased *(ii)* The analyst learns x and issues a message m *(iii)* The investor observes m, updates her beliefs and adopts an action a *(iv)* The payoffs of both agents are realised.

Equilibrium concept

The equilibrium concept is Perfect Bayesian Equlibria (PBE). An equilibrium consists of reaction function, $a(m)$, a family of message functions, $(v(\cdot \mid x))_{x\in\Re}$ and a family of belief functions $(\beta(\cdot \mid m))_{m\in\Re}$ such that

$$a \in \arg\max \int_{\Re} U^I(a,x)\beta(x \mid m)\,dx \tag{3.3}$$

$$\text{If } v(m \mid x) > 0, \text{ then } m \in \arg\max U^A(a(m),x,b) \tag{3.4}$$

$$\beta(x \mid m) = \begin{cases} \dfrac{qv(m|x)\phi(x)}{(1-q)\phi(m)+q\int_{\Re} v(m \mid x)\,\phi(x)\,dx} & if \quad m \neq x \\ \dfrac{(1-q)\phi(m)+qv(m|x)\phi(x)}{(1-q)\phi(m)+q\int_{\Re} v(m \mid x)\,\phi(x)\,dx} & if \quad m = x \end{cases} \tag{3.5}$$

3.4 Analysis

Communication games usually have many equilibria, and this game is no exception. I focus on the most responsive continuous equilibrium and derive a number of predictions about it. Section 3.7 shows that the transmission of information in the real world is consistent *with and only with* the predictions of the most responsive continuous equilibrium.

Define $A_i^j(x)$ as the function that maps states observed by analyst type i into actions under equilibrium j. Similarly define $a^j(m)$ as the reaction function under equilibrium j.

Definition1 *A* continuous *equilibrium is an equilibrium where the functions $A_i^j(x)$ are continuous. Consider two continuous equilibria, J and K, and assume that $A_i^J(x)$ and $A_i^K(x)$ are strictly increasing in $X_i^J \subseteq R$ and $X_i^K \subseteq R$ respectively. Then,*

1. *Equilibrium J is* more responsive *than equilibrium K if $X_i^K \subset X_i^J$ $\forall i$.*

2. *Equilibrium J is the* most responsive *continuous equilibrium if every other continuous equilibrium k has a corresponding X_i^k such that $X_i^k \subset X_i^J \; \forall i$.*

3. *Equilibrium J is* fully-responsive *if $X_i^J = X \equiv R \; \forall i$.*

Intuitively, in a continuous equilibrium the action does not "jump around" when the state of the world is slightly higher or lower. In a very responsive equilibrium a lot of information is being credibly transmitted by the analyst (i.e. the investor is responding to higher messages by taking higher actions). The rest of this Section shows that in the most responsive continuous equilibrium the reaction function of the investor is concave and has a lower slope when q is higher.

3.4.1 The reaction of the investor

From Section 3.3, an equilibrium is a combination of strategies and beliefs that satisfies (3.3), (3.4), and (3.5). In the following Lemma, I combine (3.3) and (3.5) to study the equilibrium reaction function.

Lemma 4 *In any equilibrium of this game the reaction of an investor to a message m' is given by:*

$$a = E[x \mid m'] = \frac{(1-q)m'\phi(m') + q\int_{\Re} xv\,(m' \mid x)\,\phi\,(x)\,dx}{(1-q)\phi(m') + q\int_{\Re} v\,(m' \mid x)\,\phi\,(x)\,dx} \tag{3.6}$$

If m' is issued using a pure strategy $\mu(\cdot)$, the reaction of the investor is given by:

$$a = E[x \mid m'] = \frac{(1-q)m'\phi(m') + q\mu^{-1}(m')\phi\left(\mu^{-1}(m')\right)}{(1-q)\phi(m') + q\phi\left(\mu^{-1}(m')\right)} \tag{3.7}$$

Lemma 4 illustrates the concept of the "credibility" of a message. Consider, for instance, the intuition of (3.7). The investor is uncertain about two things: the information of the analyst, x, and his type, "honest" or "biased". If the investor knew with certainty that the analyst issuing m' is honest, her optimal action would be $a = m'$. Conversely, if she knew that the analyst is a biased type, it would be optimal for her to make $a = \mu^{-1}(m')$. The weight that the investor gives to m' ("following literally") over $\mu^{-1}(m')$ ("undoing the bias") can therefore be regarded as the credibility that she gives to the message. (3.7) states that this credibility is equal to the Bayesian posterior probability of the message having been created by an honest analyst, $\Pr[Honest \mid m = m'] = [(1-q)\phi(m')] \,/\, [(1-q)\phi(m') + q\phi\left(\mu^{-1}(m')\right)]$. It is easy to see that this probability depends negatively on q, and positively on $[\phi(m')/\phi\left(\mu^{-1}(m')\right)]$, the relative chances that a normal distribution produced m' rather than $\mu^{-1}(m')$.

3.4.2 The analyst's problem

I now turn to the equilibrium behaviour of the analyst. When the honest analyst observes x', he issues, by assumption, a truthful message $m = x'$. The biased analyst maximises his payoff by solving the following problem:

$$\underset{m}{Max}\ 2ba - (a - x')^2$$

$$s.t.\ a = E[x \mid m] \text{ as given by (3.6)}$$

After rearranging, the first order condition of this problem can be written as:

$$(E[x \mid m] - x' - b)\frac{\partial E[x \mid m]}{\partial m} = 0 \tag{3.8}$$

Any equilibrium of this game must therefore meet conditions (3.6) and (3.8). (3.8) shows that, if the investor's action increases continuously with the message ($\frac{\partial E[x|m]}{\partial m} > 0$), in equilibrium the biased analyst must be reaching his bliss point ($E[x \mid m] = x' + b = a^A$). By assumption, altering the message has no direct cost *per se* (talk is "cheap", or more precisely, "free"), so this condition is quite intuitive. Clearly, if $E[x \mid m]$ increases continuously with m, and raising m has zero cost for the biased analyst, he will always be able to exaggerate infinitely and induce his optimum a^A.

3.4.3 Fully-responsive equilibria

Before characterising the most responsive continuous equilibrium, I show that a fully-responsive equilibrium does not exist in this game.

Proposition 6 *A* fully-responsive *equilibrium does not exist.*

Definition 1 states that in a fully-responsive equilibrium the investor always takes higher actions when the information of the analyst is higher, regardless of whether he is an honest or a biased type. The honest analyst makes $m = x$, so he cannot induce different actions under different states of the world if $\partial E[x \mid m]/\partial m = 0$ holds in some interval of the message space. Therefore a fully-responsive equilibrium can only exist if the implicit function of m on x, given by $E[x \mid m] - x - b = 0$, is defined over the entire state space. The proof of Proposition 6 shows that this is impossible. The reason is that, for a sufficiently high value of x, a value of m that meets $E[x \mid m] - x - b = 0$ cannot be found.

To understand the intuition, imagine that the message is sent using a pure strategy $\mu(\cdot)$. Using the fact that, in equilibrium, $\mu^{-1}(m) = x$, $E[x \mid m] - x - b = 0$ can be rewritten as:

$$F(m,x) = \frac{(1-q)m\phi(m) + qx\phi(x)}{(1-q)\phi(m) + q\phi(x)} - x - b = 0 \tag{3.9}$$

$$F(m,x) = m - x - b\left(1 + \frac{q\phi(x)}{(1-q)\phi(m)}\right) = 0 \tag{3.10}$$

(3.10) shows that a fully-responsive equilibrium requires the message to be always higher than the state x by an amount at least equal to b. Moreover, when x is higher the credibility of the corresponding message is lower (note that $\frac{d}{dx}[\phi(x)/\phi(x+\xi)] > 0$ if $\xi > 0$) and the "wedge" between m and x has to be higher to meet condition (3.10). The proof of Proposition 6 shows that, when x is very high, credibility is so low that it becomes impossible to find a message that will induce $a = x + b$ from the investor.

3.4.4 The most responsive equilibrium

After ruling out fully-responsive equilibria, I characterise the continuous equilibrium where more information is transmitted to the investor.

Proposition 7 *If it exists, the most responsive continuous equilibrium meets the following conditions:*

1. *A cutoff value, x^* divides the state space in two sections; a cutoff value, m^* divides the message space in two sections, and m^* is a solution to the equation $F(m, x^*) = 0$.*

2. *If the biased analyst observes $x \leq x^*$, he issues a message (in pure strategies) using a function $m = \mu^{MR}(x)$ with the following properties:*

   (a) $\frac{\partial \mu^{MR}}{\partial x}>0$

   (b) $\frac{\partial^2 \mu^{MR}}{\partial x^2}>0$

   (c) $\frac{\partial^2 \mu^{MR}}{\partial x \partial q}>0$

3. *If the investor observes $m \leq m^*$, she chooses her action using an action function, $a^{MR}(m)$ with the following properties:*

   (a) $\frac{\partial a^{MR}}{\partial m}>0$

   (b) $\frac{\partial^2 a^{MR}}{\partial m^2}<0$

   (c) $\frac{\partial^2 a^{MR}}{\partial m \partial q}<0$

4. *If the investor observes $m > m^*$, she always chooses the same action a^*, which is determined by the equation:*

$$a^* = \frac{(1-q)\phi(m^*) + q\phi(x^*)}{(1-q)(1-\Phi(m^*)) + q(1-\Phi(x^*))} \tag{3.11}$$

5. *If the biased analyst observes a value $x > x^*$, he chooses his message (in mixed strategies) using a family of functions $(\upsilon^{MR}(\cdot \mid x))_{x>x^*}$ so as to make, $\forall m > m^*$:*

$$\frac{(1-q)\, m\phi(m) + q\int_{x^*}^{\infty} x\upsilon^{MR}(m \mid x)\, \phi(x)\, dx}{(1-q)\, \phi(m) + q\int_{x^*}^{\infty} \upsilon^{MR}(m \mid x)\, \phi(x)\, dx} = a^* \tag{3.12}$$

6. *If the biased analyst observes x^*, he is indifferent between issuing a message in the responsive or the non-responsive parts of the equilibrium.*

$$a^* = \frac{(1-q)\phi(m^*)m^* + q\phi(x^*)x^*}{(1-q)\phi(m^*) + q\phi(x^*)} \tag{3.13}$$

Proposition 7 is the main theoretical result of the chapter. It states that, under the most responsive continuous equilibrium of this game, the reaction function is concave when the message is below certain value and flat if the message is too optimistic (see Figure C). Furthermore, in this equilibrium the slope of the reaction function is lower when the ex ante probability of the analyst being a biased type is higher (see Figure D). These are the predictions that will be tested in the empirical analysis below.

Intuition of the equilibrium

To understand intuitively this equilibrium consider first the convexity of the message function in the responsive part (see Figure A). The message function when $x \leq x^*$ has to satisfy (3.10). If x is higher, the analyst has to issue a higher message if he is to induce an inflated action from the investor. In fact, he has to make m higher by a larger amount than the increase in x, to compensate for the loss of credibility (as argued in Subsection 3.4.3) when he conveys more optimism in his message. Therefore when x is higher, m and x are, in equilibrium, further apart.

But how does the slope of the message function, $\mu^{MR}(x)$, depend on x? If x increases from an already high position, the analyst is more affected by subsequent losses of credibility (i.e. a decrease of $[q\phi(x) / (1-q)\phi(m)]$ reduces the action by more because m and x are further apart). Therefore, as x increases, the analyst has to exaggerate *increasingly more* to maintain a "wedge" between the investor's optimal average of x and m and his information x. Hence, the convexity of the message function.

Consider now the concavity of the reaction function in the responsive part of the equilibrium. Remember that the reaction function is the inverse of the message function ($a^A = x + b = \mu^{MR^{-1}}(m) + b$). The strict concavity in the responsive part (see Figure C when $m \leq m^*$) is the inverse of the strict convexity of the message function. Because the biased analyst exaggerates increasingly as x becomes higher, the investor discounts increasingly as the message becomes more optimistic. In other words, because the analyst anticipates that optimistic messages are less credible, the more favourable his private information is, the more he inflates them. By doing this he only strengthens the investor's presumption that she should heavily discount optimistic messages. In equilibrium the message of a biased analyst increases with his private information in a convex way, and its converse (i.e. the reaction of the investor) increases with the message only in a concave way.

The same intuition holds for the cross-derivatives of the message and the reaction function. When q is higher, the weight that the investor assigns to the analyst telling the truth becomes lower, as can be seen in (3.10). Thus m and x have to be further apart, at any state of x, if q is higher. That makes the analyst more seriously affected by the loss of credibility, and more prone to exaggeration, when x is higher. The message function has therefore a higher slope if the exogenous parameter q is higher (see Figure B). Conversely, the action function has a lower slope in its responsive part (see Figure D) if q is higher, as higher strategic optimism by an analyst with a higher q is matched by stronger discounting of the messages produced by such analyst. In other words, because investors associate optimistic forecasts relatively more with his presumed exaggeration and relatively less with the truth, an analyst more suspected to be biased is forced to exaggerate more if he wishes to inflate the investor's action. This fact is anticipated by the investor, reinforcing his disposition to react less to better news when those are conveyed by a highly suspected analyst.

Consider now the logic of the non-responsive section of the equilibrium. In this section, all messages are treated equally and the action is the average of the possible states in which an analyst would issue a message above m^*. Part 4 of Proposition 7 comes simply from rearranging this condition. I do not explicitly characterise the family of message functions in the non-responsive section. This does not imply a loss of predictions because empirically I only observe in the sample $E[x \mid m]$ and m, but not x. It is therefore not possible to test any predictions on the message function and my empirical efforts will be directed to the estimation of the action function. A valid message function, however, must leave the investor completely uninformed about where the state of the world lies (obviously, the investor knows that x lies in the interval (x^*, ∞), but not where in the interval). Because the honest analyst always reports the truth, the combined message can only be uninformative if the biased analyst "hides" this information by adopting a message function that makes the expected state, conditional on the message, always equal to a^*. Equation (3.12) displays this condition. Lastly, Part 6 states that the action function must be continuous at the intersection between the responsive and the non-responsive section.

Other equilibria

There are two other continuous equilibria in this game and I describe them in the proof of Proposition 7 (see Appendix). One of them is the standard non-responsive equilibrium, displayed in Figure E. In this equilibrium, the message of the analyst is completely uninformative, and the action of the investor coincides with the unconditional expectation of x. The other equilibrium is identical to the most responsive equilibrium of Proposition 7 except for a "flat" section in the left tail of the message space (see Figure F). It is immediate to conclude that both equilibria are less responsive than the one described in Proposition 7.

3.4.5 Existence

The most responsive equilibrium does not exist for every possible set of the exogenous parameter values (b, q, σ^2). There are two reasons for this. Firstly, the system of equations (3.11) and (3.13), which together determine the point (x^*, m^*), may not have a solution. Proposition 6 states that (3.10), and therefore (3.13), are not defined for every possible value of x. As a result the intersection with the curve (3.11) is not guaranteed, and an equilibrium may not exist.

The second reason relates to the family of message functions in the flat section of the state space. Observe the left hand side of condition (3.12). The honest analyst, who is part of the game with probability $(1 - q)$, issues messages in this section of the equilibrium with a probability $[1 - \Phi(m^*)]$. Such messages have to be uninformative to the investor (i.e. they must be hidden) for the equilibrium to exist. This is only possible if the biased analyst, who exists with probability q, issues the appropriate message when he observes a state higher than x^*. As q tends to 0, (3.12) cannot hold for every message $m > m^*$. I posit that a sufficient condition for the existence of a family of functions $v\,(m \mid x)$ that make (3.12) hold is that a state observed by a biased analyst occurs at least as often as a state observed by an honest analyst $((1-q)\left[1-\Phi(m^*)\right] \leq q\left[1-\Phi\left(x^*\right)\right])$. The conjecture below summarises this discussion.

Supposition 1 *The most responsive equilibrium exists if the exogenous parameters* $\{b, q, \sigma^2\}$ *are contained in the set* $\{[2,3], [0.2, 0.9], [0.5, 1.5]\}$

In the above conjecture, I have used numerical methods to calculate the set of parameters for which (3.11) and (3.13) have a solution and $(1-q)\left[1-\Phi(m^*)\right] \leq q\left[1-\Phi\left(x^*\right)\right]$ is satisfied.

I accept the absence of analytical conditions for the most responsive equilibrium to exist as a drawback of this model. However, it is important to emphasize that my intention is not to develop a general theory of credibility and cheap talk, but to provide an example of a possible equilibrium characterised by a number of intuitive predictions that can compared with the empirical evidence. It is this task that I undertake in the following sections.

3.5 Empirical Tests

The two main predictions of the model can be observed in Figure C and Figure D: the action function is concave and it has a lower slope if there is a greater *prior* suspicion that the analyst is a biased type.

$$\frac{\partial^2 a}{\partial m^2} < 0 \text{ and } \frac{\partial^2 a}{\partial m \partial q} < 0$$

In the model $a = E[x \mid m]$, so this is equivalent to:

$$\frac{\partial^2 E[x \mid m]}{\partial m^2} < 0 \text{ and } \frac{\partial^2 E[x \mid m]}{\partial m \partial q} < 0$$

I use two complementary strategies to estimate the reaction function empirically. Consider first the relation between $E[x \mid m]$ and m. In the sample I observe an analyst's forecast (f), the consensus (c) and hence the message ($m = f - c$), but not his private information ($x = g - c$) . It is impossible to know whether a particular analyst has deliberately raised his forecast, as x and m cannot be directly compared. By using actual earnings per share (e) I can, however, study statistically whether analysts are likely to be strategically raising their forecasts. The difference between actual earnings and the consensus at a particular moment in time is, by definition, an unbiased expectation of the private information held by an analyst at that moment: $(e - c) = E[g] - c = E[x]$. The theoretical relation between $E[x]$ and m can be captured empirically by the relation between actual earnings and forecasts.

$$E[x] = (e - c) = \varphi(f - c,\, q) + \varepsilon \tag{3.14}$$

If all analysts report their information truthfully ($m = x$) I will find[11] that $\frac{\partial(e-c)}{\partial(f-c)} = 1$, $\frac{\partial^2(e-c)}{\partial(f-c)^2} = 0$ and $\frac{\partial^2(e-c)}{\partial(f-c)\partial q} = 0$. By contrast, a finding that $\frac{\partial(e-c)}{\partial(f-c)} < 1$, $\frac{\partial^2(e-c)}{\partial(f-c)^2} < 0$ and $\frac{\partial^2(e-c)}{\partial(f-c)\partial q} < 0$ would support the theoretical model.

I proxy q for a particular analyst at moment t by the percentage of times that this analyst has forecasted above the actual value $((f-c)-(e-c) > 0)$ in the periods previous to t. Under the null hypothesis of truthtelling, an analyst incurs a high percentage of optimistic errors only by chance and a high value of the proxy for q is uninformative about his present message strategy. In my model, however, past optimistic errors are indicative of the analyst being a biased type and informative about his present behaviour.

[11] These predictions also hold even if the analyst is not perfectly informed about actual earnings, although in such case the goodness of fit of any estimation will be lower.

Equation (3.14) relates forecasts and optimal expectations of earnings *as constructed by the econometrician.* However, it does not indicate whether real investors behave in a manner consistent with the theoretical model[12]. To examine this, I study the change in the price of the shares of a company around the date on which a forecast was made $r = (\Delta P/P)$. The shares of a company increase (decrease) in price when good (bad) news about its earnings are believed by investors. I therefore study:

$$r = \pi(f - c,\ q) + \eta \tag{3.15}$$

If all analysts are truthful in their forecasts (and investors know it), I will find $\frac{\partial^2 r}{\partial (f-c)^2} = 0$ and $\frac{\partial^2 r}{\partial (f-c)\partial q} = 0$. By contrast the finding that $\frac{\partial^2 r}{\partial (f-c)^2} < 0$ and $\frac{\partial^2 r}{\partial (f-c)\partial q} < 0$ would support the theoretical model.

3.5.1 Cognitive biases and Naive Investors

It is important to stress that the assumption of strategic behaviour for both the biased analyst and the investor is essential for the predictions of the theoretical model in Section 3.3. To emphasise this, I analise in this subsection the predictions of models where the analyst suffers from "behavioural optimism" or where the average investor is "naive".

Khaneman and Lovallo [27] have suggested that some analysts view the companies that they are covering in a unique narrow frame (the same way as parents view their children as better than their neighbours'). They are unable to accept the fact that their companies

---

12 The fact that the action equals the *ex post* expected state of the world is not a general result, but a direct consequence of using a quadratic loss function centred around x. In the real world, there are many reasons why a and $E[x \mid m]$ differ. For example the investor may face non-linear transaction costs, or his valuation of the shares of the company may not increase linearly with its earnings. For this reason it is important to study the actions of investors separately.

may do badly, and therefore experience optimistic cognitive biases on them. There are different ways to study the action function under what Khanneman and Lovallo term "the insider's view". In my framework, it could be assumed that *all* analysts suffer from a cognitive bias b on the company that they follow. In this case the optimal action of investors would be to just subtract b from the message, $a = m - b$ (see Figure G).

Alternatively, it could be assumed that some analysts are unbiased ($m = x$) whereas others suffer from cognitive, honest, bias ($m = x + b$) and that the type of the analyst is unknown to the investor. Such model would be equivalent to the model in Section 3.3 with the exception that the biased analyst would be non-strategic. It is easy to see that the optimal reaction of an investor to a message m in this case would be:

$$a = E[x \mid m] = \frac{(1-q)m\phi(m) + q(m-b)\phi(m-b)}{(1-q)\phi(m) + q\phi(m-b)} \tag{3.16}$$

The plot of (3.16) (see Figure H) shows that the action function is concave along some range and then convex. The reason is that a non-strategic analyst does not anticipate the investor's discounting of more optimistic messages and therefore does not raise his forecast more when he observes a higher x. As a result, a rational investor never discounts from the message a magnitude higher than b. When the message is very optimistic, an amount very close to b has already been discounted, and the slope of the reaction function is therefore close to one.

Consider now the predictions of a model where the analyst is strategic in his biased behaviour, and the investor is "naive". Note that in this case the action of the investor and the econometrician's optimal estimate of earnings do not coincide ($a \neq E[x \mid m]$). As Figure I shows, a naive investor takes the message at face value ($a = m$). In such case it suffices for a biased analyst to add a constant bias to his message to reach his bliss point. Consequently, the optimal estimate of earnings has the same shape as in the case of the possibly biased non-rational analyst. I display this case in Figure J.

3.6 Data

Quarterly earnings per share forecasts and earnings are taken from the I/B/E/S Detail Database[13]. In order to reduce heteroskedasticity, I eliminate forecasts for firms with share prices below $5 or market capitalisations below $100 million (both in 2002 CPI-deflated dollars) and I scale forecasts and earnings by the stock price 10 days before the forecast date. I further trim the 1% extreme values of the test variables to avoid concerns that the results are driven by outliers. The results are not, however, sensitive to the choice of outlier rule.

The unit of analysis is an analyst-firm pair. The reason for this is that an analyst may be biased with respect to one firm but not with respect to another. In the future I will often use the term "analyst" for an analyst-firm pair whenever no confusion is caused.

When conducting an event study it is important to establish with some accuracy the date on which a forecast was made. I/B/E/S has reported this date with increasing precision over time, with a big improvement occurring at the beginning of the nineties, when the data-

[13] I/B/E/S has several advantages with respect to other databases. It is free of survivorship bias and (unlike Zacks) it comprises the great majority of public forecasts. Furthermore, unlike actuals from others sources like COMPUSTAT, I/B/E/S actual earnings are recorded on the same basis as I/B/E/S forecasts. According to Abarbanell and Lehavy [2] significant changes were made in the period 1989-1991 and in 1995 to systematically redefine reported earnings in order to improve its quality. Although I include in my sample some years previous to the 1995, all the empirical results remain unchanged when the sample is restricted to the period 1996-2002.

entering system was switched from manual into electronic. From 1992 almost all analyst enter their forecasts directly into the system, usually within 24 hours of providing them to clients. For this reason I limit the sample to the period 1993-2001.

I compute the consensus c as the average of all forecasts made on the last day on which a forecast occurred. Because forecasts on a specific day incorporate information contained in previous forecasts, I believe that this measure represents the best possible estimation of the prevailing consensus. However the results remain unchanged even when alternative measures are used, like the average of the last three forecasts or a weighted average of all prevailing forecasts. The test variables are $Earnings = EARNINGS - CONSENSUS = (e-c)$ and $Forecast = FORECAST - CONSENSUS = (f-c)$. Henceforth in the chapter I will refer to earnings and forecasts as the variables net of the consensus, and to gross earnings and gross forecast as the variables before the consensus is subtracted.

I compute a proxy for q using the past information on the analyst-firm pair. I first construct a variable d that equals 1 if a forecast is above actual earnings and 0 if it is equal to or below actual earnings. For every analyst-firm pair s^{th} forecast, I compute a weighted average of d from its first forecast to the forecast $s-1$. I use an *inverse-weighting* scheme to assign higher weights to more recent forecasts. If a forecast l has been issued h_l quarters before forecast s, the inverse weighting assigns weight $\omega_l = (1/h_l)\left(\sum_{k=1}^{s-1}(1/h_k)\right)$to the value d_l and the proxy for q becomes $Prior_{is} = \sum_{l=1}^{s-1}\omega_{il}d_{il}$. The inverse-weighting scheme accounts for two possible effects: a recency effect (investors remember better more recent forecast errors) and changing bias over time (more recent forecasts are more indicative of the existence of a bias in the present than forecasts in the distant past).

Lastly, I use CRSP data to assign a market reaction to every forecast. I compute investor's reaction as the abnormal return (i.e. the firm's raw return less the value-weighted market return) on the shares of the company on which a forecast was made, around the forecast date (all the main results remain unchanged when I use a two-day or a three-day window around the forecast date). $Return_{it} = \frac{P_{it}-P_{it-1}}{P_{it-1}} - \frac{MP_{it}-MP_{it-1}}{MP_{it-1}}$, where P is the share price of the company and MP is the value-weighted average price of the market.

The sample I use contains 1,178,425 forecasts, covering 7,274 analysts, 7,236 firms and 526 brokerage houses over a period of 10 years. The number of firm-analyst pairs is 105,698 and the number of firm-quarter combinations is 110,137. Table 1 reports summary statistics for the major variables used in the empirical analysis.

3.7 Results

3.7.1 Nonparametric results

In this section I do not impose any parametric form on the estimation of the conditional means of earnings and returns:

$$Earnings_{it} = \varphi(Forecast_{it},\ Prior_{it}) + \varepsilon_{it} \tag{3.17}$$

$$Return_{it} = \pi(Forecast_{it},\ Prior_{it}) + \eta_{it} \tag{3.18}$$

Pooled Regressions

I pool initially all analysts together. Figures K and L display the estimated relation between earnings, returns and forecasts. I use a nonparametric regression with a normal kernel, optimal bandwidth and 50 grid points. Both figures show that the estimated curves are concave and therefore support the predictions of the theoretical model. They also reject the predictions of models based on alternative assumptions about the rationality of the agents. Figure K also casts doubts over empirical efforts, common in the literature, based on the assumption that the relationship between earnings and forecasts is linear (see, for example, Zitzewitz [45]).

I interpret the concavity of the earnings function as evidence that analysts are positively biased, specially when they make optimistic forecasts. I interpret Figure L as evidence that investors are aware of this fact, and regard optimistic (but not pessimistic) forecasts with scepticism.

Given the big sample size, the confidence intervals are very close to the estimated conditional means, specially for the (less noisy) earnings function.

Note lastly that the slope of the earnings function is slightly higher than 1 for pessimistic forecasts, a finding that is not consistent with the theoretical predictions. Although this is a potentially interesting question, I believe that it does not invalidate the relevance of the model.

In order to understand the economic magnitude of my empirical findings, consider the example of a firm with a share price of \$40 and an underlying consensus about earnings in the following quarter of 40 cents per share. Assume that an average analyst makes a forecast of 0. Figure K shows that the optimal expectation of earnings is slightly below 0; I find from the output of the nonparametric estimation that expected earnings are in fact, -0.7 cents (the optimal expectation of gross earnings is therefore 39.3 cents per share).

Assume now that the analyst makes a forecast of 10 cents, a very optimistic forecast (a forecast of 10 cents is approximately in the 90th percentile of the forecast distribution). I find that expected earnings conditional on such forecast are 1.3 cents, that is, only 2 cents higher than if the forecast is 0.

Would a pessimistic forecast of equal magnitude have a similar absolute effect on expected earnings? Figure K shows that it would not. Quantitatively, expected earnings conditional on a forecast of -10 are -13.6 cents. I conclude that the magnitudes are significant: expected earnings change 6.5 times more if a forecast of 10 cents, in absolute value, is optimistic rather than pessimistic.

More importantly, do investors react differently to optimistic and pessimistic forecasts? The concavity of the curve in Figure L shows that they do. The standard deviation of the daily return in the sample is 3.6%. Using the results of the nonparametric regression I find that the price of the shares of a company increases by only 0.15% when an analyst makes an optimistic forecast of 10 cents on a certain day. However, they decrease by an average of 0.7%, (a magnitude more than 5 times higher, and around 20% of the daily return standard deviation) if the forecast is of the same absolute value but negative sign. The magnitudes are also economically significant.

Regressions by different values of *Prior*

In this subsection I study the effect of the prior probability that the analyst is biased on the shape of the earnings and return functions. To achieve this, I rank the observations in the sample according to the variable *Prior* and split the sample into five quintiles. Higher (lower) quintiles refer to analysts that have predicted very often above (below) actual earnings.

I run nonparametric regressions separately for each quintile and plot them together in Figures M and N. Figure M shows that optimistic forecasts imply higher earnings *only* if they are announced by an analyst who has not predicted too optimistically in the past. Indeed the estimated functions for the first two quintiles are reasonably linear and close to the 45-degree line. On the other hand, optimistic forecasts announced by analysts with high values of *Prior* are strongly biased and not related to higher earnings. The bias is however much lower when such analysts make pessimistic forecasts. Indeed I find that the conditional mean functions for the three higher quintiles are strongly concave curves.

Figure N displays the reaction of investors to the announcements made by different types of analysts. Importantly, investors react equally to *any* pessimistic forecast, regardless of who made it (note that the curves for different quintiles intersect if the news are pessimistic). However, investors react to an optimistic forecast by driving the price of the shares up *only* if that forecast is made by an analyst in a low quintile.

The difference in the reaction to optimistic news announced by different analysts is also economically significant. Assume for instance that a forecast of 10 cents is announced by an analyst in the first quintile. Using the output of the nonparametric estimations I find that expected earnings are 6.8 cents higher, and the share price increases by 0.27% (almost twice the reaction to an average analyst, and 7.5% of the daily return standard deviation). However, if the analyst announcing a forecast of 10 cents belongs to the highest quintile, expected earnings are actually *lower* (by 2 cents) and the share price changes by just 0.02% on average.

3.7.2 Parametric Results

In this section I complement the results above with the estimation of parametric nonlinear versions of (3.17) and (3.18). This allows me to introduce controls on the estimation and to account for intra-cluster serial correlation on the errors, a potentially serious problem.

Pooled Regressions

I first study the concavity of the earnings and return functions for the average analyst:

$$Earnings_{it} = \alpha + \alpha_0 Forecast_{it} + \alpha_1 ForSq_{it} + \xi_{it} \tag{3.19}$$

$$Return_{it} = \beta + \beta_0 Forecast_{it} + \beta_1 ForSq_{it} + \beta_2 LReturn_{it} + \nu_{it} \tag{3.20}$$

, where $ForSq$ accounts for Forecast Squared.

Because the return function is essentially an event study, it is important to ensure that I am capturing the reaction of investors to an individual forecast rather than to other contemporaneous firm-specific events that may affect their valuation of the firm. I add the one-day lagged abnormal return, $LReturn$, to the regression to control for this possibility. The public forecasts of securities analysts cannot react as quickly as the market to a firm-specific event, since analysts have to communicate their forecasts to their private clients at least some hours before releasing them publicly to the I/B/E/S system. As a result, introducing the lagged return should control for the effects of any event that influences the share price and also induces a revision of the firm's earnings forecast (in unreported regressions I ensure that adding up to five lags does not affect the empirical results). I also include year and quarter dummies.

I estimate (3.19) and (3.20) by Least Squares and Analyst Fixed Effects. The t statistics are constructed using cluster-adjusted standard errors, with all forecasts for the same firm-quarter grouped in one cluster. Cluster-adjusted standard errors guarantee consistency against arbitrary intra-cluster error correlation as well as arbitrary forms of heteroskedasticity across the whole sample. In the sample cluster-adjusted firm-quarter standard errors are much larger than GLS or OLS standard errors, and very similar to errors obtained by grouping forecasts by firm or by analyst. Thus I regard them as conservative estimates.

I also report the results of median regressions. Because very often financial data have very fat tails, it is important to ensure that the results are robust to outlying observations. Median regressions are more robust to outliers because they minimise the sum of absolute deviations, rather than the sum of squared deviations.

The results of the pooled parametric regressions are displayed in Table 2. The coefficients on $Forecast$ and $ForSq$ are positive and negative, respectively, across all regressions, confirming the existence of a concave relationship between the test variables. The t statistics are all significant at the 1% level. Note that the introduction of fixed effects does not alter significantly the magnitude of the coefficients. The median regression of the return function, however, produces estimates that are much lower in value than the least squares estimation (0.19 versus 0.58). Although the coefficients are still significant, this seems to support the hypothesis that observations at the tails of the distribution may have a disproportionate effect on the empirical results.

Regressions by different values of $Prior$

In this subsection I interact $Forecast$ and $ForSq$ with the prior belief that the analyst is biased. I also allow for different intercepts by including both the variable $Prior$ and its square, $PrSq$.

In the return function I introduce further controls to ensure that I am isolating the reaction of the market to individual forecasts. Because analysts sometimes release forecasts for the same firm on the same day, I control for the average forecast of the other analysts who made an announcement on the same day in which an analyst i announces his forecast. Including the control variable $AvForecast$ will help to isolate the reaction to the message of a particular analyst. I control also for the average forecast squared, $AvForSq$, and the number of forecasts that compose such average, $NumFor$. Lastly, I include the interactions of $AvForecast$ and $AvForSq$ with $NumFor$. The final specifications are as follows:

$$Earnings_{it} = \gamma + \gamma_0 Forecast_{it} + \gamma_1 ForSq_{it} + \gamma_2 Prior_{it} + \gamma_3 PrSq_{it}$$
$$+\gamma_4 Forecast_{it} * Prior_{it} + \gamma_5 ForSq_{it} * Prior_{it} + \zeta_{it} \quad (3.21)$$

$$Return_{it} = \delta + \delta_0 Forecast_{it} + \delta_1 ForSq_{it} + \delta_2 Prior_{it} + \delta_3 PrSq_{it} \quad (3.22)$$
$$+\delta_4 Forecast_{it} * Prior_{it} + \delta_5 ForSq_{it} * Prior_{it} + \sum_{j=1}^{J} \delta_j X_{jit} + \tau_{it}$$

, where X refers to the control variables discussed above.

My aim with this parametric specification is to test whether $\frac{\partial^2 Earnings}{\partial Forecast \partial Prior} = \gamma_4 + 2\gamma_5 Forecast_{it} < 0$ and $\frac{\partial^2 Return}{\partial Forecast \partial Prior} = \delta_4 + 2\delta_5 Forecast_{it} < 0$. The results are displayed in Table 3. The coefficients of the interaction between *Forecast* and *Prior* are all statistically significant.

By operating with the coefficients γ_4, γ_5, I find that the predictions of the model are largely met for the earnings function. From the output of the LS estimation, for instance, I have that $\gamma_4 = -0.23$ and $\gamma_5 = -39$. I conclude from these values that the marginal effect of *Prior* on the slope of the earnings function is negative unless the forecast is very pessimistic[14]. Similar conclusions are obtained for the FE and $Median$ specifications of the earnings function. These conclusions corroborate the findings, that the slope of the earnings function is slower for analysts with higher *Prior* unless they announce pessimistic forecasts.

[14] $-0.23 - 2 * 39 * Forecast = 0$ at the value of $Forecast = -0.003$, which is located approximately at the 15th percentile of the distribution of *Forecast*.. For values of *Forecast* above -0.003 the cross-derivative has a negative sign.

I reach similar (if quantitatively stronger) conclusions by operating the coefficients of the return function. The coefficients $\delta_4 = -0.16$ and $\delta_5 = 5.95$ in the LS estimation, for example, indicate that the slope of the return function is lower for analysts with higher values of $Prior$ unless they announce extremely bad news (i.e. in the first five percentiles of the $Forecast$ distribution).

3.8 Why are Analysts Biased?

As was mentioned in the Introduction, regulators have implemented reforms in the last years to tackle the bias in analysts' advice to investors. Choosing the right policy instrument to address such bias is clearly a complex issue that goes beyond the scope of this chapter. Making the right choice, however, requires some understanding of the reasons why analysts are biased. In this section I use the empirical strategy developed above to discriminate between two alternative hypotheses[15] that have inspired recent alternative policies.

In the belief that investment banking conflicts of interests are significant in the industry, America's securities regulators imposed in April 2003 reforms to strengthen the internal separation between the investment banking and the research divisions of big brokerage houses. There is substantial evidence that this belief is well-founded. Very often, for instance, an analyst in the research division of a big house can find that the investment bank division is counselling the very same company that he is forecasting on. Such counselling is habitually related to corporate finance activities, like initial public offering

[15] Common to both hypotheses about analysts' misaligned incentives is the belief that managers of the companies studied want them to be optimistic. Lim [31], for instance, states that "managers prefer favourable forecasts because these support higher capital valuations and, hence, their compensation levels".

("IPOs"), seasoned equity offering ("SEOs"), mergers or acquisitions. In fact, analysts at big houses are frequently involved in these activities, consulting on possible deals, joining road shows and starting coverage of prospective or current corporate clients. Since big houses make most of their profits from corporate financing, analysts compensations can be strongly determined by their contribution to their employer seizing a bigger share of this profitable business (for evidence on this, see Unger [43], quoted in Boni and Womack [10]). It is often stated in the financial media that this has a strong impact on the priorities of analysts working for firms with investment banking divisions[16].

Some authors have suggested, on the other hand, that analysts cooperate with the company's management in order to maintain and improve their access to it. Lim [31] has argued, for instance, that analysts can only have steady and timely access to relevant information through their contacts with company managers. More interestingly, Boni and Womack [10] have emphasised that one of the services that investors value most from analysts is their ability to arrange meetings for them with company executives. Both authors have argued that these may be the reasons why analysts seek to avoid painting a bleak picture of the company's future[17]. Regulation Fair Disclosure, introduced by the SEC in late 2001, to eliminate selective disclosure of material information, has been inspired by these arguments.

---

[16] For instance, Boni and Womack [10] offer survey-based evidence that buy-side professionals believe that the desire to attract investment banking business for their firm motivates analysts far more than any possible concern for the accuracy of their predictions. Anecdotal evidence is strongly consistent with this belief. As an example, one Goldman Sachs analyst was quoted as listing in an internal email his three most important goals for 2000 as "1) Get more investment banking revenue 2) Get more investment banking revenue 3) Get more investment banking revenue" (The Economist [42])

[17] Boni and Womack [10] report that 69.3% of institutional investors believe direct pressure from corporate management has a continuing effect on analysts research.

It is important to understand the different implications of the "investment banking" reason and the "management access" reason. Consider first the size of the brokerage house that employs the analyst. If the investment banking reason is true, analysts in bigger brokerage houses will be more biased than those in smaller brokerage houses. This is because only the biggest brokerage houses have an investment bank division, and can therefore expose their analysts to such conflicts of interest.

Under the "management access" reason analysts are eager to preserve managers' cooperation. There is evidence, however, that managers of a company also benefit from receiving analysts' coverage and want to encourage it. The reason is that firms covered by securities analysts are better known to investors and in general boost their share price[18]. Therefore the relationship between analysts and management can be regarded as one where the parties need each other and are eager to obtain concessions from each other. Lim [31] has suggested that concessions by analysts could include biasing the earnings' forecast.

I posit that under the "management access" reason the relative "bargaining power" of the two parties influences the extent to which analyst incentives are misaligned. For example, analysts from big brokerage houses hold a strong position, as the high visibility that they provide is strongly coveted by management. Similarly older analysts and analysts that follow more companies hold a stronger position and are therefore less biased. On the other hand, when a company is more important and more widely covered, I expect their

[18] Krigman, Shaw and Womack [30] conducted a survey among 578 firms that went public between 1993 and 1995 and conducted a SEO within three years of their IPO. They asked decision-makers at companies that had switched lead underwriters the main reasons for that decision. The main reason was revealed to be to "buy additional and influential analyst coverage from the new lead underwriter".

managers to enjoy more "bargaining power", and the forecasts made on those firms to be more biased.

I first provide preliminary evidence in support of the "management access" reason by identifying the determinants of *Prior*. I regress *Prior* on the size of the brokerage house, the number of companies that an analyst follows, the analyst's general experience, the company's size in terms of book value and market capitalisation, and the coverage that the company is receiving from analysts. Since neither *Prior* nor the exogenous variables vary within an analyst-firm-quarter cell, I use only one observation per cell and drop the rest. The results are displayed in Table 4.

I first find that *Prior* is positively correlated with analyst coverage, book value and market capitalisation and negatively correlated with size of the brokerage size, the number of companies that the analyst follows and the analyst general experience. This is consistent with the "management access" reason. I interpret the preliminary finding that analysts from bigger brokerage houses are less biased than those from smaller houses as evidence against the "investment banking" reason.

The results in Table 4 seem to indicate that the bias of an analyst is correlated to certain variables related to his "bargaining power". Using the empirical strategy developed in Section 3.7, however, I can provide a better test of the "management access" hypothesis. If this hypothesis and the theoretical model of Section 3.3 are correct, an analyst working for a small brokerage house is very likely to have misaligned incentives. As a result, such analyst will predict above actual earnings if he makes an optimistic forecast, although not as much if his forecast is pessimistic. Correspondingly, the market will react to pessimistic but not to optimistic forecasts.

I repeat regressions (3.21) and (3.22), and substitute *Prior* by each of the variables that were correlated with it in Table 4. If the hypotheses are true, I expect to find $\frac{\partial^2 Earnings}{\partial Forecast \partial X} > 0$ and $\frac{\partial^2 Return}{\partial Forecast \partial X} > 0$ for the variables of brokerage size, number of companies that the analyst follows and his general experience and $\frac{\partial^2 Earnings}{\partial Forecast \partial X} < 0$ and $\frac{\partial^2 Return}{\partial Forecast \partial X} < 0$ for the variables of coverage of the firm, and size of the firm (book value and market capitalisation). The results are displayed in Table 5.

Note first that the findings for the earnings function are largely consistent with the "management access" hypothesis. For instance, expected earnings are higher following an optimistic forecast if the analyst making it works in a big brokerage house and follows sparsely covered and small (in terms of book value) firms. The coefficients on number of companies that the analyst follows and market capitalisation also have the predicted sign but are not significant. In the case of market capitalisation this is probably due to the strong collinearity with the coverage of the firm and its book value (the correlations of market capitalisation with them is 68% and 60% respectively). The coefficient on experience has, however, the wrong sign.

The findings for the return function do not support the hypothesis. Only two coefficients are significant, and the one on book value displays the wrong sign.

I confirm these results regarding the earnings function by running nonparametric regressions separately for sub-samples of high and low levels of the test variables. The results are plotted in Figures O-T. From Figure O it can be seen that analysts from bigger houses are less biased when they announce good news than analysts from smaller houses. Similarly, analysts following widely covered (Figure R) or bigger firms (Figures S and T) are less biased when they announce good news, but not as much when they announce bad news. Figures P and Q do not seem to suggest that the variables capturing the number of firms followed by the analyst and his experience significantly affect the shape of the earnings function.

3.9 Concluding Remarks

In this chapter I have found evidence of a strategic interaction taking place in the stock market between investors and biased securities analysts. By using a simple model of cheap talk, I have obtained two intuitive predictions about the optimal reaction of investors to forecasts. I have then shown that the behaviour of investors and analysts in the real world is consistent with these predictions. At an academic level, these findings demonstrate the need to use a game-theoretical approach to study the effect of forecasts in the financial markets.

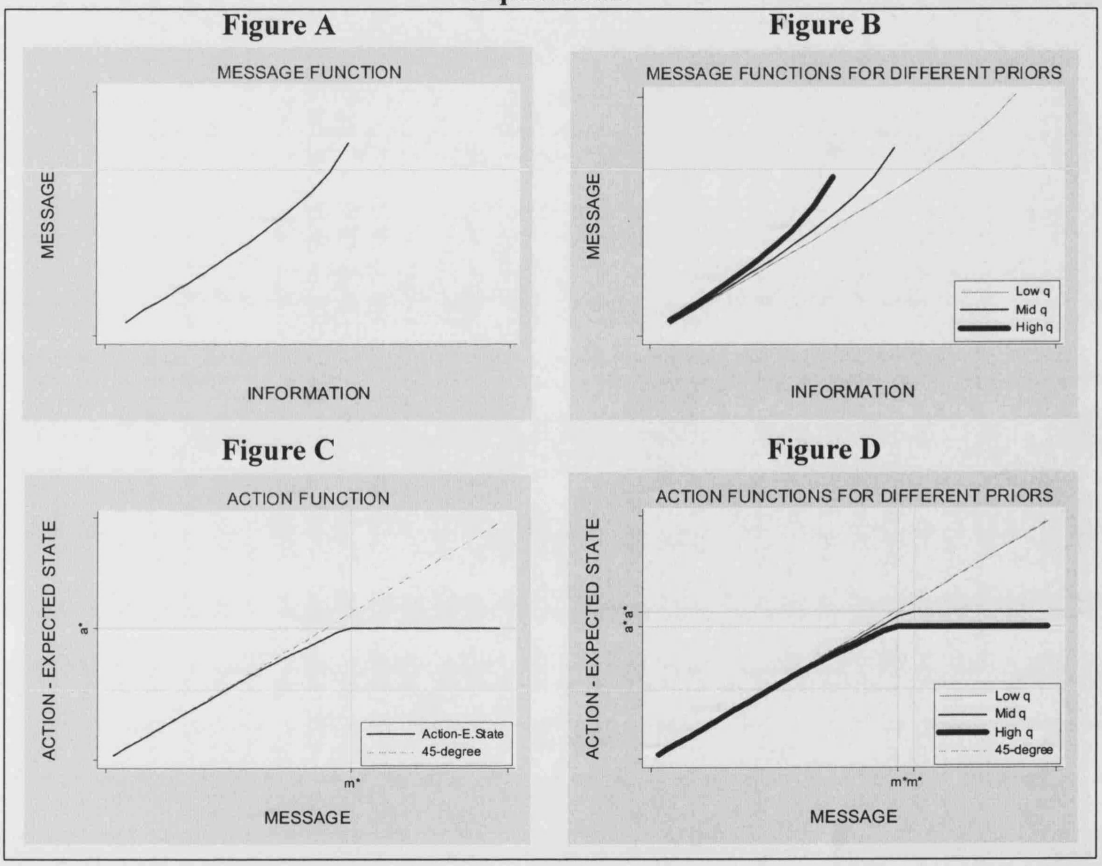
My findings also have more practical implications. In the last years, regulators in the US have introduced a number of reforms in the securities industry aimed at curbing the misalignment of incentives between analysts and investors. This chapter shows that investors *as a whole* are able to anticipate the misalignment of incentives and discount earnings forecasts accordingly. Investors are also able to discriminate across analysts, assigning less credibility to analysts with highly biased records. Contrary to what is sometimes believed, the average investor does not react naively to public announcements by analysts.

I have also shown, however, that lack of naivety on the investor's side does not imply that there is no scope for public intervention. This is easily understood by observing the flat right tail in the action function of Figures C and L. Even if investors attempt to discount the bias, the misalignment of incentives is detrimental to the efficient transmission of information. The reason is that it prevents analysts from credibly transmitting good news about the firm's earnings, and therefore obstructs the efficient transmission of information in the market.

3.10 Appendix

3.10.1 Figures A-D: Equilibrium

Equilibrium



3.10.2 Figures E-H: Other Equilibria, Cognitive Biases and Naive Investors

Other Equilibria

Figure E

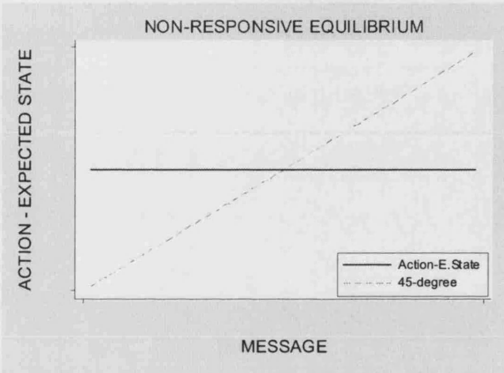
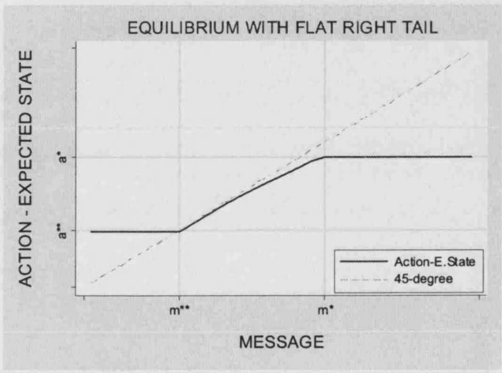


Figure F



Cognitive Biases and Naive Investors

Figure G

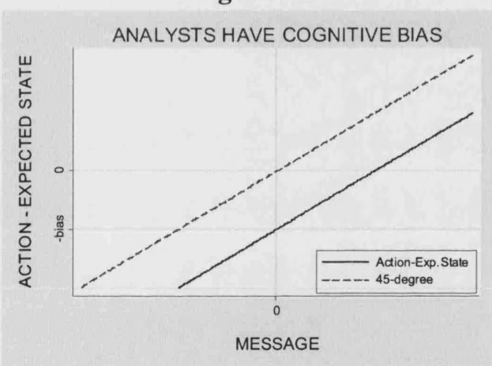


Figure H

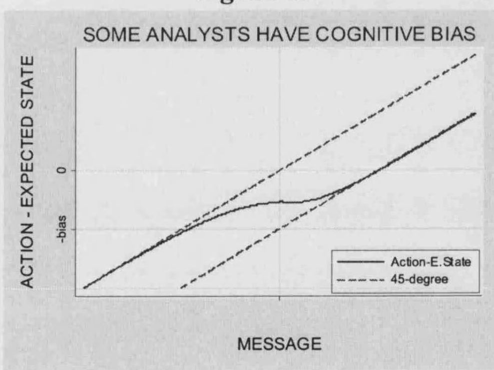


Figure I

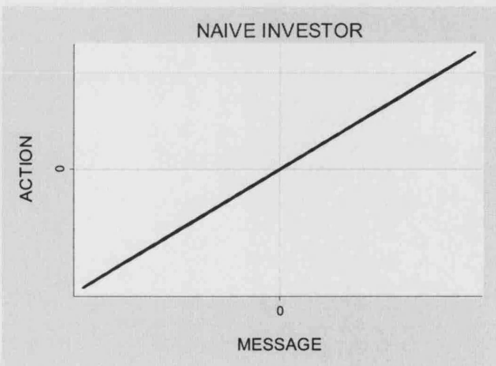
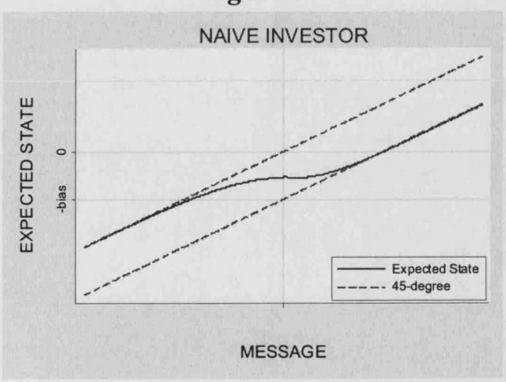


Figure J



3.10.3 Figures K-N: Nonparametric Regressions of Earnings and Returns

Nonparametric Regressions of Earnings and Returns

Figures K and L show nonparametric regressions of Earnings and Return on Forecast. The scale in the horizontal axis shows the percentiles of the *Forecast* variable. Forecast is the analyst forecast minus the prevailing consensus (consensus is the average forecast of the previous day in which a forecast was made on the same firm/quarter earnings). Earnings is the actual value of these earnings minus the consensus. Return is the abnormal return on the firm's share price on the day in which the forecast was made. Figures M and N show nonparametric regressions of Earnings and Return on Forecast for analysts with different prior probability of being biased. For each analyst I compute the variable Prior (defined in Section 3.6), and split the sample in five quintiles according to this variable. q1 (the thinner line) refers to the quintile of analysts with the lowest value of Prior. I run the regressions separately for each quintile. I use a normal kernel, optimal bandwidth (1.6*Standard Deviation of Forecast*Number of Observations) and 50 grid points. The confidence intervals are calculated with bootstrapping techniques (samples with replacement and 500 replications). The scale in the horizontal axis shows the percentiles of the Forecast variable.

Figure K

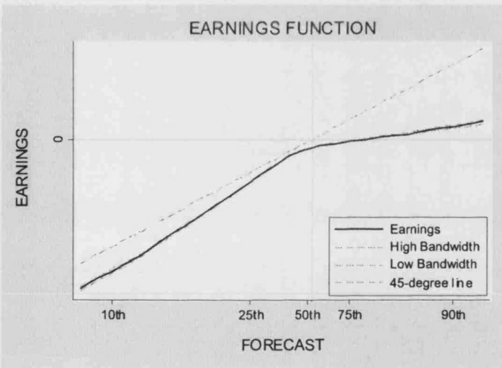


Figure L

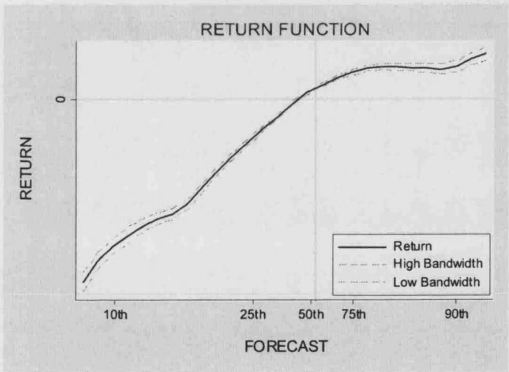


Figure M

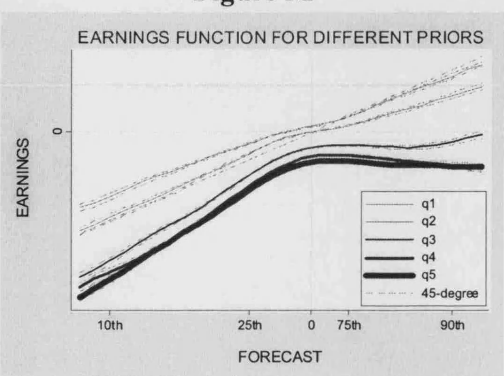
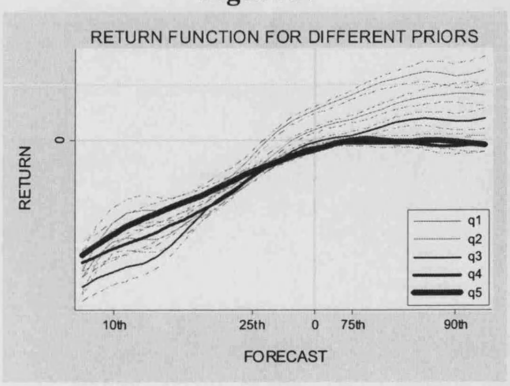


Figure N



3.10.4 Figures O-T: The Importance of Relative Bargaining Power

The Importance of Relative Bargaining Power

Figures O-T show the Earnings and Return regressions for different analyst and firm characteristics. The firm characteristics are defined in Table 4. I use the same kernel, optimal bandwidths, grid points and bootstrapping techniques as in Figures K-N.

Figure O

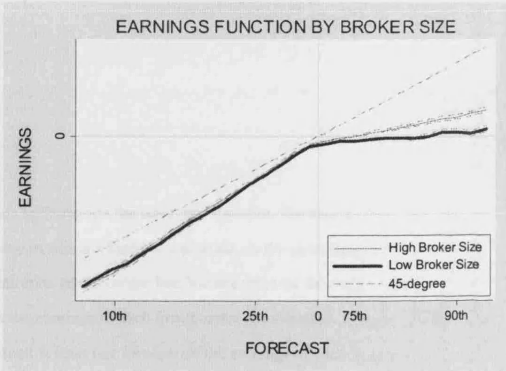


Figure P

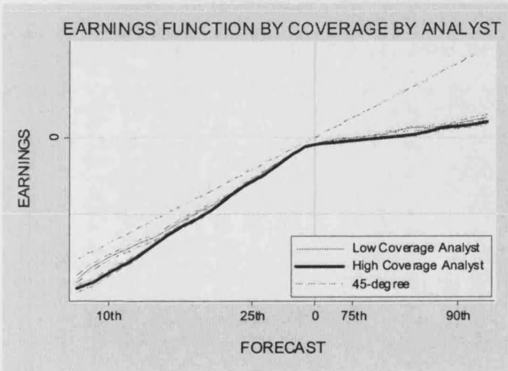


Figure Q

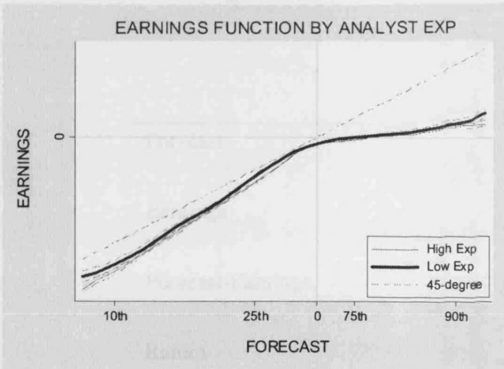


Figure R

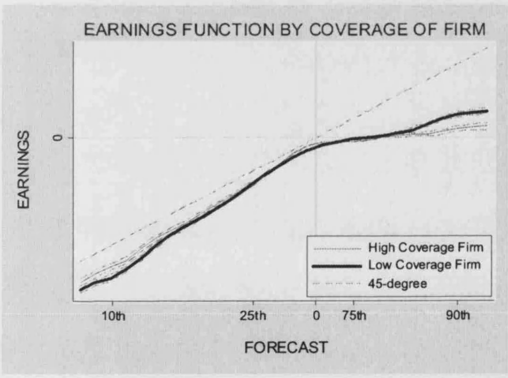


Figure S

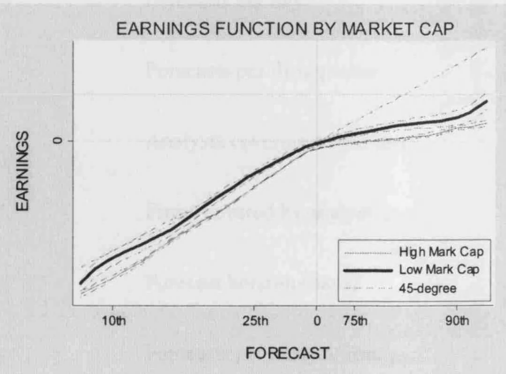
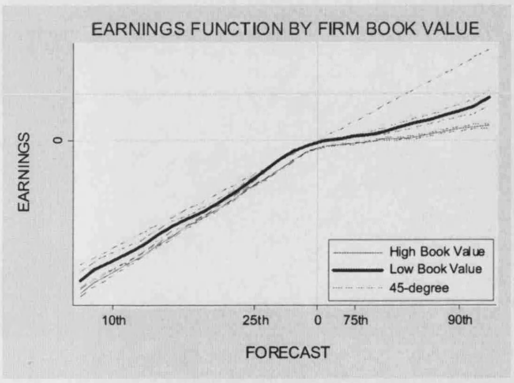


Figure T



3.10.5 Table 1: Summary Statistics

Table 1
Summary Statistics

This table reports the summary statistics. Forecast is the analyst forecast minus the prevailing consensus (consensus is the average forecast of the previous day in which a forecast was made on the same firm/quarter earnings). Earnings is the actual value of these earnings minus the consensus. Return is the abnormal return on the firm's share price on the day in which the forecast was made. Forecasts per firm/quarter is the number of forecasts that were made on the earnings of each firm/quarter combination. Similarly for Forecasts per firm/quarter/analyst. Analysts per firm/quarter is the number of analysts that issued at least one forecast on the earnings of each firm/quarter combination. Firms per analyst/quarter is the number of firms for which an analyst issued at least one forecast in a quarter. Forecast horizon is the number of days between the forecast announcement and the announcement of actual earnings.

	Mean	S.D.	Median	p25	p75
Forecast	-0.00039	0.00339	-0.00006	-0.00107	0.00064
Earnings	-0.00130	0.00688	-0.00002	-0.00190	0.00099
Forecast-Earnings	0.00090	0.00636	0	-0.00090	0.00113
Return	-0.00089	0.03635	-0.00067	-0.01639	0.01555
Forecasts per day	1.35	1.15	1	1	1
Forecasts per firm/quarter	9.78	11.14	6	3	12
Analysts covering a firm/quarter	6.07	5.44	4	2	8
Firms covered by analyst/quarter	7.38	8.70	6	3	10
Forecast horizon (days)	85.49	45.13	86	47	117
Forecasts per analyst/firm/quarter	2.12	1.22	2	1	3

3.10.6 Table 2: Pooled Earnings and Return Regressions

Table 2
Pooled Earnings and Return Regressions

This table reports the regressions of Earnings and Returns on Forecasts. These three variables are defined as in Table 1. All analyst-firm pairs are pooled together. ForSq is the Forecast Squared. LReturn is the one-day lagged Return variable. Y&Q accounts for year and quarter dummies. LS accounts for standard least squares. FE regressions include analyst-firm dummies. Median accounts for median (quantile) regressions. The Standard Errors (in parentheses) are robust and adjusted for firms-quarters correlation for the LS and FE regressions. The number of observations is 1,178,792. *** denotes significant at the 1% level.

	Earnings			Return		
	LS	FE	Median	LS	FE	Median
Forecast	0.65***	0.66***	0.76***	0.59***	0.58***	0.19***
	(0.006)	(0.005)	(0.0005)	(0.02)	(0.02)	(0.01)
ForSq	-28.2***	-14.6***	-19.4***	-16.0***	-7.31***	-9.47***
	(0.88)	(0.66)	(0.049)	(2.13)	(2.22)	(1.41)
LReturn				0.01***	0.007***	0.003***
				(0.002)	(0.001)	(0.0009)
Y&Q Ef.	Yes	Yes	No	Yes	Yes	No
Adj.R^2	16%	37%	13%	0.5%	11%	0.05%

3.10.7 Table 3: Regressions with Prior

Table 3
Regressions including the effect of the Prior probability that the analyst is biased

This table reports the regressions of Earnings and Returns on Forecasts and Prior (the prior probability that the analyst is biased). Table 1 has the definitions of the first three variables, ForSq and Y&Q Ef.. Prior is computed as the (weighted) percentage of times that the analyst has forecasted above the actual value in the past. The weights (defined in Section 3.6) are higher for more recent forecasts. PrSq is the variable Prior squared. Controls include LReturn, the average forecast by other analysts on the same day and firm/quarter on which the analyst issued his forecast, the average forecast squared, the number of forecasts that compose this average, and the interactions of the average forecasts and its square with the number of forecasts the compose the average. LS, FE and Median regressions are defined as in Table 2. The Standard Errors (in parentheses) are robust and adjusted for firms-quarters correlation for the LS and FE regressions. ***, ** and * denote significant at the 1%, 5% and 10% level. Regressions include a constant. The number of observations is 965,527.

	Earnings			Return		
	LS	FE	Median	LS	FE	Median
Forecast	0.69***	0.69***	0.84***	0.34***	0.35***	0.16***
	(0.009)	(0.007)	(0.001)	(0.03)	(0.03)	(0.02)
ForSq	-19.1***	-11.7***	-5.96***	-11.7***	3.06	-3.83
	(1.32)	(1.03)	(0.11)	(3.27)	(3.26)	(2.43)
Prior	-0.01***	-0.008***	-0.002***	-0.01***	-0.01***	-0.006***
	(0.0002)	(0.0002)	(0.00002)	(0.0006)	(0.0008)	(0.0006)
PrSq	0.009***	0.007***	0.001***	0.008***	0.009***	0.005***
	(0.0002)	(0.0002)	(0.00003)	(0.0007)	(0.0008)	(0.0007)
Prior*For.	-0.23***	-0.18***	-0.35***	-0.16**	-0.26***	-0.23***
	(0.02)	(0.02)	(0.003)	(0.08)	(0.08)	(0.075)
Prior*ForSq	-39.0***	-12.0***	-63.7***	5.95	-19.9**	-12.8*
	(3.13)	(2.63)	(0.32)	(7.99)	(8.25)	(7.06)
Y&Q Ef.	Yes	Yes	No	Yes	Yes	No
Controls	No	No	No	Yes	Yes	Yes
Adj.R^2	18%	40%	14%	0.9%	12%	0.2%

3.10.8 Table 4: The Determinants of Prior

Table 4
The Determinants of the Prior Probability that the Analyst is Biased

This table reports the regressions of the prior probability that the analyst (analyst-firm pair) is a biased type on a number of analyst and firm characteristics. Prior is defined as in Table 3. Broker Size is the number of analysts that are employed in a brokerage house in a year. Coman is the number of firms on which the analyst issues at least one forecast in a year. Ancom is the number of analysts that issue a forecast on the firm in a year. GenEx is the number of quarters in which the analyst has issued at least one forecast. Marcap is the end-of-year market capitalisation of the firm. Book is the end-of-year book value of the firm. For each year and variable I split the sample in ten deciles and, to avoid endogeneity problems, I assign every analyst-firm pair the decile of the previous year. The results are unchanged if I include the variables in absolute levels and/or in their current year values. *** denotes significant at the 1% level. The standard errors (in parentheses) are heteroskedasticity robust. The number of observations is 477,434.

	Prior	
	Coefficient	S.E.
Broker Size	-0.003***	(0.0005)
ComAn	-0.003***	(0.003)
GenExp	-0.014***	(0.002)
AnCom	0.003***	(0.0004)
MarCap	0.004***	(0.0005)
Book	0.009***	(0.0003)
Year & Quarter Effects	Yes	
Fixed Effects	Yes	
Adj. R^2	60%	

3.10.9 Table 5: The importance of relative bargaining power

Table 5

The importance of relative bargaining power

This table reports the effect on the Earnings and Return regression of interacting Forecast and ForSq (defined in Table 2) with a number of firm and analyst characteristics (defined in Table 4). The controls are defined in Table 3. S.E. (in parentheses) are robust and adjusted for firms-quarters correlation. The Standard Errors (in parentheses) are robust and adjusted for firms-quarters correlation for the LS and FE regressions. ***, ** and * denote significant at the 1%, 5% and 10% level. Regressions include a constant. The number of observations is 1,047,479.

	Earnings		Return	
	Coefficient	S.E.	Coefficient	S.E.
Broker Size*For	0.02***	(0.003)	0.03***	(0.009)
Broker Size*ForSq	2.12***	(0.40)	1.33	(0.95)
Coman*For	0.0008	(0.002)	-0.003	(0.008)
Coman*ForSq	0.50	(0.32)	0.57	(0.82)
GenExp*For	-0.005	(0.008)	0.03	(0.02)
GenExp*ForSq	-4.28***	(0.93)	-0.83	(2.30)
Ancom*For	-0.026***	(0.003)	-0.01	(0.01)
Ancom*ForSq	-1.27***	(0.41)	-1.35	(1.08)
Marcap*For	-0.005	(0.004)	0.009	(0.015)
Marcap*ForSq	-0.58	(0.51)	1.16	(1.43)
Book*For	0.001	(0.004)	0.009	(0.01)
Book*ForSq	-0.94**	(0.42)	2.35**	(1.15)
Controls	No		Yes	
Year & Quarter Effects	Yes		Yes	
Fixed Effects	Yes		Yes	
R^2	17%		1%	

3.10.10 Proofs

I use the following notation throughout the Appendix. First of all, $p = q/(1-q)$. Further, remember that $F(m,x) = 0$ is the implicit function of m on x defined by (3.10). Define $F(m,x)$ as a function with two exogenous variables, m and x. Define $F(m,x_0)$ as a function with one exogenous variable m, where x_0 is a parameter of the function. Lastly, define $F(m_0,x_0)$ as a specific value, given by the function $F(m,x)$ when it is evaluated at the values m_0 and x_0.

Proof of Lemma 4 When the investor observes a message m', his optimal action solves:

$$\underset{a}{Max} - \int (a-x)^2 \beta(x \mid m')\, dx$$
$$\underset{a}{Max}\{-a^2 + 2a \int x\beta(x \mid m')\, dx - \int x^2 \beta(x \mid m')\, dx\}$$

The optimal action is therefore $a = E[x \mid m']$.

Proof of Proposition 6 It follows from the discussion in the main text that a fully-responsive equilibrium requires $a(m)$ to be increasing. I first rule out fully-responsive equilibria where the message is issues using mixed strategies. Assume that m_1 and m_2, with $m_1 \neq m_2$, are both sent in equilibrium when $x = x_1$. If the analyst is indifferent between both m_1 and m_2, it must be that $a(m_1) = a(m_2)$. But in this case the action function is not strictly increasing and the equilibrium is not fully responsive.

Assume now that the message is issued using a pure strategy. The equilibrium message function, $m = \mu(x)$, must solve (3.10) and be defined for all x. But I show now that there exists a value of x, x_0, such that $F(m,x_0) < 0\ \forall m$. Take $x_0 = \frac{\sigma^2}{bp}$. If $m \leq x_0$, the inequality is immediate. If $m > x_0$, then $m^2 > x_0^2$. The exponential function has the property that $\exp(z) > 1 + z$. Therefore $F(m, x_0) < m - x_0 - b - bp(1 + (m^2 - x_0^2)/2\sigma^2) = m - \frac{\sigma^2}{2bp} - b - bp - \frac{bp}{2\sigma^2}m^2$. But this last expression reaches a maximum at $m_0 = \frac{\sigma^2}{bp}$, at which it is equal to $-b(1+p) < 0$. So (3.10) is not defined at the point $x_0 = \frac{\sigma^2}{bp}$. ■

Before proving Proposition 7, I use the following Lemma to derive some technical properties about (3.10).

Lemma 5 *Consider $F(m, x) = 0$. For any set of parameter values $b > 0, q > 0$ and $\sigma^2 > 0$:*

1. *There always exists a value of x, x_0, such that this equation has at least one solution m_0, and another value x_1 such that this equation has no solution.*

2. *There exists only one value of x, $\widetilde{x}$, for which this equation has a unique solution, $\widetilde{m}$. For every value $x < \widetilde{x}$ the equation has exactly two solutions, $\mu^1(x)$ and $\mu^2(x)$, with $\mu^1(x) < \widetilde{m} < \mu^2(x)$ $\forall x$.*

3. *The function $\mu^1(x)$ satisfies the following properties:*

   (a) $\frac{\partial \mu^1}{\partial x}$*>1*

   (b) $\frac{\partial^2 \mu^1}{\partial x^2}$*>0*

   (c) $\frac{\partial^2 \mu^1}{\partial x \partial q}$*>0*

Proof

Part 1

The second statement derives directly from Proposition 6. To prove the first statement, distinguish two cases:

1. If $p < 1$, make $x_0 = -2b$ and note that the function $F(m, -2b)$ is a continuous function $\forall m$, that $F(0, -2b_0) > 0$ and that $F(-2b_0, -2b_0) < 0$. It follows that there is a $m_0 \in]-2b, 0[$ such that $F(m_0, -2b) = 0$ and (3.10) has a solution.

2. If $p \geq 1$, make $x = -2bp$, and note that the function $F(m, -2bp)$ is a continuous function $\forall m$ that $F(0, -2bp) > 0$ and that $F(-2bp, -2bp) < 0$. It follows that there is a $m_0 \in]-2bp, 0[$ such that $F(m_0, -2bp) = 0$ and (3.10) has a solution.

Part 2

I first show that if (3.10) has a solution for x_0, then it has exactly two solutions for every $x < x_0$. Call m_0 to the solution of (3.10) for $x = x_0$ and make $k = m_0 - x_0 > 0$. Take a value $x_1 < x_0$ and note that the function $F(m, x_1)$ is continuous and twice differentiable $\forall m$, and that $F(x_1, x_1) < 0$ and $F(x_1 + k, x_1) > 0$. By L'Hôpital rule, note that:

$$\lim_{m \to +\infty} \frac{m - x_1}{1 + p\exp((m^2 - x_1^2)/2\sigma^2))} = \frac{1}{mp\exp((m^2 - x_1^2)/2\sigma^2))} = 0$$

It follows that there is a value m_2 such that $\forall m \geq m_2$,

$m - x_1 < b\,(1 + p\exp((m^2 - x_1^2)/2\sigma^2)))$. For those values of m, $F(m, x_1) < 0$. In particular, note that $F(m_2, x_1) < 0$ and that $m_2 > x_1 + k$. So I have found three values, $x_1 < x_1 + k < m_2$ such that $F(x_1, x_1) < 0$, $F(x_1 + k, x_1) > 0$ and $F(m_2, x_1) < 0$.

Therefore there are two values m_{11} and m_{12}, such that $x_1 < m_{11} < x_1 + k < m_{12} < m_2$ and $F(m_{11}, x_1) = F(m_{12}, x_1) = 0$. This shows that (3.10) has *at least* two solutions for every $x < x_0$.

I now show that (3.10) has *only* two solutions for every $x < x_0$. Assume that there is a third value, m_{13}, such that $F(m_{13}, x_1) = 0$. In that case there are two values, $m_3 \in (m_{11}, m_{13})$ and $m_4 \in (m_{13}, m_{12})$ such that $\frac{\partial F(m_3,x_1)}{\partial m} = \frac{\partial F(m_4,x_1)}{\partial m} = 0$, and another value, $m_5 \in (m_3, m_4)$ such that $\frac{\partial^2 F(m_5,x_1)}{\partial m^2} = 0$. But it is easy to see that $\frac{\partial^2 F(m,x_1)}{\partial m^2} < 0\ \forall m$, hence a contradiction and (3.10) has *exactly* two solutions.

From Proposition 6 and Part 1 of this Lemma, I have that the real set $H = \{x : \exists m / F(m,x) = 0\}$ (henceforth H) is non-empty and bounded from above. I also know that if x_0 belongs to H, then every $x < x_0$ also belongs to it. It follows that H has a supreme value, $\widetilde{x}$, and it is of the form $(-\infty, \widetilde{x}[$ or $(-\infty, \widetilde{x}]$. I finish the proof of Part 2 by showing that equation (3.10) has a unique solution for the supreme value $\widetilde{x}$.

I first show that *at least* one solution exists. Since $\widetilde{x}$ is the supreme of H, I can build a strictly increasing sequence $\{x_n\}_{n=1}^{\infty}$ such that $\lim\limits_{n\to\infty} x_n = \widetilde{x}$, and for every x_n there exist two values $m_{n,1}$ and $m_{n,2}$, with $m_{n,1} < m_{n,2}$ and $(x_n, m_{n,1})$ and $(x_n, m_{n,2})$ satisfy (3.10). Call $m_n = m_{n,1}$ and note that the sequence $\{m_n\}_{n=1}^{\infty}$ is bounded. It is bounded from below, since $\forall n$, $m_n > x_n > x_I$, where x_I is the first term of the sequence that converges to $\widetilde{x}$. It is also bounded from above. To see that, make $s = \max(|x_I|, |\widetilde{x}|)$. Since $\lim\limits_{m\to+\infty} \frac{m+s}{1+p_0 \exp((m^2-s^2)/2\sigma^2))} = 0$, there is a m_0 such that $\forall m \geq m_0$, $\frac{m+s}{1+p_0 \exp((m^2-s^2)/2\sigma^2))} < b$. Therefore, $\forall m \geq m_0$ and any n I have that $F(m, x_n) < F(m, s) < 0\ \forall n$ and the sequence is bounded from above by the value m_0.

The sequence $\{m_n\}_{n=1}^{\infty}$ is bounded so I can build a subsequence $\{m_{n_k}\}_{k=1}^{\infty}$that converges to a certain $\widetilde{m}$. Therefore the sequence $\{(x_{n_k}, m_{n_k})\}_{k=1}^{\infty}$ converges to $(\widetilde{x}, \widetilde{m})$.

Consider the function of two variables $F(x, m)$, which is continuous at $(\widetilde{x}, \widetilde{m})$. Since $\lim_{k\to+\infty}(x_{n_k}, m_{n_k}) = (\widetilde{x}, \widetilde{m})$, $\lim_{k\to+\infty} F(x_{n_k}, m_{n_k}) = F(\widetilde{x}, \widetilde{m})$. Since $F(x_{n_k}, m_{n_k}) = 0$, then $F(\widetilde{x}, \widetilde{m}) = 0$ and $m = \widetilde{m}$ is a solution to (3.10).

Lastly I show that the solution $\widetilde{m}$ is unique. Consider the function $F(m, \widetilde{x})$, which is continuous and twice differentiable. Assume that there are two values $\widetilde{m}_1 < \widetilde{m}_2$ that satisfy $F(\widetilde{m}_1, \widetilde{x}) = F(\widetilde{m}_2, \widetilde{x}) = 0$. If that is the case, there will be a value $m_3 \in]\widetilde{m}_1, \widetilde{m}_2[$ such that $\frac{\partial F(m_3, \widetilde{x})}{\partial m} = 0$. Since $\frac{\partial^2 F(m, \widetilde{x})}{\partial m^2} < 0$, $F(m, \widetilde{x})$ has a relative maximum at m_3, and since $F(\widetilde{m}_1, \widetilde{x}) = F(\widetilde{m}_2, \widetilde{x}) = 0$, then $F(m_3, \widetilde{x}) > 0$. Consider now the continuous function of x, $F(m_3, x)$. From the argument above, $F(m_3, \widetilde{x}) > 0$. On the other hand, $F(m_3, m_3) < 0$ and $m_3 > \widetilde{x}$. Consequently there must be a $x_{00} \in]\widetilde{x}, m_3[$ such that $F(m_3, x_{00}) = 0$. But if that is the case $\widetilde{x}$ is not the supreme of H, which is a contradiction.

Part 3

Assume that $x_1 < \widetilde{x}$ and $F(m_1, x_1) = F(m_2, x_1) = 0$. If $m_1 < m_2$, then note that$\frac{\partial F(m_1, x_1)}{\partial m} > 0$ and $\frac{\partial F(m_2, x_1)}{\partial m} < 0$. The proof is as follows. It was shown in the previous Part that $\frac{\partial^2 F(m, x_1)}{\partial m^2} < 0 \; \forall m$ and that there is a point $m_3 \in (m_1, m_2)$ such that $\frac{\partial F(m_3, x_1)}{\partial m} = 0$. It follows immediately that $\frac{\partial F(m_1, x_1)}{\partial m} > 0$ and $\frac{\partial F(m_2, x_1)}{\partial m} < 0$.

Note also that $\frac{\partial F(\widetilde{m}, \widetilde{x})}{\partial m} = 0$. The proof is as follows. It was shown in the previous Part that I can take a strictly increasing sequence $\{x_n\}_{n=1}^{\infty}$ that converges to $\widetilde{x}$ and two sequences $\{m_{n1}\}_{n=1}^{\infty}$ and $\{m_{n2}\}_{n=1}^{\infty}$ that converge to $\widetilde{m}$ such that $m_{n1} < m_{n2}$ and

$\{x_n, m_{n1}\}$ and $\{x_n, m_{n2}\}$ are solutions to (3.10). Assume that $\frac{\partial F(\widetilde{m},\widetilde{x})}{\partial m} > 0$, and consider the function $F(x,m)$, which is continuous at $(\widetilde{x},\widetilde{m})$. Since $\lim_{n\to+\infty}(x_n, m_{n2}) = (\widetilde{x},\widetilde{m})$, $\lim_{n\to+\infty} F(x_n, m_{n2}) = F(\widetilde{x},\widetilde{m}) \leq 0$, which is a contradiction. Similarly assuming $\frac{\partial F(\widetilde{m},\widetilde{x})}{\partial m} < 0$ leads to a contradiction, so necessarily $\frac{\partial F(\widetilde{m},\widetilde{x})}{\partial m} = 0$.

I proceed now to prove the different Subparts of Part 3.

Part 3.a Differentiating (3.10):

$$\frac{\partial \mu^1}{\partial x} = \frac{\frac{\partial F(m,x)}{\partial x}}{\frac{\partial F(m,x)}{\partial m}} = \frac{1 - \frac{bpx}{\sigma^2}\exp\left(\left(m^2 - x^2\right)/2\sigma^2\right)}{1 - \frac{bpm}{\sigma^2}\exp\left(\left(m^2 - x^2\right)/2\sigma^2\right)}$$

Since $\mu^1(x) < \mu^2(x)$, then $\frac{\partial F(m,x)}{\partial m} > 0$. Since $m > x$, it also follows that $\frac{\partial F(m,x)}{\partial x} > 0$ and $\frac{\partial \mu^1}{\partial x} > 1$.

Part 3.b Differentiating (3.10) twice and rearranging:

$$\frac{\partial^2 \mu^1}{\partial x^2} = \frac{\exp\left(-x^2/2\sigma^2\right)\frac{pb}{\sigma^2}}{\frac{\partial F(m,x)}{\partial m}}\{(\left(\frac{\partial m}{\partial x}\right)^2 - 1) + \frac{1}{\sigma^2}(m\left(\frac{\partial m}{\partial x}\right) - x)^2\}$$

The numerator and the denominator of the fraction are positive numbers. The term in brackets is positive, since $\frac{\partial m}{\partial x} > 1$, which completes the proof.

Part 3.c Differentiating (3.10) twice over x and over p.

$$\frac{\partial^2 m}{\partial x \partial p} = \frac{\left(\frac{b}{\sigma^2}\exp(\frac{m^2-x^2}{2\sigma^2})\right)\left(pb\exp(\frac{m^2-x^2}{2\sigma^2})(1 - \frac{bpx}{\sigma^2}\exp(\frac{m^2-x^2}{2\sigma^2})) + m - x\right)}{(1 - \frac{bpm}{\sigma^2}\exp(\frac{m^2-x^2}{2\sigma^2}))^3}$$

All elements of the fraction are positive, and therefore $\frac{\partial^2 m}{\partial x \partial p} > 0$. Furthermore $p = q/(1-q)$ so $\frac{\partial^2 m}{\partial x \partial q}$ has the same sign as $\frac{\partial^2 m}{\partial x \partial p}$. ■

Proof of Proposition 7 Any equilibrium must satisfy (3.6) and (3.8). From the proof of Proposition 6, $a(m)$ is increasing in an interval $[m_1, m_2]$ only if every message in this interval is produced in pure strategies. So in every region where $a(m)$ is increasing (3.10) must hold. But Proposition 6 states that if x is very high, no message satisfies (3.10). Therefore $\frac{\partial E[x|m]}{\partial m} = 0$ in equilibrium if x is higher than certain value which, for the moment, I call x^*. Assume that $\frac{\partial E[x|m]}{\partial m} > 0$ if $x \leq x^*$. Call $(\infty, x^*]$ and (x^*, ∞) the responsive and the non-responsive parts of the equilibrium. Lemma 5, shows that (3.10) has exactly two solutions, $\mu^1(x)$ and $\mu^2(x)$, with $\mu^1(x) < \mu^2(x)$. But $\mu^2(x)$ is bounded from below and cannot be part of an equilibrium. So in the responsive part of the equilibrium the message strategy must be $\mu^1(x)$. So $\mu^{MR}(x) = \mu^1(x)$, and Parts 2.a, 2.b and 2.c of Proposition 7 follow directly from Parts 3.a, 3.b, 3.c of Lemma 5.

Call $m^* = \mu^{MR}(x^*)$ and note that the message space can be divided in a responsive part, $(\infty, m^*]$ and a non-responsive part, (m^*, ∞). This proves Part 1.

In the responsive section $E[x \mid m] - x - b = 0$, so the biased analyst is reaching his bliss point. Therefore $a = \left[\mu^{MR}\right]^{-1}(m) + b$ and Parts 3.a, 3.b, and 3.c of this proposition follow from Parts 2.a, 2.b and 2.c.

Consider now the non-responsive part. Call a^* the value of the "flat" action when $m > m^*$. But $m > m^*$ can be issued only if the analyst is honest and $x > m^*$ or the analyst is biased and $x > x^*$. a^* therefore averages over all the possible values of x in which the equilibrium message is above m^*.

$$a^* = P\left[Honest \mid m > m^*\right] * E\left[x \mid m > m^*,\ Honest\right] \\ + P\left[Biased \mid m > m^*\right] * E\left[x \mid m > m^*,\ Biased\right]$$

$$a^* = \frac{\int_{m^*}^{\infty} \phi(x)\,dx * \int_{m^*}^{\infty} x\phi(x)\,dx + \int_{x^*}^{\infty} \phi(x)\,dx * \int_{x^*}^{\infty} x\phi(x)\,dx}{\int_{m^*}^{\infty} \phi(x)\,dx + \int_{x^*}^{\infty} \phi(x)\,dx} \tag{3.23}$$

Rearranging (3.23) it is easy to find (3.11). That proves Part 4. Part 5 comes directly from (3.6).

Note that $a(m)$ is continuous in the responsive part of the equilibrium, as shown in Part 3. The non-responsive part is flat, and therefore continuous. So the equilibrium is continuous if the action does not "jump" at the point x^*. (3.11) capture this condition and determines with (3.10) the point (x^*, m^*).

Lastly I comment on the other continuous equilibria of the model of Section 3.3, and show that they are less responsive than the one described in Proposition 7.

It is immediate to see that there can be an equilibrium where $a = E[x] = 0$ and m solves $(1-q)\, m\phi(m) + q \int xv(m \mid x)\, \phi(x)\, dx = 0\ \forall m$. Because in this equilibrium the action does not vary with the message, it is clearly less responsive that the one identified in Proposition 7.

Note next that an equilibrium where the action function is increasing in two intervals $[m_1, m_2]$ and $[m_3, m_4]$ but flat in the interval $[m_2, m_3]$, where $m_1 < m_2 < m_3 < m_4$, does not exist. The reason is the following. If the action function is flat between m_2 and m_3, then $a(m_2) = a(m_3)$. If the action function is increasing in the intervals $[m_1, m_2]$ and $[m_3, m_4]$, then from (3.8) it must be that the biased analyst is reaching his bliss point both when he announces m_2 and when he announces m_3. This is only possible if either $a(m_2) < a(m_3)$ which is an obvious contradiction, or m_2 and m_3 are both sent when the biased analyst observes the same x. But from Lemma 5 that the implicit function of m on x determined by $F(m, x) = 0$ is uniquely determined, which creates a contradiction.

The only possibility left is now for an equilibrium with an increasing section in the middle of the message space and two flat sections on the lower and upper extremes. Lastly, note that an equilibrium is possible where:

1. The action function is strictly increasing in the interval $[m^{**}, m^{*}]$ and it is characterised by Part 3 of Proposition 7.

2. The action is flat to the right of m^{*} with a^{*} determined by (3.11) and (3.12) holding for any $m > m^{*}$.

3. The action is flat to the left of m^{**}, with a^{**} characterised by:

$$a^{**} = \frac{q\phi\left(x^{**}\right) + (1-q)\phi\left(m^{**}\right)}{q\Phi\left(x^{**}\right) + (1-q)\Phi\left(m^{**}\right)}$$

and the condition below holding for any $m < m^{**}$:

$$\frac{(1-q)\, m\phi\left(m\right) + q\int_{-\infty}^{m^{**}} xv\left(m \mid x\right)\phi\left(x\right)dx}{(1-q)\,\phi\left(m\right) + q\int_{-\infty}^{m^{**}} v\left(m \mid x\right)\phi\left(x\right)dx} = a^{**}$$

4. An analyst that observes x^{**} is indifferent between the responsive and the non-responsive part of the equilibrium:

$$a^{**} = \frac{q\phi(x^{**})x^{**} + (1-q)\phi(m^{**})m^{**}}{q\phi(x^{**}) + (1-q)\phi(m^{**})}$$

and, similarly, (3.13) holds for x^{*}.

Figure F shows the shape of this equilibrium. It is immediate to realise that it is identical to the most responsive equilibrium except for an extra flat part to the left of the message m^{**}. The conclusion that this is a less responsive equilibrium is therefore obvious. ■

3.11 References

Abarbanell, J., and R. Lehavy, 2002 "Biased Forecasts or Biased Earnings? The Role of Reported Earnings in Explaining Apparent Bias and Over/Underreaction in Analysts' Earnings Forecasts" Unpublished Manuscript.

Abarbanell, J., and R. Lehavy, 2000 "Differences in Commercial Database Reported Earnings: Implications for Inferences Concerning Analyst Forecast Rationality, the Association between Prices and Earnings, and Firm Reporting Discretion", Unpublished Manuscript.

Austen-Smith, D. 1990 "Information Transmission in Debate", *American Journal of Political Science*, 34:124-152.

Avery, C. and M. Meyer 2003 "Designing Hiring and Promotion Procedures when Evaluators are Biased" Unpublished Manuscript.

Banerjee, A., 1992 "A Simple Model of Herd Behaviour," *Quarterly Journal of Economics*, 107: 797-817.

Badrinath, S.G. and S. Wahal, 2002 "Momentum Trading by Institutions" *Journal of Finance*, 57: 2449-2478.

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

Bikhchandani, S., D. Hirshleifer and I. Welch, 1992 "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economiy*, 100:992-1026.

Boni, L. and K. Womack, 2002 "Wall-Street's Credibility Problem: Misaligned Incentives and Dubious Fixes?" Wharton Financial Institutions Center Working Paper.

Boni, L. and K. Womack, 2002 "Solving the Sell-Side Research Problem: Insights from Buy-Side Professionals" Unpublished Manuscript.

Brown, L., 1998 "Managerial Behaviour and the Bias in Analysts' Earnings Forecasts", Unpublished Manuscript.

Chen, Q. and W. Jiang, 2003 "Analysts' Weighting of Private and Public Information", Unpublished Manuscript.

Crawford, V., 2002 "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions", Unpublished Manuscript.

Crawford, V. and J. Sobel, 1982 "Strategic Information Transmission", *Econometrica,* 50(6):1431-1452.

Dechow, P., A. Hutton and R. Sloan, 1998 "The Relation between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings" Unpublished Manuscript.

De Garidel-Thoron, T., and M. Ottaviani, 2000 "The Economics of Advice", Unpublished Manuscript.

Fisher, A. and R. Heinkel, 2003 "Reputation and Managerial Truth-Telling as Self-Insurance", Unpublished Manuscript.

Dessein, W., 2002 "Authority and Communication in Organizations", *Review of Economic Studies,* 69.

Francis, J. and D. Philbrick, 1993 "Analysts' Decisions as Products of a Multi-Task Environment" *Journal of Accounting Research,* 31:216-230.

Francis, J. and L. Soffer, 1997 "The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecasts Revisions" *Journal of Accounting Research,* 35:193-211

Francis, J. and R. Willis, 2001 "An Alternative Test for Self-Selection in Analysts' Forecasts" Unpublished Manuscript.

Gilligan, T. and K. Krehbiel, 1989 "Asymmetric Information and Legislative Rules with a Heterogeneous Committee", *American Journal of Political Science,* 33:459-490.

Gilligan, T. and K. Krehbiel, 1989 "Organization of Informative Committees by a Rational Legislature", *American Journal of Political Science,* 34:531-564.

Gu, Z. and J. Wu, 2001 "Earnings Skewness and Analysts Forecasts Bias", Unpublished Manuscript.

Henriques, D. 2000 "News in the Age of Money", *Columbia Journalism Review,* November/December Issue

Jackson, A., 2003 "Trade Generation, Reputation and Sell-Side Analysts", Unpublished Manuscript

Khaneman, D. and D. Lovallo, 1993 "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking", *Management Science,* 39: 17-31.

Keane, M. and D. Runkle, 1998 "Are Financial Analysts' Forecasts of Corporate Profits Rational?", *Journal of Political Economy,* 106: 768-805.

Krishna, V. and J. Morgan, 2001 "Asymmetric Information and Legislative Rules: Some Amendments", *American Political Science Review,* 95:435-452.

Krigman, L., W. Shaw and K. Womack, 2001 "Why do firms switch underwriters?," *Journal of Financial Economics,* 60: 245-284

Lim, T., 2001. "Rationality and Analysts' Forecast Bias," *Journal of Finance,* 56: 369-385.

Lin, H. and M. McNichols, 1998 "Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations", *Journal of Accounting and Economics,* 25:101-127.

Malmedier, U. and D. Shanthikumar, 2003 "Are Investors Naive about Incentives?". Unpublished Manuscript

McNichols, M. and P. O'Brien, 1997 "Self-Selection and Analyst Coverage" *Journal of Accounting Research,* 35:167-199

Michaely, R. and K. Womack, 1999. "Conflict of interest and the credibility of underwriter analyst recommendations," *Review of Financial Studies,* 12: 653-686.

Morgan, J. and P. Stocken, 2003 "An Analysis of Stock Recommendations", *Rand Journal of Economics,* 34: 183-203.

Prendergast, C. and L. Stole, 1996 "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy,* 104: 1105-1034.

Ottaviani, M. and P. Sorensen, 2002, "The Strategy of Professional Forecasting", Unpublished Manuscript.

Ottaviani, M. and F. Squintani, 2002 "Non-Fully Strategic Information Transmission", Unpublished Manuscript.

Scharfstein, D. and J. Stein, 1990 "Herd Behaviour and Investment," *American Economic Review,* 80: 465-479.

Sobel, J., 1985 "A Theory of Credibility," *Review of Economic Studies,* 52: 557-573.

The Economist, 11th of April, 2003.

Unger, L. (US Securities and Exchange Commission Acting Chairman), July 31 2001 "Testimony to Committee on Financial Services".

Welch, I., 2000 "Herding among Security Analysts", *Journal of Financial Economics,* 58: 369-396.

Zitzewitz, E., 2001a "Measuring Herding and Exaggeration by Equity Analysts and Other Opinion Sellers," Unpublished Manuscript.

Zitzewitz, E. 2001b "Opinion-producing Agents: Career Concerns and Exaggeration," Unpublished Manuscript.