Financial Analysts and Collective Reputation: Theory and Evidence*

Stefano Bonini[†], Filippo Pavesi[‡] and Massimo Scotti[§]

This draft: January 2011

Abstract

This paper studies the impact of collective reputation on the reporting strategy of experts that face conflicts of interest. The framework we propose applies to different settings involving decision makers that rely on experts for making informed decisions, in particular we consider sell-side financial analysts. We find that collective reputation has a non-monotonic effect on the degree of information revelation. In general, truthful revelation is more likely to occur when there is more uncertainty on the average ability of analysts as a group. In particular, above a certain threshold, an increase in collective reputation always makes truthful revelation more difficult to achieve. We test this theory by creating an index on the market's perception of the general reliability of analysts' recommendations. The empirical analysis provides evidence that collective reputation plays a role in determining the behaviour of analysts independently of their individual reputation.

*The authors acknowledge financial support from SDA Bocconi University. We thank Pierpaolo Battigalli, Paolo Colla, Amil Dasgupta, Giovanna Iannantuoni, Andrea Pratt, Jean-Charles Rochet, and Guido Tabellini. We also thank seminar participants at the SERC 2009 in Singapore and the Asset 2009 Conference in Istanbul for helpful comments. We thank Hakon Sandset for excellent research assistance. The ideas expressed in this paper are those of the authors and do not necessarily reflect the position of the authors' respective institutions. Any errors remain our own.

†Corresponding author. Department of Finance, Bocconi University, Via Roentgen 1, 20133, Milan, Italy. Email: stefano.bonini@unibocconi.it

‡Department of Economics, University of Milan-Bicocca, Piazza dell'Ateneo 1, 20126, Milan, Italy and Department of Public Policy, Central European University (CEU), Nador Utca 11, Budapest, Hungary. Email: filippo.pavesi@unimib.it

§School of Finance and Economics, University of Technology, Sydney, PO Box 123 Broadway NSW 2007 Australia. Email: massimo.scotti@uts.edu.au

Introduction

An extensive body of research has shown that equity analysts significantly impact shares market prices. However a similarly large evidence has also shed light on the extensive conflict of interests of sell-side analysts. In 2002 the SOX act among other issues, explicitly addressed sell-side analysts conflicts by imposing strict independence rules to safeguard investors from biased, overstated or misleading recommendations. These provisions though have proven limitedly effective as there is ample evidence that sell-side reports are still largely biased (Walker 2010; Hong and Kacperczyck, 2010; Goedhart et al. 2010). In this paper we try to provide a new view on this problem arguing that analysts incentives in providing accurate information is influenced not only by their individual reputational concerns but also by the prevailing market view of the industry as a whole, i.e. the collective reputation of analysts perceived by the market. In particular we try to capture the collective reputation effect on experts' incentives to accurately reveal information by developing a theoretical model and testing it through an innovative index measuring the reputation of the equity analyst industry.

Our results indicate that the market view of analysts' reputation does indeed influence the equity research accuracy, conditional and unconditional on the analysts individual reputation.

Equity analysts are expert information providers to investors and individuals at large. Their information can take either the form of a direct recommendation to follow a specific course of action (e.g. "buy" recommendation) or the form of a forecast that individuals use to inform their decisions (e.g. a "target price"). In all cases, the value of the expert's information relies on at least two components. First, the presumed ability of the analyst to gather accurate information about an unobserved state of the world upon which the success of a specific action depends. Second, the presumption that the expert truthfully reports his information. Independently from his/her individual ability, the value that decision-makers attribute to such information may depend on a degree of collective reputation, which is an assessment of the overall performance of analysts as a group. Collective reputation plays a role whenever individual ability is unknown and past individual behaviour cannot be perfectly observed. In addition past behaviour of the group to which the individual expert belongs, conditions the group's present behavior and can therefore be used to predict how the individual expert will behave. This implies that the incentives of individuals that are recognized as being part of a group are influenced by the group's reputation. On the other hand experts often face incentives that are not fully compatible with truthful revelation of their information. In particular, there are situations where professional information providers have a clear bias in favor of reporting over-optimistically (or over-pessimistically) on some unknown state of the world upon which receivers must base their decisions.[1] In all these cases, they face a conflict of interest with the party that eventually uses the information. Reputation acquisition is typically regarded as a mechanism that can offset this bias and mitigate the negative effects associated with conflicts of interest.[2] However the idea that analyst behaviour may be influenced by their collective reputation rather than or in addition to their individual reputation has not been considered by the literature.

In this paper we model a reporting environment where an expert is concerned about the collective reputation of experts as a group, but at the same time receives some form of compensation whenever he manages to induce the receivers to believe that the world is in one specific state. We analyze the effect of these contrasting objectives on the expert's incentives to truthfully reveal his information.

The nature of the bias we consider is such that regardless of the initial beliefs of decision-makers, an expert always has an incentive to induce them to attribute greater probability to a particular state of the world. For example, even if public information regarding a particular state of the world (such as the future performance of a particular stock) happens to be pessimistic, we assume the expert always benefits from convincing those who rely on his advice that things are not as bad as they think. A feature of our model is that the bias is increasing both in the uncertainty on the state of the world, since the expert's recommendation has a greater impact on the receivers' beliefs when public information on the state is less precise, as well as on the level of collective reputation, since decision makers attribute more weight to the advice of experts that have established a group reputation.

We show that, despite the bias, reputation is still effective in reducing the incentives to misreport. The nature of the most informative equilibrium in our setting is qualitatively similar to that of a reputational cheap talk model a la Ottaviani and Sorensen (2006), where conflicts of interest are not present. As in Ottaviani and Sorensen (2006), reputational concerns fail to be an effective disciplining device only when public infor-

[1] There is a large body of literature showing evidence that affiliated analysts have an optimism bias resulting from their involvement in the investment banking activity of their brokerage house (Michaely and Womack (1999), Barber et al. (2006, 2007)).

[2] Stickel (1992), Mikhail, Walther, and Willis (1999), Hong and Kubik (2003), Fang and Yasuda (2009) all document that reputation has a disciplining effect on analyst behavior.

mation is characterized by little uncertainty. In these cases, experts disregard their private information and conform to public information fearing that any contrarian signal they receive is probably incorrect.

Our main theoretical result is that improvements in collective reputation may have negative effects on information revelation. We show that a variation in the share of experts with high quality information (i.e., a higher level of initial collective reputation) has a non-monotonic effect on the incentives to truthfully reveal information and therefore on the level of informational efficiency. In particular, an increase in this share leads to less misreporting as long as the initial fraction of better-informed experts is not too high. However, beyond a certain threshold any increase in initial collective reputation results in a decrease in informational efficiency. Intuitively, when initial reputation is high, experts have less scope for reputation acquisition and at the same time face greater incentives to be over-optimistic, since decision makers attribute more weight to the advice of well established experts.

We test the theoretical predictions by developing an innovative, simple measure analysts' collective reputation. Following Tetlock (2008), Tetlock et al. (2008) and Loughran and MacDonald (2010) we applied a "Bag-of-words" approach to a large number of business press articles to extract the market view on the equity analysis industry that we use as a proxy of collective reputation. Using analysts forecast accuracy to capture analysts bias, we show that the market sentiment does indeed impact the magnitude and sign of prediction errors. Controlling for individual reputation as measured by the quality of the bank analysts are working for, we show that the high quality individual reputation forecasters are more sensitive to changes in collective reputation whereas lower quality banks are limitedly or not affected by the overall market view. Furthermore, our results indicate that non top banks show a consistently higher overestimation in forecasts as measured by much higher predicted, future, returns.

The paper is structured as follows: Section 1 presents the theoretical model; Section 2.describes the reputation index methodology and the prediction error measures. Section 4 describes data collection; Section 5 presents the results; Section 6 concludes.

1 The Model

An expert is called upon to provide information to a pool of individuals who have to make a forecast about the state of world. The state of the world w is either high or low, i.e., $w \in \{h, l\}$, and all players hold the same prior belief θ that the state is h. At the beginning of the game, the expert observes a private and non-verifiable signal $s_i \in \{s_h, s_l\}$ about the true state, whose accuracy depends on the expert's ability t. We assume that the expert is either good or bad, i.e., $t \in \{g, b\}$, and that ability affects the accuracy of the signal as follows:

$$\Pr(s_h|t = g, w = h) = \Pr(s_l|t = g, w = l) = p, \; p \in (1/2, 1) \quad (1)$$

$$\Pr(s_h|t = b, w = h) = \Pr(s_l|t = b, w = l) = z, \; z \in (1/2, p] \quad (2)$$

Therefore, both types of experts can count on an informative (yet imperfect) signal, with the good type having a more accurate signal than a bad type. We assume that neither the expert nor the receivers know the expert's type, and all players hold the same prior belief α that the expert is good.[3] We interpret α as the prior collective reputation of the expert, in other words the ex-ante probability that a decsion maker faces of being matched with a good expert.

After observing the signal, the expert chooses a report that is publicly released in the form of a costless binary message $m_j \in \{m_h, m_l\}$. Receivers observe message m_j and revise their beliefs about the true state of the world. We denote with $\widehat{\theta}_{\alpha,m_j} \equiv \Pr(w = h|m_j)$, the receivers' posterior belief that the state of the world is h, given that message m_j was sent by an expert when prior reputation is α. As we will see, in an equilibrium where some information is transmitted, the higher the collective reputation of the expert, the more the receivers trust the message sent. The subscript α highlights this relationship.

At the end of the game, the true state of the world is revealed and together with the message of the expert is used by the receivers to revise their beliefs about the collective ability.[4] We denote with $\widehat{\alpha}_{w,m_j} \equiv \Pr(t = g|w, m_j)$, the receivers' posterior belief that the experts are good upon observing state w and message m_j. We interpret $\widehat{\alpha}_{w,m_j}$ as the new level of collective reputation acquired by the expert category at the end of the game.

[3] This assumption is without loss of generality as far as the key results of paper are concerned, and makes the analysis more tractable. Assuming that the expert knows his own type does not affect the nature of the results. All the results hold for $z = 1/2$, however we make use of informative signals of bad types of experts, $z \in (1/2, p]$ when analyzing variations in $(p - z)$ in section (xx).

[4] In fact, in our model the receivers perform the task of forecasting the state of the world and the the ability of experts as a whole. Notice that we do not explicitly model the payoff of the receivers. Instead, we follow the approach of Ottaviani and Sorensen (2006) and implicitly assume that receivers are rewarded for accurately forcasting both the state of the world and the ability of the expert.

To model the fact that a single expert is concerned about the collective reputation that experts as a group have established for being a valuable providers of information and the contemporaneous existence of conflicts of interest, we construct a psychological game where the payoff of the single expert depends positively on the receivers' posterior beliefs $\widehat{\theta}_{\alpha,m_j}$ and $\widehat{\alpha}_{w,m_j}$, as follows:

$$\pi(m_j) = k\widehat{\theta}_{\alpha,m_j} + (1-k)\widehat{\alpha}_{w,m_j},\ k \in [0,1] \tag{3}$$

The component $\widehat{\alpha}_{w,m_j}$ captures the reputational concerns of the expert.[5] The component $\widehat{\theta}_{\alpha,m_j}$ gives the expert an incentive to inflate the receivers' belief that the state is h, and thus creates a conflict of interest with the receivers, since the expert now has a bias in favor of information that increases the receivers' perception that the state is h.[6] Finally, the parameter $k \in [0,1]$ weighs these two components and can be seen as a measure of the severity of conflicts of interest. The structure and the parameters of the game (with the sole exception of the expert's signal) are common knowledge.[7]

Notice that interpreting h and l respectively as favorable and unfavorable states for the receivers, the model represents the over-optimism bias that has been discussed both in the finance literature on sell side analysts and in the political science literature on government agencies' forecasts.[8] For the sake of exposition, in the remainder of the paper we will adopt this interpretation and refer to the expert's bias as to the over-optimism bias.

1.1 Equilibrium Analysis

In this section, we analyze the incentives of an expert to truthfully report his information and characterize the most informative equilibrium.[9]

[5] This reduced form to account for reputational concerns is widely adopted in studies that model the reputation of experts and managers (see for example Sharfstein and Stein (1990), Ottaviani and Sorensen (2006) and Gentzkow and Shapiro (2006)).

[6] See Battigalli and Dufwenberg (2009) for an analysis of extensive-form psychological games.

[7] It is worth noticing that since also k is common knowledge, we do not address the case when receivers are uncertain about the incentives of the expert (see Sobel (1985), Benabou and Laroque (1992), Morgan and Stocken (2003) for a formal analysis of the case when there is uncertainty about the expert's incentives).

[8] Assuming that the expert has an interest in inflating the receivers' belief about the state being h, is without loss of generality. Our setup is well suited for analyzing a more general setting, where the expert has an incentive to manipulate the receivers' beliefs in a desired direction.

[9] Our model presents the well-known problem of equilibrium multiplicity that is common to any cheap-talk game. A babbling equilibrium where all messages are taken to be meaningless and ignored always exists.

At the moment of sending message m_j, the true state of the world is unknown to the expert. The expert uses his signal s_i to compute the expected impact of message m_j on his reputation, as follows:

$$E\left(\widehat{\alpha}_{w,m_j}|s_i\right) = \Pr(w = h|s_i)\widehat{\alpha}_{h,m_j} + \Pr(w = l|s_i)\widehat{\alpha}_{l,m_j}$$

Therefore, the expected payoff of the expert from sending message m_j reads:

$$E\left(\pi(m_j)|s_i\right) = k\widehat{\theta}_{\alpha,m_j} + (1-k)E\left(\widehat{\alpha}_{w,m_j}|s_i\right)$$

Before analyzing the incentives of an expert to truthfully report his information, it is convenient to gain an intuition of the tensions involved in the reporting decision of the expert. In any equilibrium where some information is transmitted we have that $\widehat{\theta}_{\alpha,m_h} > \widehat{\theta}_{\alpha,m_l}$.[10] This introduces an incentive to report message m_h and represents a threat to truthtelling whenever signal s_l is received. In fact, the presence of reputational concerns counterbalances this over-optimism bias. As long as $k \in (0,1)$, the expert has to trade off the temptation of sending m_h with the negative effects that this message might have on collective reputation in case the message turns out to be incorrect.

The equilibrium concept we use is that of Perfect Bayesian Equilibrium (PBE). The expert will truthtfully report signal s_i if and only if the expected payoff of truthtelling is greater than the payoff of reporting a message that is different from the signal received. Thus, a truthtelling equilibrium exists if and only if for every $i, j \in \{h, l\}$, $E\left(\pi(m_i)|s_i\right) \geq E\left(\pi(m_i)|s_j\right)$, or equivalently:

$$k\widehat{\theta}_{\alpha,m_l} + (1-k)E\left(\widehat{\alpha}_{w,m_l}|s_l\right) \geq k\widehat{\theta}_{\alpha,m_h} + (1-k)E\left(\widehat{\alpha}_{w,m_h}|s_l\right) \quad (4)$$

$$k\widehat{\theta}_{\alpha,m_h} + (1-k)E\left(\widehat{\alpha}_{w,m_h}|s_h\right) \geq k\widehat{\theta}_{\alpha,m_l} + (1-k)E\left(\widehat{\alpha}_{w,m_l}|s_h\right) \quad (5)$$

In a truthtelling equilibrium, posterior collective reputation takes on only two possible values, which we denote with $\underline{\alpha}$ and $\overline{\alpha}$, where:

$$\underline{\alpha} \equiv \widehat{\alpha}_{l,m_h} = \widehat{\alpha}_{h,m_l}$$
$$\overline{\alpha} \equiv \widehat{\alpha}_{h,m_h} = \widehat{\alpha}_{l,m_l}$$

[10] Since the expert's signals are informative, in any equilibrium where signals are truthfully reported with some positive probability, the messages of the expert contain some information.

with $\overline{\alpha} > \alpha > \underline{\alpha}$.[11] Making a correct evaluation increases collective reputation from its initial level α to the higher level $\overline{\alpha}$. Making a wrong evaluation decreases reputation from α to the lower level $\underline{\alpha}$. In the rest of the paper we denote $(\overline{\alpha}-\underline{\alpha})$ as the reputational reward of providing a correct evaluation. This allows us to write conditions (4) and (5) in the following way:

$$k\left(\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}\right) \leq (1-k)(\overline{\alpha}-\underline{\alpha})\left(1 - 2\Pr\left(w = h|s_l\right)\right) \tag{6}$$

$$k\left(\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}\right) \geq (1-k)(\overline{\alpha}-\underline{\alpha})\left(1 - 2\Pr\left(w = h|s_h\right)\right) \tag{7}$$

For each of the above conditions, we refer to the left hand side as the benefit of providing a high message, and to the right hand side as the expected reputational gain of sending a low message. Notice that the right hand side of (6) represents the expected reputational gain of truthtelling when receiving a low signal, while the right hand side of (7) represents the expected reputational gain of misreporting when receiving a high signal.

We now establish that when experts have reputational concerns some information can be transmitted. The most informative equilibrium is reminiscent of Ottaviani and Sorensen (2001, 2006) as described in the following proposition:

Proposition 1 *For $k \in [0,1)$, the most informative equilibrium is separating (i.e., fully revealing) for $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ and pooling (i.e., uninformative) for $\theta \notin [\underline{\theta}, \underline{\theta}]$.*

(Proof: see Appendix)

For an intuition of Proposition 1, first notice that Lemma 1 implies that when θ is very low (high), receivers expect the economy to be in state l (h) regardless of the message sent by the expert. As a result, the net gain from inflating the beliefs of the receivers by sending m_h instead of a m_l, is very small and the choice of the expert is mainly driven by reputational concerns. However, reputational concerns make truthtelling impossible when the prior is relatively extreme. In these cases, the expert may believe that any contrarian signal he receives is probably incorrect. Being worried about the adverse impact of ex-post incorrect messages on reputation, he disregards his private information and reports the signal that is more likely to be correct ex-post. As the ex-ante probability that the true state is h increases, the expected reputational gain of reporting the low message decreases independently from the signal received. This

[11] We show this result in the Appendix.

conservative behavior on the part of the expert exists as long as the expert has some concerns about collective reputation (i.e., for $k < 1$).

On the other hand, Proposition 1 also highlights how truthful revelation occurs for interior values of θ. As illustrated in Lemma 1, in these cases conflicts of interest play a greater role with respect to the limit cases when θ approaches 0 or 1. Therefore, reputational concerns are still somewhat effective in inducing truthtelling behavior, even in the presence of conflicts of interest. Indeed, Proposition 1 suggests that the nature of the most informative equilibrium in the presence of over-optimism bias ($k \in (0, 1)$), is not qualitatively different from the case when conflicts of interest are absent and the expert is solely concerned about his reputation ($k = 0$).

1.2 Comparative Statics: Variation in Collective Reputation (α)

In this section, we examine how variations in collective reputation affect the most informative equilibrium of Proposition 1. What we are interested in is how changes α affect the truthtelling region $\left[\underline{\theta}, \overline{\theta}\right]$, as measured by the difference $\overline{\theta} - \underline{\theta}$. With a slight abuse of terminology, we refer to any increase (decrease) in $\overline{\theta} - \underline{\theta}$ as to an increase (decrease) in informational efficiency. To gain further insight into our findings, we carry out numerical analysis which we refer to in presenting the results.

The key finding is that significantly different results arise when conflicts of interest are present ($k \in (0, 1)$), as opposed to the case when conflicts of interest are absent ($k = 0$). For the sake of exposition, it is convenient to define some properties of the truthtelling equilibrium in the case when $k = 0$:

Remark 2 *Let $\underline{\theta}^*$ and $\overline{\theta}^*$ denote the threshold values for an expert with no conflicts of interest (i.e., $k = 0$). Then, $\underline{\theta}^* = 1 - [\alpha p + (1 - \alpha)z]$ and $\overline{\theta}^* = \alpha p + (1 - \alpha)z$.*

(Proof: see Appendix)

The previous remark suggests that in the absence of conflicts of interest, the truthtelling region is symmetrically centered around $\theta = \frac{1}{2}$, and expands as α, p and z increase. In particular, $\underline{\theta}^*$ ($\overline{\theta}^*$) is decreasing (increasing) in α, p and z.

We next analyze how variations in prior reputation affect informational efficiency. As a first step, we focus on the relationship between α and the different payoff components of the expert as described in the following remark:

Remark 3 *(i) The benefit of sending a high report,* $\left(\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}\right)$ *is increasing in initial collective reputation* α*; (ii) The reputational reward of being recognized as a good expert,* $\overline{\alpha} - \underline{\alpha}$ *is strictly concave in* α*, with* $(\overline{\alpha} - \underline{\alpha}) = 0$ *for* $\alpha = 0, 1$*.*

(Proofs: see Appendix)

The benefit of sending a high report increases with the level of collective reputation. Higher levels of prior reputation imply that there are greater chances that the expert facing the decision makers received a more accurate signal. Therefore, his message has a greater impact on the beliefs of decision makers. The way $(\overline{\alpha} - \underline{\alpha})$ changes in response to variations in the initial level of collective reputation instead reflects the common idea that individuals sluggishly change their mind in response to new evidence when they already hold a strong prior belief about something or somebody. On the contrary, new information typically leads to larger swings in beliefs when the level of uncertainty is high.

The previous remark suggests that above a certain level of α, the reputational reward of providing a correct evaluation, becomes negligible with respect to the benefit of sending a high report (indeed, the difference between these two components grows larger as α increases). As a result, above a threshold level of α the expert's bias in favor of the high message becomes stronger and actually increases with α. This makes both truthtelling thresholds $\overline{\theta}$ and $\underline{\theta}$ decrease with α, reflecting the idea that, as α grows larger, the expert has a stronger incentive to report a high message for any level of θ.[12]

Remark 2 bears a deeper consequence as far as the impact of reputation on informational efficiency in concerned. As α increases above a certain threshold, the difference between $(\overline{\alpha} - \underline{\alpha})$ and $\left(\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}\right)$ grows larger (with the former in fact progressively shrinking to zero), meaning that the reporting incentives of the expert are increasingly dominated by his interest to sway the beliefs of decision makers in favor of state h. As a result, for relatively large values of α, the benefit of sending the high message, irrespectively of the signal observed, dominates the expected reputational gain of making a correct evaluation, thus reducing informational efficiency. This effect clearly intensifies

[12] At $\theta = \underline{\theta}$ an expert that has received a *high signal* is indifferent between reporting a high message and reporting a low message. Ceteris paribus, an increase in α breaks this indifference in favour of the high message, which in fact implies that at $\theta = \underline{\theta}$ the expert is now truthfully reporting the high signal (i.e. the new truthtelling threshold, say $\underline{\theta}$', is lower than the initial one, $\underline{\theta}$). On the other hand, at $\theta = \overline{\theta}$ an expert that has received a *low signal* is indifferent between reporting a high message and reporting a low message. Again, ceteris paribus, an increase in α breaks this indifference in favour of the high message, implying that at $\theta = \overline{\theta}$ the expert is now pooling on the high signal (i.e. the new truthtelling threshold, say $\overline{\theta}$', is lower than the initial one, $\overline{\theta}$).

as α approaches to 1 (in this limit case, the truthtelling region becomes an empty set).

A similar reasoning applied to the case when initial reputation is below a certain threshold suggests that an increase in α leads to an expansion of the truthtelling region when α is indeed below a certain threshold. The following proposition summarizes the previous reasoning:

Proposition 4 *There always exist: (i) a level of initial reputation α above which an increase in α reduces informational efficiency (i.e., $(\overline{\theta} - \underline{\theta})$); (ii) a level of initial reputation α below which an increase in α increases informational efficiency (i.e., $(\overline{\theta} - \underline{\theta})$ increases).*

(Proof: see Appendix)

The result in Proposition 2 contrasts with the case of no conflicts of interest ($k = 0$), where an increase in reputation always translates into an improvement of informational efficiency.[13] Now, a further increase in prior reputation above a certain threshold (i.e., a reduction of uncertainty on expert ability) makes the truthtelling space shrink.[14]

Numerical analysis illustrates how both $\overline{\theta}$ and $\underline{\theta}$ are hump-shaped in α.

INSERT FIGURE 1 HERE

Furthermore, the threshold level of α above which an increase in prior reputation leads to a stronger bias towards h is a relatively intermediate value (i.e., close to 1/2). Thus this effect cannot be considered as a limit case that sets in only for extreme values of initial reputation. Prior reputation therefore has a non-monotonic effect on informational efficiency when conflicts of interest are present. Notice that for extreme values of α informational efficiency tends to zero. In other words, a very high level of reputation is as bad as a very low level of initial reputation as far as informational efficiency is concerned.

1.3 Empirical Test

Our model suggests that when the market for analysts is populated by a large share of well established analysts, less information will be contained in financial reports i.e.

[13] Notice from Remark (When $k = 0$ the truthtelling region monotonically expands from $2z-1$ (when $\alpha \to 0$) to $2p-1$ (when $\alpha \to 1$) and the greatest amount of information is transmitted when $\alpha \to 1$.

[14] This result bears some resemblance to Holmstrom (1999), that considers a dynamic setting with moral hazard.

forecast accuracy should be lower. Similarly, when analysts collective reputation is particularly negative due, possibly, to prior scandals or cross-sectional failures in providing reliable information, analysts should have less incentive in delivering efficiently information to the market through accurate forecast. Finally, in intermediate scenarios when there is uncertainty on the quality of the equity research industry quality, then analysts should have an incentive to increase information revelation through accurate forecasts.

A caveat applies to these predictions: as shown by Leone and Wu (2007), Fang and Yasuda (2009), analysts care about individual reputation and better quality analysts should have more incentives to accurately and fully reveal information even when the market collective view of the industry is low. Hence, we expect that individual reputation may mitigate the information reduction effect of a pro-tempore very high or very low collective reputation of the industry.

In the following sections we test these prediction by developing a previously unavailable index of collective reputation. We further develop a measure of individual reputation of the bank by ranking all research firms on two items: the number of published reports and the number of analysts working for that bank. As Hong and Kacperczyck (2010) show, top quality research institutions show robust evidence of being larger, more productive and populated by higher ranking professionals as measured by the Morninigstar ranking of individual analysts.

2 Index development

Testing the empirical predictions of our model requires a measure of collective reputation of the equity analysts industry. Unfortunately such a measure is not available through standard sources which provide time-series data only for standard quantitative measures such as number of analysts or the size of the Investment Banking Industry. Most of the reputation measures available such as Morningstar or Carter-Manaster are either individual rankings of analysts that can and have been used as proxies for individual reputation. or aggregate (and very limitedly time-varying) measures of banks' reputation. We address this issue by computing an innovative index developed according to an increasingly popular approach, i.e. content analysis of newspapers and magazines articles. A growing field in Finance is now devoting attention on how qualitative information in newspapers, annual reports, internet message boards and analyst reports is processed by market participants. Tetlock (2007) measures the interaction of the daily Wall Street Journal article "Abreast of the Market" and share returns. He focuses exclusively on the negative wording in the article and how this predicts market prices and trading volume. He concludes that a large negative sentiment in the newspaper article leads to decreasing share returns the following couple of days. Tetlock et al. (2008) extend this study by examining how the Wall Street Journal and the Dow Jones Newswires articles can be used to predict individual firms' accounting earnings and stock returns. They focus on all firms in the S&P 500 index. Their findings show that the linguistic media content captures otherwise hard to quantify aspects of firms fundamentals, which investors quickly incorporate in share prices.

While these two papers focus exclusively on negative wording in newspapers, other contributions have incorporated both positive and negative wording. Davis et al. (2006) examine whether managers' use of optimistic or pessimistic language in earning press releases delivers differential information to the market relating to the expected future firm performance by investigating the market response conditional on the "tone" of the language. They conclude that managers' use of the language is an important tool and that the market meaningfully responds to the selection of optimistic vs. pessimistic tone. Li (2006) creates an index to measure the risk sentiment of 10-K filings in the United States to examine the implications of the risk sentiment in these reports for future earnings and stock returns. He finds that firms with a large increase in the risk sentiment have more negative changes in earnings the following year. A common feature of these studies is that qualitative/linguistic information is gathered at the company level, thus providing relatively large amount of press items. Differently from previous studies we need to develop an industry-level measure of reputation, which requires carefully selecting press contents and collecting only industry-wide relevant articles.

2.1 Index construction methodology

2.1.1 Quantifying qualitative information

In order to quantify textual information, such as newspaper articles, we must devise a quantitative representation of the text. We achieve this goal by adopting the "Bag-of-Words" approach (Harris, 1954) which assumes that a document is a bag of individual and separate word, not taking grammar or order of the word into account. This means the phrase "home made" is, from this perspective, equivalent to "made home". Following this approach it is possible to develop a document terms matrix using either all words in the text or the count of pre-specified words. The challenge is then to translate this matrix into a meaningful conceptual representation of the story (Tetlock et al. 2008). In order to do this we collapse the document term matrix into three predetermined categories, financial positive words, financial negative words and the Harvard IV-4 negative psychological dictionary. By doing this we assume that all words within the three categories are equally informative and that all words not included in these three categories are uninformative for our case.

The words within each category, Financial Positive (finpos), Financial Negative (finneg) and Harvard IV-4 negative psychological dictionary (Harvard 4neg) are set by three recognized dictionaries for these categories . This removes the need for subjective judgement in setting what positive and negative words are. It also makes the study more replicable and transparent. Harvard IV-4 Negative psycho-social dictionary is widely recognized and used in text-analysis research . However, the Harvard IV-4 dictionary is developed mainly for psychology and sociology. Several commentators[15] have argued that this approach may result in an excessively pessimistic estimate of the true content of public news. Accordingly, we complement our analysis by using also the Loughran and McDonald (2010) dictionary (henceforth LM dictionary) that has been developed explicitly for business and finance applications. Therefore we develop two separate indices adopting alternatively the LM dictionary and the Harvard IV-4 dictionary.

2.1.2 Indexing Methodology

Consistent with previous literature, we create a simple measure of the market sentiment by computing a bag-of-words word count in all newspaper articles in the Financial Times (FT) and the Wall Street Journal (WSJ) in the time frame 1995 – 2009. These two sources are widely distributed and read by financial professionals and should represent the most representative source of the market, regulator and investors opinions on financial market topics and, a fortiori of equity analysts. Using the Harvard IV-4 and LM dictionaries, we process all articles with a WordStat 6.0 custom-designed code in order to extract the reputation-relevant information. This provides us with a document term matrix sorted by case number, and gives us the finneg, finpos and Harvard 4neg variables (i.e. number of words counted in each category for each case). In order to measure the sentiment of each case we model the following two approaches:

[15] See Loughran and MacDonald (2010) for a detailed discussion of this potential bias.

$$\text{Sentiment}=\frac{\text{No. Positive words} - \text{No. Negative words}}{\text{No. Positive words} + \text{No. Negative words}} \tag{8}$$

$$\text{SentimentAdjusted} = \frac{\text{Sentiment- } \mu\text{Sentiment}}{\sigma\text{Sentiment}} \tag{9}$$

The first index methodology (Sentiment) conveys intuitive information: when the index is greater than zero, the number of positive words in an article outweighs the number of negative ones, thus conveying a positive sentiment. In order to compare cases of different sizes we divide by the total number of positive and negative words. In the second approach we standardize the index by its long-term mean and standard deviation to allow for non-stationarity.

We collected articles as follows: firstly we perform a query in the Factiva and Lexis Nexis databases filtering all articles in the Wall Street Journal and The Financial Times which contained the keywords “Equity Analyst(s)”, “Sell-side analyst(s)”, “Research Department”, “Equity research” or “Analyst research”. The query returned 6,169 articles published during the time period 1 January 1995 to 31 December 2009.

However, a large number of hits returned opinions by equity analyst on certain shares rather than true opinions on the sentiment on equity analyst. We therefore performed an additional manual screening of all articles which left 48 relevant articles published in the Wall Street Journal and 231 relevant articles in the Financial Times, totalling 279 articles published in the time period. The distribution of articles between Wall Street Journal and the Financial Times is skewed, where the FT has published over 80% of the relevant articles. However, the Wall Street Journal is a newspaper focused mainly on neutral/factual news stories, while the Financial Times publish discussions, leaders and opinions. This skewness is therefore to be expected. We reckon that this manual sorting may introduce a potential bias. However, while some bias is unavoidable, since the sorting criterion was rather unequivocal we are confident that results are sufficiently robust.

Data reported in Table 1 show an average of 18.6 articles per year with a significant increase in relevant articles after 1998.

INSERT TABLE 1 HERE

From 1995 to 1998 we record a limited number of articles that suggests that the market sentiment towards analysts in that period was largely neutral. We see a peak in the number of articles in 2003 with 65 articles. The peak in 2002-2003 is probably due to the discussion on equity analysts after the IT-bubble and the SOX act.

Figure 2 reports a graphical representation of the sentiment index calculated according to the the LM dictionary and the Harvard IV-4 dictionary approach.

INSERT FIGURE 2 HERE

The two methodologies yield very similar results with a correlation of 0.91. Selecting quarterly observations instead of yearly observations, returns a similar pattern with a fractional increase in correlation to 0.92. This evidence suggests that the selection of any of the two methodologies should not affect results of the regression analysis.

Figure 1 shows a high level of volatility in the sentiment index between 1995 and 1998. As shown in Table 1, these three years record a low number of articles. The interpretation of this result is twofold: on the one hand a low number of articles may imply limited attention to the general issue of analysts quality and conflict of interest which is consistent with the prevailing (anecdotal) market view in the late '90s; on the other hand, a low number of articles is also consistent with a neutral view on the matter. Both dictionaries show a peak in 2000, which is also around the time of the internet bubble, followed by a sharp decline the following years possibly due to the discussion about the independence of the equity analysts. This negative sentiment continues until 2005 and 2006 where, again after a period of strong growth in the share market, the general sentiment turns again positive. The variability in the index the 2007 – 2009 might be explained by a lot of uncertainty in the market following the inception of the financial crisis. These figures indicate potential bias due to the extreme skewness in the distribution of the consituent of the index. To minimize as possible the effect on the empirical tests we opt for restricting the analyses in the following section to the window 01/01/1999 to 12/31/2007.

2.2 Equity analyst accuracy and the reputation sentiment index

We test our theoretical predictions by analyzing 110,564 analyst target prices issued by 1,306 analysts working for 296 different banks, on 3,048 companies listed on the NYSE from 01/01/1999 to 12/31/2007. We choose to focus on analyst target prices rather than earnings forecasts or qualitative recommendation because of the greater degree of variation of target prices and the much larger information content of target prices as opposed to the earnings estimate and recommendations.[16] In fact at any given time, a target price should be the analyst's best estimate of the expected future price of a stock. Target prices are not a measure of a company's fair value but are actually measures of the fair value subjectively adjusted by each analyst for exogenous factors such as market momentum, liquidity or industry factors. As such, we believe that target prices are the optimal instrument to test the information revelation by equity analysts In particular we measure the information content of target prices by computing their accuracy given by the degree of proximity of the share price to the target following the methodology developed in Bonini et al. (2010) This approach identifies two metrics δ_2 and δ_4 capturing respectively: the proximity of the share price to the target at any point in time over the prediction window of the target and the proximity of the share price to the target at the end of the forecast windows. Formally we compute:

$$\delta_2 = \left(\left(\frac{TP_t}{P_m} \right) - 1 \mid TP_t > P_t ; 1 - \left(\frac{TP_t}{P_m} \right) \mid TP_t < P_t \right) \tag{10}$$

$$\delta_4 = \left(\left(\frac{TP_t}{P_{t+n}} \right) - 1 \mid TP_t > P_t ; 1 - \left(\frac{TP_t}{Pt+n} \right) \mid TP_t < P_t \right) \tag{11}$$

where:

t: report issue date by analyst i on company j

P_t: stock market price at the research report publication date t

TP_t: target price given by analyst at the research report publication date t

P_m: maximum/minimum price level within the prediction time horizon.

$t+n$: date of the subsequent report issued by firm i on company j or the end of the prediction time horizon

As shown by Asquith et al. (2005) and Bonini et al. (2010) the δ_2 measure is theoretically sound but of almost no value to investors as it can be assessed only ex-post. Differently the δ_4 measure can be easily included in an investment strategy as it is anchored to a fixed, ex-ante known point in time. Figure 3 shows the joint distribution of δ_2 and δ_4 accuracy measures on both the indices. Index is the index computed through the LM dictionary and Indexharvard is the index computed using the Harvard

[16] See Walker et al. (2009), Bonini et al. (2010), Da and Schaumberg (2008).

IV-4 dictionary. The four panels show the market reputation on equity analysts and the equity research accuracy on a quarterly basis between quarter 1 (Q1) 1999 and quarter 4 (Q4) 2007.

INSERT FIGURE 3 HERE

As expected we obtain a weak correlation between the δ_2 measure and the two indices. Differently, when looking at the bottom panels in Figure 2 we observe a much stronger relationship. It seems that a positive market sentiment on analysts increases the δ_4 forecast error. These two graphs suggest that a high positive sentiment on equity analysts increases the forecast errors of equity analysts.

3 Empirical results

The descriptive analysys provides a first support to the theoretical predictiions. In this section we formally test the effects of collective and individual reputation on analysts accuracy. We begin by running the following univariate regression:

$$\delta_i = \alpha + \beta index^* + \varepsilon$$

Where i is the selected accuracy measure and $index*$ is either the index using financial dictionary (Index) and the H4neg dictionary (Indexharvard). Table 2 and 3 show the results for Index and Indexharvard, respectively. In all four regression models parameters are significant at the 1% level, indicating that the market sentiment does influence the accuracy of the analyst forecasts. As expected from figure 2 both indices affect δ_4 more than δ_2. The β for the indexes in δ_4 is also in line with our expectations suggesting that the better the sentiment (i.e. high and positive index) the higher is the ex-post forecast error. Similarly, as predicted by our model and shown in figure 1 a poor sentiment (i.e. low or negative index) actually decreases the forecasts error.

INSERT TABLE 2 HERE
INSERT TABLE 3 HERE

To further support the univariate analysis results we introduce three control variables. In particular:

- Implicit return (IR). This is the return of a share that is implicitly communicated by equity analysts when issuing a target price. It is defined as the target price when the report is issued over the market price of the share when the report is issued.

$$IR = \frac{TP_t}{P_t} - 1 \tag{12}$$

- A momentum measure (*marketindex*) calculated as the return of the S&P 500 market index over the prediction window
- A dummy variable if the research report is issued by a US top-ten bank (topbank).

We accordingly run the following multivariate regression:

$$\delta_i = \alpha + \beta_1 Index^* + \beta_2 IR + \beta_3 marketindex + \beta_4 topbank + \varepsilon \tag{13}$$

Table 4 and 5 report the results. All models are again highly significant with extremely high F-statistics and R^2. The estimated parameters for both indices are significant for the two accuracy metrics are Surprisingly, when adopting the δ_4 accuracy metric the R^2 drops to 0,2955, indicating that the model does not capture all elements of dependent variable. The regressions also confirm or results in regression (3) for both δ_2 and δ_4, even though the index variable is not significant for the δ_4 regression using the financial dictionary. We also see that the coefficients are lower, indicating that they have a low impact on the dependent variable. The positive IR coefficient in all regressions shows that there is a positive relationship with implicit returns and forecast error. The positive coefficient also indicates that generally the top 10 US banks will have higher forecast errors both during a report period (δ_2) and at the end date (δ_4), even though the coefficient is not statistically significant in the δ_2 regressions. We will look further into individual banks below.

INSERT TABLE 4 AND 5 HERE

We continue by sorting the index into four quartiles, where the first (lower) quartile is a very negative sentiment, while the fourth (higher) quartile provides a very good sentiment. We then run the following regression:

$$\delta_i = \alpha + \beta_1 qrtIndex * + \beta_2 IR + \beta_3 marketindex + \beta_4 topbank + \varepsilon \tag{14}$$

where qrtIndex* is the quartile of the index. It is worth noting that since the two indices are strongly correlated we can avoid to run separate regressions for each Index, hence we show results computed using the LM Index only. The results are reported in table 6

.

INSERT TABLES 6 HERE

The output of this regression confirms the previous results indicating that the level of collective reputation prevailing in the market does indeed influence analysts' behaviour and their forecast errors. However, the results so far are somewhat conflicting. They indicate that the sentiment index holds a slight inverse relationship with the δ_2 error while it holds a direct relationship with the δ_4 error. As we argued in the mode, this result can be interpreted as the possible joint effect of individual reputation in determining analysts behaviour.

3.1 Individual and Collective reputation

The regressions above show some puzzling differences between analyst reports issued by Top Banks as opposed to non Top banks and according to the choice of the accuracy metric. We try to further shed light on this evidence by separating the data into two groups, one including top ten US-banks and one with everybody else. We accordingly run the following regression:

$$\delta_i = \alpha + \beta_1 Index * + \beta_2 IR + \beta_3 marketindex + \varepsilon \tag{15}$$

Results are consistently more significant for top-banks than for non-top banks, in particular when looking at the δ_4 metric.

INSERT TABLES 7 AND 8 HERE

The index parameter is positive and significant for the δ_4 measure suggesting that when reputation is very high, top-bank analysts issue more overestimated forecasts tend to reveal less information as predicted by our theoretical model. As suggested by previous studies non-top analysts try to step up in the reputation scale by being significantly (but uninformatively) bolder. Their incentive for truthful information revelation is smaller as their forecast are option-like opportunities to move into the top-tier club. as such they are less affected by changes in the collective reputation and consistently overshoot their predictions.

Using the quartile index specification confirms the results and supports the view that individual and collective reputation play joint but somehow opposite roles in determining the level of analysts accuracy.

INSERT TABLE 9 HERE

4 Conclusions

Collective reputation plays an important role in shaping the incentives of experts that face conflicts of interest driven by an over-optimism bias. The main result of our model is that reputation has a non-monotonic effect on information transmission, and greater uncertainty on expert ability is associated with more information revelation. In other words, when a category of experts has established a reputation for providing valuable information, single experts may have strong incentives to release biased reports, much like when experts are generally seen as. It is precisely the uncertainty on ability, that creates greater incentives for individual experts to truthfully reveal their information, in order increase the collective reputation of their group. Once this standing has been attained, the over-optimism bias tends to prevail over the reputational losses that experts may incur, by erroneously forecasting a future state of the world.

These results suggest an empirical implication for the case of sell-side financial analysts. In a situation where collective reputation is particularly high, we should observe analyst target prices to exhibit greater prediction errors with respect to other market scenarios characterized by more uncertainty on collective reputation. Absent any established measure of collective reputation, we create a simple measure to capture the analysts reputation perceived by the market. Following Tetlock (2008), Tetlock et al. (2008) and Loughran and MacDonald (2010) we applied a "Bag-of-words" approach to a large number of business press articles to extract the market view on the equity analysis industry that we use as a proxy of collective reputation. We then applied this index to a large sample of over 110,000 equity target prices to examine the effect of fluctuations of the analysts collective reputation on the analysts forecasting errors. Our theoretical model predicts that the market sentiment should affect the prediction errors of equity analysts and that negative or positive sentiments should affect the accuracy of analysts in different ways.

Our results indicate that the market sentiment does indeed impact the magnitude and sign of prediction errors. Controlling for individual reputation as measured by the quality of the bank analysts are working for, we show that the high quality individual reputation forecasters are more sensitive to changes in collective reputation whereas lower quality banks are limitedly or not affected by the overall market view. Furthermore, our results indicate that non top banks show a consistently higher overestimation in target prices as measured by much higher implicit returns.

This paper provides a first look on how the market sentiment affects equity analysts. However, we reckon the possibility of significant improvements to our analysis. First, we expect to be able to further improve the quality of our indices by expanding the sources to other financial related news distributors such as financial magazines (e.g. The Economist and Business Week), financial sections of non-financial newspapers such as The New York Times or measuring the view of small investors through internet message boards. Secondly, the selection criteria can be further improved by creating more stringent sorting rules and/or document-term matrices, as in Tetlock et al. (2008). Finally, it would be interesting to explore additional feature of the effects of collective reputation on market participants such as industry, market or cross-country differences, the determinants of collective reputation and the impact that reputation at the analyst level may have on forecasting accuracy. We leave these issues for future research.

5 Appendix

Expert's Posterior Beliefs.

$$\begin{aligned}
\Pr(w &= h|s_h) = \frac{\theta(\alpha p + (1-\alpha)z)}{\theta(\alpha p + (1-\alpha)z) + (1-\theta)(\alpha(1-p) + (1-\alpha)(1-z))} \\
\Pr(w &= l|s_h) = 1 - \Pr(w = h|s_h) \\
\Pr(w &= h|s_l) = \frac{\theta(\alpha(1-p) + (1-\alpha)(1-z))}{\theta(\alpha(1-p) + (1-\alpha)(1-z)) + (1-\theta)(\alpha p + (1-\alpha)z)} \\
\Pr(w &= l|s_l) = 1 - \Pr(w = h|s_l)
\end{aligned}$$

■

Posterior Reputations under Truthtelling. In a truthtelling equilibrium the expert reports the signal he has observed. Therefore:

$$\widehat{\alpha}_{w,m_j} \equiv \Pr(t = g|w, m_j) = \begin{cases} \frac{\alpha p}{\alpha p + (1-\alpha)z} & \text{for } (w = h,\ j = h), (w = l,\ j = l) \\ \frac{\alpha(1-p)}{\alpha(1-p)+(1-\alpha)(1-z)} & \text{for } (w = h,\ j = l), (w = l,\ j = h) \end{cases}$$

Let $\overline{\alpha} \equiv \frac{\alpha p}{\alpha p+(1-\alpha)z}$ and $\underline{\alpha} \equiv \frac{\alpha(1-p)}{\alpha(1-p)+(1-\alpha)(1-z)}$. Then for $\alpha \in (0,1)$, $p \in \left(\frac{1}{2}, 1\right)$ and $z \in \left[\frac{1}{2}, p\right)$:

$$\overline{\alpha} - \underline{\alpha} = \frac{\alpha p}{\alpha p + (1-\alpha)z} - \frac{\alpha(1-p)}{\alpha(1-p) + (1-\alpha)(1-z)} = \frac{\alpha(1-\alpha)(p-z)}{(1-\alpha(p-z)-z)(\alpha(p-z)+z)} > 0$$

■

Lemma 5 *In a truthtelling equilibrium, the benefit of sending a high message,* $k\left(\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}\right)$ *satisfies the following properties: a) it is strictly positive for* $\theta \in (0,1)$ *and equal to zero for* $\theta = 0, 1$*; b) it is strictly concave in* θ *with a maximum at* $\theta = \frac{1}{2}$*.*

Proof of Lemma 1. Since $k \in [0,1]$, we can analyze $f(\theta) \equiv \widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}$. In a truthtelling equilibrium the expert reports the signal he has observed. Therefore:

$$\widehat{\theta}_{\alpha,m_j} \equiv \Pr(w = h|m_j) = \Pr(w = h \mid s_j) = \begin{cases} \frac{\theta(\alpha p+(1-\alpha)z)}{\theta(\alpha p+(1-\alpha)z)+(1-\theta)(\alpha(1-p)+(1-\alpha)(1-z))} & \text{for } j = h \\ \frac{\theta(\alpha(1-p)+(1-\alpha)(1-z))}{\theta(\alpha(1-p)+(1-\alpha)(1-z))+(1-\theta)(\alpha p+(1-\alpha)z)} & \text{for } j = l \end{cases}$$

With a bit of algebra we obtain:

$$
\begin{aligned}
f(\theta) &\equiv \widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l} = \\
&= \frac{\theta(-1+\theta)\left(-1+2\left(\alpha(p-z)+z\right)\right)}{\left(\theta\left(2\left(\alpha(p-z)+z\right)-1\right)-\left(\alpha(p-z)+z\right)\right)\left(1+\theta\left(2\left(\alpha(p-z)+z\right)-1\right)\right)-\left(\alpha(p-z)+z\right)}
\end{aligned}
$$

Let $q \equiv \alpha(p-z)+z$. Then, $f(\theta) = -\frac{\theta(1-\theta)(2q-1)}{(2q\theta-\theta-q)(1+2q\theta-\theta-q)}$. Notice that for $\alpha \in (0,1)$, $p \in \left(\frac{1}{2},1\right)$ and $z \in \left[\frac{1}{2},p\right)$, we have that $\frac{1}{2} < q < 1$. Then:

$$f(\theta) > 0 \text{ for } 0 < \theta < 1$$

$$f(\theta) = 0 \ \text{ for } \theta = 0,1$$

$$\frac{\partial f(\theta)}{\partial \theta} = -\frac{q(1-q)(2q-1)(2\theta-1)}{\left(2q\theta-\theta-q\right)^2\left(1+2q\theta-\theta-q\right)^2}\begin{cases} > 0 & \text{for } 0 < \theta < \frac{1}{2} \\ = 0 & \text{for } \theta = \frac{1}{2} \\ < 0 & \text{for } \frac{1}{2} < \theta < 1 \end{cases}$$

$$\frac{\partial^2 f(\theta)}{\partial \theta^2} = 2q(1-q)(2q-1)\left(\frac{1}{(2q\theta-\theta-q)^3} - \frac{1}{(1+2q\theta-\theta-q)^3}\right) < 0 \ \text{ for } 0 < \theta < 1$$

■

Lemma 6 *The expected reputational gain of sending the low message,* $(1-k)(\overline{\alpha}-\underline{\alpha})\left(1-2\Pr\left(w=h|s_i\right)\right)$ *satisfies the following properties: a) it is positive at* $\theta = 0$ *and negative at* $\theta = 1$ *for* $i = h, l$*; b) it is strictly decreasing in* θ *for* $i = h, l$*; it is strictly concave in* θ *for* $i = l$ *and strictly convex in* θ *for* $i = h$.

Proof of Lemma 2. Let $g(\theta) \equiv (1-k)(\overline{\alpha}-\underline{\alpha})1-2\Pr\left(w=h|s_l\right)$ and $v(\theta) \equiv (1-k)(\overline{\alpha}-\underline{\alpha})\left(1-2\Pr\left(w=h|s_h\right)\right)$. Using the values of $\overline{\alpha}$, $\underline{\alpha}$, $\Pr\left(w=h|s_l\right)$ and $\Pr\left(w=h|s_h\right)$ we obtain:

$$
\begin{aligned}
g(\theta) &= \frac{(1-k)(1-\alpha)\alpha(p-z)(-\theta+\alpha(p-z)+z)}{(-1+\alpha(p-z)+z)(\alpha(p-z)+z)(\alpha(-1+2\theta)(p-z)-z+\theta(-1+2z))} && \text{(RHS of (6))} \\
v(\theta) &= \frac{(1-k)\alpha(1-\alpha)(p-z)(-1+\theta+\alpha(p-z)+z)}{(-1+\alpha(p-z)+z)(\alpha(p-z)+z)(1+\alpha(-1+2\theta)(p-z)-z+\theta(-1+2z))} && \text{(RHS of (7))}
\end{aligned}
$$

Let $q \equiv \alpha(p-z)+z$. Then, $g(\theta) = \frac{\alpha(p-q)(\theta-q)}{q(1-q)(2q\theta-\theta-q)}$ and $v(\theta) = \frac{\alpha(p-q)(1-\theta-q)}{q(1-q)(2\theta q-\theta-q+1))}$. Notice that for $\alpha \in (0,1)$, $p \in \left(\frac{1}{2},1\right)$ and $z \in \left[\frac{1}{2},p\right)$, we have that $\frac{1}{2} < z < q < p < 1$. Then:

$$
\begin{aligned}
& g(\theta)\begin{cases} >0 & \text{for } 0<\theta<q \\ =0 & \text{for } \theta=q \\ <0 & \text{for } q<\theta<1 \end{cases} \\
& g(0) = \frac{\alpha(p-q)}{q(1-q)} > 0,\ g(1) = -\frac{\alpha(p-q)}{q(1-q)} < 0 \\
& \frac{\partial g(\theta)}{\partial \theta} = -\frac{2\alpha(p-q)}{(q+\theta-2q\theta)^2} < 0 \ \text{ for } 0<\theta<1 \\
& \frac{\partial g^2(\theta)}{\partial \theta^2} = -\frac{4\alpha(p-q)(2q-1)}{(q+\theta-2q\theta)} < 0 \ \text{ for } 0<\theta<1
\end{aligned}
$$

$$
\begin{aligned}
& v(\theta)\begin{cases} >0 & \text{for } 0<\theta<1-q \\ =0 & \text{for } \theta=1-q \\ <0 & \text{for } 1-q<\theta<1 \end{cases} \\
& v(0) = \frac{\alpha(p-q)}{q(1-q)} > 0,\ v(1) = -\frac{\alpha(p-q)}{q(1-q)} < 0 \\
& \frac{\partial v(\theta)}{\partial \theta} = -\frac{2\alpha(p-q)}{(-1+q+\theta-2q\theta)^2} < 0 \ \text{ for } 0<\theta<1 \\
& \frac{\partial v^2(\theta)}{\partial \theta^2} = \frac{4\alpha(p-q)(2q-1)}{(1-q-\theta+2q\theta)^3} > 0 \ \text{ for } 0<\theta<1
\end{aligned}
$$

$$
g(\theta)-v(\theta) = \frac{2\alpha(p-q)(2q-1)(1-\theta)\theta}{q(1-q)(1-q-\theta+2q\theta)(q+\theta-2q\theta))} > 0 \ \text{ for } 0<\theta<1
$$

■

Proof of Proposition 1. Consider the two conditions for truthtelling:

$$
\begin{aligned}
k[\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}] &\le (1-k)(\overline{\alpha}-\underline{\alpha})\left[1-2\Pr\left(w=h|s_l\right)\right] && \text{(A1)} \\
k[\widehat{\theta}_{\alpha,m_h} - \widehat{\theta}_{\alpha,m_l}] &\ge (1-k)(\overline{\alpha}-\underline{\alpha})\left[1-2\Pr\left(w=h|s_h\right)\right] && \text{(A2)}
\end{aligned}
$$

We first prove that for every value of $\alpha \in (0,1)$, $k \in [0,1)$, $p \in \left(\frac{1}{2},1\right)$ and $z \in \left[\frac{1}{2},p\right)$, there exist $\underline{\theta} \in [0,1]$ and $\overline{\theta} \in [0,1]$ such that for $\theta \in \left[\underline{\theta},\overline{\theta}\right]$ conditions (A1) and (A2) are satisfied simultaneously. Consider condition (A1) first. Using lemmas 1 and 2, we can write (A1) as follows:

$$-\frac{k\theta(1-\theta)(2q-1)}{(2q\theta-\theta-q)\,(1+2q\theta-\theta-q)} \leq \frac{(1-k)\alpha(p-q)(\theta-q)}{(1-q)q(2q\theta-\theta-q)}$$

Notice that $\frac{1}{2} \leq z < q < p < 1$. Thus, for $\theta \in (0,1)$, $2q\theta-\theta-q < 0$ and (A1) is equivalent to:

$$\frac{k\theta(1-\theta)(2q-1)}{1+2q\theta-\theta-q} \leq -\frac{(1-k)\alpha(p-q)(\theta-q)}{(1-q)q} \tag{A3}$$

Finally, let $h(\theta) = -\frac{k\theta(1-\theta)(2q-1)}{2q\theta-\theta-q}$ and $r(\theta) = \frac{(1-k)\alpha(p-q)(\theta-q)}{(1-q)q}$, and notice that:

a) $r(0) > h(0) = 0$, $r(1) < h(1) = 0$

b) $r(\theta)$ is a negatively sloped straight line.

c) $h(\theta)$ is non-negative, continuous, and strictly concave for $\theta \in (0,1)$.

Properties a), b) and c) imply that there exists a unique $\overline{\theta} \in (0,1)$ such that for any $\theta < \overline{\theta}$ (A3) (and therefore (A1)) are satisfied.

Focusing on condition (A2) and following the same steps above, we can prove the existence and uniqueness of a $\underline{\theta} \in (0,1)$ such that, for any $\theta > \underline{\theta}$, (A2) is satisfied. From lemma 2 we know that for $\theta \in (0,1)$ the RHS of condition (A1) is strictly greater than the RHS of condition (A2). This result, together with the uniqueness of $\underline{\theta}$ and $\overline{\theta}$ implies that $\overline{\theta} > \underline{\theta}$. Therefore, (A1) and (A2) are simultaneously satisfied for $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$.

Finally, notice that a babbling equilibrium where the expert sends m_h with probability π and m_l with probability $1-\pi$ irrespectively of the signal observed always exists. In this case all messages are taken to be meaningless and ignored: $\widehat{\theta}_{\alpha,m_j} = \theta$ for any $i = h, l$, and $\widehat{\alpha}_{w,m_j} = \alpha$ for any $w = h, l$ and $j = h, l$, making the expert indifferent between the two messages. ■

Corollary 7 *For condition (A1),* $\frac{\partial RHS}{\partial \theta}\Big|_{\theta=\overline{\theta}} > \frac{\partial LHS}{\partial \theta}\Big|_{\theta=\overline{\theta}}$. *For condition (A2),* $\frac{\partial RHS}{\partial \theta}\Big|_{\theta=\underline{\theta}} > \frac{\partial LHS}{\partial \theta}\Big|_{\theta=\underline{\theta}}$.

Proof of Corollary 1. The result in Corollary 1 is an immediate consequence of uniqueness of $\overline{\theta}$ and $\underline{\theta}$, together with the properties in lemma 1 and lemma 2. In words, the RHS of (A1) always intersects the LHS from above. The same is true for condition (A2). ■

Proof of Remark 1. When $k = 0$, condition (A3) boils down to $0 \leq \alpha(p-q)(\theta-q)$. The associated equation has solution $\theta = q = \alpha p + (1-\alpha)z \equiv \overline{\theta}^*$. The value of $\underline{\theta}^*$ is obtained in the same way from condition (A2) ■

Proof of Remark 2. Let $q = \alpha\,(p-z) + z$, where $z < q < p$. Notice that:

(i) $\frac{\partial\left(\widehat{\theta}_{\alpha,m_h}-\widehat{\theta}_{\alpha,m_l}\right)}{\partial\alpha} = \theta(1-\theta)\,(p-z)\left(\frac{1}{(q+\theta(1-2q))^2}+\frac{1}{(1-q-\theta(1-2q))^2}\right) > 0$ for any $\alpha \in (0,1)$.

(ii) $\frac{\partial(\overline{\alpha}-\underline{\alpha})}{\partial\alpha} = \frac{(p-z)\left(\alpha^2(p-1)p+(\alpha-1)^2z-(\alpha-1)^2z^2\right)}{(q-1)^2q^2}$; Notice that: $\frac{\partial(\overline{\alpha}-\underline{\alpha})}{\partial\alpha} = 0 \Leftrightarrow \alpha_0 = \frac{z-z^2-\sqrt{pz-p^2z-pz^2+p^2z^2}}{p^2-p+z-z^2}, \alpha_1 = \frac{z-z^2+\sqrt{pz-p^2z-pz^2+p^2z^2}}{p^2-p+z-z^2}$, where $\alpha_1 < 0 < \alpha_0 < 1$.

$\frac{\partial^2(\overline{\alpha}-\underline{\alpha})}{\partial\alpha^2} = 2\,(p-z)\left(-\frac{(1-p)(1-z)}{(1-q)^3}-\frac{pz}{q^3}\right) < 0$ for $\alpha \in (0,1)$. Therefore, for $\alpha \in (0,1)$, $\overline{\alpha}-\underline{\alpha}$ is strictly concave with a maximum at $\alpha = \alpha_0$. ■

Lemma 8 *(i) The benefit of sending a high report,* $\left(\widehat{\theta}_{\alpha,m_h}-\widehat{\theta}_{\alpha,m_l}\right)$ *is increasing in initial reputation* α*; (ii) The reputational reward of being recognized as a good expert,* $\overline{\alpha}-\underline{\alpha}$ *is strictly concave in* α*, with* $(\overline{\alpha}-\underline{\alpha}) = 0$ *for* $\alpha = 0, 1$.

Proof of Lemma 3. Consider condition (A1) and notice that: (i) For $\alpha \to 0$, $LHS_1 \to \frac{k\theta(2z-1)(1-\theta)}{(2z\theta-\theta-z)(2z\theta-\theta-z+1)}$ and $RHS_1 \to 0$; thus, for $\alpha \to 0$, $\overline{\theta} \to 0$; (ii) For $\alpha \to 1$, $LHS_1 \to \frac{k\theta(2p-1)(1-\theta)}{(2p\theta-\theta-p)(2p\theta-\theta-p+1)}$ and $RHS_1 \to 0$; thus, for $\alpha \to 1 : \overline{\theta} \to 0$.

Now notice that $\overline{\theta}$ is positive and continuous for $\alpha \in (0,1)$. This, together with (i), (ii) imply that : There exist an $\alpha' \in (0,1)$ such that for $\alpha \in (0,\alpha')$, $\frac{\partial\overline{\theta}}{\partial\alpha} > 0$; There exist an $\alpha'' \in (0,1)$ such that for $\alpha \in (\alpha',1)$, $\frac{\partial\overline{\theta}}{\partial\alpha} < 0$.

A similar argument applies to condition (A2) to show that: (iii) For $\alpha \to 0$, $\underline{\theta} \to 0$; (iv) For $\alpha \to 1$, $\underline{\theta} \to 0$. Again, continuity and the fact that $\underline{\theta}$ is positive for any $\alpha \in (0,1)$ imply that: There exist an $\alpha^+ \in (0,1)$ such that for $\alpha \in (0,\alpha^+)$, $\frac{\partial\underline{\theta}}{\partial\alpha} > 0$; There exist an $\alpha^{++} \in (0,1)$ such that for $\alpha \in (\alpha^+,1)$, $\frac{\partial\underline{\theta}}{\partial\alpha} < 0$. ■

Proof of Proposition 2. From the results in the proof of lemma 3 we have that: (i) For $\alpha \to 0$, $\overline{\theta}-\underline{\theta} \to 0$; (ii) For $\alpha \to 1$, $\overline{\theta}-\underline{\theta} \to 0$. Since $\overline{\theta}-\underline{\theta}$ is positive for any value of $\alpha \in (0,1)$, by continuity there exist a value of $\alpha \in (0,1)$ below which $\overline{\theta}-\underline{\theta}$ is increasing in α, and a value of $\alpha \in (0,1)$ above which $\overline{\theta}-\underline{\theta}$ is decreasing in α. ■

References

[1] Alesina, Alberto, and Nouriel Roubini [with Gerald D. Cohen] (1997). "Political Cycles and the Macroeconomy", Cambridge, MA: MIT Press.

[2] Banks, Jeffrey S., and Barry R. Weingast (1989). "The Political Control of Bureaucracies under Asymmetric Information", *American Journal of Political Science*, vol. 36, pp 509–24.

[3] Barber, B., Lehavy, R., McNichols, M., Trueman, B. (2006). "Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations", *Journal of Accounting and Economics*, vol. 41, pp. 87–117.

[4] Barber, B., Lehavy, R., Trueman, B. (2007). "Comparing the stock recommendation performance of investment banks and independent research firms", *Journal of Financial Economics*, vol. 85, pp. 490-517.

[5] Battigalli, Pierpaolo, and Martin Dufwenberg (2009). "Dynamic Psychological Games", *Journal of Economic Theory*, vol. 144, pp. 1-35.

[6] Benabou, Roland, and Guy Laroque (1992). "Using Privileged Information to Manipulate Markets: Insiders, Gurus and Credibility", *Quarterly Journal of Economics*, vol. 107, pp. 921-58.

[7] Bendor, Jonathan, Serge Taylor, and Roland Van Gaalen (1985). "Bureaucratic Expertise versus Legislative Authority: A Model of Deception and Monitoring in Budgeting", *American Political Science Review*, vol. 70, pp. 1041–1060.

[8] Bonini, S., Zanetti, L., Bianchini, R. & Salvi, A. (2010) Target Price Accuracy in Equity Research Journal of Business Finance & Accounting, June 2010

[9] Brandenburger, A., and B. Polak (1996). "When managers cover their posteriors: making the decisions the market wants to see", *RAND Journal of Economics*, vol. 27, pp. 523–541.

[10] Carlsen, Fredrik (1999). "Inflation and Elections: Theory and Evidence for Six OECD Economies", *Economic Inquiry* (January), pp. 119–35.

[11] Carpenter, Daniel P. (2001). "The Forging of Bureaucratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies", 1862–1928. Princeton, NJ: Princeton University Press.

[12] Crawford, Vincent P., and John Sobel (1982). "Strategic Information Transmission", *Econometrica*, vol. 50, pp. 1431-51.

[13] Dasgupta, Amil, and Andrea Prat (2008). "Information Aggregation in Financial Markets with Career Concerns", *Journal of Economic Theory*, vol. 143, pp. 83-113.

[14] Davis A., Piger J., Sedor L. (2006) Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases, Working paper, Federal Reserve Bank of St. Louis, available at "http://research.stlouisfed.org/wp/2006/2006-005.pdf"

[15] Fang, Lily, and Ayako Yasuda (2009). "The Effectiveness of Reputation as a Disciplinary Mechanism in Sell-side Research", forthcoming in *Review of Financial Studies*.

[16] Gentzkow, Matthew, and Jesse M. Shapiro (2006). "Media Bias and Reputation", *Journal of Political Economy*, vol. 114, pp. 280-316.

[17] Grossman, Sanford, and Joseph Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets", *American Economic Review*, vol. 70, pp. 393-408.

[18] Harris, Zellig (1954). "Distributional Structure". Word 10 (2/3): 146–62

[19] Heclo, Hugh (1975). "OMB and the Presidency - The Problem of 'Neutral Competence'", *The Public Interest*, vol. 38, pp. 80–99.

[20] Holmström, B. (1999). "Managerial Incentive Problems: A Dynamic Perspective", *Review of Economic Studies*, vol. 66, pp. 169–182.

[21] Hong H. and Kacperczyck M., (2010) "Competition and Bias", *Quarterly Journal of Economics*, Fortchoming

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

[23] Kreps, David M., and Robert Wilson, (1982). "Reputation and Imperfect Information", *Journal of Economic Theory*, vol. 27, pp. 253-79.

[24] Leone, A. and Wu, J. S., (2007), "What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts," Working Paper, Simon School of Business Rochester University

[25] Lizzeri, Alessandro (1999)."Information Revelation and Certification Intermediaries", *Rand Journal of Economics*, vol. 30, pp. 214-231.

[26] Loughran T., and McDonald B., (2010) When is a liability not a liability? Textual analysis, dictionaries and 10-Ks, Journal of Finance Forthcoming

[27] Goedhart M., Raj R., and Saxena A., (2010), "Equity Analysts: still to bullish", McKinsey Quarterly, April 2010

[28] Michaely, Roni, and Kent L. Womak (1999). "Conflict of Interest and the Credibility of Underwriter Analyst Recommendations," *Review of Financial Studies*, vol. 12, pp. 653-686.

[29] Mikhail, M.B., Walther, B.R., Willis, R.H. (1999). "Does forecast accuracy matter to security analysts?", *Accounting Review*, vol. 74, pp. 185–200.

[30] Milgrom, Paul, and John Roberts (1982). "Predation, Reputation and Entry Deterrence", *Journal of Economic Theory*, vol. 27, pp. 280-312.

[31] Morgan, John, and Phillip C., Stocken (2003). "An Analysis of Stock Recommendations", *RAND Journal of Economics, vol.* 34, pp. 183-203.

[32] Morris, Stephen (2001). "Political Correctness", *Journal of Political Economy*, vol. 109, pp. 231-265.

[33] Ottaviani, Marco, and Peter N. Sorensen (2000). "Herd Behavior and Investment: Comment", *American Economic Review*, vol. 90, pp. 695-703.

[34] Ottaviani, Marco, and Peter N. Sorensen (2001). "Information aggregation in debate: who should speak first?", *Journal of Public Economics, vol.* 81, pp. 393-421.

[35] Ottaviani, Marco, and Peter N. Sorensen (2006). "Reputational Cheap Talk", *Rand Journal of Economics*, vol. 37, pp. 155-175.

[36] Rourke, Francis E. (1992). "Responsiveness and Neutral Competence in American Bureaucracy",.*Public Administration Review*, vol. 52, pp. 539–46.

[37] Scharfstein, David S., and Jeremy C. Stein (1990. "Herd Behavior and Investment", *American Economic Review*, vol. 80, pp. 465-479.

[38] Sobel, J. (1985). “A Theory of Credibility", *Review of Economic Studies*, vol. 52, pp. 557–573.

[39] Stickel, S.E. (1992). "Reputation and Performance Among Security Analysts", *Journal of Finance*, vol. 47, pp. 1811–1836.

[40] Tetlock P. C. (2007) Giving Content to Investor Sentiment: The Role of Media in the Stock Market- Journal of Finance LXII (3), 1139 – 1168

[41] Tetlock P.C., Saar-Tscechansky M., & Macskassy S. (2008) More than words: Quantifying language to measure firms' fundamentals Journal of Finance LXIII (3), 1437 – 1467

[42] Trueman, Brett (1994). "Analyst Forecasts and Herding Behavior", *Review of Financial Studies*, vol. 7, pp. 97-124.

[43] Weatherford, M. Stephen (1987). “How Does Government Performance Influence Political Support?", *Political Behavior*, vol. 9, pp. 5–28.

[44] Welch, Ivo (1992). "Sequential Sales, Learning, and Cascades", *Journal of Finance*, vol. 47, pp. 695-732.

[45] Wilson, James Q. (1989). "Bureaucracy: What Government Agencies Do and Why They Do It", New York: Basic Books.

Table 1

Indices descriptive Statistics

This table reports summary statistics for the two indices. The first column reports the total number of collected articles; the second column reports the total number of words included in the collected articles; columns 3-5 report the breakdown of words in positive, negative and uncertain words from a financial perspective, the sixth column reports the number of negative words according to the Harvard IV 4 dictionary; columns 7 reports the estimated values of the Index computed according to the LM (2008) methodology, column 8 report the same index adjusted by standard deviation, column 9 and 10 report the estimated and standard deviation adjusted values of the index computed according to the harvard methodology.

Year	totalarticles	totalwords	finneg	finpos	finuncert	hneg4	LM Index	LM Index ADJ	Harvard Index	Harvard index ADJ
1995	2	4724	126	30	51	210	-0.71	-1.89	-0.75	-1.56
1996	3	3082	23	27	41	95	-0.41	1.24	-0.56	1.76
1997	2	2237	41	9	16	81	-0.73	-2.07	-0.80	-2.42
1998	2	1951	34	17	26	81	-0.56	-0.33	-0.65	0.11
1999	9	8456	180	82	83	358	-0.52	0.02	-0.63	0.55
2000	8	9534	158	96	70	373	-0.41	1.23	-0.59	1.18
2001	28	17105	376	154	146	716	-0.54	-0.18	-0.65	0.23
2002	54	34726	741	321	278	1635	-0.52	0.06	-0.67	-0.21
2003	65	40546	966	372	315	1926	-0.55	-0.24	-0.68	-0.29
2004	25	18274	389	147	159	785	-0.58	-0.52	-0.68	-0.43
2005	9	5728	83	47	49	209	-0.47	0.53	-0.63	0.46
2006	18	10985	171	112	100	518	-0.42	1.15	-0.64	0.26
2007	16	8401	202	85	85	411	-0.54	-0.17	-0.66	0.04
2008	18	12836	201	122	107	541	-0.43	0.97	-0.63	0.47
2009	20	15023	295	140	133	702	-0.51	0.20	-0.67	-0.14
Average	18.6	12907.2	265.733	117.4	110.6	576.07	-0.53	-1.3E-07	-0.6593882	-4E-07

Table 2

Univariate regression on the LM Index

The regression shows the effect that collective reputation has on equity analyst accuracy. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index computed by using the LM financial dictionary. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent variable	
	δ_2	δ_4
Intercept	0.1382	0.5918
	(0.0069)***	(0.0415)***
Index	-0.0353	0.1167
	(0.0085)***	(0.0444)***
Adj R^2	0.0001	0.0001
F-Statistic	17.10***	6.92***
Observations	110,354	110,354

Table 3

Univariate regression on Harvard Index

The regression shows the effect market that collective reputation has on equity analyst accuracy. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index computet by using the Harvard H4neg dictionary. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent variable	
	δ_2	δ_4
Intercept	0.1311	0.6575
	(0.0080)***	(0.0415)***
Indexharvard	-0.0493	0.2748
	(0.0128)***	(0.0666)***
Adj R2	0.0001	0.0001
F-Statistic	14.84***	17.01***
Observations	110,354	110,354

Table 4

Multivariate Control Regressions on the LM Index

The regression shows the effect market sentiment has on equity analyst accuracy. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index computed by using the LM financial dictionary. IR is the implicit return, marketindex is the market index and topbank a dummy variable showing if the report is issued by a US top 10 bank. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent Variable	
	δ_2	δ_4
Intercept	0.3122	-0.6163
	(0.0191)***	(0.1950)***
Index	-0.0385	0.046
	(0.0037)***	-0.0378
IR	0.7182	2.2497
	(0.0010)***	(0.0104)***
Marketindex	-0.0002	0.0069
	(0.0001)***	(0.0002)***
topbank	0.023	0.0282
	(0.0056)***	-0.057
Adj R2	0.8158	0.2955
F-Statistic	E***	11 570***
Observations	110,354	110,354

Table 5

Multivariate Control Regressions on the Harvard Index

The regression shows the effect that collective reputation has on equity analyst accuracy. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index computed by using the Harvard H4neg dictionary. IR is the implicit return, marketindex is the market index and topbank a dummy variable showing if the report is issued by a US top 10 bank. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and ***

	Dependent Variable	
	δ_2	δ_4
Intercept	0.2904	-0.5186
	(0.0197)***	(0.2009)***
IndexHarvard	-0.0589	0.1327
	(0.0056)***	(0.0574)**
IR	0.7182	2.2495
	(0.0012)***	(0.0104)***
Marketindex	-0.0002	0.0064
	(0.0001)***	(0.0002)***
topbank	0.0246	0.0254
	(0.0056)***	-0.057
Adj R2	0.8158	0.2955
F-Statistic	E***	11 570***
Observations	110,354	110,354

Table 6

Mukltivariate quartile regressions

The regression shows the effect market sentiment has on equity analyst accuracy. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index using the LM financial dictionary. IR is the implicit return, marketindex is the market index and topbank a dummy variable showing if the report is issued by a US top 10 bank. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent Variable	
	δ_2	δ_4
Intercept	0.3873	-0.7347
	(0.0189)***	(0.1926)***
Qrt LM Index	-0.0219	0.0479
	(0.0024)***	(0.025)*
IR	0.7182	2.2495
	(0.0010)***	(0.0104)***
Marketindex	-0.0002	0.0006
	(0.0000)***	(0.0002)***
topbank	0.0235	0.0281
	(0.0056)***	-0.057
Adj R2	0.8158	0.2955
F-Statistic	E***	11 571***
Observations	110,354	110,354

Table 7

Differences between the topbanks and nontopbanks - LM index

The regression shows the effect that collective reputation has on equity analyst accuracy. We separate it into the two categories Topbank and Non-topbank. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index using the LM financial dictionary. IR is the implicit return and marketindex is the market index. k. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent Variable			
	δ_2		δ_4	
	Topbank	Non-topbank	Topbank	Non-topbank
Intercept	0.283	0.3475	-0.1345	-0.8976
	(0.0239)***	(0.0273)***	-0.1149	(0.3111)***
Index	-0.0354	-0.0407	0.0634	0.0387
	(0.0047)***	(0.0052)***	(0.0228)***	-0.0598
IR	0.7344	0.7132	17,980	2.3881
	(0.0017)***	(0.0013)***	(0.0081)***	(0.0147)***
Marketindex	-0.0002	-0.0003	0.0003	0.0009
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0003)***
Adj R2	0.8171	0.8156	0.5366	0.2777
F-Statistic	63,213***	E***	16,388***	8,702***
Observations	42,458	67,869	42,458	67,869

Table 8

Differences between the topbanks and nontopbanks - Harvard Index

The regression shows the effect that collective reputation has on equity analyst accuracy. We separate it into the two categories Topbank and Non-topbank. δ_2 and δ_4 is the analyst accuracy and index is the sentiment index computed by using the Harvard H4neg dictionary. IR is the implicit return and marketindex is the market index. k. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent Variable			
	δ_2		δ_4	
	Topbank	Non-topbank	Topbank	Non-topbank
Intercept	0.2726	0.3176	-0.05177	-0.7918
	(0.0197)***	(0.2820)***	-0.1183	(0.3203)**
IndexHarvard	-0.0442	-0.0714	0.1318	0.1393
	(0.0067)***	(0.0840)***	(0.0323)***	-0.0954
IR	0.7344	0.7132	1.7979	2.3881
	(0.0017)***	(0.0013)***	(0.0081)***	(0.0147)***
Marketindex	-0.0002	-0.0002	0.0003	0.0008
	(0.0001)***	(0.0001)***	(0.0001)***	(0.0003)***
Adj R2	0.817	0.8156	0.5367	0.2773
F-Statistic	63,189***	E***	16,395***	8,702***
Observations	42,458	67,869	42,458	67,869

Table 9

Differences between the topbanks and nontopbanks using quartiles

This table reports regression results for the effect that collective reputation has on equity analyst accuracy, controlling for individual banks reputation. We divide our sample in two sub-groups: Topbank and Non-topbank where ToBank status is granted to banks ranking in the Top10 underwriting activity ranking published by The banker. δ_2 and δ_4 is the analyst accuracy, IR is the implicit return and marketindex is the market index. Robust standard errors are reported in parentheses. Significance at 10%, 5% and 1% level is denoted by by *, ** and *** respectively.

	Dependent Variable			
	δ_2		δ_4	
	Topbank	Non-topbank	Topbank	Non-topbank
Intercept	0.3506	0.4282	-0.2755	-10,127
	(0.2351)***	(0.0271)***	(0.1130)**	(0.3081)***
Qrt LM Index	-0.0197	-0.0235	0.0517	0.0495
	(0.0032)***	(0.0347)***	(0.0151)***	-0.0394
IR	0.7345	0.7133	1.7978	2.3881
	(0.0017)***	(0.0013)***	(0.0081)***	(0.0147)***
Marketindex	-0.0002	-0.0002	0.0003	0.0009
	(0.0000)***	(0.0000)****	(0.0000)***	(0.0003)***
Adj R2	0.817	0.8156	0.5366	0.2777
F-Statistic	63,182***	E***	16,391***	8,702***
Observations	42,458	67,896	42,458	67,896

Figure 1

Information efficiency and reputation

This figure illustrates the degree of information efficiency conditional on different realizations of the reputation parameter α

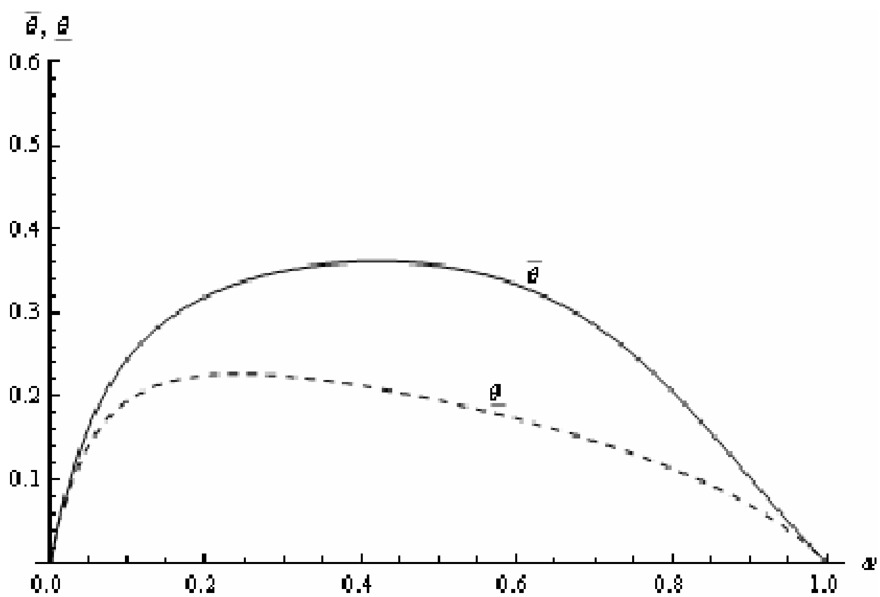

Figure 2

Sentiment indices

This figure plots the estimated sentiment Indices calculated according to the Findic and Harvard IV4 dictionary approach. Both indices have been estimated by scxaling the index for its standard deviation.

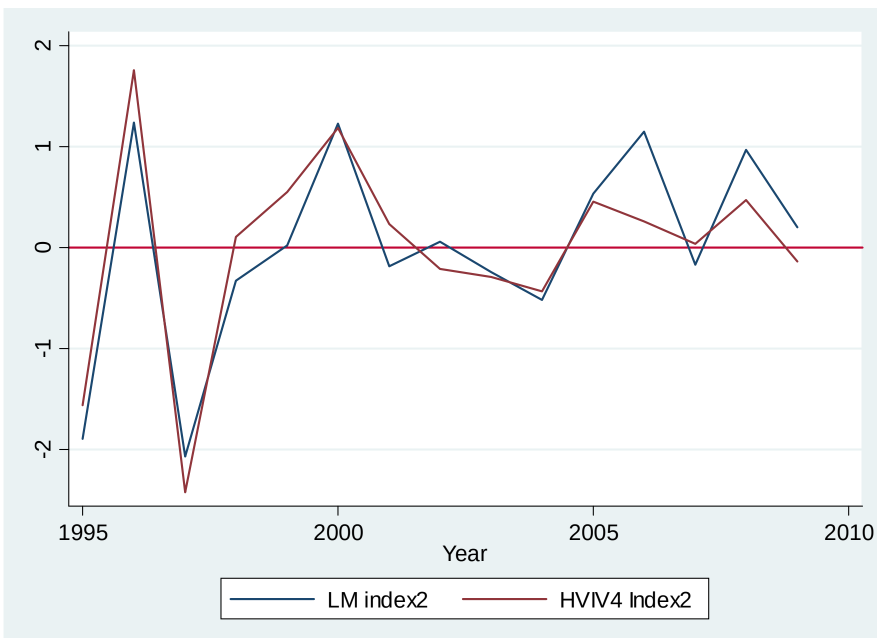

Figure 3

Correlation between indices and forecast accuracy

These figure plots the mean forecast error measured alternatively by the δ_2 or δ_4 measure against the two different Index metrics. IN particular the right side of the figure reports the graphical correlation between the two forecast accuracy metrics and the Index computed following LM (2008), while the left side of the figure plots the two different measures of accuracy against the index computed following the harvard IV4 methodology.