Universidad del Rosario

Essays On Equity Research and Informational Decisions

Ph.D. Dissertation

José Gabriel Astaiza-Gómez

Supervisors:

Professor Jorge Flórez Acosta, Colegio Mayor de Nuestra Señora del Rosario.

Professor Karoll Gómez Portilla, Universidad Nacional de Colombia.

Colegio Mayor de Nuestra Señora del Rosario

Facultad de Economía

Doctorado en Economía

June 21, 2021

Contents

Acknowledgments

I thank Karoll Gómez and Jorge Flórez for their thorough readings and commentaries on my reports and also Sébastian Pouget for his constructive discussions with me. Also, I thank professors Carlos Castro, Carlos Cañon, Çağatay Kayi, Guillem Roig, Mounu Premm, Santiago Sautúa and Hugo Ramírez, as well as researchers Catalina Franco, Wilber Baires, Andrés Cárdenas and Juan Carlos Zambrano for their insightful comments. The feedback received during the Industrial Organization Workshop at Universidad del Rosario as well as during the Finance Workshops at Toulouse School of Economics, CESA, Universidad Javeriana and Universidad del Rosario, have been very valuable.

The agility of the administrative stuff of la Facultad de Economía to provide me with information and their eagerness to help when I most needed it, made my work easier at Universidad del Rosario as well as outside. My sister María Alejandra and her husband Andrés, were very welcoming at their home in Bogotá, which I found very convenient. My parents, always patient to me, offered me their wise advice and their comfortable house in Popayán, which I took advantage of during the pandemic.

Chapter 1

Signal Strength, Conflicting Signals and Beliefs Updating: Evidence From Sell-Side Analysts' Forecasts

Abstract: I empirically study how confirmatory bias works for strong and contradictory signals using data on sell-side analysts. I first model an agent who is prone to confirmatory bias and whose task is to value a stock based on a signal, and introduce the effects of the signal strength by relaxing Rabin and Schrag's (1999) assumption of a constant bias severity. Afterwards I use target prices to measure forecast bias and the growth in Earnings Per Share as signals, and regress analysts' forecast bias over different deciles of favorable signals interacted with prior negative forecast bias in a dynamic panel data model. I find that analysts do not react positively to favorable signals when the prior is pessimistic, except for sufficiently strong signals which cause analysts to issue more optimistic target prices.

Key words: Bayesian updating, misinterpretation, pessimistic prior, optimistic forecast.

1.1 Introduction

People have the tendency to misinterpret ambiguous evidence in ways that confirm their existing beliefs and the literature shows that this confirmation bias is present in the beliefs updating process of experts of varying professions such as psychiatrists and financial analysts. In particular, the empirical evidence shows that sell-side financial analysts not only do they have (commercial) incentives to issue inflated forecasts, but also misinterpret firms' earnings when these are contrary to their prior beliefs (Pouget, Sauvagnat and Villeneuve, 2017). It is interesting that, while there is strong evidence on the existence of confirmatory bias in our beliefs updating process when signals are contradictory, there is a lack of research about the impact of the extremeness of these ambiguous signals, as the experimental literature shows that people overestimate the representativeness of small samples and have the tendency to focus on the strength of the signals (Tversky and Kahneman, 1971; Griffin and Tversky, 1992).

In this paper, I refine the understanding on beliefs updating with ambiguous signals by taking into account the strength of the signals. To be more concrete, I study whether sell-side analysts' forecast bias react to the signal strength when the signal is contradictory to prior beliefs. Based on the theoretical model of Barberis, Shleifer and Vishny (1998), I model the dynamics of beliefs updating, for an analyst with confirmatory bias, whose bias is decreasing with respect to the signal extremeness, and for a non biased investor. My model offers new testable predictions on how changes in forecast bias are affected by the extremeness of signals, received in the presence of confirmatory bias. The different beliefs about the future between the biased analyst and the investor, imply that the bias (i.e. the difference between the analyst's projected stock price and the realized stock price) is likely to move in the direction of the analyst's prior belief when the signal is contradictory, but is more likely to move in the direction of the signal as the signal extremeness increases. I use data on stock prices and price forecasts to measure analysts' bias, and the growth in Earnings Per Share as signals, and regress analysts' forecast bias over different deciles of past high signals interacted with prior negative forecast bias. I find negative reactions of forecast bias to favorable contradictory signals, except for signals above the eighth decile for which analysts react positively, and these results are robust to a change in my proxy for prior beliefs. Moreover, I also carry out a similar exercise using forecasts revisions instead of forecast bias and find that when analysts have a negative prior belief, they revise their forecasts upward for sufficiently strong high signals.

This paper adds to two strands of the literature. First, my paper adds to the research on sell-side analysts (SSAs) by showing that they issue biased forecasts because of their cognitive biases and not only because of the well known fact that they have commercial incentives which leads them to issue positively biased forecasts. Second, and adding to the behavioral literature, I enrich the understanding of confirmatory bias by incorporating the extremeness of the evidence in the processing of contradictory signals.

Confirmatory bias is a concept that has a precise meaning within the context of my model, and under certain scenarios it is consistent with Pouget, Sauvagnat and Villeneuve (2017) who are the first to bring the concept of confirmatory bias to the financial literature. It is defined in the broadest possible sense as the misinterpreting of a current signal, when it is different to the most probable signal according to previous beliefs[1]. As such, my estimates of confirmatory bias are not directly comparable to those obtained in earlier studies using different methods. Perhaps the most significant limitation of my approach, is that I cannot capture all possible scenarios of contradictory signals present in my model, in my econometric specification. Since analysts' beliefs are not observable, my empirical findings on forecast bias can be regarded as consistent with the testable hypotheses derived from my model, without claims of being concluding evidence on a unique explanation about the underlying process of analysts' beliefs updating.

There are eight sections (including the introduction) in this paper. In section two I expose the literature related to analysts' optimism and their reaction to earnings, as well as the behavioral literature. Later on, in section three, I present a theoretical model of beliefs updating in a dynamic framework where the intensity of the signal is important and there is confirmation bias. In sections four and five I explain the variables and the econometric specification respectively. In section six I present the results on forecast bias, in section seven I directly test for confirmation bias using forecasts revisions, and finally, in section eight I conclude.

1.2 Related Literature

My paper is directly related to the literature on behavioral finance and specifically on confirmation bias. Rabin and Schrag (1999) are the first to formalize this cognitive bias and Pouget, Sauvagnat and Villeneuve (2017) are the first to study how confirmatory bias affects asset pricing in financial markets. In particular, the latter propose a model in which perfectly rational agents are endowed with the true probability model, while biased agents are endowed with a different probability model in which they may ignore information that is inconsistent with their prior views. Using data on earnings and forecast revisions, they find that analysts are less likely to revise their forecasts upward (downward) after a positive (negative) signal, for signals that are contrary to analysts' prior beliefs. These pioneer studies provide the basis to formally analyze confirmatory bias, from which I build a model where the signal intensity is incorporated in the beliefs updating of the biased agent. Interestingly, Griffin and Tversky (1992) show

[1] An interesting different path is that of Sadler (2021). In my model, there is a fixed set of states and a contradictory signal is contradictory in reference to the agent's updated probabilities. The path of Sadler (2021), is to model how an agent decides on the support of the prior, without saying anything about what probability to assign to the different states. In that work, there is a characterization of various agents who face the problem of deciding which events to accept as "observed" when there is different information from various sources and direct verification is infeasible. In that setting, the idea that two claims or observations are contradictory, signifies that the intersection of two events (two collections of states) is empty, and it is not related to a probability over the states.

experimental evidence suggesting that people tend to focus on the strength or extremeness of signals and less on their predictive validity, which is different than ignoring or misinterpreting signals. Also, Tversky and Kahneman (1971) conclude that "people (even those trained in statistics) make radical inferences on the basis of small samples."

In order to obtain testable predictions, I use a Markov model where the valuation the analyst makes, depends on the expectations of an infinite series of earnings, taking elements from Barberis, Shleifer and Vishny (1998). The model of the authors, although motivated by the psychological evidence in Griffin and Tversky (1992), is not directly derived from a formal definition of cognitive biases, since unfortunately, the psychology literature does not establish quantitatively what kind of information is strong (or mild) and what kind of information is of high (or low) predictive validity. Nevertheless, by modeling biased agents, they are capable of generating series of stock valuations that show overreaction and underreaction.

I analyze forecast bias with a particular interest in the beliefs updating process with cognitive biases, separating signals into several levels of strength, but analyzing forecast bias is an old idea in the literature, most of all because of the interest in studying analysts' commercial incentives. SSAs are hired by brokerage firms whose revenue comes mainly from trading commissions and the analysts' income is linked to these commissions. When issuing forecasts, analysts do care about their reputation and thus about their accuracy (see Jackson, 2005; Mikhail, Walther and Willis, 1999; Groysberg, Healy and Maber, 2011). Still, since the income of analysts is linked to the revenues of their brokerage firms, then they try to increase the trading volume of the stocks they cover by issuing positively biased (optimistic) forecasts (Cowen, Groysberg and Healy, 2006; Jackson, 2005). Consistent with this, DeBondt and Thaler (1990) find evidence of optimism and overreaction in analysts' forecasts from a linear regression of current earnings changes over current forecast bias in earnings. However, as Abarbanell and Bernard (1992) point out about these results, the most overly optimistic consensus forecasts are for those companies with the weakest past performance[2]. Easterwood and Nutt (1999) build on Abarbanell and Bernard (1992) by incorporating dummies for quartiles of earnings changes, separating values of performance into low, normal, and high groups and run a linear regression of current forecast bias on past performance interacted with the dummies. Using consensus forecasts, they find that analysts overreact to prior signals in the upper quartile, and that analysts underreact to prior signals in the lower quartile. Overall, both results are consistent with systematic optimism.

The signal itself is a multidimensional object, from which I analyze its strength only. For instance, opposite information from different sources oblige us to take decisions on which signals to process and which to discard, before assigning probabilities or including them in our beliefs updating (Sadler, 2021).

[2]Similarly, Ali, Klein and Rosenfeld (1992) found that optimism is "most pronounced for firms that previously reported negative annual earnings."

Also, there are behaviors that appear in social or collective situations only, such as those explained by social transmission bias (Hirshleifer, 2020). In addition, there are differences in neural circuits across investors when processing information related to buying and selling (Vieito, Da Rocha, and Rocha, 2015; De Martino et al., 2013). My paper is mainly on confirmatory bias. This is an individual bias (as opposed to a socially induced bias) and is based on the idea that the individual may ignore a signal. This contrasts with the experimental evidence that shows that people have the tendency to focus on the strength of the signals (Tversky and Kahneman, 1971; Griffin and Tversky, 1992). Then the signal strength is directly related to confirmatory bias. As my paper uses observational data, unfortunately it does not speak on the actual neural processes (one of the dimensions) that cause beliefs updating. I follow the tradition of using earnings as signals, which is the most important piece of information for long-term (not intra-day) equity valuation. This variable does not have variability related to the credibility of the source (another dimension) so this research does not speak to how variation in credibility influences forecast bias.

1.3 A Model of Forecast Bias With Signal Intensity and Confirmatory Bias

There is a market with one stock that pays out 100% of the firm's earnings as dividends, and there are three agents: a representative, risk-neutral investor with discount rate δ, a representative financial analyst with discount rate δ that has the task of valuing the stock, and the state of nature that decides on firm earnings, $N_t = N_{t-1} + y_t$, where y_t is a shock in earnings. As there are not different investors with heterogeneous beliefs but only a representative investor that captures the consensus, the stock price at a time t is equivalent to the net present value of the investor's expectations on future earnings.

The Nature

There are two states of nature, namely, a mean-reverting world (state R) and a trending world (state T) which are described by the following Markov processes:

Table 1.1: Transition Matrices of Earnings Shocks

State R

	$y_{t+1} = +y$	$y_{t+1} = -y$
$y_t = +y$	π_L	$1 - \pi_L$
$y_t = -y$	$1 - \pi_L$	π_L

State T

	$y_{t+1} = +y$	$y_{t+1} = -y$
$y_t = +y$	π_H	$1 - \pi_H$
$y_t = -y$	$1 - \pi_H$	π_H

where $y_t \in \{-y, +y\}$, $-y \in \mathbb{R}_{<0}$, $+y \in \mathbb{R}_{>0}$. Under the mean-reverting state R, a negative (positive) shock will be followed by another negative (positive) shock with a low probability of $\pi_L < \pi_H$. In particular, I will assume $0 < \pi_L < \frac{1}{2} < \pi_H < 1$. Under state T of trending earnings, a negative (positive) earnings shock will be followed by another negative (positive) earnings shock with a high probability of π_H. The process controlling the switching from one state to another is also a Markov process as in the following matrix:

Table 1.2: Transition Matrix Between States R and T

	$State_{t+1} = R$	$State_{t+1} = T$
$State_t = R$	$1 - \lambda_1$	λ_1
$State_t = T$	λ_2	$1 - \lambda_2$

where $Sate_t \in \{R, T\}$, stands for the state of nature, λ_1 is the probability of going from state R to state T, and λ_2 is the probability of going from state T to state R. For ease of exposition, I will focus in the case where $\lambda_1 + \lambda_2 < 1$ which means that λ_1 and λ_2 are small in the sense that there is rarely a transitions from one state to another. Also, I will assume that $\lambda_1 < \lambda_2$. Since the unconditional (steady-state) probability of state R is $\mathbb{P}(R) = \dfrac{\lambda_2}{\lambda_1 + \lambda_2}$ (this calculation is shown in the appendix), the condition $\lambda_1 < \lambda_2$ implies that the investor thinks of State R as being, on average, more likely than State T.

The Investor

The investor is a Bayesian agent, i.e. he observes the evidence and updates his prior. In particular, I will denote p_t^R as the investor's beliefs of being in state R, which depends on his signal and past beliefs: $p_t^R = \mathbb{P}(State_t = R | y_t, y_{t-1}, p_{t-1}^R)$. As the stock price P_t equals the investor's estimate of the stock value,

$$P_t = \mathbb{E}_t \left\{ \frac{N_{t+1}}{1+\delta} + \frac{N_{t+2}}{(1+\delta)^2} + \dots \right\} \tag{1.1}$$

where, $\mathbb{E}_t$ is the expected value operator, the stock price at each t satisfies

$$P_t = \frac{N_t}{\delta} + y_t(\alpha_1 + \alpha_2 p_t^R) \tag{1.2}$$

where α_1 and α_2 are constant and depend only on of π_L, π_H, λ_1 , λ_2 and δ as described in the appendix. This is one o the main results in Barberis, Shleifer and Vishny (1998).

The Analyst

I am interested in modeling how forecasts deviate from realized stock prices. I assume that the analyst is small in the sense that his forecast does not affect the stock price. Although this is a strong assumption, it allows me to focus on the effects of cognitive biases on forecasts, in a clean manner without the complexity of a feedback process between the stock price and the forecast bias. In order to isolate my analysis from aspects related to risk preferences, I assume also that the analyst is risk neutral with respect to future earnings when estimating intrinsic values. Furthermore, the analyst uses the same Markov chains as the investor. There is confirmatory bias in the beliefs updating process of the analyst in the sense that he misinterprets, with a positive probability, a signal that is inconsistent with his prior beliefs. To be more concrete, after the realization of a signal y_t, the analyst sees a distorted signal σ_t, that is dependent on his distorted prior beliefs $\hat{p}^R_{t-1}$ in the following manner

$$\sigma_t = y_t \mathbb{1}_{((\mathbb{P}(R)-\hat{p}^R_{t-1})(\sigma_{t-1}*y_t)>0)} + z_t y_t \mathbb{1}_{((\mathbb{P}(R)-\hat{p}^R_{t-1})(\sigma_{t-1}*y_t)\leq 0)} \tag{1.3}$$

In the above equation, $\mathbb{1}_{(\cdot)}$ is an indicator function that takes the value of one whenever the condition in its subscript is met and $z_t \in \{-1, 1\}$ is a random variable that is independent across t and identically distributed. The fact that $z_t \in \{-1, 1\}$ indicates that the analyst misreads conflicting evidence as confirming evidence as in Rabin and Schrag (1999), which is alternative to the definition of confirmatory bias of Pouget, Sauvagnat and Villeneuve (2017), in which $z_t \in \{0, 1\}$, indicating that the biased agent tends to ignore evidence that conflicts with his beliefs. In addition, at each t, z_t is distributed $\mathbb{P}(z = -1) = \epsilon$ i.e. the severity of the bias is summarized by ϵ. The intuition is the following: if the analyst (mis)beliefs that signals are following a negative trend, i.e. the analyst underestimates the probability of being in the reverting world ($(\mathbb{P}(R)-\hat{p}^R_{t-1}) > 0$) and previously perceived a negative signal, $\sigma_{t-1} = -y$, then, a current positive signal $y_t = +y$ constitutes a positive contradictory signal. As a contradictory signal activates the right element of the right-hand side of equation 1.3, then he perceives $\sigma_t = -y$ with probability ϵ. Similarly, $y_t = -y$ preceded by $\sigma_{t-1} = +y$ constitutes a negative contradictory signal when prior beliefs are such that $(\mathbb{P}(R) - \hat{p}^R_{t-1}) > 0$. In a particular case, when $\epsilon = 1$ the analyst is completely biased. For instance, for a positive contradictory signal (i.e. $y_t = +y$, $\sigma_{t-1} = -y$, $(\mathbb{P}(R) - \hat{p}^R_{t-1}) > 0$), if $\epsilon = 1$, then the perceived signal $\sigma_t = (+y)(-1) = -y$ is negative with probability 1. Notice that I am not requiring $\mathbb{P}(R)$ to equal a specific value (e.g. a number close to $\frac{1}{2}$), as the psychology literature does not delve into the values consistent with confirmation bias.

There are two additional contradictory signals, both of them possible when the analyst believes that the world is mean-reverting: the condition $\sigma_{t-1} = -y$ and $y_t = -y$ (the analyst expected $+y$ with a high probability $(1 - \pi_L)$), and the condition $\sigma_{t-1} = +y$ and $y_t = +y$ (the analyst expected $-y$ with a high probability $(1 - \pi_L)$).

In order to introduce the effects of the signal strength, I relax Rabin and Schrag's (1999) assumption that $\epsilon \in [0,1]$ is constant and instead I assume it is a decreasing function of the observed magnitude of y. That is, the stronger the signal, the smaller the severity of confirmation bias[3]. The analyst updates his beliefs $\hat{p}_t^R$ according to his distorted perceptions, that is, he estimates $\hat{p}_t^R = \mathbb{P}(State_t = R|\sigma_t, \sigma_{t-1}, \hat{p}_{t-1}^R)$ according to the following processes:

Table 1.3: Transition Matrices of Perceived Earnings Shocks

State R

	$\sigma_{t+1} = +y$	$\sigma_{t+1} = -y$
$\sigma_t = +y$	π_L	$1 - \pi_L$
$\sigma_t = -y$	$1 - \pi_L$	π_L

State T

	$\sigma_{t+1} = +y$	$\sigma_{t+1} = -y$
$\sigma_t = +y$	π_H	$1 - \pi_H$
$\sigma_t = -y$	$1 - \pi_H$	π_H

The analyst's estimate of the stock value for t is the distorted net present value of future earnings

$$V_t = a + \mathbb{E}_t \left\{ \frac{\hat{N}_{t+1}}{1+\delta} + \frac{\hat{N}_{t+2}}{(1+\delta)^2} + \dots \right\} \tag{1.4}$$

where $\hat{N}_t = N_{t-1} + \sigma_t$ is the perception of the value of earnings and $a > 0$ captures the empirical fact that analysts inflate their estimates due to trading incentives.

1.3.1 Model Results

Proposition one tells us that the analyst's estimation of the stock intrinsic value depends on his misinterpretations of current signals.

Proposition 1*: The biased forecast satisfy*:

$$V_t = a + \frac{\hat{N}_t}{\delta} + \sigma_t(\alpha_1 + \alpha_2 \hat{p}_t^R) \tag{1.5}$$

where $\hat{N}_t = N_{t-1} + \sigma_t$, $a > 0$ *and* α_1 *and* α_2 *are or the form*

$$\alpha_1 = \frac{1}{\delta} \left\{ \boldsymbol{\gamma_0'}(1+\delta) \left[\left[\boldsymbol{I}(1+\delta) - \boldsymbol{Q} \right]^{-1} \boldsymbol{Q} \boldsymbol{\gamma_1} \right] \right\}$$

$$\alpha_2 = \frac{1}{\delta} \left\{ \boldsymbol{\gamma_0'}(1+\delta) \left[\left[\boldsymbol{I}(1+\delta) - \boldsymbol{Q} \right]^{-1} \boldsymbol{Q} \boldsymbol{\gamma_2} \right] \right\}$$

[3] One example of a density that is a function of an absolute value and is decreasing is the Laplace distribution.

with

$$\boldsymbol{Q} = \begin{bmatrix} (1-\lambda_1)\pi_L & (1-\lambda_1)(1-\pi_L) & \lambda_2\pi_L & \lambda_2(1-\pi_L) \\ (1-\lambda_1)(1-\pi_L) & (1-\lambda_1)\pi_L & \lambda_2(1-\pi_L) & \lambda_2\pi_L \\ \lambda_1\pi_H & \lambda_1(1-\pi_H) & (1-\lambda_2)\pi_H & (1-\lambda_2)(1-\pi_H) \\ \lambda_1(1-\pi_H) & \lambda_1\pi_H & (1-\lambda_2)(1-\pi_H) & (1-\lambda_2)\pi_H \end{bmatrix}$$

$$\boldsymbol{\gamma_0}' = (1,-1,1,-1)$$

$$\boldsymbol{\gamma_1'} = (0,0,1,0)$$

$$\boldsymbol{\gamma_2'} = (1,0,-1,0)$$

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The intuition of the proof is the following: first, I find the convergence of earnings; second I find an expression for expected signals; and third, I find the convergence for the series of expected earnings shocks. More precisely, the series of perceived earnings converges to $\frac{\hat{N}_t}{\delta}$ and thus

$$V_t = a + \frac{\hat{N}_t}{\delta} + \frac{1}{\delta}\left\{\mathbb{E}_t(\sigma_{t+1}) + \frac{\mathbb{E}_t(\sigma_{t+2})}{1+\delta} + \frac{\mathbb{E}_t(\sigma_{t+3})}{(1+\delta)^2} + \dots\right\}$$

The expression for expected signals is $\mathbb{E}_t(\sigma_{t+j}|\Phi_t) = \sigma_t\boldsymbol{\bar{\gamma}'Q^j\hat{q}}^t + (-\sigma_t)\underline{\boldsymbol{\gamma}}\boldsymbol{'Q^j\hat{q}}^t$ where $\boldsymbol{Q}$ is a Markov chain containing the probabilities of going from the pair $(State_t, \sigma_t)$ to the pair $(State_{t+1}, \sigma_{t+1})$; $\boldsymbol{\hat{q}}^t = \left[\hat{p}_t^R,\ 0,\ 1-\hat{p}_t^R,\ 0\right]'$; $\bar{\boldsymbol{\gamma}}' = (1,0,1,0)$; and $\underline{\boldsymbol{\gamma}}' = (0,1,0,1)$. And the series of expected shocks converges to

$$\sigma_t\left\{\boldsymbol{\gamma_0'}(1+\delta)\left[[\boldsymbol{I}(1+\delta)-\boldsymbol{Q}]^{-1}\boldsymbol{Q\gamma_1}\right] + \boldsymbol{\gamma_0'}(1+\delta)\left[[\boldsymbol{I}(1+\delta)-\boldsymbol{Q}]^{-1}\boldsymbol{Q\gamma_2}\right]\hat{p}_t^R\right\}$$

See the proof in the appendix.

Proposition 2: *In face of a conflicting signal, the size of the forecast bias depends on the realizations of confirmation bias (z_t).*

Proof of Proposition 2:

The forecast bias $B_t = V_t - P_t$ is equivalent to

$$B_t = a + \frac{\sigma_t - y_t}{\delta} + \alpha_1(\sigma_t - y_t) + \alpha_2(\sigma_t \hat{p}_t^R + y_t p_t^R)$$

As a conflicting signal activates the right term of the right hand side of equation 1.3, in face of a contradictory signal B_t is equivalent to

$$B_t = a + \frac{z_t y_t - y_t}{\delta} + \alpha_1(z_t y_t - y_t) + \alpha_2(z_t y_t \hat{p}_t^R + y_t p_t^R) \tag{1.6}$$

Then B_t depends on the realized value of z_t.

1.3.2 Novel Empirical Predictions

***Proposition** 3: In face of the prior belief of a trend of negative signals and a positive current signal, the analyst is more inclined to bias his forecast in the direction of the last signal when the signal strength increases.*

Proof of Proposition 3:

Assume a positive contradictory signal: $\mathbb{P}(R) > \hat{p}_{t-1}^R, \sigma_{t-1} = (-y), y_t = +y$. In this case, the forecast bias is

$$B_t = a + \frac{z_t(+y) - (+y)}{\delta} + \alpha_1[z_t(+y) - (+y)] + \alpha_2[z_t(+y)\hat{p}_t^R + (+y)p_t^R] \tag{1.7}$$

When looking at equation 1.7, we can consider *two cases*: when $z_t = 1$ and when $z_t = -1$. Consider first, the case when the effects of confirmatory bias are realized ($z_t = -1$):

$$B_t = a + \frac{(-1)(+y) - (+y)}{\delta} + \alpha_1[(-1)(+y) - (+y)] + \alpha_2[(-1)(+y)\hat{p}_t^R + (+y)p_t^R]$$

$$B_t = a - \frac{2}{\delta}(+y) - \alpha_1 2(+y) + \alpha_2[p_t^R - \hat{p}_t^R](+y)$$

case in which the change in the bias is

$$\frac{\partial B_t}{\partial(+y)} = -2\left[\frac{1}{\delta} + \alpha_1\right] - \alpha_2[\hat{p}_t^R - p_t^R] \tag{1.8}$$

and there exist parameters for which the change in the bias is negative:

$$\frac{\partial B_t}{\partial(+y)} < 0 \iff -2\left[\frac{1}{\delta} + \alpha_1\right]\alpha_2^{-1} < \hat{p}_t^R - p_t^R$$

Second, consider the case when the effects of confirmatory bias are not realized ($z_t = 1$):

$$B_t = a + \frac{(+y) - (+y)}{\delta} + \alpha_1[(+y) - (+y)] + \alpha_2[(+y)\hat{p}_t^R + (+y)p_t^R]$$

$$B_t = a + \alpha_2[\hat{p}_t^R + p_t^R](+y)$$

case in which the change in the bias is greater than zero, for changes in the positive signals

$$\frac{\partial B_t}{\partial(+y)} = \alpha_2[\hat{p}_t^R + p_t^R] \tag{1.9}$$

That is, for $z_t = 1$ we have that $\frac{\partial B_t}{\partial(+y)} > 0$. As the probability $1 - \epsilon$ of $z_t = 1$ augments for lager values of $+y$, the probability of observing $\frac{\partial B_t}{\partial(+y)} > 0$ increases for more extreme values of $+y$.

Proposition 4: *In face of the prior belief of a trend of positive signals and a negative current signal, the analyst is more inclined to bias his forecast in the direction of the last signal when the signal strength increases.*

Proof of Proposition 4:
Assume a negative contradictory signal: $\mathbb{P}(R) > \hat{p}_{t-1}^R, \sigma_{t-1} = (+y), y_t = (-y)$. In this case, the forecast bias is

$$B_t = a + \frac{z_t(-y) - (-y)}{\delta} + \alpha_1[z_t(-y) - (-y)] + \alpha_2[z_t(-y)\hat{p}_t^R + (-y)p_t^R] \tag{1.10}$$

Again, we can consider *two cases* when looking at equation 1.10. First consider the case when the effects of confirmatory bias are realized ($z_t = -1$):

$$B_t = a + \frac{(-1)(-y) - (-y)}{\delta} + \alpha_1[(-1)(-y) - (-y)] + \alpha_2[(-1)(-y)\hat{p}_t^R + (-y)p_t^R]$$

$$B_t = a - \frac{2}{\delta}(-y) - \alpha_1 2(-y) + \alpha_2[p_t^R - \hat{p}_t^R](-y)$$

case in which the change in the bias is

$$\frac{\partial B_t}{\partial(-y)} = -2\left[\frac{1}{\delta} + \alpha_1\right] - \alpha_2[\hat{p}_t^R - p_t^R] \tag{1.11}$$

and there exist parameters for which the change in the bias is negative:

$$\frac{\partial B_t}{\partial(-y)} < 0 \iff -2\left[\frac{1}{\delta} + \alpha_1\right]\alpha_2^{-1} < \hat{p}_t^R - p_t^R$$

Second, the case when the effects of confirmatory bias are not realized ($z_t = 1$):

$$B_t = a + \frac{(-y) - (-y)}{\delta} + \alpha_1[(-y) - (-y)] + \alpha_2[(-y)\hat{p}_t^R + (-y)p_t^R]$$

$$B_t = a + \alpha_2[\hat{p}_t^R + p_t^R](-y)$$

case in which the change in the bias is larger than zero for changes in the negative signals

$$\frac{\partial B_t}{\partial(-y)} = \alpha_2[\hat{p}_t^R + p_t^R] \tag{1.12}$$

That is, for $z_t = 1$ we have that $\dfrac{\partial B_t}{\partial(-y)} > 0$.

As the probability $1-\epsilon$ of $z_t = 1$ augments for more extreme values of $-y$, the probability of observing $\dfrac{\partial B_t}{\partial(-y)} > 0$ increases for more extreme values of $-y$.

Testable Predictions: From propositions 3 and 4, my testable predictions are

- Given a negative prior and a current positive earnings shock, the change in forecast bias is negative for less extreme shocks, and positive for more extreme shocks.
- Given a positive prior and a current negative earnings shock, the change in forecast bias is positive for less extreme shocks, and negative for more extreme shocks.

An alternative, and maybe simpler manner to write down the predictions is the following: when $sign(prior) \neq sign(latest\ signal)$

- $sign(\Delta Forecast\ Bias) = sign(prior)$ if the signal is not extreme
- $sign(\Delta Forecast\ Bias) = sign(latest\ signal)$ if the signal is extreme

1.4 Data and Variables

1.4.1 Data

My sample consists of firms included in the CRSP stock index which currently is composed of 3586 securities traded on NYSE, Amex or NASDAQ. For each firm in the sample, I observe the quarterly series of Earnings Per Share (EPS), as well as daily data on its stock price, market capitalization and number of analysts' recommendations. In addition, I observe daily data on the consensus target price, which is the average forecast of the stock price for the next 12 months from the analysts who cover that stock, and excludes forecasts older than three months when it is calculated. Also, I observe daily data on the "News Heat - Daily Max Readership" index of Bloomberg. This index is constructed by Bloomberg based upon the "number of times each article is read by its users, as well as the number of times users search for news for a specific stock" (Ben-Rephael, Da and Israelsen, 2017) and takes higher values for higher levels of readers activity. Furthermore, I observe quarterly data on Corporate Profits of the U.S. National Income and Product Accounts as well as the quarterly forecasts on Corporate Profits from the Survey of Professional Forecasters[4].

1.4.2 Variables

My dependent variable is the quarterly forecast bias in terms of optimism in target prices. For each firm i and quarter t, I calculate the forecast bias as

$$y_{i,t} = \frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$$

where $TP_{i,t-4}$ is the consensus forecast issued at the end of the quarter $t-4$ (see figure 1.1) for the next 4 quarters on stock i and $P_{i,t}$ is the stock price at the end of the quarter t[5]. Forecasts on stock prices express analysts' opinions about the stock market in the most direct and intuitive manner without the statistical problems associated to earnings management when using forecasts on earnings or operating cash flows to capture optimism.

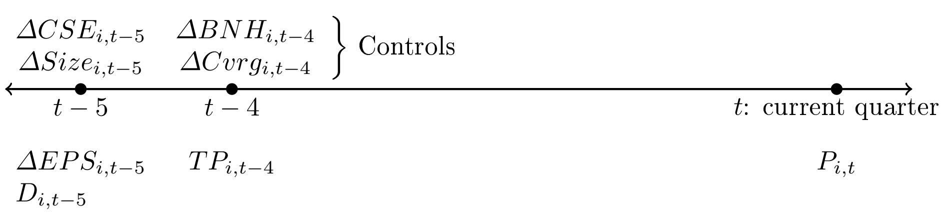


Figure 1.1: Timing of signals, forecasts, realized prices and controls

In order to measure the signal, which is the independent variable of interest in the model, I use growth in Earnings Per Share (ΔEPS). I calculate the signal as the change in Earnings Per Share ($\Delta EPS_{i,t-5}$) scaled by the stock price, i.e. $\Delta EPS_{i,t-5} = \frac{EPS_{i,t-5} - EPS_{i,t-6}}{P_{i,t-6}}$. I use the lag $t-5$ because when issuing a forecast at time $t-4$, analysts observe a past signal at $t-5$. It is important to note that this variable is not only sequentially exogenous to forecast bias but also strictly exogenous in the sense that analysts' forecasts on stock prices do not affect the present or the future realized earnings for any company. Reported earnings are determined by the accounting revenues and costs of firms and not by stock forecasts.

[4] The Survey of Professional Forecasters, conducted by the Federal Reserve Bank of Philadelphia, includes panelists affiliated to different industries such as Universities, Manufacturers, Investment Advisors and Insurance Companies among others.

[5] Notice that $y_{i,t}$ is a very intuitive measure of optimism since it equals the difference between the projected growth in price $\frac{TP_{i,t-4}}{P_{t-4}}$ and the realized growth $\frac{P_{i,t}}{P_{i,t-4}}$

Given that stock prices aggregate information from market participants, managers might use stock prices as a source of information to take decisions about corporate investments when this prices convey new information to managers (see e.g. Chen, Goldstein and Jiang, 2007; and Fishman and Hagerty, 1989). Nevertheless, as analysts forecasts are biased and are not determined by the aggregate decisions of market participants, it is not likely that managers use analysts' target prices to take decisions, and thus it is not likely that analysts' forecasts on stock prices affect realized earnings and free cash flows in the present or the near future, or that SSAs' forecasts affect macroeconomic performance.

To measure different levels of signal intensity that allow me to distinguish, for example, a high signal from an extremely high signal or a low signal from an extremely low signal, I group them by cross-sectional deciles. As in Easterwood and Nutt (1999), I do not use standard deviations as a reference or z-scores to measure high and low signals but quantiles. The standard deviation as a measure of dispersion can bring an idea of how far is an observation from the mean and, in a symmetric distribution with a kurtosis around 3, how many observations are between the mean and a threshold. Therefore, in such a distribution, the standard deviation is useful as a reference to separate the sample in groups. Given that the cross-sectional distributions of EPS Growth are leptokurtic and not always symmetrical (i.e. some are skewed to the left and some to the right; see the Appendix), using the deciles (and not the mean as a reference point) allow me to group the sample in sets of high and low values with the same number of observations, notwithstanding how different are the skewness and kurtosis between cross-sectional distributions. As shown in table 1.14 in the Appendix, the seventh deciles are within 0.001 standard deviations and 0.04 standard deviations from the means, and in some distributions the seventh decile is less than the mean. Thus, using e.g. one standard deviation to the left for low signals and one to the right for high signals, would result in groups not only with few observations but also with a considerable different number of observations between the high and low signals. Consequently, high and low signals are defined in terms of their relative position with respect to other observed signals, and not in terms of the distance from a hypothetical value such as the average[6] signal.

I classify the observations of signals in the right tail of the distribution in weak (W), medium (M) and strong (S) high or favorable signals (see figure 1.2). To do so, I construct the dummies $H^W_{i,t-5}$ which takes the value of one for observations of $\Delta EPS_{i,t-5}$ greater or equal than the seventh cross-sectional decile, $H^M_{i,t-5}$ for the eighth decile and $H^S_{i,t-5}$ for the ninth decile. Similarly, I classify the left tail of the distribution of signals in weak (W), medium (M) and strong (S) low or unfavorable signals by calculating the dummies $L^W_{i,t-5}$, which takes the value of one for observations of $\Delta EPS_{i,t-5}$ lower or equal than the third cross-sectional decile, $L^M_{i,t-5}$ for the second decile and $L^S_{i,t-5}$ for the first decile. Therefore, in order to estimate the effects of high signals on forecast bias, I calculate high signals as $\Delta EPS_{i,t-5}(H^W_{i,t-5})$,

[6] The average is a value that might not correspond to any actual value. For instance, if the observations set is $\{0, 1\}$, its average of 0.5 is not an observation.

$\Delta EPS_{i,t-5}(H^M_{i,t-5})$ and $\Delta EPS_{i,t-5}(H^S_{i,t-5})$. Also, in order to estimate the effects of low signals on forecast bias I calculate low signals as $\Delta EPS_{i,t-5}(L^W_{i,t-5})$, $\Delta EPS_{i,t-5}(L^M_{i,t-5})$ and $\Delta EPS_{i,t-5}(L^S_{i,t-5})$. Together, statistically positive estimates (positive reaction) on high signals and statistically negative estimates on low signals, is consistent with systematic optimism. In addition, increasing analysts' positive reaction upon the intensity of signals is consistent with Griffin and Tversky (1992).

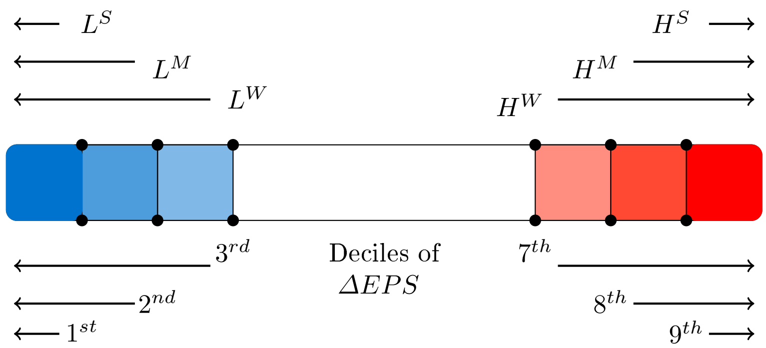


Figure 1.2: Dummies for Cross-Sectional Deciles of Signals

I estimate high contradictory signals as $\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$, $\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$ and $\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$ (see figure 1.3) where $D_{i,t-5}$ is a dummy that takes the value of one for a negative (pessimistic) forecast bias on stock prices, i.e. for $y_{i,t-5} < 0$. Since forecasts are issued at $t-4$, analysts observe a past signal at $t-5$ and past pessimism on stock prices is verified when $y_{i,t-5} = \frac{TP_{i,t-9} - P_{i,t-5}}{P_{i,t-9}}$ is negative. Alternatively and as a robustness check, from the forecasts of Corporate Profits, $D_{i,t-5}$ takes the value of one whenever the difference between the forecast (issued at $t-6$ for $t-5$) and its realized value (at $t-5$) is less than zero. This is in line with the practice of financial analysts of using aggregate Earnings Per Share or the Corporate Profits from the National Income and Product Accounts to predict movements in stock markets[7].

In addition, I use $upward\_high_{i,t}$ as my dependent variable which is a dummy that takes the value of one whenever $TP_{i,t} - TP_{i,t-1} > 0$ (upward forecast revision) and $\Delta EPS_{i,t-1}$ is above the 7th cross-sectional decile. Also, I calculate $downward_{i,t-1}$ which is a dummy that takes the value of one whenever there is a downward forecast revision ($TP_{i,t-1} - TP_{i,t-2} < 0$), and interact it with different levels of signal strength in order to capture contradictory signals, that is, I calculate $downward_{i,t-1}(\Delta EPS_{i,t-1}H^W_{i,t-1})$, $downward_{i,t-1}(\Delta EPS_{i,t-1}H^M_{i,t-1})$ and $downward_{i,t-1}(\Delta EPS_{i,t-1}H^S_{i,t-1})$.

[7]For instance, the Cyclically Adjusted Price-Earnings Ratio (CAPE) popularized by Shiller, Campbell and Greenspan in 1996, equals the level of a stock market index divided by the 10-year average of aggregate earnings per share. Moreover, the forecasting ability of the CAPE model improves when using Corporate Profits instead of accounting earnings (Siegel, 2016).

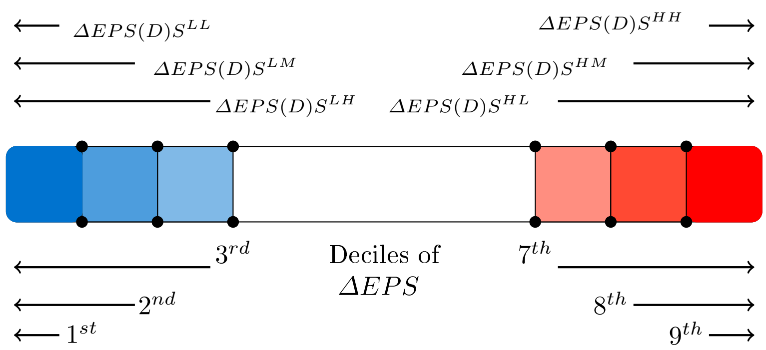


Figure 1.3: Signals Grouped by Deciles Interacted with $D_{i,t-5}$

To control for reputational incentives, I follow Butler and Saraoglu (1999) and calculate for each stock the standardized (across stocks) value of the squared forecast bias $[y_{i,t-5}]^2 = \left[\frac{TP_{i,t-9} - P_{i,t-5}}{P_{i,t-9}}\right]^2$. If analysts consider their reputation when issuing forecasts, they should correct their past relative inaccuracy (relative to the analysts that follow other stocks) and thus past values of the cross-standardized squared error ($CSE_{i,t-5}$) should be negatively related to present forecast bias. Following the literature I control for the size of the company using the log of market capitalization ($Size_{i,t-5}$), and also for analyst coverage ($\Delta Cvrg_{i,t-4}$). Since not all reports have a target price[8] I use the change in the number of recommendations to measure coverage (see e.g. Niehaus and Zhang, 2010). Also, considering that information seeking is fundamental to stock price formation (Grossman and Stiglitz, 1976; De Long et al., 1990) and to analysts' precision (Fischer and Stocken; 2010, Hayes, 1998), I use the change in the quarterly average of the Bloomberg's measure for user activity at the terminals, "News Heat - Daily Max Readership" ($\Delta BNH_{i,t-4}$), to capture and control for the information gathering by stock market participants. Notice that while $Size_{i,t-5}$ and $\Delta EPS_{i,t-5}$ are firm characteristics, $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are characteristics of the informational environment at the moment of issuing a forecast. The summary statistics are in table 1.4.

[8]The description of the variable "Analyst Recommendation" provided by Bloomberg affirms that "[r]ecommendations may or may not have a target price associated with them." Asquith, Mikhail and Au (2005) analyzed more than 1000 analyst reports from 11 different investment banks covering 46 industries' during 1997 - 1999 and found that, while all reports included a stock recommendation, only 72.6% contained a target price.

Table 1.4: Summary Statistics.

The dependent variable y is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the target price on stock i for the next 4 quarters; CSE equals the standardized value of $\left[\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}\right]^2$; $Size$ corresponds to the log of market capitalization; $\Delta Cvrg$ equals the first differences of the number of analysts' recommendations; ΔBNH equals the first differences of the quarterly average of Bloomberg's *News Heat – Daily Max Readership*; and ΔEPS is $\frac{EPS_{i,t}-EPS_{i,t-1}}{P_{i,t-1}}$. The period goes from the second quarter of 2006 to the fourth quarter of 2016.

	CSE	$Size$	$\Delta Cvrg$	ΔBNH	y	ΔEPS
Min	-0.253	-4.538	-10.00	-4.00	-2.0773	-3.2625
1st Qu.	-0.142	5.402	0.00	0.00	-0.14	-0.0058
Median	-0.075	6.790	0.00	0.00	0.092	0.0002
Mean	-0.003	6.792	0.0829	0.0064	0.174	-0.0038
3rd Qu.	-0.032	8.139	0.00	0.00	0.401	0.0064
Max.	47.593	13.494	40.00	4.00	4.055	0.7653
N	91205	106609	105663	105663	90335	106845

1.4.3 Descriptive Statistics

In line with the literature on SSAs I drop the extreme values of the sample of optimism eliminating the top 0.5% and the bottom 0.5% of ΔEPS and y. The final sample consists of 3169 stocks included in the CRSP index with time-series from the second quarter of 2006 to the fourth quarter of 2016. From table 1.4 we can see that the mean and the median of the forecast bias are positive as we would expect from the empirical literature. Also, from table 1.4 and figure 1.5 we see that the distribution has a right tail that is longer than the left tail which Abarbanell and Lehavy (2003) call the tail asymmetry, and means that "far more extreme forecast errors of greater absolute magnitude are observed in the ex-post 'optimistic' tail of the distribution than in the 'pessimistic' tail." In my sample, the average forecast bias is larger than the median, the third quartile is more than three times the size observed for the first quartile, and the maximum is larger than the absolute value of the minimum.

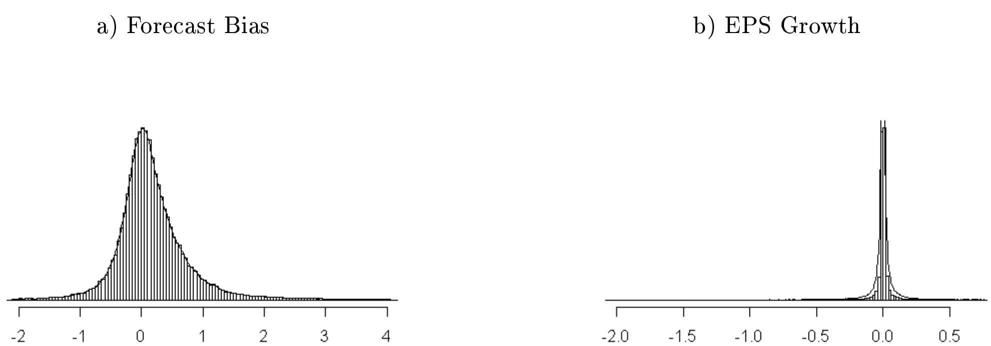


Figure 1.4: Histograms of Forecast Bias and EPS Growth. *Forecast Bias* and *EPS* Growth are calculated as $y_{i,t} = \frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$ and $\Delta EPS_{i,t} = \frac{EPS_{i,t}-EPS_{i,t-1}}{P_{i,t-1}}$ respectively.

The concern about the tail asymmetry is that, the likelihood of an observation of forecast bias falling into the tail asymmetry could be conditional on realizations of economic variables and thus, "differences in the manner in which researchers implicitly or explicitly weight observations that fall into these asymmetries contribute to inconsistent conclusions concerning analyst bias and inefficiency."

Table 1.5: Summary Statistics of Forecast Bias Partitioned by EPS Growth

Forecast Bias and *EPS Growth* are calculated as $y_{i,t} = \frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$ and $\Delta EPS_{i,t} = \frac{EPS_{i,t} - EPS_{i,t-1}}{P_{i,t-1}}$ respectively. * refers to the percentage of optimistic cases relative to the total number of observations; ** refers to the number of pessimistic forecasts divided by the number of optimistic forecasts.

	Cases in the Bottom 25% of ΔEPS		Cases in the Middle of ΔEPS		Cases in the Top 25% of ΔEPS	
	y	ΔEPS	y	ΔEPS	y	ΔEPS
Min	-2.071	-3.2625	-2.077	-0.0058	-2.071	0.0064
1st Qu.	-0.105	-0.0440	-0.144	-0.0016	-0.163	0.0104
Median	0.164	-0.0185	0.065	0.0002	0.098	0.0187
Mean	0.265	-0.0652	0.123	0.0003	0.197	0.0495
3rd Qu.	0.543	-0.0101	0.324	0.0023	0.456	0.0434
Max.	4.030	-0.0058	4.048	0.0064	4.055	0.7653
N	21399	26712	47674	53421	21262	26712
$y > 0$ $freq^*$.	0.6571		0.5818		0.5958	
$\frac{\#pessim}{\#optim}^{**}$	0.5210		0.7177		0.6768	

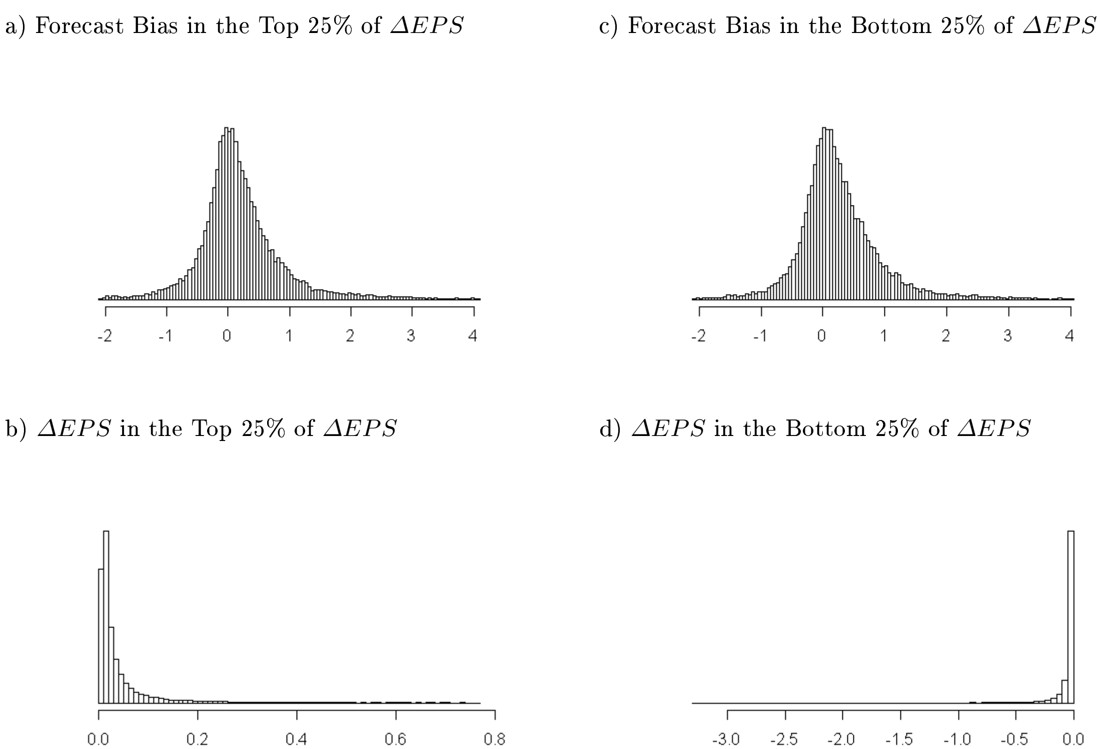


Figure 1.5: Histograms of Forecast Bias and EPS Growth. *Forecast Bias* and *EPS Growth* are calculated as $y_{i,t} = \frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$ and $\Delta EPS_{i,t} = \frac{EPS_{i,t} - EPS_{i,t-1}}{P_{i,t}}$ respectively.

There are no indications of Abarbanell and Lehavy's (2003) remark If I partition my sample by the set of cases in the top 25%, the bottom 25% and the observations in the middle of ΔEPS (see table 1.5). Both the median and the mean of optimism in each group are positive, the percentages of optimistic forecasts are all around 60% and the ratios of pessimistic to optimistic forecasts are all between 0.52 and 0.72. In contrast, Abarbanell and Lehavy (2003) found, for a skewed distribution of optimism in consensus EPS forecasts, that the median of optimism given negative (positive) earnings changes, were positive (negative), the percentages of optimistic forecasts were 50% (34%, a difference of 16 pp) and the ratios of the number of pessimistic to the number of optimistic forecasts were 0.81 (1.83, the double for positive earnings changes).

1.5 Empirical Strategy

1.5.1 Estimation

In this section I explain my empirical strategy used test my novel theoretical predictions. I specify the following model:

$$y_{i,t} = c_i + \sum_{j=1}^{4} \rho_j y_{i,t-j} + \boldsymbol{x}_{i,t-5}\boldsymbol{\delta} + \boldsymbol{w}_{i,t-4}\boldsymbol{\gamma} + \lambda_t + u_{i,t} \tag{1.13}$$

where c_i is a firm-level unobserved effect, λ_t is a time effect common to all firms, $u_{i,t}$ is the error term with $t = 1, ..., T$ and $i = 1, ..., N$ and the other variables are defined as in section 1.4. The vector of explanatory variables $\boldsymbol{x}_{i,t-5}$ of dimension 1x17 is composed by $\Delta EPS_{i,t-5}$, $D_{i,t-5}$, $L^S_{i,t-5}$, $L^M_{i,t-5}$, $L^W_{i,t-5}$, $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ as well as by their interactions. It also includes $Size_{i,t-5}$ and the cross-standardized squared error (CSE) at $t-5$ since at t past inaccuracy is observable only from $t-5$[9]. The vector $\boldsymbol{w}_{i,t-4}$ of 1x2 includes $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ which are not firm characteristics but are variables associated to the informational environment. The vectors of parameters $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are of dimensions 17x1 and 2x1 respectively.

In equation 1.13, statistically positive values of the parameters on the interactions $\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$, $\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$ and $\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$ would show that analysts bias their forecasts in the direction of a high contradictory signal for different levels of intensity. For clarity of exposition, I express equation 1.13 as:

$$\begin{aligned} y_{i,t} &= c_i + \textstyle\sum_{j=1}^{4} \rho_j y_{i,t-j} + \delta_1 \Delta EPS_{i,t-5} + \delta_2 D_{i,t-5}\Delta EPS_{i,t-5} \\ &+\delta_{HW}(D_{i,t-5}\Delta EPS_{i,t-5}H^W_{i,t-5}) + \delta_{HM}(D_{i,t-5}\Delta EPS_{i,t-5}H^M_{i,t-5}) + \delta_{HS}(D_{i,t-5}\Delta EPS_{i,t-5}H^S_{i,t-5}) \\ &+\boldsymbol{x}_{i,t-5}\boldsymbol{\delta} + \boldsymbol{w}_{i,t-4}\boldsymbol{\gamma} + \lambda_t + u_{i,t} \end{aligned}$$

where $\boldsymbol{x}_{i,t-5}$ and $\boldsymbol{\delta}$ include the variables and the parameters, respectively, that are not explicit in the

[9] Notice that the squared error at $t-5$ is $SE_{i,t-5} = \left[\frac{TP_{i,t-9} - P_{i,t-5}}{P_{i,t-9}}\right]^2$

equation. If analysts are issuing forecasts that are rational in the classical sense, then all estimates should not be statistically significant. If given a negative prior, analysts are more optimistic when signals are high whatever their strength, then $\delta_1+\delta_2+\delta_{HW}>0, \quad \delta_1+\delta_2+\delta_{HM}>0$ and $\delta_1+\delta_2+\delta_{HS}>0$. If, given a negative prior, strong high signals are ignored (which resembles confirmatory bias), then $\delta_1+\delta_2+\delta_i\leq 0$ for some $i\in\{HW,HM,HS\}$. Also, if the signal strength helps updating analysts beliefs, then $\delta_i\neq 0$ for some $i\in\{HW,HM,HS\}$ and $\delta_{HW}<\delta_{HM}, \quad \delta_{HW}<\delta_{HS}$. That is, as the strength of the high signal increases, analysts bias their forecasts in the direction of the signal to a greater extent, notwithstanding their pessimistic prior.

A detailed explanation on identification is provided in the appendix (page 37).

1.6 Results

Traditionally, rational forecasts are considered to have forecast errors with an unconditional mean of zero (unbiasedness), a zero mean conditional on current and past values of the forecasted variable (efficiency), and zero correlation with other variables in the information set (Ackert and Hunter, 1995; Eastwood and Nutt, 1999; Lim, 2001; Keane and Runkle, 1998; Abarbanell and Bernard, 1992). Therefore, the estimates of a statistical model in which forecast bias is a function of past observed biases and public information, should not be statistically significant if analysts are issuing rational forecasts. In table 1.6, I report the results from a regression that shows the basic relation between signals and forecast bias where column (1) shows the results of a specification with two autoregressive terms and column (2) includes one autoregressive term. Although both specifications show that the instruments are valid according to the Sargan test, the inclusion of two autoregressive terms augments the probability that there is no second-order serial correlation of the residuals (see the m-statistics), i.e. that the instruments are valid. This first regression shows a coefficient on $\mathit{\Delta EPS}_{i,t-5}$ significantly positive, consistent with DeBondt and Thaler (1990) and that $y_{i,t}$ is autocorrelated with an estimate such that $0<\rho<1$ which does not meet the classical rational expectations hypothesis since the bias is to some extent predictable from past observed biases and public information.

Table 1.6: Panel Regression of Forecast Bias on Earnings Per Share Growth

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The control variables are the following: $CSE_{i,t-5}$ equals the standardized value of $\left[\frac{TP_{i,t-9}-P_{i,t-5}}{P_{i,t-9}}\right]^2$; $Size_{i,t-5}$ corresponds to the log of market capitalization; $\Delta Cvrg_{i,t-4}$ equals the first differences of the number of analysts' recommendations; and $\Delta BNH_{i,t-4}$ equals the first differences of the quarterly average of Bloomberg's *News Heat − Daily Max Readership*. In both specifications there are 5 more instruments than regressors. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for $y_{i,t-1}$ are the lags from $t-2$ to $t-4$ in column (1) and from $t-2$ to $t-3$ in column (2). The instruments for the other variables are the first differences of themselves.***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)
$y_{i,t-1}$	0.6806***	0.6311***
	(0.0315)	(0.0267)
$y_{i,t-3}$	0.0311***	—
	(0.0098)	—
$\Delta EPS_{i,t-5}$	0.1032***	0.0932***
	(0.0318)	(0.0289)
Controls	yes	yes
N	70252	71414
(Sargan) $\chi^2(5)$	8.0540	6.6235
(p-value)	(0.1533)	(0.2502)
First-order m-statistic	-11.0217	-12.6965
(p-value)	(0.0000)	(0.0000)
Second-order m-statistic	-0.2153	-1.8137
(p-value)	(0.8295)	(0.0697)

Although not reported, the coefficient on coverage changes is significant and negative, which is consistent with the research that show that stocks with higher analyst following have more accurate analysts' reports (Merkley, Michaely and Pacelli, 2017; Wieland 2011; Hong, Lim and Stein, 2000). The estimate of the parameter on $\Delta BNH_{i,t-4}$ is negative (although not significant), sign that is consistent with the idea that more informed investors have more precise signals which induces analysts to issue more accurate and less optimistic reports. For instance, Fischer and Stocken (2010) theoretically show that, as investors receive a more precise signal, analysts make more precise forecasts in order to increase the investors' responsiveness and gain credibility which suggests that variables that capture informed trading may serve as a control for reputational incentives. In addition, $CSE_{i,t-5}$ is statistically negative consistent with analysts that, after knowing their inaccuracy relative to the analysts that follow other stocks, try to correct their past relative inaccuracy in line with Butler and Saraoglu (1999). The statistically positive coefficient on $Size_{i,t-5}$ is in line with Hayes (1998), who theoretically proposes that firm size incentivize analysts to follow the stock whenever they have favorable views about it.

In table 1.7 I report the estimates of equation 1.13 using the signal dummies $L^W_{i,t-5}$ and $H^W_{i,t-5}$ which take the value of one for signals that are less or equal than the third decile and greater or equal than the seventh decile respectively. Also $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. Columns (1), (3), (5) and (7) include controls and columns (2), (4), (6) and (8) do not include the endogenous controls. Columns (1) and (2) show the estimates of the model that includes the signals $\Delta EPS_{i,t-5}$ and the dummies without interactions. Columns (3) and (4) also include $D_{i,t-5}$ interacted with $L^W_{i,t-5}$, $H^W_{i,t-5}$ and $\Delta EPS_{i,t-5}$. Columns (5) and (6) add high ($\Delta EPS_{i,t-5}H^W_{i,t-5}$) and low signals ($\Delta EPS_{i,t-5}L^W_{i,t-5}$), and columns (7) and (8) add high and low signals interacted with $D_{i,t-5}$.

As seen in columns (7) and (8), the estimates show that, given the observed pessimism at the time of observing a past signal ($D_{i,t-5} = 1$), analysts react negatively to high (contradictory) signals (see figure 1.6). An increase in a past high signal given a past pessimistic prior, is not associated to an increase in optimism. More specifically, $\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;H^W=1} = -0.2$ in both columns. Also, the estimates show that analysts react positively to low signals given a negative prior. To be more concrete, $\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;L^W=1}$ is either 0.0167 or 0.0385. In addition, the positive estimates on $H^W_{i,t-5}$ show that the average forecast bias is positive (optimistic) and statistically different for those stocks with a high signal compared to the forecast bias on the stocks with signals in the middle. The negative estimates on $D_{i,t-5}L^W_{i,t-5}$ show that the average forecast bias is negative (pessimistic) whenever the signal was low, and it is statistically different with respect to those stocks with signals in the middle.

Table 1.7: Panel Regression of Forecast Bias on Low, High and High Contradictory Signals. Past Pessimism in Stock Prices as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummy $L^W_{i,t-5}$ takes the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd cross-sectional decile. The dummy $H^W_{i,t-5}$ takes the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th cross-sectional decile. The dummy $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous to $u_{i,t}$. In the specifications that include the non-exogenous controls (columns (1), (3), (5) and (7)) there are 5 more instruments than regressors and the specifications without these controls have one more instrument than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$y_{i,t-1}$	0.7168***	0.6886***	0.7172***	0.6888***	0.7232***	0.6930***	0.7240***	0.6932***
	(0.0332)	(0.0268)	(0.0332)	(0.0268)	(0.0336)	(0.0269)	(0.0336)	(0.0269)
$y_{i,t-3}$	0.0446***	0.0377***	0.0445***	0.0376***	0.0443***	0.0371***	0.0443***	0.0370***
	(0.0112)	(0.0010)	(0.0112)	(0.0010)	(0.0112)	(0.0099)	(0.0112)	(0.0100)
$\Delta EPS_{i,t-5}$	0.0522	0.0509	0.0560	0.0543	0.3162	0.3010	−1.2238	−0.9614
	(0.0358)	(0.0332)	(0.0390)	(0.0360)	(0.7536)	(0.6389)	(1.0784)	(0.9240)
$D_{i,t-5}$	−0.0095**	−0.0078**	−0.0023	−0.0014	−0.0033	−0.0022	−0.0048	−0.0033
	(0.0044)	(0.0038)	(0.0048)	(0.0041)	(0.0049)	(0.0041)	(0.0050)	(0.0042)
$L^H_{i,t-5}$	−0.0081**	−0.0078**	−0.0017	−0.0019	−0.0030	−0.0031	−0.0038	−0.0039
	(0.0038)	(0.0034)	(0.0053)	(0.0047)	(0.0054)	(0.0048)	(0.0055)	(0.0048)
$H^L_{i,t-5}$	0.0159***	0.0146***	0.0207***	0.0187***	0.0151***	0.0134***	0.0141**	0.0124**
	(0.0039)	(0.0035)	(0.0054)	(0.0049)	(0.0055)	(0.0050)	(0.0055)	(0.0050)
$D_{i,t-5}L^W_{i,t-5}$			−0.0166**	−0.0153**	−0.0178**	−0.0163***	−0.0145**	−0.0129**
			(0.0068)	(0.0059)	(0.0070)	(0.0059)	(0.0073)	(0.0062)
$D_{i,t-5}H^W_{i,t-5}$			−0.0114	−0.0098	−0.0084	−0.0069	−0.0046	−0.0031
			(0.0070)	(0.0062)	(0.0071)	(0.0062)	(0.0074)	(0.0064)
$D_{i,t-5}\Delta EPS_{i,t-5}$			−0.0293	−0.0208	−0.077	−0.0645	3.8521***	3.1644***
			(0.0844)	(0.0728)	(0.0852)	(0.0736)	(1.4306)	(1.2059)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$					−0.3938	−0.3748	1.1412	0.8816
					(0.7535)	(0.6394)	(1.0774)	(0.9238)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$					−0.0208	−0.0199	1.5300	1.2551
					(0.7574)	(0.6411)	(1.0829)	(0.9264)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$							−3.8454***	−3.1259***
							(1.4418)	(1.2092)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$							−4.0393***	−3.3579***
							(1.4387)	(1.2119)
Controls	Yes	No	Yes	No	Yes	No	Yes	No
N	66966	66966	66966	66966	66966	66966	66966	66966
(Sargan) χ^2	7.0558	0.4759	6.9770	0.4974	7.1498	0.4741	7.0761	0.4934
(p-value)	(0.2165)	(0.4903)	(0.2223)	(0.4806)	(0.2097)	(0.4911)	(0.2150)	(0.4824)
First-order m-statistic	-10.3798	-21.7398	-10.3496	-21.7398	-10.0674	-21.7772	-10.0137	-21.7726
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Second-order m-statistic	0.3062	0.4104	0.2724	0.3861	0.3046	0.4163	0.2843	0.3930
(p-value)	(0.7594)	(0.6815)	(0.7853)	(0.6994)	(0.7607)	(0.6772)	(0.7762)	(0.6943)

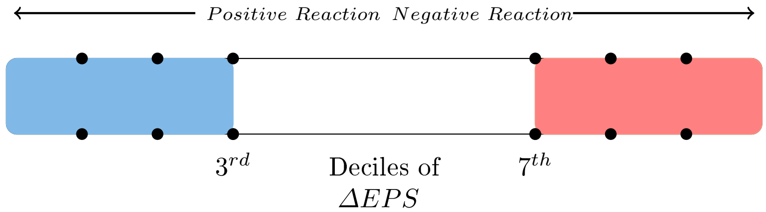


Figure 1.6: Results on Favorable and Unfavorable Signals Interacted with $D_{i,t-5}$

In order to verify if the strength of favorable signals counteract the effects of negative prior beliefs, I calculate the interactions disaggregating favorable and unfavorable signals by deciles. I show a short version of the results in table 1.8 (see the complete set of estimates in page 40). The negative estimates on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{W}_{i;t-5})$ and the positive estimates on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{S}_{i,t-5})$ indicate that, given a pessimistic prior ($D_{i,t-5} = 1$), analysts do not react positively to favorable signals consistent with confirmation bias (see figure 1.7), but they do so when the signal is above the eighth decile which is consistent with Griffin and Tversky (1992). More precisely,

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;H^{W}=1} = -1.75$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;H^{M}=1} = 3.85$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;H^{S}=1} = 5.03$$

In both regressions, with and without controls, the estimates are nondecreasing functions of the signal. As shown in columns (1) and (2), the estimates on the high contradictory signals increase as we go from the seventh decile to the ninth decile. For instance, in the specification with controls of column (7), the coefficient on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{W}_{i;t-5})$ is statistically negative, the one on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{M}_{i;t-5})$ is statistically zero and the estimate on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{S}_{i;t-5})$ is statistically positive. That is, as the strength of the high contradictory signal increases, the parameters go from statistically negative to statistically positive. Given their pessimistic prior, the estimate on the signals above the seventh decile, is negative enough to result in a negative reaction. Also, the estimates on $\Delta EPS_{i,t-5}(D_{i,t-5}H^{M}_{i;t-5})$ and $\Delta EPS_{i,t-5}(D_{i,t-5}H^{S}_{i;t-5})$ show that analysts bias their forecast in the direction of the signals above the ninth decile, notwithstanding their pessimistic prior. The fact that the values of the estimates are increasing in the deciles show that the strength of the high signal counteracts the effects of a negative prior as expected from Griffin and Tversky (1992).

With respect to unfavorable signals given a negative prior, the estimates show that reaction is nondecreasing with respect to the strength of the signal, consistent with Griffin and Tversky (1992). Specifically, in column (7) the results are

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;L^W=1} = 0.2457$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;L^M=1} = 3.8447$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D=1;L^S=1} = 3.8447$$

Table 1.8: Panel Regression of Forecast Bias on High Contradictory Signals Grouped by Deciles. Past Pessimism in Stock Prices as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummies $L^W_{i,t-5}$, $L^M_{i,t-5}$ and $L^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd, 2nd and 1st cross-sectional deciles respectively. The dummies $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. The dummy $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous. In the specifications that include the non-exogenous controls there are 5 more instruments than regressors and the specifications without these controls have one more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

	(1)	(2)
$\Delta EPS_{i,t-5}$	-1.2301	-0.9643
	(1.0801)	(0.9239)
$D_{i,t-5}\Delta EPS_{i,t-5}$	3.8447^{***}	3.1541^{***}
	(1.4305)	(1.2046)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^S_{i,t-5})$	-0.0782	-0.1271
	(0.5189)	(0.4707)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^M_{i,t-5})$	-0.1563	-0.0362
	(1.0403)	(0.8987)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$	-3.5990^{*}	-2.9476^{*}
	(1.8950)	(1.6303)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$	-5.6271^{***}	-5.1155^{***}
	(1.9495)	(1.6470)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$	0.3704	0.8009
	(0.9680)	(0.7636)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$	1.1764^{*}	0.9165
	(0.6971)	(0.6103)
Controls	Yes	No
N	66966	66966
Sargan (χ^2)	7.0700	0.5120
Arellano-Bond for AR(1)	-9.9550^{***}	-21.7800^{***}
Arellano-Bond for AR(2)	0.2653	0.3754

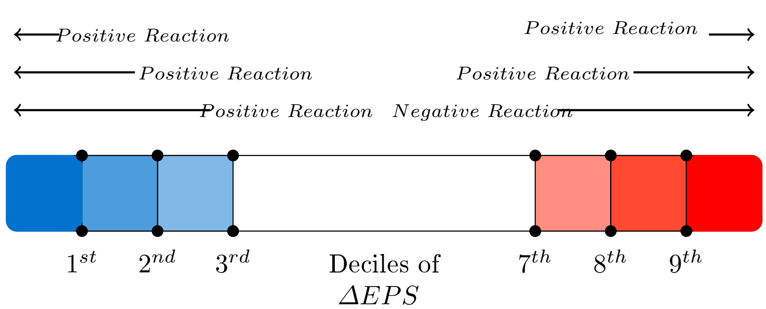


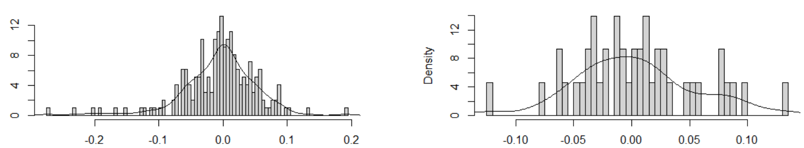
Figure 1.7: Results on Favorable and Unfavorable Signals by Deciles Interacted with D

Although not reported, the results of all controls were similar as those of table 1.6 in terms of significance and sign for both specifications. These results point out that forecast accuracy improves with higher analyst coverage, that firm size is positively associated to favorable views about the stock and that analysts partially correct their past relative inaccuracy. Additionally, the negative relation between the news seeking at Bloomberg Terminals and optimism suggests that variables that capture informed trading may serve as a control for reputational incentives (see e.g. Fischer and Stocken, 2010)

1.6.1 Robustness Checks. Using Forecasts on Corporate Profits.

As a robustness check, I now use the negative forecast bias in Corporate Profits as a proxy for prior beliefs in line with the practice of financial analysts of using aggregate Earnings Per Share or the Corporate Profits from the National Income and Product Accounts to predict movements in stock markets e.g. using Shiller's Cyclically Adjusted Price-Earnings Ratio (CAPE) . As Siegel (2016) found, the forecasting ability of the CAPE model improves when using Corporate Profits instead of reported GAAP[10] earnings. Therefore, $D_{i,t-5}$ takes the value of one whenever the difference between the forecast (issued at $t-6$ for $t-5$) and its realized value (at $t-5$) is less than zero.

First note that forecast errors in Corporate Profits are symmetric: forecast errors in the ex-post optimistic tail of the distribution are not more extreme (in absolute magnitude) than errors in the 'pessimistic' tail, and vice-versa.



1969Q1-2017Q4 2006Q2-2016Q4

Figure 1.8: Distribution of Errors on Corporate Profits (Scaled by Realized Values)

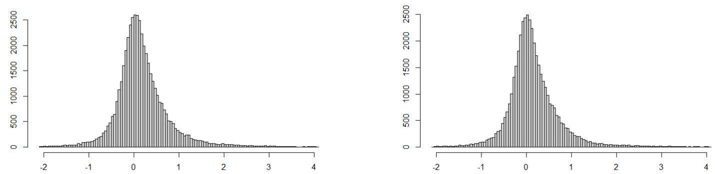
Table 1.9: Summary Statistics. Forecast Error in Corporate Profits.

Period	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
1969Q1-2017Q4	−0.006	0.059	−0.274	−0.034	0.025	0.190
2006Q2-2016Q4	0.003	0.052	−0.125	−0.032	0.026	0.132

Source: Author's calculations with data of the **Survey of Professional Forecasters** conducted by the National Bureau of Economic Research.

Also notice, that the distribution of forecast bias on stock prices is similar for the negative and the non negative values of errors in Corporate Profits.

[10] GAAP stands for "generally accepted accounting principles."



$(CPForecast_{t-6} - CP_{t-5}) \geq 0$ $(CPForecast_{t-6} - CP_{t-5}) < 0$

Figure 1.9: Distribution of Forecast Bias.

Table 1.10: Summary Statistics On Forecast Bias

	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$D_{t-5} = 0$	0.188	0.614	−2.077	−0.135	0.418	4.055
$D_{t-5} = 1$	0.158	0.582	−2.073	−0.146	0.382	4.052

The results, in table 1.16 shown in the appendix, are similar to those obtained before for favorable signals. For $D_{i,t-5} = 1$, analysts react negatively to favorable signals. Also, the average forecast bias is optimistic for those stocks with a high signal and statistically different than the forecast bias on the stocks with signals in the middle. The average forecast bias is pessimistic whenever the signal was low, and it is statistically different with respect to those stocks with signals in the middle.

The results of disaggregating the signals by deciles are in table 1.11 (see the complete set of estimates in page 40). In both regressions with and without controls, the estimates on high contradictory signals are nondecreasing with respect to the intensity of the signal, although are not statistically significant. With respect to low signals given a negative prior, the corresponding estimates go from statistically negative to statistically positive. In the regression with controls, the estimates increase as the signal lowers.

Table 1.11: Panel Regression of Forecast Bias on High Contradictory Signals Grouped by Deciles. Past Pessimism in Corporate Profits as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummies $L^W_{i,t-5}$, $L^M_{i,t-5}$ and $L^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd, 2nd and 1st cross-sectional deciles respectively. The dummies $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. The dummy $D_{i,t-5}$ takes the value of one whenever the consensus forecast on Corporate Profits issued at $t-6$ is less than the actual Corporate Profits at $t-5$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous. In the specifications that include the non-exogenous controls there are 5 more instruments than regressors and the specifications without these controls have one more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)
$\Delta EPS_{i,t-5}$	-0.9591	-1.0581
	(0.0101)	(0.0092)
$D_{i,t-5}\Delta EPS_{i,t-5}$	2.4906	2.4872*
	(1.5501)	(1.3656)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$	0.9178	1.0022
	(1.1373)	(1.0116)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$	1.2385	1.3133
	(1.1422)	(1.0166)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^S_{i,t-5})$	1.5260**	1.3124**
	(0.6020)	(0.5424)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^M_{i,t-5})$	1.8070	1.8638*
	(1.1194)	(1.0568)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$	-5.9489***	-5.7569***
	(2.1191)	(1.8895)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$	-2.1517	-1.4759
	(1.8235)	(1.6332)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$	-0.9154	-1.1275
	(0.8432)	(0.7446)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$	0.5007	0.0858
	(0.4678)	(0.4366)
Controls	Yes	No
N	70252	70941
(Sargan) χ^2	7.9669	0.0126
First-order m-statistic	-10.5142***	-24.6369***
Second-order m-statistic	-0.1764	1.0275

1.6.2 High Contradictory Signals and Low Contradictory Signals

In the previous sections I showed the relation of forecast bias with favorable signals given a negative prior. In this section, I analyze unfavorable contradictory signals, that is, low signals, by deciles, when $D_{i,t-5} = 0$. In table 1.12 I show the results of a panel regression of forecast bias on high contradictory and low contradictory signals grouped by deciles. In column (1), the dummy $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. In column (2), the dummy $D_{i,t-5}$ takes the value of one whenever the consensus forecast on Corporate Profits issued at $t-6$ is less than the actual Corporate Profits at $t-5$.

Table 1.12: Panel Regression of Forecast Bias on High Contradictory Signals and Low Contradictory Signals Grouped by Deciles.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummies $L^W_{i,t-5}$, $L^M_{i,t-5}$ and $L^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd, 2nd and 1st cross-sectional deciles respectively. The dummies $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. In column (1), the dummy $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. In column (2), the dummy $D_{i,t-5}$ takes the value of one whenever the consensus forecast on Corporate Profits issued at $t-6$ is less than the actual Corporate Profits at $t-5$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous. The specifications include the non-exogenous controls and count with 5 more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)
$y_{i,t-1}$	0.728***	0.687***
	(0.034)	(0.032)
$y_{i,t-3}$	0.045***	0.031***
	(0.011)	(0.010)
$\Delta EPS_{i,t-5}$	−1.294	−0.953
	(1.080)	(1.133)
$D_{i,t-5}$	−0.005	−0.037
	(0.005)	(0.100)
$L^W_{i,t-5}$	−0.016**	−0.012
	(0.007)	(0.008)
$H^W_{i,t-5}$	0.006	0.008
	(0.008)	(0.008)
$D_{i,t-5}L^W_{i,t-5}$	−0.002	−0.024*
	(0.011)	(0.012)
$D_{i,t-5}H^W_{i,t-5}$	0.015	0.003
	(0.013)	(0.012)
$D_{i,t-5}\Delta EPS_{i,t-5}$	3.902***	2.488
	(1.431)	(1.545)
$\Delta EPS_{i,t-5}L^S_{i,t-5}$	0.536	0.039
	(0.492)	(0.495)
$\Delta EPS_{i,t-5}L^M_{i,t-5}$	2.599**	2.053*
	(1.057)	(1.064)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$	−1.943	−1.193
	(1.578)	(1.582)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$	2.623*	1.449
	(1.572)	(1.811)
$\Delta EPS_{i,t-5}H^M_{i,t-5}$	−0.486	0.545
	(1.028)	(1.246)
$\Delta EPS_{i,t-5}H^S_{i,t-5}$	−0.527	−0.758*
	(0.401)	(0.420)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^S_{i,t-5})$	−0.564	1.529**
	(0.719)	(0.775)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^M_{i,t-5})$	−2.630*	−0.202
	(1.504)	(1.571)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$	−0.677	−3.927
	(2.230)	(2.394)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$	−6.548***	−2.213
	(2.264)	(2.304)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$	0.747	−1.576
	(1.400)	(1.492)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$	1.651**	1.222*
	(0.811)	(0.628)
Controls	*Yes*	*Yes*
N	3,169	3,169
(Sargan) χ^2	7.041791	8.022714
(p-value)	0.21755	0.15499
First-order m-statistic	−10.01454	−10.58612
(p-value)	0.0000	0.0000
Second-order m-statistic	0.3393307	−0.1574762
(p-value)	0.73436	0.87487

First notice that the results with respect to favorable contradictory signals are similar to those seen in previous sections, in the sense that as the strength of the favorable evidence increases, analysts bias their forecasts in the direction of the signal to a greater extent, notwithstanding their pessimistic prior. For instance, from column (1) we can see that

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=1;H^W_{-5}=1} = 0 + 3.9 + 2.62 - 6.55 = -0.03$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=1;H^M_{-5}=1} = 0 + 3.9 + 0 + 0 = 3.9$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=1;H^S_{-5}=1} = 0 + 3.9 + 0 + 1.65 = 5.55$$

and the corresponding results from column (2) are 0.0, 0.0 and 0.464 respectively. The results with respect to unfavorable contradictory signals indicate that analysts do not react positively to low contradictory signals, except for those below the second decile. In column (1) the results are

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=0;L^W_{-5}=1} = 0$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=0;L^M_{-5}=1} = 2.6$$

$$\left.\frac{\partial y}{\partial \Delta EPS}\right|_{D_{-5}=0;L^S_{-5}=1} = 0$$

and the results in column (2) are 0.0, 2.053, and 0.0 respectively.

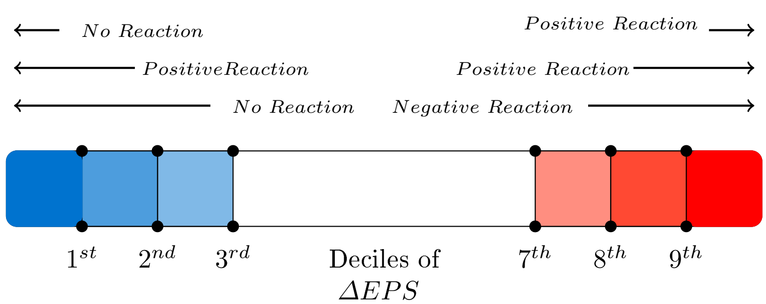


Figure 1.10: Results on Favorable and Unfavorable Contradictory Signals.

1.7 Empirical Evidence On Confirmatory Bias

Up to now I have shown that forecast bias is a function of the signal strength and past beliefs, claiming that the empirical evidence on forecast bias is consistent with what is expected from my model of confirmatory bias and signal strength, and this is because I do not observe analysts' true beliefs. For the purpose of testing the hypothesis of confirmation bias in target prices, in this section I follow Pouget, Sauvagnat and Villeneuve's (2017) definition of confirmatory bias. The intuition of this definition is that, if the current common signal and the past belief updating have the same sign, the biased agent uses the common signal to form his beliefs. If they are of different sign, the biased agent may bias the current signal to form his beliefs. In other words, biased agents "may ignore a signal when this signal is inconsistent with their prior beliefs."

One of the implications of such definition is that an analyst tends to revise his belief less often when signals have a sign that is different from the sign of his last belief revision. Thus, I use a probability model to test the null hypothesis, that analysts revise their forecasts in the direction of the latest signal, irrespective of their prior beliefs about the prospect of the stock, i.e. that they are rational. In particular I test the hypothesis that analysts revise their forecasts upward when latest signal is high, irrespective of whether they revised their target price downward. To this end, I regress the dummy $upward\_high_{i,t}$ over $downward_{i,t-1}(\Delta EPS_{i,t-1}H^W_{i,t-1})$, $downward_{i,t-1}(\Delta EPS_{i,t-1}H^M_{i,t-1})$ and $downward_{i,t-1}(\Delta EPS_{i,t-1}H^S_{i,t-1})$. The results are in table 1.13. Column (1) shows the results of a logit model, column (2) shows the estimates of a logit model with individual fixed effects of Stammann, Heiss, and McFadden (2016), column (3) shows ols estimates, and column (4) shows the Arellano-Bond estimates of a dynamic linear regression.

Table 1.13: Forecast Revisions.

The dependent variable $upward\_high_{i,t}$ is a dummy that takes the value of one whenever $TP_{i,t} - TP_{i,t-1} > 0$ and $\Delta EPS_{i,t-1}$ is above the 7th cross-sectional decile; $downward_{i,t-1}$ is a dummy that takes the value of one whenever $TP_{i,t-1} - TP_{i,t-2} < 0$; $\Delta EPS_{i,t-1}$ is calculated as $\frac{EPS_{i,t-1} - EPS_{i,t-2}}{P_{i,t-2}}$. The dummies $H^W_{i,t-1}$, $H^M_{i,t-1}$ and $H^S_{i,t-1}$ take the value of one whenever $\Delta EPS_{i,t-1}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. Column (1) shows the results of a logit model, column (2) shows the estimates of a logit model with individual fixed effects of Stammann, Heiss, and McFadden (2016), column (3) shows ols estimates, and column (4) shows the Arellano-Bond estimates of a dynamic linear regression. The instruments for $upward\_high_{i,t-1}$ are its the lags from $t-2$ to $t-14$ in the dynamic regression. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis for the dynamic model.

	(1)	(2)	(3)	(4)
$upward\_high_{i,t-1}$				0.042^{***}
				(0.006)
$upward\_high_{i,t-2}$				0.007
				(0.006)
$downward_{i,t-1}(\Delta EPS_{i,t-1}S^{HL}_{i,t-1})$	$-10,826.120$	$-8,611$	-14.236^{***}	-5.984^{***}
	(188, 247.500)	(161, 000)	(0.723)	(0.490)
$downward_{i,t-1}(\Delta EPS_{i,t-1}S^{HM}_{i,t-1})$	6, 493.822	5, 152	10.475^{***}	5.033^{***}
	(200, 347.700)	(175, 500)	(0.754)	(0.489)
$downward_{i,t-1}(\Delta EPS_{i,t-1}S^{HH}_{i,t-1})$	2, 744.735	1, 856	3.370^{***}	0.893^{***}
	(69, 876.340)	(71, 360)	(0.235)	(0.112)
Controls	*Yes*	*Yes*	*Yes*	*Yes*
First-order m-statistic				37.2013
(p-value)				0.0000
Second-order m-statistic				0.4956
(p-value)				0.6202

As we see from the table, although only in the linear models it is possible to reject the null hypothesis of rational analysts, in all specifications the estimates of the low favorable signals are negative, indicating confirmation bias, but as the signal strength increases, the estimates turn to positive indicating that the signal strength is important in the belief updating process as in Griffin and Tversky (1992).

1.8 Conclusions

In this research I study how SSAs' stock price forecasts react to the signal strength when this signal is contradictory and there is a negative prior. Assuming that the severity of the confirmation bias is decreasing with respect to the signal intensity, I first model an agent who updates his beliefs dynamically and is prone to confirmatory bias. My model offers novel empirical predictions with regards to the relationship between forecast bias and signal intensity in the presence of confirmation bias which I test with a dynamic panel data model. Using data on stock target prices and earnings per share, I find evidence of negative reactions to favorable contradictory signals, except for signals above the eighth decile for which analysts react positively. Following Griffin and Tversky (1992), if analysts tend to focus on the strength or extremeness of evidence, and not so much in its predictive validity, then they bias their forecasts in accordance to the direction the signals. Furthermore, from forecast revisions, I find that when analysts have a negative prior belief they revise their forecasts upward for sufficiently strong high signals, indicating that analysts pay attention to the signal strength.

What are the observable consequences for market outcomes? A model of beliefs updating with extreme signals and confirmatory bias with direct testable implications on stock prices (returns, volatility) is one approach to this question in further research. In addition, as from experimental evidence we know that not only the signal strength matters, but also its predictive validity, an important related question for further research is if investors classify evidence according to its predictive validity and react correspondingly.

Appendix

Distributions

Table 1.14: Cross-Sectional Distributions of EPS Growth

The column $z - score$ $(3rd)$ shows the standardized value of the 3th decile of EPS Growth for the cross-sectional distribution at each quarter, and the column $z - score$ $(7th)$ shows the standardized value of the 7th decile. $Kurtosis$ is the fourth standardized moment minus 3, and $Skewness$ is the third standardized moment.

Date	Mean	S.D.	3th Decile	z-score (3rd)	7th Decile	z-score (7th)	Kurtosis	Skewness
2006-06-30	0.006	0.425	−0.002	−0.019	0.004	−0.005	2,318.185	44.219
2006-09-29	0.010	0.213	−0.003	−0.059	0.003	−0.029	641.239	22.547
2006-12-29	0.004	0.132	−0.004	−0.061	0.004	−0.004	467.300	17.194
2007-03-30	−0.005	0.261	−0.004	0.004	0.003	0.031	795.939	−22.587
2007-06-29	0.004	0.125	−0.002	−0.046	0.005	0.005	362.405	4.758
2007-09-28	0.001	0.263	−0.003	−0.017	0.003	0.007	1,828.280	36.886
2007-12-31	0.0003	0.211	−0.004	−0.022	0.003	0.015	448.802	1.981
2008-03-31	0.016	0.341	−0.005	−0.062	0.004	−0.036	1,030.631	28.520
2008-06-30	0.002	0.200	−0.003	−0.026	0.006	0.016	360.903	−5.908
2008-09-30	0.001	0.450	−0.006	−0.015	0.004	0.005	1,764.046	36.940
2008-12-31	−0.064	2.071	−0.014	0.024	0.002	0.032	2,638.528	−50.492
2009-03-31	0.379	15.035	−0.006	−0.026	0.012	−0.024	2,740.020	52.247
2009-06-30	0.067	1.614	−0.004	−0.044	0.010	−0.035	2,249.096	45.571
2009-09-30	0.022	0.917	−0.002	−0.026	0.009	−0.014	1,374.135	32.409
2009-12-31	−0.003	0.341	−0.005	−0.007	0.006	0.025	338.807	−11.199
2010-03-31	0.007	0.684	−0.004	−0.017	0.006	−0.002	588.470	2.591
2010-06-30	0.036	1.115	−0.002	−0.034	0.007	−0.026	1,472.658	36.352
2010-09-30	−0.023	1.206	−0.003	0.016	0.006	0.023	2,351.513	−47.506
2010-12-31	0.001	0.419	−0.005	−0.013	0.005	0.010	425.421	−2.719
2011-03-31	0.004	0.580	−0.005	−0.014	0.005	0.002	972.196	2.762
2011-06-30	0.003	0.179	−0.002	−0.028	0.005	0.015	422.642	−2.422
2011-09-30	−0.014	0.755	−0.003	0.015	0.005	0.025	2,345.710	−47.769
2011-12-30	−0.057	2.944	−0.007	0.017	0.004	0.021	2,417.493	−48.982
2012-03-30	0.023	3.483	−0.004	−0.008	0.006	−0.005	1,436.331	22.120
2012-06-29	−0.014	1.313	−0.003	0.009	0.005	0.015	2,076.744	−42.784
2012-09-28	0.001	0.282	−0.004	−0.017	0.004	0.011	544.679	4.771
2012-12-31	−0.002	0.364	−0.005	−0.007	0.004	0.018	642.109	−5.154
2013-03-28	−0.0003	0.539	−0.004	−0.007	0.005	0.010	975.542	−4.814
2013-06-28	0.009	0.138	−0.002	−0.079	0.005	−0.028	264.846	13.658
2013-09-30	−0.024	1.493	−0.003	0.014	0.004	0.019	2,181.122	−45.875
2013-12-31	0.010	1.004	−0.004	−0.014	0.003	−0.007	1,911.371	40.966
2014-03-31	0.013	0.372	−0.004	−0.046	0.003	−0.027	1,191.037	32.927
2014-06-30	−0.861	31.867	−0.001	0.027	0.005	0.027	1,756.798	−40.912
2014-09-30	0.941	44.771	−0.002	−0.021	0.004	−0.021	2,284.957	47.822
2014-12-31	−0.009	0.390	−0.005	0.011	0.003	0.031	990.839	−27.337
2015-03-31	−0.0001	0.138	−0.005	−0.035	0.003	0.023	221.669	−4.908
2015-06-30	0.001	0.113	−0.002	−0.024	0.004	0.030	299.644	5.248
2015-09-30	−0.050	2.345	−0.003	0.020	0.003	0.023	2,213.986	−47.005
2015-12-31	0.053	2.921	−0.005	−0.020	0.004	−0.017	2,144.598	45.955
2016-03-31	0.019	0.685	−0.005	−0.035	0.004	−0.023	1,227.180	28.097
2016-06-30	0.008	0.308	−0.001	−0.031	0.006	−0.008	938.433	16.955
2016-09-30	0.005	0.212	−0.003	−0.038	0.004	−0.005	363.691	4.439
2016-12-30	0.001	0.170	−0.005	−0.036	0.003	0.011	406.175	15.840

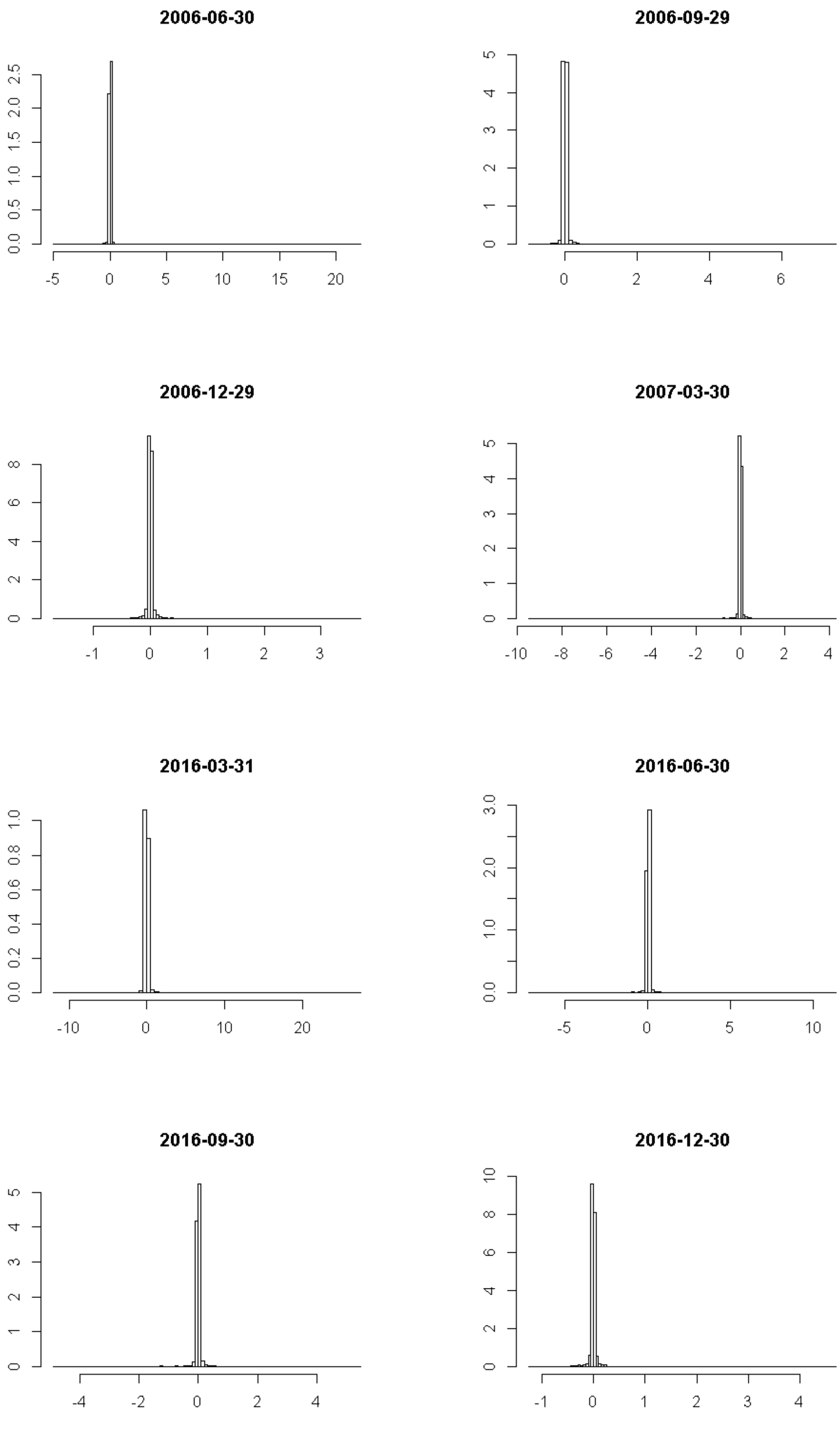


Figure 1.11: Cross-Sectionals Distributions for Quarters 2006Q2, 2006Q3, 2006Q4, 2007Q1, 2016Q1,2016Q2, 2016Q3 and 2016Q4.

Identification

As explained in section 1.4, $\Delta EPS_{i,t-5}$ and its deciles are not affected by analysts' forecasts at t. In addition, since $D_{i,t-5}$, $CSE_{i,t-5}$ and $Size_{i,t-5}$ are included at $t-5$, they can be considered exogenous in equation 1.13. That is, the vector $\boldsymbol{x}_{i,t-5}$ is exogenous. In contrast, it is very likely that there is feedback between optimism (which contains target prices at $t-4$) and analysts' coverage ($\Delta Cvrg_{i,t-4}$) and news seeking ($\Delta BNH_{i,t-4}$). Thus, the vector $\boldsymbol{w}_{i,t-4}$ is not exogenous in equation 1.13. Nevertheless, it is reasonable to say that $\boldsymbol{w}_{i,t-4}$ is sequentially exogenous in the sense that forecasts at $t-4$ do not affect the coverage or news seeking before $t-4$. I will use this to estimate the parameters.

In equation 1.13, the random effects assumption of independence between the regressors and the unobserved term is invalid. Additionally, as Nickell (1981) showed, the fixed effects estimator of ρ_1 is biased and inconsistent for a fixed T and $N \to \infty$ since the within-transformed lagged dependent variable and the within-transformed error are correlated. Thus, I use a difference transformation of equation 1.13 to eliminate the unobserved c_i:

$$\Delta y_{i,t} = \sum_{j=1}^{4} \rho_j \Delta y_{i,t-j} + \Delta \boldsymbol{x}_{i,t-5}\boldsymbol{\delta} + \Delta \boldsymbol{w}_{i,t-4}\boldsymbol{\gamma} + \Delta\lambda_t + \Delta u_{i,t} \tag{1.14}$$

where $\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$ and the parameter vector of interest $\boldsymbol{\delta}$ is the same of equation 1.13. That is, estimating equation 1.14 gives the parameters of equation 1.13. In equation 1.14, $\Delta u_{i,t}$ and $\Delta y_{i,t-1}$ are, by construction, correlated, i.e. $\Delta y_{i,t-1}$ brings an endogeneity problem so ordinary least squares estimators are inconsistent. Nevertheless, with $u_{i,t}$ that are $i.d.d$ we have that $\mathbb{E}[(u_{i,t} - u_{i,t-1})(y_{i,t-2})] = 0$ which suggests that we can use an instrumental variables approach.

As proposed by Arellano and Bond (1991)[11], all lags of y from $t-2$ are potential instruments for $\Delta y_{i,t-1}$ that solve the endogeneity problem in equation 1.14. In addition, for $\boldsymbol{w}_{i,t-4}$ we have $\mathbb{E}[(u_{i,t} - u_{i,t-1})(w_{i,t-6})] = 0$ and I can use all lags of $\boldsymbol{w}_{i,t}$ from $t-6$ as potential instruments for $\Delta \boldsymbol{w}_{i,t-4}$. From this set of potential valid instruments, I must choose a number of them that brings an overidentified model in order to test their validity. The validity of these instruments, nonetheless, do require that the error terms $u_{i,t}$ of equation 1.13 are not serially correlated, that is, it requires $\mathbb{E}[\Delta u_{i,t} \Delta u_{i,t-2}] = 0$ for which I use the Arellano and Bond's (1991) m-statistic to test the second-order residual serial correlation coefficient, and Sargan test of over-identifying restrictions. Finally, the exogeneity of the signals allows us to have consistent estimators of their corresponding parameters in equation 1.14 without using its lagged values as instruments and to give them a causality interpretation.

[11] I do not make use of the orthogonal deviations proposed by Arellano and Bover (1995) to eliminate c_i because I can expect that ρ_1 does not approach to 1. Although, individually, stock prices and target prices may be integrated of order one, notice that $y_{i,t}$ can be expressed as the difference between $\frac{TP_{i,t}}{P_t}$ and $\frac{P_{i,t+4}}{P_t}$ both of which are growth rates and a linear combination of two $I(0)$ variables is $I(0)$. In this regard, Brav and Lehavy (2003) found that the target price-to-stock price ratio seems graphically stationary and that target and stock prices are cointegrated.

As in Arellano and Bond (1991), the potential valid instruments increase as t increases. For instance, at $t = 6$ the moment conditions are:

$$\mathbb{E}[(u_{i,6} - u_{i,5})(y_{i,4})] = 0$$

$$\mathbb{E}[(u_{i,6} - u_{i,5})(y_{i,3})] = 0$$

$$\mathbb{E}[(u_{i,6} - u_{i,5})(y_{i,2})] = 0$$

$$\mathbb{E}[(u_{i,6} - u_{i,5})(y_{i,1})] = 0$$

$$\mathbb{E}[(u_{i,6} - u_{i,5})(y_{i,0})] = 0$$

and therefore, for $\Delta y_{i,6}$ I can use some or all of the instruments from the set $\{y_{i,4}, y_{i,3}, ..., y_{i,0}\}$. For $t = 7$ the moment conditions are

$$\mathbb{E}[(u_{i,7} - u_{i,6})(y_{i,5})] = 0$$

$$\mathbb{E}[(u_{i,7} - u_{i,6})(y_{i,4})] = 0$$

$$\vdots$$

$$\mathbb{E}[(u_{i,7} - u_{i,6})(y_{i,0})] = 0$$

and the set of potential valid instruments for $\Delta y_{i,7}$ is therefore $\{y_{i,5}, y_{i,4}, ..., y_{i,0}\}$. Similarly, for each $t \geq 6$ the instruments for $\Delta \boldsymbol{w}_{i,t-4}$ are the lags of $\boldsymbol{w}_{i,t}$ from $t-6$. This variable is instrumented since the forecast bias $y_{i,t}$ is constructed using the price forecast issued at $t-4$ which, as explained in the beginning of this section, is correlated with $\Delta \boldsymbol{w}_{i,t-4}$. Finally, $\Delta \boldsymbol{x}_{i,t-5}$ is used as its own instrument.

To develop the estimators, I express equation 1.13 as

$$y_i = c_i + X_i \boldsymbol{\beta} + u_i \tag{1.15}$$

and equation 1.14 as

$$\Delta y_i = \Delta X_i \boldsymbol{\beta} + \Delta u_i \tag{1.16}$$

where $\Delta X_i \equiv (\Delta y_{i,-1}, \Delta y_{i,-2}, \Delta y_{i,-3}, \Delta y_{i,-4}, \Delta \boldsymbol{x}_{i,-5}, \Delta \boldsymbol{w}_{i,-4}, d)$, d is a time dummy and $\boldsymbol{\beta}$ is the vector of parameters $(\rho_1, \rho_2, \rho_3, \rho_4, \boldsymbol{\delta}, \boldsymbol{\gamma}, \lambda)'$. Notice that the vector $\boldsymbol{\beta}$ of equation 1.16 is the same as the vector in equation 1.15 so that I am estimating the parameters of the equation in levels through the equation in differences. Remembering that $\boldsymbol{w}_{i,t-4}$ is instrumented using its lags from $t-6$, let

$$\Delta u_i = \begin{bmatrix} \Delta u_{i,6} \\ \vdots \\ \Delta u_{i,T} \end{bmatrix}$$

and let the block diagonal matrix Z_i be the instrument matrix:

$$\begin{bmatrix} [y_{i,0},...,y_{i,4},\boldsymbol{w}_{i,0},\boldsymbol{\Delta x}_{i,1},d_6] & 0 & \dots & 0 \\ 0 & [y_{i,0},...,y_{i,5},\boldsymbol{w}_{i,0},\boldsymbol{w}_{i,1},\boldsymbol{\Delta x}_{i,2},d_7] & & 0 \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & [y_{i,0},...,y_{i,T-2},\boldsymbol{w}_{i,0},...,\boldsymbol{w}_{i,T-6},\boldsymbol{\Delta x}_{i,T-5},d_T] \end{bmatrix}$$

Therefore, the set of moment conditions can be expressed as $\mathbb{E}[Z_i'\Delta u_i] = \mathbb{E}[Z_i'(\Delta y_i - \Delta X_i\boldsymbol{\beta})] = 0$ and the GMM estimators of the parameters can be found by solving:

$$\min_{\boldsymbol{\beta}} \left[\sum_{i=1}^{N}[Z_i'(\Delta y_i - \Delta X_i\boldsymbol{\beta})\right]' A \left[\sum_{i=1}^{N}[Z_i'(\Delta y_i - \Delta X_i\boldsymbol{\beta})\right]$$

where A is the weighting matrix of the moments, which must satisfy $A = \mathbb{E}[Z_i'\Delta u_i \Delta u_i' Z_i]^{-1}$ so that we get the most efficient estimator. Notice that with error terms in levels that are *i.i.d.* we have that

$$\mathbb{E}[\Delta u_i \Delta u_i'] = \sigma_u^2 G = \sigma_u^2 \begin{bmatrix} 2 & -1 & 0 & \dots \\ -1 & 2 & & 0 \\ 0 & \ddots & \ddots & -1 \\ \vdots & 0 & -1 & 2 \end{bmatrix}$$

Thus, it is possible to first solve for $\boldsymbol{\beta}$ utilizing $\hat{A} = [\sum_{i=1}^{N} Z_i' G Z_i]^{-1}$ and estimate $\Delta\hat{u}_i$. Afterwards, we can estimate A as

$$A^* = [\sum_{i=1}^{N} Z_i'\Delta\hat{u}_i\Delta\hat{u}_i' Z_i]^{-1}$$

Therefore, the GMM estimator can be computed as:

$$\boldsymbol{\beta}_{GMM} = (\Delta X' Z A^* Z' \Delta X)^{-1}(\Delta X' Z A^* Z' \Delta y)$$

$$= \left[\left(\sum_{i=1}^{N} \Delta X_i' Z_i\right) A^* \left(\sum_{i=1}^{N} Z_i' \Delta X_i\right)\right]^{-1} \left[\left(\sum_{i=1}^{N} \Delta X_i' Z_i\right) A^* \left(\sum_{i=1}^{N} Z_i' \Delta y_i\right)\right]$$

Table 1.15: Panel Regression of Forecast Bias on High Contradictory Signals Grouped by Deciles. Past Pessimism in Stock Prices as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummies $L^W_{i,t-5}$, $L^M_{i,t-5}$ and $L^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd, 2nd and 1st cross-sectional deciles respectively. The dummies $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. The dummy $D_{i,t-5}$ takes the value of one for $y_{i,t-5} < 0$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous. In the specifications that include the non-exogenous controls (columns (1), (3), (5) and (7)) there are 5 more instruments than regressors and the specifications without these controls have one more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$y_{i,t-1}$	0.7168***	0.6886***	0.7172***	0.6888***	0.7232***	0.6930***	0.7238***	0.6929***
	(0.0332)	(0.0268)	(0.0332)	(0.0268)	(0.0336)	(0.0269)	(0.0337)	(0.0269)
$y_{i,t-3}$	0.0446***	0.0377***	0.0445***	0.0376***	0.0443***	0.0371***	0.0442***	0.0369***
	(0.0112)	(0.0010)	(0.0112)	(0.0010)	(0.0112)	(0.0099)	(0.0112)	(0.0099)
$\Delta EPS_{i,t-5}$	0.0522	0.0509	0.0560	0.0543	0.3162	0.3010	−1.2301	−0.9643
	(0.0358)	(0.0332)	(0.0390)	(0.0360)	(0.7536)	(0.6389)	(1.0801)	(0.9239)
$D_{i,t-5}$	−0.0095**	−0.0078**	−0.0023	−0.0014	−0.0033	−0.0022	−0.0046	−0.0033
	(0.0044)	(0.0038)	(0.0048)	(0.0041)	(0.0049)	(0.0041)	(0.0050)	(0.0042)
$L^W_{i,t-5}$	−0.0081**	−0.0078**	−0.0017	−0.0019	−0.0030	−0.0031	−0.0039	−0.0039
	(0.0038)	(0.0034)	(0.0053)	(0.0047)	(0.0054)	(0.0048)	(0.0055)	(0.0048)
$H^W_{i,t-5}$	0.0159***	0.0146***	0.0207***	0.0187***	0.0151***	0.0134***	0.0140**	0.0124**
	(0.0039)	(0.0035)	(0.0054)	(0.0049)	(0.0055)	(0.0050)	(0.0056)	(0.0050)
$D_{i,t-5}L^W_{i,t-5}$			−0.0166**	−0.0153**	−0.0178**	−0.0163***	−0.0134	−0.0116
			(0.0068)	(0.0059)	(0.0070)	(0.0059)	(0.0095)	(0.0084)
$D_{i,t-5}H^W_{i,t-5}$			−0.0114	−0.0098	−0.0084	−0.0069	0.0085	0.0092
			(0.0070)	(0.0062)	(0.0071)	(0.0062)	(0.0113)	(0.0097)
$D_{i,t-5}\Delta EPS_{i,t-5}$			−0.0293	−0.0208	−0.077	−0.0645	3.8447***	3.1541***
			(0.0844)	(0.0728)	(0.0852)	(0.0736)	(1.4305)	(1.2046)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$					−0.3938	−0.3748	1.1477	0.8845
					(0.7535)	(0.6394)	(1.0791)	(0.9237)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$					−0.0208	−0.0199	1.5363	1.2579
					(0.7574)	(0.6411)	(1.0846)	(0.9263)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^S_{i,t-5})$							−0.0782	−0.1271
							(0.5189)	(0.4707)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^{LM}_{i,t-5})$							−0.1563	−0.0362
							(1.0403)	(0.8987)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$							−3.5990*	−2.9476*
							(1.8950)	(1.6303)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$							−5.6271***	−5.1155***
							(1.9495)	(1.6470)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$							0.3704	0.8009
							(0.9680)	(0.7636)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$							1.1764*	0.9165
							(0.6971)	(0.6103)
Controls	Yes	No	Yes	No	Yes	No	Yes	No
N	66966	66966	66966	66966	66966	66966	66966	66966
(Sargan) χ^2	7.0558	0.4759	6.9770	0.4974	7.1498	0.4741	7.0700	0.5120
(p-value)	(0.2165)	(0.4903)	(0.2223)	(0.4806)	(0.2097)	(0.4911)	(0.2155)	(0.4743)
First-order m-statistic	-10.3798	-21.7398	-10.3496	-21.7398	-10.0674	-21.7772	-9.9550	-21.7800
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Second-order m-statistic	0.3062	0.4104	0.2724	0.3861	0.3046	0.4163	0.2653	0.3754
(p-value)	(0.7594)	(0.6815)	(0.7853)	(0.6994)	(0.7607)	(0.6772)	(0.7908)	(0.7073)

Table 1.16: Panel Regression of Forecast Bias on Low, High and High Contradictory Signals. Past Pessimism in Corporate Profits as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummy $L^W_{i,t-5}$ takes the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd cross-sectional decile. The dummy $H^W_{i,t-5}$ takes the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th cross-sectional decile. The dummy $D_{i,t-5}$ takes the value of one whenever the consensus forecast on Corporate Profits issued at $t-6$ is less than the actual Corporate Profits at $t-5$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous to $y_{i,t}$. In the specifications that include the non-exogenous controls (columns (1), (3), (5) and (7)) there are 5 more instruments than regressors and the specifications without these controls have one more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$y_{i,t-1}$	0.6775***	0.7224***	0.6787***	0.7229***	0.6836***	0.7262***	0.6840***	0.7263***
	(0.0313)	(0.0212)	(0.0314)	(0.0212)	(0.0317)	(0.0212)	(0.0317)	(0.0212)
$y_{i,t-3}$	0.0311***	0.0356***	0.0311***	0.0356***	0.0309***	0.0352***	0.0310***	0.0352***
	(0.0097)	(0.0094)	(0.0097)	(0.0094)	(0.0097)	(0.0093)	(0.0097)	(0.0093)
$\Delta EPS_{i,t-5}$	0.0485	0.0415	0.0782*	0.0614	0.2593	0.1319	−0.9495	−1.0560
	(0.0343)	(0.0334)	(0.0472)	(0.0452)	(0.7319)	(0.6550)	(1.1340)	(1.0117)
$D_{i,t-5}$	−0.0280	0.0975**	−0.0339	0.1029***	−0.0347	0.1016***	−0.0365	0.1010***
	(0.0988)	(0.0387)	(0.0998)	(0.0388)	(0.0998)	(0.0388)	(0.1005)	(0.0388)
$L^W_{i,t-5}$	−0.0096***	−0.0124***	−0.0036	−0.0056	−0.0050	−0.0070	−0.0049	−0.0069
	(0.0037)	(0.0034)	(0.0052)	(0.0049)	(0.0052)	(0.0049)	(0.0053)	(0.0050)
$H^W_{i,t-5}$	0.0153***	0.0106***	0.0183***	0.0142***	0.0137***	0.0097**	0.0138***	0.0098**
	(0.0037)	(0.0035)	(0.0051)	(0.0048)	(0.0052)	(0.0049)	(0.0053)	(0.0050)
$D_{i,t-5}L^W_{i,t-5}$			−0.0132*	−0.0148**	−0.0130*	−0.0146**	−0.0121	−0.0138*
			(0.0071)	(0.0066)	(0.0071)	(0.0066)	(0.0077)	(0.0070)
$D_{i,t-5}H^W_{i,t-5}$			−0.0058	−0.0074	−0.0049	−0.0066	−0.0039	−0.0058
			(0.0072)	(0.0067)	(0.0073)	(0.0067)	(0.0073)	(0.0067)
$D_{i,t-5}\Delta EPS_{i,t-5}$			−0.0746	−0.0499	−0.0833	−0.0566	2.5352	2.5034*
			(0.0970)	(0.0898)	(0.0976)	(0.0900)	(1.5466)	(1.3648)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$					−0.3025	−0.1893	0.9080	1.0006
					(0.7319)	(0.6549)	(1.1343)	(1.0119)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$					0.0285	0.1325	1.2351	1.3169
					(0.7339)	(0.6572)	(1.1393)	(1.0169)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$							−2.6228*	−2.5649*
							(1.5515)	(1.3681)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$							−2.6140*	−2.5523*
							(1.5538)	(1.3693)
Controls	Yes	No	Yes	No	Yes	No	Yes	No
N	70252	70941	70252	70941	70252	70941	70252	70941
(Sargan) χ^2	8.0680	0.0063	8.0239	0.0047	8.1240	0.0070	8.1156	0.0076
(p-value)	(0.1525)	(0.9367)	(0.1549)	(0.9452)	(0.1495)	(0.9350)	(0.1500)	(0.9304)
First-order m-statistic	-11.0414	-24.5654	-10.8990	-24.5680	-10.6283	-24.6111	-10.6097	-24.6105
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Second-order m-statistic	-0.2203	0.9647	-0.2224	0.9655	-0.1774	0.9920	-0.1745	1.0005
(p-value)	(0.8256)	(0.3347)	(0.8240)	(0.3343)	(0.8592)	(0.3212)	(0.8615)	(0.3171)

Table 1.17: Panel Regression of Forecast Bias on High Contradictory Signals Grouped by Deciles. Past Pessimism in Corporate Profits as Prior.

The dependent variable $y_{i,t}$ is calculated as $\frac{TP_{i,t-4}-P_{i,t}}{P_{i,t-4}}$; $TP_{i,t-4}$ is the consensus target price on stock i for the next 4 quarters; the signal $\Delta EPS_{i,t-5}$ is calculated as $\frac{EPS_{i,t-5}-EPS_{i,t-6}}{P_{i,t-6}}$. The dummies $L^W_{i,t-5}$, $L^M_{i,t-5}$ and $L^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is lower or equal than the 3rd, 2nd and 1st cross-sectional deciles respectively. The dummies $H^W_{i,t-5}$, $H^M_{i,t-5}$ and $H^S_{i,t-5}$ take the value of one whenever $\Delta EPS_{i,t-5}$ is higher or equal than the 7th, 8th and 9th cross-sectional deciles respectively. The dummy $D_{i,t-5}$ takes the value of one whenever the consensus forecast on Corporate Profits issued at $t-6$ is less than the actual Corporate Profits at $t-5$. All specifications include $Size_{i,t-5}$ and $CSE_{i,t-5}$ which are exogenous. In the specifications that include the non-exogenous controls (columns (1), (3), (5) and (7)) there are 5 more instruments than regressors and the specifications without these controls have one more instruments than regressors. The instruments for $y_{i,t-1}$ are its the lags from $t-2$ to $t-4$. The instruments used for $\Delta Cvrg_{i,t-4}$ and $\Delta BNH_{i,t-4}$ are their lags from $t-6$ to $t-8$. The instruments for the other variables are the first differences of themselves. ***, ** and * indicate significance at the 0.01, 0.05 and 0.1 levels, respectively. Robust standard errors of Windmeijer (2005) in parenthesis.

Varible	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$y_{i,t-1}$	0.6775***	0.7224***	0.6787***	0.7229***	0.6836***	0.7262***	0.6859***	0.7269***
	(0.0313)	(0.0212)	(0.0314)	(0.0212)	(0.0317)	(0.0212)	(0.0319)	(0.0212)
$y_{i,t-3}$	0.0311***	0.0356***	0.0311***	0.0356***	0.0309***	0.0352***	0.0309***	0.0352***
	(0.0097)	(0.0094)	(0.0097)	(0.0094)	(0.0097)	(0.0093)	(0.0097)	(0.0093)
$\Delta EPS_{i,t-5}$	0.0485	0.0415	0.0782*	0.0614	0.2593	0.1319	−0.9591	−1.0581
	(0.0343)	(0.0334)	(0.0472)	(0.0452)	(0.7319)	(0.6550)	(1.1370)	(1.0114)
$D_{i,t-5}$	−0.0280	0.0975**	−0.0339	0.1029***	−0.0347	0.1016***	−0.0390	0.1019***
	(0.0988)	(0.0387)	(0.0998)	(0.0388)	(0.0998)	(0.0388)	(0.1008)	(0.0388)
$L^W_{i,t-5}$	−0.0096***	−0.0124***	−0.0036	−0.0056	−0.0050	−0.0070	−0.0054	−0.0075
	(0.0037)	(0.0034)	(0.0052)	(0.0049)	(0.0052)	(0.0049)	(0.0053)	(0.0050)
$H^W_{i,t-5}$	0.0153***	0.0106***	0.0183***	0.0142***	0.0137***	0.0097**	0.0135**	0.0095*
	(0.0037)	(0.0035)	(0.0051)	(0.0048)	(0.0052)	(0.0049)	(0.0053)	(0.0050)
$D_{i,t-5}L^W_{i,t-5}$			−0.0132*	−0.0148**	−0.0130*	−0.0146**	−0.0291***	−0.0290***
			(0.0071)	(0.0066)	(0.0071)	(0.0066)	(0.0109)	(0.0101)
$D_{i,t-5}H^W_{i,t-5}$			−0.0058	−0.0074	−0.0049	−0.0066	−0.0021	−0.0092
			(0.0072)	(0.0067)	(0.0073)	(0.0067)	(0.0101)	(0.0092)
$D_{i,t-5}\Delta EPS_{i,t-5}$			−0.0746	−0.0499	−0.0833	−0.0566	2.4906	2.4872*
			(0.0970)	(0.0898)	(0.0976)	(0.0900)	(1.5501)	(1.3656)
$\Delta EPS_{i,t-5}L^W_{i,t-5}$					−0.3025	−0.1893	0.9178	1.0022
					(0.7319)	(0.6549)	(1.1373)	(1.0116)
$\Delta EPS_{i,t-5}H^W_{i,t-5}$					0.0285	0.1325	1.2385	1.3133
					(0.7339)	(0.6572)	(1.1422)	(1.0166)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^S_{i,t-5})$							1.5260**	1.3124**
							(0.6020)	(0.5424)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^M_{i,t-5})$							1.8070	1.8638*
							(1.1194)	(1.0568)
$\Delta EPS_{i,t-5}(D_{i,t-5}L^W_{i,t-5})$							−5.9489***	−5.7569***
							(2.1191)	(1.8895)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^W_{i,t-5})$							−2.1517	−1.4759
							(1.8235)	(1.6332)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^M_{i,t-5})$							−0.9154	−1.1275
							(0.8432)	(0.7446)
$\Delta EPS_{i,t-5}(D_{i,t-5}H^S_{i,t-5})$							0.5007	0.0858
							(0.4678)	(0.4366)
Controls	Yes	No	Yes	No	Yes	No	Yes	No
N	70252	70941	70252	70941	70252	70941	70252	70941
(Sargan) χ^2	8.0680	0.0063	8.0239	0.0047	8.1240	0.0070	7.9669	0.0126
(p-value)	(0.1525)	(0.9367)	(0.1549)	(0.9452)	(0.1495)	(0.9350)	(0.1581)	(0.9106)
First-order m-statistic	-11.0414	-24.5654	-10.8990	-24.5680	-10.6283	-24.6111	-10.5142	-24.6369
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Second-order m-statistic	-0.2203	0.9647	-0.2224	0.9655	-0.1774	0.9920	-0.1764	1.0275
(p-value)	(0.8256)	(0.3347)	(0.8240)	(0.3343)	(0.8592)	(0.3212)	(0.86)	(0.3042)

Steady State Probabilities

Let $p^n_{i,j}$ be the probability of going from state i to state j after n periods. The unconditional probability of state j of a Markov chain, is the steady-state probability defined as:

$$Lim_{n\to\infty}p^n_{i,j} = \mathbb{P}(j)$$

which exists for any irreducible Markov chain. The Markov process controlling the switching between states R and T is

	$State_{t+1} = R$	$State_{t+1} = T$
$State_t = R$	$1 - \lambda_1$	λ_1
$State_t = T$	λ_2	$1 - \lambda_2$

with $\lambda_1 + \lambda_2 < 1$. Notice that the transition matrix, which I will call Ω, is irreducible, that is, there exists some n for which $p^n_{i,j} > 0$ for all i and j (all states communicate). Denoting $P^T = [\mathbb{P}(R) \quad \mathbb{P}(T)]$, the steady-state equations are $P^T \Omega = P^T$ and $\mathbb{P}(R) + \mathbb{P}(T) = 1$ which is equivalent to

$$[\mathbb{P}(R) \quad \mathbb{P}(T)] \begin{bmatrix} 1-\lambda_1 & \lambda_1 \\ \lambda_2 & 1-\lambda_2 \end{bmatrix} = [\mathbb{P}(R) \quad \mathbb{P}(T)]$$

$$\mathbb{P}(R)(1-\lambda_1) + \mathbb{P}(T)\lambda_2 = \mathbb{P}(R)$$

$$\mathbb{P}(R)\lambda_1 + \mathbb{P}(T)(1-\lambda_2) = \mathbb{P}(T)$$

$$\mathbb{P}(R) = 1 - \mathbb{P}(T)$$

And plugging $\mathbb{P}(R) = 1 - \mathbb{P}(T)$, the solution is:

$$\mathbb{P}(T) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

$$\mathbb{P}(R) = 1 - \mathbb{P}(T) = \frac{\lambda_2}{\lambda_1 + \lambda_2}$$

Proofs

Proof of Proposition 1:

The transition matrices for the pairs $(State_{t+j}, y_{t+j})$ are

	$(S_{t+1} = R, \sigma_{t+1} = +y)$	$(S_{t+1} = R, \sigma_{t+1} = -y)$
$(State_t = R, \sigma_t = +y)$	$(1-\lambda_1)\pi_L$	$(1-\lambda_1)(1-\pi_L)$
$(State_t = R, \sigma_t = -y)$	$(1-\lambda_1)(1-\pi_L)$	$(1-\lambda_1)\pi_L$

	$(State_{t+1} = R, \sigma_t = +y)$	$(State_{t+1} = R, \sigma_t = -y)$
$(State_t = T, \sigma_t = +y)$	$\lambda_2\pi_L$	$\lambda_2(1-\pi_L)$
$(State_t = T, \sigma_t = -y)$	$\lambda_2(1-\pi_L)$	$\lambda_2\pi_L$

	$(S_{t+1}=T,\sigma_{t+1}=+y)$	$(S_{t+1}=T,\sigma_{t+1}=-y)$
$(State_t=R,\sigma_t=+y)$	$\lambda_1\pi_H$	$\lambda_1(1-\pi_H)$
$(State_t=R,\sigma_t=-y)$	$\lambda_1(1-\pi_H)$	$\lambda_1\pi_H$

	$(S_{t+1}=T,\sigma_{t+1}=+y)$	$(S_{t+1}=T,\sigma_{t+1}=-y)$
$(State_t=T,\sigma_t=+y)$	$(1-\lambda_2)\pi_H$	$(1-\lambda_2)(1-\pi_H)$
$(State_t=T,\sigma_t=-y)$	$(1-\lambda_2)(1-\pi_H)$	$(1-\lambda_2)\pi_H$

From the first to the last, denote this transition matrices as $\boldsymbol{W_1}$, $\boldsymbol{W_2}$, $\boldsymbol{W_3}$ and $\boldsymbol{W_4}$ respectively:

$$\boldsymbol{W_1} = \begin{bmatrix} (1-\lambda_1)\pi_L & (1-\lambda_1)(1-\pi_L) \\ (1-\lambda_1)(1-\pi_L) & (1-\lambda_1)\pi_L \end{bmatrix} \boldsymbol{W_2} = \begin{bmatrix} \lambda_2\pi_L & \lambda_2(1-\pi_L) \\ \lambda_2(1-\pi_L) & \lambda_2\pi_L \end{bmatrix}$$

$$\boldsymbol{W_3} = \begin{bmatrix} \lambda_1\pi_H & \lambda_1(1-\pi_H) \\ \lambda_1(1-\pi_H) & \lambda_1\pi_H \end{bmatrix} \boldsymbol{W_4} = \begin{bmatrix} (1-\lambda_2)\pi_H & (1-\lambda_2)(1-\pi_H) \\ (1-\lambda_2)(1-\pi_H) & (1-\lambda_2)\pi_H \end{bmatrix}$$

and also let

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{W_1} & \boldsymbol{W_2} \\ \boldsymbol{W_3} & \boldsymbol{W_4} \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} (1-\lambda_1)\pi_L & (1-\lambda_1)(1-\pi_L) & \lambda_2\pi_L & \lambda_2(1-\pi_L) \\ (1-\lambda_1)(1-\pi_L) & (1-\lambda_1)\pi_L & \lambda_2(1-\pi_L) & \lambda_2\pi_L \\ \lambda_1\pi_H & \lambda_1(1-\pi_H) & (1-\lambda_2)\pi_H & (1-\lambda_2)(1-\pi_H) \\ \lambda_1(1-\pi_H) & \lambda_1\pi_H & (1-\lambda_2)(1-\pi_H) & (1-\lambda_2)\pi_H \end{bmatrix}$$

in which $\lambda_1\pi_H = \mathbb{P}(State_{t+j} = T, \sigma_{t+j} = \sigma_t | State_{t+j-1} = R, \sigma_{t+j-1} = \sigma_t)$. In Markov chains, each element of the j^{th} power of $\boldsymbol{Q}$, i.e. each element of $\boldsymbol{Q^j}$, gives the probability after j steps. More precisely, the entry corresponding to row r and column c of $\boldsymbol{Q^j}$ gives the probability that the Markov chain starting in the realization r will be in c after j steps. Notice that $\boldsymbol{Q}$ is a regular Markov chain i.e. some power of the transition matrix has only positive elements (any transition matrix that has no zeros determines a regular Markov chain). Also recall that for this type of chain, it is true that long-range predictions are independent of the starting state.

The analyst's forecast is

$$V_t = a + \mathbb{E}_t \left\{ \frac{\hat{N}_{t+1}}{1+\delta} + \frac{\hat{N}_{t+2}}{(1+\delta)^2} + \dots \right\}$$

Notice that the expected values at period t of the perceived earnings are equivalent to

$$\begin{aligned}
\mathbb{E}_t(\hat{N}_{t+1}) &= \hat{N}_t + \mathbb{E}_t(\sigma_{t+1}) \\
\mathbb{E}_t(\hat{N}_{t+2}) &= \hat{N}_t + \mathbb{E}_t(\sigma_{t+1}) + \mathbb{E}_t(\sigma_{t+2}) \\
&\vdots \\
\mathbb{E}_t(\hat{N}_{t+j}) &= \hat{N}_t + \mathbb{E}_t(\sigma_{t+1}) + \mathbb{E}_t(\sigma_{t+2}) + \cdots + \mathbb{E}_t(\sigma_{t+j})
\end{aligned}$$

then, after introducing expectations and rearranging terms in the forecast equation

$$\begin{aligned}
V_t = a \quad & + \frac{\hat{N}_t}{(1+\delta)^1} + \frac{\hat{N}_t}{(1+\delta)^2} + \frac{\hat{N}_t}{(1+\delta)^3} + \cdots + \frac{\hat{N}_t}{(1+\delta)^j} + \frac{\hat{N}_t}{(1+\delta)^{j+1}} + \cdots \\
& + \frac{\mathbb{E}_t(\sigma_{t+1})}{(1+\delta)^1} + \frac{\mathbb{E}_t(\sigma_{t+1})}{(1+\delta)^2} + \frac{\mathbb{E}_t(\sigma_{t+1})}{(1+\delta)^3} + \cdots + \frac{\mathbb{E}_t(\sigma_{t+1})}{(1+\delta)^j} + \frac{\mathbb{E}_t(\sigma_{t+1})}{(1+\delta)^{j+1}} + \cdots \\
& + \frac{\mathbb{E}_t(\sigma_{t+2})}{(1+\delta)^2} + \frac{\mathbb{E}_t(\sigma_{t+2})}{(1+\delta)^3} + \cdots + \frac{\mathbb{E}_t(\sigma_{t+2})}{(1+\delta)^j} + \frac{\mathbb{E}_t(\sigma_{t+2})}{(1+\delta)^{j+1}} + \cdots \\
& + \frac{\mathbb{E}_t(\sigma_{t+3})}{(1+\delta)^3} + \cdots + \frac{\mathbb{E}_t(\sigma_{t+3})}{(1+\delta)^j} + \frac{\mathbb{E}_t(\sigma_{t+3})}{(1+\delta)^{j+1}} + \cdots \\
& \vdots \\
& + \frac{\mathbb{E}_t(\sigma_{t+j})}{(1+\delta)^j} + \frac{\mathbb{E}_t(\sigma_{t+j})}{(1+\delta)^{j+1}} + \cdots
\end{aligned}$$

and so on. Notice that

$$\begin{aligned}
\frac{\hat{N}_t}{(1+\delta)} + \frac{\hat{N}_t}{(1+\delta)^2} + \frac{\hat{N}_t}{(1+\delta)^3} + \cdots + \frac{\hat{N}_t}{(1+\delta)^j} &= \frac{\hat{N}_t[1 + (1+\delta)^1 + (1+\delta)^2 + (1+\delta)^3 + \cdots + (1+\delta)^{j-1}]}{(1+\delta)^j} \\
&= \hat{N}_t \left[\frac{1-(1+\delta)^j}{(1-(1+\delta))} (1+\delta)^{-j} \right] \\
&= \frac{\hat{N}_t}{\delta} \left[1 - \frac{1}{(1+\delta)^j} \right]
\end{aligned}$$

which converges to $\dfrac{\hat{N}_t}{\delta}$ when $j \to \infty$. After a similar procedure for the expectations of perceived signals, we can express equation 1.4 as

$$V_t = \quad a + \frac{\hat{N}_t}{\delta} + \frac{1}{\delta} \left\{ \mathbb{E}_t(\sigma_{t+1}) + \frac{\mathbb{E}_t(\sigma_{t+2})}{1+\delta} + \frac{\mathbb{E}_t(\sigma_{t+3})}{(1+\delta)^2} + \ldots \right\} \tag{A.1}$$

In order to calculate $\mathbb{E}_t(\sigma_{t+1})$, define Φ_t as the analyst's information set at time t consisting of the perceived earnings shocks $(\sigma_0, \sigma_1, \ldots, \sigma_t)$ and beliefs updating, which can be summarized as $(\sigma_t, \hat{p}_t^R)$, and also denote the joint probabilities, conditional on Φ_t, as

$$\begin{aligned}
\hat{p}_1^{t+j} &= Pr(State_{t+j} = R, \sigma_{t+j} = \sigma_t | \Phi_t) \\
\hat{p}_2^{t+j} &= Pr(State_{t+j} = R, \sigma_{t+j} = -\sigma_t | \Phi_t) \\
\hat{p}_3^{t+j} &= Pr(State_{t+j} = T, \sigma_{t+j} = \sigma_t | \Phi_t) \\
\hat{p}_4^{t+j} &= Pr(State_{t+j} = T, \sigma_{t+j} = -\sigma_t | \Phi_t)
\end{aligned}$$

In addition, let $\hat{\boldsymbol{p}}^{t+j} = (\hat{p}_1^{t+j}, \hat{p}_2^{t+j}, \hat{p}_3^{t+j}, \hat{p}_4^{t+j})'$, $\bar{\boldsymbol{\gamma}}' = (1,0,1,0)$, $\underline{\boldsymbol{\gamma}}' = (0,1,0,1)$, $\boldsymbol{\gamma_1'} = (0,0,1,0)$, $\boldsymbol{\gamma_2'} = (1,0,-1,0)$ and

$$\begin{bmatrix} \hat{p}_t^R \\ 0 \\ 1-\hat{p}_t^R \\ 0 \end{bmatrix} = \hat{\boldsymbol{q}}^t$$

Then we can express, for instance $\mathbb{P}(\sigma_{t+j} = \sigma_t | \Phi_t)$ as $\hat{p}_1^{t+j} + \hat{p}_3^{t+j} = \bar{\boldsymbol{\gamma}}' \hat{\boldsymbol{p}}^{t+j}$ and

$$\begin{bmatrix} \hat{p}_1^{t+j} \\ \hat{p}_2^{t+j} \\ \hat{p}_3^{t+j} \\ \hat{p}_4^{t+j} \end{bmatrix} = \boldsymbol{Q^j} \begin{bmatrix} \hat{p}_t^R \\ 0 \\ 1-\hat{p}_t^R \\ 0 \end{bmatrix} = \boldsymbol{Q^j} \hat{\boldsymbol{q}}^t$$

Importantly, I can use the expressions $\mathbb{P}(\sigma_{t+j} = \sigma_t | \Phi_t) = \bar{\boldsymbol{\gamma}}' \boldsymbol{Q^j} \hat{\boldsymbol{q}}^t$ and $\mathbb{E}_t(\sigma_{t+j} | \Phi_t) = \sigma_t \bar{\boldsymbol{\gamma}}' \boldsymbol{Q^j} \hat{\boldsymbol{q}}^t + (-\sigma_t) \underline{\boldsymbol{\gamma}}' \boldsymbol{Q^j} \hat{\boldsymbol{q}}^t$. Concretely, write equation A.1 as

$$V_t = \quad a + \frac{\hat{N}_t}{\delta} + \frac{1}{\delta} \left\{ \sigma_t \bar{\boldsymbol{\gamma}}' \boldsymbol{Q^1} \hat{\boldsymbol{q}}^{\boldsymbol{t}} - \sigma_t \underline{\boldsymbol{\gamma}}' \boldsymbol{Q^1} \hat{\boldsymbol{q}}^{\boldsymbol{t}} + \frac{\sigma_t \bar{\boldsymbol{\gamma}}' \boldsymbol{Q^2} \hat{\boldsymbol{q}}^{\boldsymbol{t}} - \sigma_t \underline{\boldsymbol{\gamma}}' \boldsymbol{Q^2} \hat{\boldsymbol{q}}^{\boldsymbol{t}}}{(1+\delta)^1} + \frac{\sigma_t \bar{\boldsymbol{\gamma}}' \boldsymbol{Q^3} \hat{\boldsymbol{q}}^{\boldsymbol{t}} - \sigma_t \underline{\boldsymbol{\gamma}}' \boldsymbol{Q^3} \hat{\boldsymbol{q}}^{\boldsymbol{t}}}{(1+\delta)^2} + \dots \right\}$$

Notice that we can factorize σ_t and $(\bar{\gamma}' - \underline{\gamma}')$. Denote $\boldsymbol{\gamma_0}' = \bar{\gamma}' - \underline{\gamma}' = (1,-1,1,-1)$. Then, we can re-write the above expression as

$$V_t = \quad a + \frac{\hat{N}_t}{\delta} + \sigma_t \frac{1}{\delta} \left[\boldsymbol{\gamma_0'} \boldsymbol{Q^1} \hat{\boldsymbol{q}}^{\boldsymbol{t}} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^2} \hat{\boldsymbol{q}}^{\boldsymbol{t}}}{(1+\delta)^1} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^3} \hat{\boldsymbol{q}}^{\boldsymbol{t}}}{(1+\delta)^2} + \dots \right]$$

Also notice that $\boldsymbol{Q^j} \hat{\boldsymbol{q}}^{\boldsymbol{t}} = \boldsymbol{Q^j} \boldsymbol{\gamma_1} + \boldsymbol{Q^j} \boldsymbol{\gamma_2} \hat{p}_t^R$. Thus, the above expression is

$$\begin{aligned}
V_t &= a + \frac{\hat{N}_t}{\delta} + \sigma_t \frac{1}{\delta} \left[\frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^1} \boldsymbol{\gamma_1}}{(1+\delta)^0} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^1} \boldsymbol{\gamma_2} \hat{p}_t^R}{(1+\delta)^0} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^2} \boldsymbol{\gamma_1}}{(1+\delta)^1} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^2} \boldsymbol{\gamma_2} \hat{p}_t^R}{(1+\delta)^1} + \dots \right] \\
V_t &= a + \frac{\hat{N}_t}{\delta} + \sigma_t \frac{1}{\delta} \left[\frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^1} \boldsymbol{\gamma_1}}{(1+\delta)^0} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^2} \boldsymbol{\gamma_1}}{(1+\delta)^1} + \dots + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^1} \boldsymbol{\gamma_2} \hat{p}_t^R}{(1+\delta)^0} + \frac{\boldsymbol{\gamma_0'} \boldsymbol{Q^2} \boldsymbol{\gamma_2} \hat{p}_t^R}{(1+\delta)^1} + \dots \right]
\end{aligned} \tag{A.2}$$

Notice that the sum of the terms next to $\boldsymbol{\gamma_1}$ satisfy

$$\frac{\boldsymbol{\gamma_0' Q^1 \gamma_1}}{(1+\delta)^0} + \frac{\boldsymbol{\gamma_0' Q^2 \gamma_1}}{(1+\delta)^1} + \cdots = \boldsymbol{\gamma_0'}(1+\delta)\left[\frac{1}{1+\delta}\left(\boldsymbol{I} + \frac{\boldsymbol{Q^1}}{(1+\delta)^1} + \frac{\boldsymbol{Q^2}}{(1+\delta)^2} + \cdots\right)\boldsymbol{Q^1 \gamma_1}\right] \tag{A.3}$$

where $\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. We now need the following two lemmas and the following theorem of Markov chains:

Lemma 1: If all eigenvalues r of matrix $\boldsymbol{A}$ satisfy $|r| < 1$, then $(\boldsymbol{I} - \boldsymbol{A})^{-1}$ is well-defined and satisfies $(\boldsymbol{I} - \boldsymbol{A})^{-1} = \boldsymbol{I} + \boldsymbol{A} + \boldsymbol{A^2} + \ldots$

Lemma 2: Let $\boldsymbol{A}$ be a nonnegative matrix with Perron-Frobenius eigenvalue r_0. Then the matrix $(r\boldsymbol{I} - \boldsymbol{A})^{-1}$ is well-defined and positive if and only if $r > r_0$.

Theorem (Fundamental Matrix): For any finite Markov chain $\boldsymbol{W}$, the fundamental matrix $\boldsymbol{M} = (\boldsymbol{I} - \boldsymbol{W})^{-1}$ is well-defined and positive. Its elements can be computed from the converging series $\boldsymbol{M} = \boldsymbol{I} + \boldsymbol{W} + \boldsymbol{W^2} + \ldots$.

Denote $\hat{\boldsymbol{Q}} = \frac{\boldsymbol{Q}}{(1+\delta)}$. If $(1+\delta) > 0$ is larger than the Perron-Frobenius eigenvalue of $\boldsymbol{Q}$, then $\hat{\boldsymbol{Q}}$ has all eigenvalues less than 1 in absolute value. By lemma 1,

$$\frac{1}{(1+\delta)}(\boldsymbol{I} - \hat{\boldsymbol{Q}})^{-1} = \frac{1}{(1+\delta)}\left(\boldsymbol{I} + \frac{\boldsymbol{Q}}{(1+\delta)} + \frac{\boldsymbol{Q}^2}{(1+\delta)^2} + \ldots\right)$$

And by the properties of inverse matrices

$$\frac{1}{(1+\delta)}(\boldsymbol{I} - \hat{\boldsymbol{Q}})^{-1} = \left[(1+\delta)\boldsymbol{I} - \boldsymbol{Q}\right]^{-1}$$

The expression $\left[(1+\delta)\boldsymbol{I} - \boldsymbol{Q}\right]^{-1}$ exists and is positive since every term in the series expansion is nonnegative. This is in fact the proof of lemma 2. By the theorem, we know that that the fundamental matrix of $\frac{\boldsymbol{Q}}{(1+\delta)}$ is well-defined and positive. The proof of the theorem follows from lemmas 1 and 2.

Then we can denote

$$\alpha_1 = \frac{1}{\delta}\left\{\boldsymbol{\gamma_0'}(1+\delta)\left[\left[\boldsymbol{I}(1+\delta) - \boldsymbol{Q}\right]^{-1}\boldsymbol{Q\gamma_1}\right]\right\}.$$

After a similar process for the sum of the terms next to $\boldsymbol{\gamma_2}$ in equation A.2 we have

$$\alpha_2 = \frac{1}{\delta}\left\{\boldsymbol{\gamma_0'}(1+\delta)\left[\left[\boldsymbol{I}(1+\delta)-\boldsymbol{Q}\right]^{-1}\boldsymbol{Q}\boldsymbol{\gamma_2}\right]\right\}$$

and thus

$$V_t = a + \frac{\hat{N}_t}{\delta} + \sigma_t(\alpha_1 + \alpha_2 \hat{p}_t^R)$$

1.9 References

Abarbanell, J. S. and Bernard, V. L. (1992). Tests of analysts' overreaction/underreaction to earnings information as an Explanation for Anomalous Stock Price Behavior. *The Journal of Finance*, 47(3), 1181 - 1207.

Abarbanell, J. and Lehavy, R. (2003). Biased Forecasts or Biased Earnings? The Role of Reported Earnings in Explaining Apparent Bias and Over/Underreaction in Analysts' Earnings Forecasts. *Journal of Accounting and Economics*, 36, 105 - 146.

Ackert, L. F., and Hunter, W. C. (1995). Rational Expectations and Security Analysts' Earnings Forecasts. *Financial Review*, 30(3), 427-443.

Ali, A., Klein, A. and Rosenfeld J. (1992). Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS. *Accounting Review*, 67(1), 183 - 198.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277 - 294.

Arellano, M. and Bover, O. (1995). Another Look at the Instrumental Variables Estimation of Error-Component Models. *Journal of Econometrics*, 68, 29 - 51.

Asquith, P., M. B. Mikhail and A. S. Au. (2005). Information Content of Equity Analyst Reports. *Journal of Financial Economics*, 75(2), 245 - 82.

Barberis, N., Shleifer, A. and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3), 307-343.

Beaver, W. H. (2002). Perspectives on Recent Capital Market Research. *The Accounting Review*, 77(2), 453 - 474.

Ben-Rephael, A., Da, Z. and Israelsen, R. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *Review of Financial Studies*, 30(9), 3009-3047.

Beyer, A. and Guttman, I. (2011). The Effect of Trading Volume on Analysts' Forecast Bias. *The Accounting Review*, 86(2), 451 - 481.

Bonini, S., Zanetti, L., Bianchini, R., and Salvi, A. (2010). Target Price Accuracy in Equity Research. *Journal of Business Finance & Accounting*, 37, 1177 - 1217.

Bradshaw, M. (2002). The Use of Target Prices to Justify Sell-Side Analysts' Stock Recommendations. *Accounting Horizons*, 16(1), 27 - 41.

Bradshaw, M. T., Brown, L. D., and Huang, K. (2013). Do Sell-Side Analysts Exhibit Differential Target Price Forecasting Ability? *Review of Accounting Studies*,18, 930 - 955.

Brav, A., and Lehavy, R. (2003). An Empirical Analysis of Analysts' Target Prices: Short-Term Informativeness and Long-Term Dynamics. *The Journal of Finance*, 58(5) , 1933 - 1967.

Butler, K.C., and Saraoglu, H. (1999). Improving Analysts' Negative Earnings Forecasts. *Financial Analysts Journal*, 48 - 56.

Cen, L., Hilary, G., and Wei, K. (2013). The Role of Anchoring Bias in the Equity Market: Evidence from Analysts' Earnings Forecasts and Stock Returns. *Journal of Financial and Quantitative Analysis*, 48(1), 47 - 76.

Clement, M. B., Hales, J. and Xue, Y. (2011). Understanding Analysts' Use of Stock Returns and Other Analysts' Revisions When Forecasting Earnings. *Journal of Accounting and Economics*, 51, 279 - 299.

Cowen, A., Groysberg, B. and Healy, P. (2006). Which Types of Analyst Firms Are More Optimistic? *Journal of Accounting and Economics*, 41, 119 - 146.

Chen, Q., Goldstein, I., and Jiang, W. (2007). Price Informativeness and Investment Sensitivity to Stock Price. *The Review of Financial Studies*, 20(3), 619-650.

Da, Z., Hong, K. P. and Lee, S. (2016). What Drives Target Price Forecasts and Their Investment Value? *Journal of Business Finance & Accounting*, 43(3) & (4), 487 - 510.

DeBondt, W. F. M. and Thaler, R. (1990). Do Security Analysts Overreact? *The American Economic Review*, 80(2), 52 - 27.

DeLong, B., Shleifer, A., Summers, L., and Waldmann, R. (1990). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, 45(2), 379 - 395.

Easterwood, J. C. and Nutt, S. R. (1999). Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or SystematicOptimism? *The Journal of Finance*, 54(5), 1777 - 1797.

Fischer, P. E. and Stocken, P. C. (2010). Analyst Information Acquisition and Communication. *The Accounting Review*, 85(6), 1985 - 2009.

Fishman, M. J., and Hagerty, K. M. (1989). Disclosure Decisions by Firms and the Competition for Price Efficiency. *The Journal of Finance*, 44(3), 633-646.

Givoly, D., Hayn, C. and Lehavy, R. (2009). The Quality of Analysts' Cash Flow Forecasts. *The Accounting Review*, 84(6), 877 - 1911.

Grossman, S. J., and Stiglitz, J. E. (1976). Information and Competitive Price Systems. *The American Economic Review*, 66(2), 246 - 253.

Gleason, C. A., Bruce Johnson, W., and Li, H. (2013). Valuation Model Use and the Price Target Performance of Sell Side Equity Analysts. *Contemporary Accounting Research*, 30(1), 80-115.

Griffin, D. and Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive psychology*, 24(3), 411-435.

Groysberg, B., Healy, P. M., and Maber, D. A. (2011). What Drives Sell Side Analyst Compensation at High Status Investment Banks? *Journal of Accounting Research*, 49(4), 969-1000.

Hayes, R. M. (1998). The Impact of Trading Commission Incentives on Analysts' Stock Coverage Decisions and Earnings Forecasts. *Journal of Accounting Research*, 36(2), 299 - 320.

Hong, H., Lim, T., and Stein, J. C. (2000). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance*, 55(1), 265 - 295.

Jackson, A. R. (2005). Trade Generation, Reputation, and Sell-Side Analysts. *The Journal of Finance*, 60(2), 673 - 717.

Kartik, N., Ottaviani, M. ans Squintani, F. (2007). Credulity, Lies, and Costly Talk. *Journal of Economic Theory*, 134, 93 - 116.

Koenker, R.W. and d'Orey (1987). Computing Regression Quantiles. *Applied Statistics*, 36, 383 - 393.

Keane, M. P., and Runkle, D. E. (1998). Are Financial Analysts' Forecasts of Corporate Profits Rational? *Journal of Political Economy*, 106(4), 768-805.

Lim, T. (2001). Rationality and analysts' forecast bias. *The Journal of Finance*, 56(1), 369-385.

Loh, R. K. and Stulz, R. M. (2018), Is Sell-Side Research More Valuable in Bad Times? *The Journal of Finance*, 73(3), 959-1013.

Merkley, K., Michaely, R. and Pacelli, J. (2017). Does the Scope of the Sell-Side Analyst Industry Matter? An Examination of Bias, Accuracy, and Information Content of Analyst Reports. *The Journal of Finance*, 72(3), 1285 - 1334.

Mikhail, M. B., Walther, B. R., and Willis, R. H. (1999). Does Forecast Accuracy Matter to Security Analysts? *The Accounting Review*, 74(2), 185-200.

Mikhail, M. B., Walther, B. R., and Willis, R. H. (2007). When Security Analysts Talk, Who Listens? *The Accounting Review*, 82(5), 1227 - 1253.

Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2(2), 175.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, 49, 1417 - 1426.

Niehaus, G. and Zhang, D. (2010). The Impact of Sell-Side Analyst Research Coverage on an Affiliated Broker's Market Share of Trading Volume. *Journal of Banking & Finance*, 34, 776 - 787.

Pouget, S., Sauvagnat, J. and Villeneuve, S. (2017). A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets. *The Review of Financial Studies*, 30(6), 2066 - 2119.

Rabin, M., and J. Schrag. (1999). First Impressions Matter: A Model of Confirmatory Bias. *Quarterly Journal of Economics*, 114, 37 - 82.

Sadler, E. (2021). A Practical Guide to Updating Beliefs From Contradictory Evidence. *Econometrica*, 89(1), 415-436.

Siegel, J. J. (2016). The Shiller CAPE Ratio: A new Look. *Financial Analysts Journal*, 72(3), 41-50.

Stammann, A., Heiss, F. and McFadden, D. (2016). Estimating Fixed Effects Logit Models with Large Panel Data. Working paper.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2), 105.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance*, 57(1), 233-264.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25-51.

Wooldrige, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Chapter 2

The Effects of Investors' Information Acquisition On Sell-Side Analysts Forecast Bias

Abstract: In this research I empirically study the effects of information acquisition by investors or traders on analysts' forecast bias. Based on the theoretical literature on sell-side analysts, I argue that forecast bias is correlated to investors' information gathering in two opposite directions. On the one hand, higher levels of reading activities about individual firms by investors induce analysts to issue more optimistic forecasts if the potential for trading is higher. On the other hand, higher levels of reading activities about individual firms by investors help them identify opportunistic behaviors and thus to discipline analysts. I find that investors' information acquisition is positively related to analysts' optimism when the potential for trading is larger, and negatively related to optimism when investors are more likely to identify inflated forecasts. Together, these results suggest that information acquisition is not only correlated to analysts' optimism but also that its effect does not work trivially and solely in one direction but it activates two different incentives in analysts' decisions.

Key words: trading incentives, analyst's credibility, responsive investors, naive investors.

2.1 Introduction

In many situations, decision makers turn to experts for information and advice. However, experts' preferences or incentives are often not perfectly aligned with those of decision makers, and experts may engage in opportunistic reporting (Meng, 2014). In particular, sell-side analysts tend to produce more inflated forecasts (DeBondt and Thaler, 1990) in order to induce more trading and thus more income for their brokerage houses (Jackson, 2005), as analysts' preferences are misaligned with those of investors. In this paper, I empirically test the hypothesis that the amount of firm-level information collected by investors has two opposite potential effects on analyst forecast bias. I argue that, the analyst acts strategically so as to take advantage of investor heterogeneity i.e. of the simultaneous presence of naive and sophisticated investors, but is restrained from systematically inducing a greater number of trades by making over optimistic forecast because he also has to address reputation effects especially from sophisticated investors.

There are differences in investors' ability to establish the quality of analysts' reports, and thus, investors react rather heterogeneously to the reports issued by analysts. Particularly, small investors trade more in response to optimistic reports than to pessimistic reports (Mikhail, Walther and Willis, 2007), while institutional investors trade more in response to conservative analysts (Hugon and Muslu, 2010) and assign votes to more accurate analysts in the *Institutional Investor*'s All-American Research Team ranking (Groysberg, Healy and Maber, 2011; Stickel, 1992). This heterogeneity in reaction to reports, has been already incorporated in the theoretical literature, showing that sell-side analysts should react strategically to investors' responses to their reports (Fischer and Stocken, 2010; Kartik, Ottaviani and Squintani, 2007). In particular, Kartik, Ottaviani and Squintani (2007) (hereafter KOS) study if an analyst incur in opportunistic behaviors when there exist investors who are heterogeneous in their ability to establish the quality of analysts' reports, and show that the presence of some naive investors induce inflated forecasts.

To study the relation between heterogeneous sophistication and inflated forecasts, KOS propose a cheap-talk model where an analyst has private information about the performance of a firm upon which he issues a report, and the state of nature decides on the true state of the world. Since it is not the main purpose of this model to study the analyst's decision on the quality of his private information but to formalize the idea that the presence of some naive investors induce inflated forecasts, it is assumed in the model that the private information of the analyst, equals the true state of the world regarding the firm's performance and that there is not a public signal. The analyst is interested in the average response of a pool of investors and there is a fraction of naive investors who cannot formulate equilibrium beliefs about the true state of the world from the analyst report, but they formulate a dis-equilibrium estimate of the true state of the world. Meanwhile, the fraction of strategic investors, formulate an equilibrium estimate and take action accordingly.

Lemma 1 and Theorem 1 of KOS state that there are conditions for which there exists an equilibrium where the analyst sends an inflated report (higher than the true state of the world) in order to induce a larger response from naive investors, even when among investors there are sophisticated ones. In other words, it is feasible a situation in which there are optimistic reports which “still reveal precise information to strategic receivers, while deceiving naive receivers.” When investors are more responsive to forecasts, this model tells us that analysts will issue more optimistic reports, and thus, I argue that greater information acquisition on stocks with greater trading potential leads to higher analyst optimism.

The cheap talk literature has studied how concerns about the future, affect communication, including reputational and career concerns (e.g. Ottaviani and Sørensen, 2006; Ely and Välimäki, 2003; Jullien and Park, 2014). However, here the question is if information acquisition by investors plays a significant role in limiting opportunistic reporting. Fischer and Stocken (2010) (hereafter FS) show that, when the quality or precision of the analyst report is common knowledge, higher investors' information acquisition induces more precise analysts' reports. Unfortunately, investors do not have a precise idea about analyst's forecast precision, but they can establish when analysts' reports are consistent with the firm's performance. Therefore, greater information acquisition on stocks for which investors are able to recognize forecast quality, should be associated to more precise - less biased analyst forecasts.

Letting the quality of the report being of common knowledge in the model of FS, provides results on how communication incentives affect the analyst's decisions on the quality of his private information upon which he issues a report. In the cheap-talk model of investors' information acquisition and forecast precision with truthful communication proposed by FS, there is an analyst who issues a report with a binary outcome, *bad* or *good*, upon his private information about the performance of a firm. Then, an investor takes action, based on both the analyst' report and a public signal (e.g. a corporate report) about the state of the world regarding the firm's performance. In this setting, the authors show that an increase in the informativeness of investors' signals, requires that the analyst augments the quality of his private information in order to maintain the investor responsive to the analyst's report[1].

In this paper, I empirically test the hypothesis that the amount of firm-level information collected by investors has two opposite potential effects on analyst forecast bias. First, analysts may issue more optimistic (inflated) forecasts for stocks to which investors are paying more attention when there is greater potential to generate trading. Second, analysts may issue less inflated forecasts to investors who have sought for more firm-level information when they have a greater ability to establish whether the analysts'

[1] This is Corollary 6 of FS.

forecasts are consistent with the firm's prospects or consistent with opportunistic behavior, so that the higher the acquisition of information at the firm level by investors, the lower the optimism in analysts' forecasts. In other words, the amount of firm-level information collected by investors is important to discipline analysts. I find that forecast bias is higher when firm-level information acquisition by investors increases, for stocks with more potential for trading businesses. Moreover, forecast bias decreases when investors acquire more firm-level information, for stocks followed by investors with a greater potential of identifying inflated forecasts. Thus, this research highlights that higher investors' reading activities are not related to forecast bias in one exclusive direction, and may deteriorate or improve the decision-making process of naive investors. Also, this study shows that the effects of investors' information acquisition is not trivial as one would expect: investor sophistication and information acquisition are not equivalent concepts, and the mechanism through which the differences in investor sophistication induce greater forecast bias, is information acquisition by less sophisticated investors.

This paper adds to the empirical literature studying analyst's trade-offs of career concerns and short-term benefits (e.g. Fang and Yasuda, 2009) and is also related with research on private investor's information acquisition (e.g. Ben-Rephael, Da and Israelsen, 2017; Chi and Shanthikumar, 2018). Furthermore, this paper adds to research studying the role of institutional investors on forecast accuracy (e.g. Ljungqvist et al., 2007). To study investors with a greater ability to establish when analysts' reports are consistent with the firm's performance, and following the research showing that institutional investors are more sophisticated (Hilary and Hsu, 2013; Boehmer and Kelley, 2009), I use stocks in the holdings of investment managers with more than 100 million USD in equity under management. Furthermore, I use firms in the Financials sector, which I argue, issue larger amounts of hard information. Also, I measure information acquisition by investors based on the level of activity at Bloomberg Terminals, calculating the quarterly changes of daily averages the "News Heat - Daily Max Readership" index of Bloomberg. In order to analyze stocks with a greater potential of trading, I use firms in the Consumer Goods sector, which I argue, are more likely to call the attention of a wider public and have characteristics that match a story of undervaluation in analysts' reports, thus stimulating the interest of a large number of investors. In addition, as non-robot EDGAR users tend to be retail investors (Chi and Shanthikumar, 2018; Loughran and McDonald, 2017; Asthana, Balsam and Sankaraguruswamy, 2004), and the FINRA Foundation's survey results show that many users of open access sources of financial information are likely to be naive, which is in line with the fact that small investors are more responsive to optimistic reports (Mikhail, Walther and Willis, 2007), studying the changes in searches at EDGAR for these firms, allows me to analyze the differential influence of investors' information acquisition on analyst forecast bias, for investors with a higher potential to generate trading. Therefore, I also calculate the quarterly changes of the number of non-robot downloads of EDGAR filings for each stock.

There are six sections in this paper including the introduction. In section two I expose the theoretical and empirical literature related to inflated forecasts, sophistication and investor reaction to analyst reports. Afterwards, in section three, I describe the data and the variables. In sections four and five I explain the empirical strategy and the results respectively. Finally, in section six I conclude.

2.2 Empirical Strategy

I test for two possible stories. First, the amount of firm-level information in the hands of investors may help reduce forecast bias, since analysts must issue more precise forecasts to keep investors responsive to their reports, when they have sought for more firm-level information. Then, the higher the acquisition of information at the firm level by investors, the lower the bias in analysts' forecasts. By contrast, if greater amounts of firm-level information sought by investors is interpreted by analysts as investors having more interest on some stocks, then analysts may issue more optimistic (inflated) forecasts in order to generate more commissions for their brokerage houses.

My specification is the following:

$$\begin{aligned} y_{i,t} = \quad & c_i + \textstyle\sum_{j=1}^{4} \rho_j y_{i,t-j} \\ & + \tau_0 \Delta BNH_{i,t-4} \\ & + \tau_1 (D_{i,CG} \times \Delta BNH_{i,t-4}) \\ & + \tau_2 (D_{i,F} \times \Delta BNH_{i,t-4}) \\ & + \boldsymbol{x}_{i,t-5} \boldsymbol{\theta} + \boldsymbol{h}_{i,t-4} \boldsymbol{\omega} + \lambda_t + \varepsilon_{i,t} \end{aligned} \tag{1}$$

where c_i is a stock-level unobserved effect, λ_t are time effects common to all firms and $\varepsilon_{i,t}$ is the error term. My dependent variable is the quarterly forecast bias in terms of optimism in target prices. For each firm i and quarter t, I calculate the forecast bias as

$$y_{i,t} = \frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$$

where $TP_{i,t-4}$ is the consensus forecast or target price, issued at the end of the quarter $t-4$ for the next 4 quarters on stock i and $P_{i,t}$ is the stock price at the end of the quarter t. Notice that $y_{i,t}$ is a very intuitive measure of optimism since it equals the difference between the projected growth in price $\frac{TP_{i,t-4}}{P_{t-4}}$ and the realized growth $\frac{P_{i,t}}{P_{i,t-4}}$.

As changes in aggregate economic activity may induce changes in forecast optimism, I inspect the behavior of the market forecast bias through time. In figure 2.1, I show periods of low economic activity (shaded areas), estimated as the quarters during which the Chicago Fed National Activity Index (CFNAI) went from positive to negative[2]. For each quarter, I calculate the market (equally weighted) forecast bias as the cross-sectional average of the forecast bias (black solid line). As we can see from the figure, there are no trends in the aggregate bias or patterns of behavior during quarters of lower economic activity.

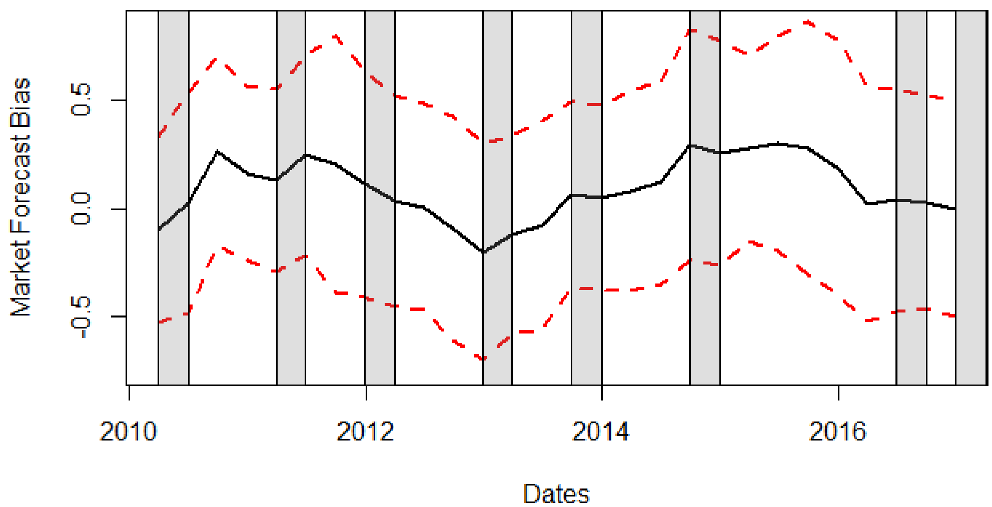


Figure 2.1: Market forecast bias calculated as the cross-sectional average of the forecast biases. Red dashed lines represent one standard deviation from the mean. Shaded areas are quarters during which the Chicago Fed National Activity Index (CFNAI) went from positive to negative until the first quarter of 2017.

2.2.1 Information Acquisition

My independent variable of interest $\Delta BNH_{i,t-4}$, corresponds to changes in investors' information gathering, and equals the dynamic change in the quarterly average of the "News Heat - Daily Max Readership" index of Bloomberg:

$$\Delta BNH_{i,t-4} = BNH_{i,t-4} - BNH_{i,t-5}$$

where $BNH_{i,t-4}$ is the daily average of the Bloomberg's index during quarter $t-4$. I argue that activities at Bloomberg terminals capture the information gathering by investors. The "News Heat - Daily Max Readership" index (BNH) is constructed by Bloomberg based upon the "number of times each article is read by its users, as well as the number of times users search for news for a specific stock" (Ben-Rephael,

[2]The CFNAI is a "monthly index designed to gauge overall economic activity and related inflationary pressure" in the U.S. It is constructed to have an average value of zero and a standard deviation of one. A negative index reading corresponds to growth below trend.

Da and Israelsen, 2017) and takes higher values for higher levels of readers activity going from 0 to 4. As documented by Ben-Rephael, Da, and Israelsen (2017), as of August 26, 2016, around 80% of Blommberg Terminal users worked in financial industries (including banking, asset management, and institutional financial services) with 32% of the job titles being portfolio managers or traders, 19% presidents or directors, and only 17% being analysts, including buy-side (who use sell-side analysts valuations and advise portfolio managers privately only) and sell-side ones. I show in tables 2.1 and 2.2 that this index is not correlated to analysts forecast revisions. In table 2.1 I show the tests of correlation between forecast updates ($TP_{i,t} - TP_{i,t-1}$) and ΔBNH, from which we cannot find evidence of correlation. In addition, in table 2.2 I show the results of a dynamic linear probability model, in which the dependent variable is a dummy $D_{i,t}^{update}$ that takes the value of one whenever the changes in stock price forecasts are less than zero. While the relationship between the probability of a downward forecast revision and $\Delta eps_{i,t-1}$ is negative, from this regression is not possible to reject the null hypothesis of zero correlation between forecast updates and $\Delta BNH_{i,t}$ or $\Delta BNH_{i,t-1}$. Added to the fact that most of Bloomberg Terminal users are in jobs related to the buy side (portfolio managers, traders, etc.), these statistics reinforce the idea that the " News Heat" index captures the information gathering by investors.

Table 2.1: Correlations Between Forecast Changes and ΔBNH
Forecast changes are $TP_{i,t} - TP_{i,t-1}$.

	sample estimates	t	p-value	95% percent confidence interval
$\Delta BNH_{i,t}$	0.005910115	1.2643	0.2061	-0.0032 - 0.0151
$\Delta BNH_{i,t-1}$	-0.007173707	-1.5347	0.1249	-0.0163 - 0.001988

Table 2.2: Linear Probability Model.
Dependent: forecast update $D_{i,t}^{update}$. In the difference equation, the valid instruments for $D_{i,t-1}^{update}$ are $D_{i,t-2}^{update}$ and $D_{i,t-3}^{update}$; for $\Delta BNH_{i,t}$ its lags from $t-2$ to $t-5$. The Arellano-Bond m-statistics for first and second-order autocorrelation indicate a good fit with valid instruments. *Note:* *p<0.1; **p<0.05; ***p<0.01.

	Dependent variable:
	$D_{i,t}^{update} = 1 \quad for \quad TP_{i,t} - TP_{i,t-1} < 0$
$D_{i,t-1}^{update}$	0.128***
	(0.012)
$\Delta eps_{i,t-1}$	-0.106**
	(0.043)
$\Delta BNH_{i,t}$	0.195
	(0.157)
$\Delta BNH_{i,t-1}$	0.039
	(0.037)
A-B m-statistic (1st): -24.6873 (*p-value* $\leq$ 2.22e-16)	
A-B m-statistic (2nd): -0.9461 (*p-value* $=$ 0.34407)	

I also look at the behavior of the BNH through time in figure 2.2. Specifically, for each day of the available data, I calculate the cross-sectional average of the BNH or the market (equally weighted) BNH

index. Also, in figure 2.2 I show the Cboe Volatility Index (VIX), which is designed to produce a measure of constant, 30-day expected volatility of the U.S. stock market. Since the index for news seeking could be capturing changes in the uncertainty and volatility of the stock market, it is interesting to analyze whether these two variables hold an obvious relation. As we can see from the figure, there are no suggestive patterns between the stock market volatility and the BNH index. Moreover, for each time series, the Augmented-Dickey Fuller test rejects the hypothesis that these are non-stationary at the 1% level of significance, and the Granger test does not reject the hypothesis that there is no Granger-causality with lags of 30 and 120 days.

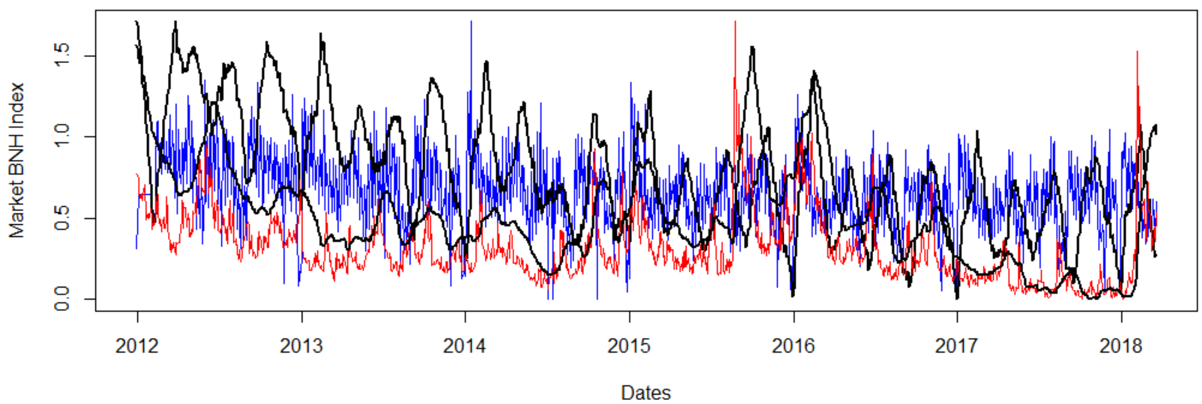


Figure 2.2: Daily market BNH (blue), Cboe Volatility Index, VIX (red), and 30-day moving averages (black) from 2011-12-30 to 2018-03-20.

I include among my dependent variables $\Delta edgar_{i,t-4}$, which is the quarterly change of the number of non-robot downloads of EDGAR filings from SEC.gov for each stock. More precisely, for each stock, I add up the number of downloads registered in one day. With the daily data, I use the average of downloads of the respective quarter for the corresponding stock. I identify the tickers of the stocks from their CIK codes[3] and name this variariable as $edgar_{i,t}$, thus, the quarterly change I use is

$$\Delta edgar_{i,t-4} = edgar_{i,t-4} - edgar_{i,t-5}$$

In the following table I show as before, the results of a dynamic linear probability model, in which the dependent variable is a dummy $D_{i,t}^{update}$ that takes the value of one whenever the changes in stock price forecasts are less than zero. Similar to the regression in table 2.2, the estimate on $\Delta eps_{i,t-1}$ is statistically negative, and the estimates on $\Delta edgar_{i,t}$ and $\Delta edgar_{i,t-1}$ are not statistically different than zero.

[3]https://www.sec.gov/include/ticker.txt

Table 2.3: Linear Probability Model.

Dependent: forecast update $D^{update}_{i,t}$. In the difference equation, the valid instruments for $D^{update}_{i,t-1}$ are $D^{update}_{i,t-2}$ and $D^{update}_{i,t-3}$; for $\Delta edgar_{i,t}$ its lags from $t-2$ to $t-5$. The Arellano-Bond m-statistics for first and second-order autocorrelation indicate a good fit with valid instruments. *Note:* *p<0.1; **p<0.05; ***p<0.01.

	Dependent variable:
	$D^{update}_{i,t} = 1 \quad for \quad TP_{i,t} - TP_{i,t-1} < 0$
$D^{update}_{i,t-1}$	0.131^{***}
	(0.009)
$\Delta eps_{i,t-1}$	-0.118^{**}
	(0.057)
$\Delta edgar_{i,t}$	-0.0002
	(0.0004)
$\Delta edgar_{i,t-1}$	-0.0001
	(0.0002)
A-B m-statistic (1st): -32.41349 (*p-value* $\leq$ 2.22e-16)	
A-B m-statistic (2nd): -0.671849 (*p-value* $=$ 0.50168)	

I use $\Delta edgar_{i,t-4}$ in order to capture information acquisition by non-robot non-institutional investors who are likely to be less sophisticated than those captured by Bloomberg searches[4]. The Investor Survey of the FINRA Foundation's 2018 National Financial Capability Study shows that while it is true that free online services provide many investors (44% of respondents) with information potentially useful for their decisions, many of the users of these open access sources of financial information are likely to be more naive or less sophisticated, as 13.4% of them think they do not pay any kind of fee for investing, 9.75% do not know how much they pay and 28,7% do not know whether any of their investment accounts allow them to make purchases on margin. The corresponding numbers, for those who reported using paid subscription services diminish to 10.1%, 5.2% and 12.7%. Da, Engelberg and Gao (2011), argue that institutional investors access information services that are more sophisticated, such as Reuters or Bloomberg terminals, whereas less sophisticated investors are more likely to obtain financial information from free sources, which is in line with Ben-Rephael, Da and Israelsen (2017) who find that their measure of institutional investor attention, based on Bloomberg searches, leads retail attention but not vice versa. The analysis of the traffic statistics at the EDGAR system carried out by Loughran and McDonald (2017), shows that non-robot investors mostly request filings of widely followed companies such as Facebook. This is consistent with research showing that retail trading is correlated to EDGAR activity (Chi and Shanthikumar, 2018; Asthana, Balsam and Sankaraguruswamy, 2004) and suggests that the counting of non-robot downloads of EDGAR filings captures information acquisition by less sophisticated investors, relative to Bloomberg Terminal users.

[4] I do not claim that all Bloomberg Terminal users are sophisticated according to an absolute measure, or that all EDGAR users are naive according to an absolute rule, but that non-robot EDGAR users tend to be less sophisticated relative to Bloomberg users.

In figure 2.3 I present daily percentages of robot downloads of EDGAR filings from 2006-05-11 to 2015-12-31 for the aggregate of stocks. Interestingly, this percentage is high and as we move forward in time, not only the percentage increases but it becomes less volatile[5]. In table 2.4 I show the summary statistics for this same percentage of robot downloads from which we can see that most of the traffic in the SEC.gov web page corresponds to non-human downloads. This shows the importance of filtering out the data on robot downloads.

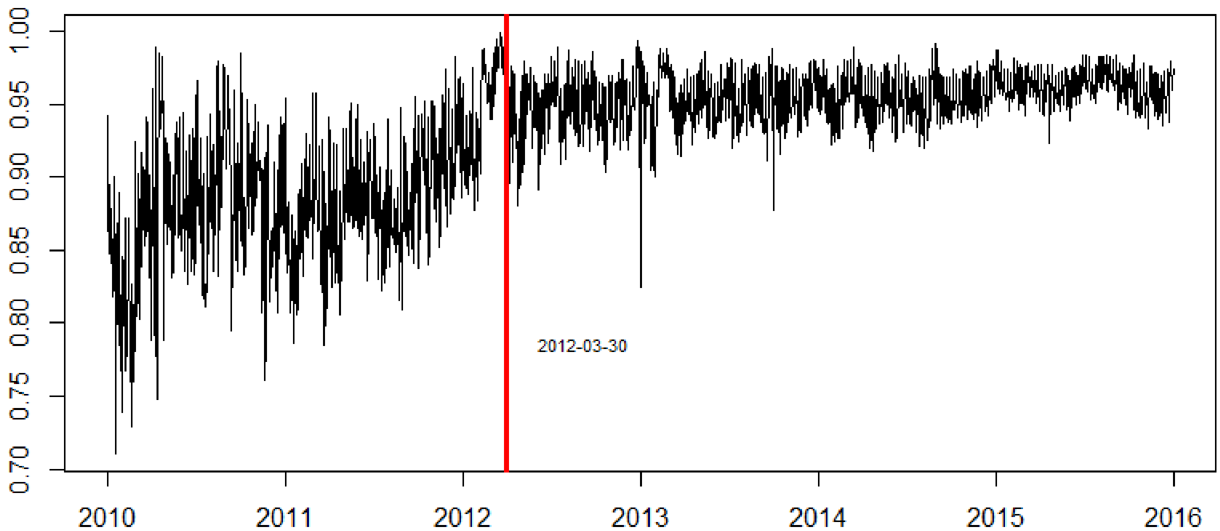


Figure 2.3: Daily percentages of robot downloads of EDGAR filings from 2006-05-11 to 2015-12-31 for the aggregate of the stocks provided by The Software Repository for Accounting and Finance of the University of Notre Dame. Files from 2005-09-24 to 2006-05-10 were lost or damaged. Robot downloads for specific stocks are not available.

Table 2.4: Summary Statistics On The Percentage of Robot Downloads of EDGAR Filings. Daily from 2006-05-11 to 2015-12-31.

Calculated as the number of robot downloads divided by the sum of non-robot and robot downloads.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.4899	0.8126	0.9035	0.8763	0.9520	0.9994

Author's calculations with data from The Software Repository for Accounting and Finance of the University of Notre Dame.

[5]Including a dummy that takes the value of one for periods after 2012, interacted with human downloads at EDGAR, do not change the results.

2.2.2 Investors' Recognition of Inflated Forecasts

In my specification, I use a sector dummy for financial firms ($D_{i,F}$), according to the sector indexes developed by the CRSP, which are based on the Industry Classification Benchmark (ICB) of FTSE International to assign companies to sectors. I define the dummy as

$$D_{i,F} = \begin{cases} 1 & if \quad stock\ i \in CRSP\ Financials\ Index \\ 0 & otherwise \end{cases}$$

Financial firms are heavily regulated and supervised (Goldsmith-Pinkham, 2016; Hugonnier and Morellec, 2017; Gunther and Moore, 2003), among others, by the Federal Reserve, Federal Deposit Insurance Corporation (FDIC), the Office of the Comptroller of the Currency (OCC) and the National Credit Union Administration (NCUA), who asses firms' financial health (safety and soundness)[6] and activities related to (anti)money laundering and consumer protection legislation. Since financial institutions are required to issue larger amounts of hard information on their performance, it may be easier for investors to identify those sell-side analysts issuing inflated forecasts for these firms.

Moreover, following the empirical literature showing that institutional investors are more sophisticated (Hilary and Hsu, 2013; Boehmer and Kelley, 2009), as an additional proxy to capture the presence of investors with a higher ability to identify opportunistic behavior, I use the dummy $D_{i,13F}$ which equals one whenever stock i is included in the list of section 13(f) securities reported in 2017Q4. The list of section 13(f) securities reports those securities in the holdings of investment managers with more than 100 million USD in equity under management[7].

$$D_{i,13F} = \begin{cases} 1 & if \quad stock\ i \in list\ of\ section\ 13(f)\ securities \\ 0 & otherwise \end{cases}$$

2.2.3 Stocks With Trading Potential

Furthermore, stocks issued by firms in the Consumer Goods sector ($D_{i,CG}$) are more likely to call the attention and be held by a wider public or have a major part in investors' portfolios, and thus are more likely to have a higher trading potential. As I show in table 2.5 Panel A, market capitalization of firms in the Consumer Goods sector is larger than the market capitalization of firms in Financials and the overall sample. Larger firms are more widely held and stimulate the interest of a large number of investors with more potential transactions business (Atiase, 1985; Bushan, 1989). Also, larger firms receive more

[6] The supervisory authorities ask for information related to capital adequacy (available capital v.s risk-weighted credit exposures), asset quality (loan's quality), management's ability to ensure the safe operation, earnings, liquidity and sensitivity to particular risk exposures. See CAMELS ratings.

[7] See https://www.sec.gov/fast-answers/answers-form13fhtm.html

coverage in the business media (Fang and Peress, 2009), and media coverage catches investors attention (Solomon, Soltes and Sosyura, 2014; Engelberg and Parsons, 2011; Tetlock, Saar-Tsechansky and Macskassy, 2008) impacting financial decision making, independently of the information conveyed (Kaniel and Parham, 2017)[8]. In addition, analysts can more easily justify inflated forecasts on stock prices issued by firms in the Consumer Goods sector. In table 2.5 Panels B and C we can observe that firms in the Consumer Goods sector not only have better fundamentals (greater return on assets) but also their stocks register, on average, lower Price-to-Earnings ratios which fit in a story of undervaluation in analysts' reports.

$$D_{i,CG} = \begin{cases} 1 & if \quad stock\ i \in CRSP\ US\ Consumer\ Goods\ Index \\ 0 & otherwise \end{cases}$$

Table 2.5: Summary Statistics for Two Sectors (Not Exhaustive)

Panel A. Market Capitalization.
Calculated as shares outstanding times the last price. Reported in millions.

	Median	Mean	N^o Firms in CRSP
Financials	2,273.26	8,860.91	759
Consumer Goods	2,510.4	13,065.9	322
All	2,266.4	10,502.6	2,542

Panel B. PE ratios. Calculated as $\frac{P_t}{EPS_t}$.

	Median	Mean	N^o Firms in CRSP
Financials	58.14	71.63	759
Consumer Goods	59.22	50.78	322
All	59.10	67.46	2,542

Panel C. Return on assets. Calculated as $\frac{Earnings_t}{Assets_t}$.

	Median	Mean	N^o Firms in CRSP
Financials	0.033432	0.009414	759
Consumer Goods	0.04145	0.03428	322
All	0.03721	0.01321	2,542

Panel D. Forecast bias. Calculated as $\frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$.

	Median	Mean	N^o Firms in CRSP
Financials	0.03796	0.09962	759
Consumer Goods	0.01851	0.05577	322
All	0.04082	0.09685	2,542

Author's calculations.

[8] Also, firms covered by the media show stronger momentum (Hillert, Jacobs and Müller, 2014) and firm size is related to favorable prospects (Ramnath, Rock and Shane, 2008; Hayes, 1998; McNichols and OBrien, 1997).

2.2.4 Controls

In my specification I control for variables of firm performance such as changes in earnings per share and return on assets. More explicitly, in equation 1, the vectors $\boldsymbol{x}_{i,t-5}$ and $\boldsymbol{h}_{i,t-4}$ contain information on earnings, return on assets, stock returns, volatilities and firm size. I denote these variables using lower-case letters as $\Delta eps_{i,t-4}$ and $roa_{i,t-4}$, which correspond to changes in Earnings Per Share scaled by the stock price, i.e. $\frac{EPS_{i,t-4}-EPS_{i,t-5}}{P_{i,t-5}}$ and the firms' Return on Assets ($\frac{Earnings_{i,t-4}}{Assets_{i,t-4}}$) respectively. I additionally include firm size as conventionally measured in the financial literature, i.e. as the log of market capitalization ($size_{i,t-4}$), also past stock returns ($return_{i,t-4}$) and volatility on returns ($sd_{i,t-4}$) estimated as the quarterly standard deviation of daily returns. Notice that $\Delta eps_{i,t-4}$ and $roa_{i,t-4}$ are variables related to fundamental or intrinsic value; $size_{i,t-4}$ is related to trading potential at the firm level; $return_{i,t-4}$ is a variable related to momentum; and $sd_{i,t-4}$ is related to stock risk. Thus, $\boldsymbol{x}_{i,t-5}$ contains either $\Delta eps_{i,t-5}$ or $roa_{i,t-5}$; $\boldsymbol{h}_{i,t-4}$ contains $return_{i,t-4}$, $sd_{i,t-4}$ and $size_{i,t-4}$; and $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ are vectors of parameters on regressors[9].

The term $\Delta BNH_{i,t-4}$ is endogenous, since $y_{i,t}$ is a function of $TP_{i,t-4}$ and higher target prices issued by analysts at $t-4$ may influence investors' attention at $t-4$. Thus, $\Delta BNH_{i,t-4}$ is correlated with the error term ($\varepsilon_{i,t}$). In addition, because of the unobserved fixed effect, I estimate the model in differences, in which, by construction, $\Delta\varepsilon_{i,t}$ and $\Delta y_{i,t-1}$ are correlated. As is standard in dynamic panel data, I deal with the endogeneity problems using instrumental variables and Arellano and Bond's (1991) 2SGMM estimators. In particular, the set of potential valid instruments for $y_{i,t-1}$ is composed of its first lag $y_{i,t-2}$ and higher and the set for $\Delta BNH_{i,t-4}$ are $\Delta BNH_{i,t-5}$ and higher orders, whose validity I test using Arellano and Bond's (1991) m-statistic. Also notice that other regressors are exogenous, since quarterly reported earnings and assets are not determined by stock forecasts but by the accounting revenues, costs and purchases of firms, and thus earnings per share and return on assets are exogenous to forecast bias. Similarly, it is not likely that the amount of information gathered by investors about a firm, during quarter t, affects the financial statements of that firm at the same quarter t.

2.3 Data

For 2542 firms included in the Center for Research in Security Prices (CRSP) stock index from the first quarter of 2010 to the fourth quarter of 2017, I observe the quarterly series of Earnings Per Share (EPS) and Return On Assets (ROA) of each firm, as well as daily data on its stock price and market capital-

[9] Analysts' optimism is linked to past signals of fundamental value such as changes in Earnings Per Share (Da, Hong and Lee, 2016; Bradshaw, 2002; Easterwood and Nutt, 1999; Abarbanell and Bernard, 1992; Ali, Klein and Rosenfeld, 1992) and also to market outcomes such as past stock returns (Ali, Klein and Rosenfeld, 1992; Abarbanell, 1991) and volatility on returns (Aslan and Kumar, 2017; Lim, 2001). In addition, the empirical literature shows that analysts' decide to report forecasts selectively, based on whether the firm has a favorable prospect.

ization. Also, I identify Financials and Consumer Goods firms, according to the Industry Classification Benchmark (ICB) of FTSE International which counts with 10 sector indexes in total. In the sector of Consumer Goods sector there are firms producing household goods (clothing, electronics, etc.), automobiles, foods and beverages; among Financials, there are banks and firms related to insurance, asset management and real estate investment trusts (REITs). In addition, I observe daily data on the consensus target price, which is the average forecast of the stock price for the next 12 months from the analysts who cover that stock, and excludes forecasts older than three months when it is calculated. Forecasts on stock prices express analysts' opinions about the stock market in the most direct and intuitive manner[10], without the statistical problems that raise from earnings management[11] when using earnings forecasts or operating cash flows[12] to capture optimism.

Moreover, I use information on the number of non-robot downloads of EDGAR filings through SEC.gov. The Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system is a public database with free access used at the U.S. Securities and Exchange Commission (SEC) which allow users to search for financial information and operations of public companies, mutual funds and exchange-traded funds among others. The Division of Economic and Risk Analysis (DERA) constructed the EDGAR log file data set containing statistics on user access to the SEC.gov website and is "intended to provide insight into the usage of publicly accessible EDGAR company filings". Among other variables, the EDGAR log file data, which counts with available information from 2003, includes the IP addresses, dates, times, browser type, Central Index Key (CIK) codes and whether the user self-identified as a crawler. [13] I use a filtered version of the data from The Software Repository for Accounting and Finance[14] of the University of Notre Dame, with information availability from 2006 to 2015 which is an extension of the information used in the paper of Loughran and McDonald (2017). This data have been filtered to eliminate damaged files[15], irrelevant entries, or those with missing CIK, accession number, IP, or date. Also, robot downloads (those with more than 49 downloads from a single IP within a single day or self identified as a web crawler), or with a server code larger or equal than 300, or records of traffic on the index page of a set of documents (e.g. index.htm) have been filtered out. The information set I use in this paper have a size of 8.55 GB and

[10]This is supported by Asquith, Mikhail and Au (2005) who find that "the market reaction to price target revisions is stronger than that of an equal percentage change in earnings forecasts."

[11]Earnings management refers to the fact that "[m]anagement can improve or impair the quality of financial statements through the exercise of discretion over accounting numbers" (Beaver, 2002), e.g. estimation of accruals. Therefore, "some 'errors' in the distribution of [analyst] forecast errors may arise only because the forecast was inappropriately benchmarked with reported [manipulated] earnings, when in fact the analyst had targeted a different earnings number" (Abarbanell and Lehavy, 2003).

[12]Givoly, Hayn and Lehavy (2009) find that "cash flow forecasts appear to be a naïve extension of analysts' earnings forecasts."

[13]In the SEC website, www.sec.gov/dera/data/edgar-log-file-data-set.html, there are 2,880 zip folders in the EDGAR log file data set for the period 2010 - 2017, one for each day, and each folder contains a "csv" file with the statistics on SEC.gov website traffic and a "README" file documenting the variables. Each file (day) downloaded directly from the SEC website with the statistics on internet search traffic, can include more than a million entries. The uncompressed data set for the period 2003-2015 consisting of 4,839 daily files, takes 1.73 terabytes.

[14]https://sraf.nd.edu/

[15]All files from 2005-09-24 to 2006-05-10 were labeled by the SEC as "lost or damaged" (Loughran and McDonald, 2017).

includes daily data, from 2010 to 2015, of the number of (exclusively) non-robot downloads for each stock, identified with the CIK number, with their respective dates. The advantage of using the EDGAR log file data set from The Software Repository for Accounting and Finance is that robot downloads of financial information have been filtered out, filtering that is not possible to carry out when analyzing data related to traffic on other free sources of information.

I also gather data on institutional investors from the list of section 13(f) securities, issued by the U.S Securities and Exchange Commission[16]. These are quarterly reports available in pdf format, from which I use the report of 2017Q4. I identify the 504 stocks in my sample that are included in this report, which counts with more than 17,000 individual securities.

I do not make use of market measures of informed trading such as the Probability of Informed Trading (PIN) of Easley et al. (1996) or those based on price impact (Sadka, 2006) since nowadays, market measures of informed trading convey information about the high-frequency-trading world, where traders are silicon, not human (O'Hara, 2015). Sell-side analysts, who rely on firm fundamentals or accounting measures to value stocks, do not make forecasts to be used in high frequency trading or even in intra-day trading, which is mostly done by computers. They make forecasts to be used in investment decisions of longer horizons, carried out by humans looking for the true intrinsic value of stocks. As O'Hara (2015) points out, over the time intervals of high frequency trading, the information in market measures is not just asset-related (investment-related[17]) but mostly order-related (speculation-related[18]) and the basic unit of market information is orders.

2.4 Results

In table 2.6, I show the results of dynamic models, using $D_{i,CG} \times \Delta BNH_{i,t-4}$ and $D_{i,F} \times \Delta BNH_{i,t-4}$ to capture differentials in analysts incentives. For stocks issued by firms in the Consumer Goods sector ($D_{i,CG} = 1$), the change in analyst optimism ($\partial y_{i,t}$) given a change in information acquisition ($\partial \Delta BNH_{i,t-4}$) is statistically positive ($0.036 = 0.051 - 0.015$), suggesting that, when firm information gathering by the public increases, analysts are more likely to increase their bias for larger firms (higher potential for trading) which at the same time have a lower PE ratio (see table 2.5 Panel B). Interestingly, the estimates on the parameters on $D_{i,F} \times \Delta BNH_{i,t-4}$, although not statistically significant, are negative. This is consistent with the idea that, when firm information gathering by the public increases, analysts

[16] https://www.sec.gov/divisions/investment/13flists.htm

[17] As defined by Fisher (1930), an investment is the buying of future income streams. More precisely, "the value of any property, or rights to wealth, is its value as a source of income and is found by discounting that expected income."

[18] Speculation is the trading of an asset for reasons not related to its fundamentals or to its ability to generate future income (see e.g. Zhang and Yao, 2016). As defined in Tirole (1982), people "exhibit speculative behavior if the right to resell [an] asset makes them willing to pay more for it than they would pay if obliged to hold it forever."

are more likely to reduce their bias for firms that issue larger amounts of hard information on their performance which makes it easier for investors to identify those sell-side analysts incurring in opportunistic behavior.

Table 2.6: Results. Arellano-Bond Estimators.

Optimism $y_{i,t}$ is the dependent variable. In the difference equation, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $\Delta BNH_{i,t-4}$ are $\Delta BNH_{i,t-6}$ and $\Delta BNH_{i,t-7}$; for the interactions between $\Delta BNH_{i,t-4}$ and sector dummies, are their lags from $t-6$ to $t-8$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

$y_{i,t-1}$	0.490^{***}	$y_{i,t-1}$	0.496^{***}
	(0.027)		(0.028)
$roa_{i,t-5}$	0.216^{***}	$\Delta eps_{i,t-5}$	0.033
	(0.057)		(0.035)
$return_{i,t-4}$	-0.133^{***}	$return_{i,t-4}$	-0.132^{***}
	(0.021)		(0.021)
$sd_{i,t-4}$	-0.046^{***}	$sd_{i,t-4}$	-0.048^{***}
	(0.011)		(0.011)
$size_{i,t-4}$	0.423^{***}	$size_{i,t-4}$	0.425^{***}
	(0.024)		(0.024)
$\Delta BNH_{i,t-4}$	-0.015	$\Delta BNH_{i,t-4}$	-0.016
	(0.016)		(0.016)
$D_{i,F} \times \Delta BNH_{i,t-4}$	-0.026	$D_{i,F} \times \Delta BNH_{i,t-4}$	-0.027
	(0.023)		(0.023)
$D_{i,CG} \times \Delta BNH_{i,t-4}$	0.051^{*}	$D_{i,CG} \times \Delta BNH_{i,t-4}$	0.051^{*}
	(0.029)		(0.029)
Observations	39690	Observations	39690
Sargan Test	10.1158	Sargan Test	9.9947
A-B m-statistic (1st)	-13.8195^{***}	A-B m-statistic (1st)	-13.8228^{***}
A-B m-statistic (2nd)	0.6856	A-B m-statistic (2nd)	0.6914

Tables 2.7 and 2.8 show results on the within estimation and the Arellano-Bond estimation, without interactions with dummies, in order to analyze the average effect of investor information acquisition on the whole sample, and to analyze the importance of including lagged dependent variable among the regressors. While in both models, the panel data with fixed effects and the dynamic panel data, the estimates on $\Delta BNH_{i,t-4}$ are negative, only in the dynamic model are statistically significant. These results point out that, when investors increase their gathering or acquisition of financial information at the firm level, analysts reduce their positive bias. In addition, the Sargan Test as well as the tests for first and second order autocorrelation developed in Arellano and Bond (1991), show the validity of the instruments used in the dynamic model. These results also show the importance of including past values of forecast bias in the specification: as the estimates on $y_{i,t-1}$ are positive, analyst bias has some persistence.

Table 2.7: Results. Arellano-Bond and Fixed Effects Estimators.

Optimism $y_{i,t}$ is the dependent variable. In both equations there are fixed and time effects. In the difference equation of the dynamic model, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $\Delta BNH_{i,t-4}$ are $\Delta BNH_{i,t-6}$ and $\Delta BNH_{i,t-7}$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

	A-B (2SGMM)	*Fixed Effects (Within)*
$y_{i,t-1}$	0.492^{***}	
	(0.027)	
$roa_{i,t-4}$	0.035	-0.077^{***}
	(0.080)	(0.021)
$roa_{i,t-5}$	0.218^{***}	0.046^{***}
	(0.056)	(0.017)
$return_{i,t-4}$	-0.130^{***}	-0.515^{***}
	(0.020)	(0.009)
$sd_{i,t-4}$	-0.046^{***}	0.040^{***}
	(0.010)	(0.007)
$size_{i,t-4}$	0.421^{***}	0.308^{***}
	(0.023)	(0.005)
$\Delta BNH_{i,t-4}$	-0.025^{**}	-0.001
	(0.011)	(0.004)
Observations	39690	45766
	Sargan Test: 1.5777	R^2: 0.1215
	A-B m-statistic (1st): -13.8266***	Adjusted R^2: 0.0691
	A-B m-statistic (2nd): 0.6616	F Statistic: 995.343*** (df = 6; 43191)

Table 2.8: Results. Arellano-Bond and Fixed Effects Estimators.

Optimism $y_{i,t}$ is the dependent variable. In both equations there are fixed and time effects. In the difference equation of the dynamic model, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $\Delta BNH_{i,t-4}$ are $\Delta BNH_{i,t-6}$ and $\Delta BNH_{i,t-7}$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

	A-B (2SGMM)	*Fixed Effects (Within)*
	(1)	(2)
$y_{i,t-1}$	0.498^{***}	
	(0.027)	
$\Delta eps_{i,t-4}$	0.024	-0.141^{***}
	(0.043)	(0.026)
$\Delta eps_{i,t-5}$	0.047	-0.062^{**}
	(0.040)	(0.026)
$return_{i,t-4}$	-0.130^{***}	-0.515^{***}
	(0.021)	(0.009)
$sd_{i,t-4}$	-0.048^{***}	0.040^{***}
	(0.011)	(0.007)
$size_{i,t-4}$	0.424^{***}	0.307^{***}
	(0.024)	(0.005)
$\Delta BNH_{i,t-4}$	-0.026^{**}	-0.001
	(0.011)	(0.004)
Observations	39690	45766
	Sargan Test: 1.4667	R^2: 0.1218
	A-B m-statistic (1st): -13.8712^{***}	Adjusted R^2:0.069448
	A-B m-statistic (2nd): 0.6905	F Statistic: 998.251*** (df = 6; 43191)

The results on the dynamic panel data models (column (1) in tables 2.7 and 2.8) show that return on assets is a better performance variable than changes in earnings per share to explain analysts forecasts, and that performance values before $t-4$ (the date forecasts are issued) are better than values at $t-4$, which is sensible since financial statements take time to be elaborated and are not reported instantly by firms. Furthermore, the results show that stock returns, return volatility and firm sizes are important explanatory variables of analyst forecast bias.

Now, in table 2.9, I show the results of using the stocks in the 13(f) list ($D_{i,13F}$) in order to capture those stocks in the holdings of institutional investors, and the non-robot downloads of EDGAR filings to capture those stocks grabbing the attention of investors with a lower ability to identify opportunistic behaviors. The positive and significant estimates on $D_{i,CG} \times \Delta edgar_{i,t-4}$ suggest that analysts tend to be more optimistic for stocks with a higher trading potential. In addition, for stocks issued by financial firms in the holdings of institutional investors ($D_{i,F} = D_{i,13F} = 1$) , the change in forecast bias ($\partial y_{i,t}$) given a change in information acquisition ($\partial \Delta BNH_{i,t-4}$) is statistically negative ($-0.002 = -0.067 + 0.066 + 0.077 - 0.078$), consistent with the idea that analysts are more likely to reduce their bias for stocks followed by investors with a higher ability to identify inflated forecasts. These results indicate that information acquisition activate two different analysts' incentives. In addition, the estimates on firm size indicate that larger firms are likely to be more popular and to catch more investors attention, similar to Atiase (1985) and Fang and Peress (2009), thus inducing analysts to issue more optimistic forecasts.

Table 2.9: Results. Arellano-Bond Estimators Using EDGAR Filings Downloads and The 13(f) list. Optimism $y_{i,t}$ is the dependent variable. In the difference equation, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $\Delta BNH_{i,t-4}$ are $\Delta BNH_{i,t-6}$ and $\Delta BNH_{i,t-7}$; for $\Delta edgar_{i,t-4}$ are $\Delta edgar_{i,t-6}$ and $\Delta edgar_{i,t-7}$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

$y_{i,t-1}$	0.498^{***}
	(0.034)
$roa_{i,t-5}$	0.292^{***}
	(0.054)
$return_{i,t-4}$	-0.133^{***}
	(0.025)
$sd_{i,t-4}$	-0.038^{***}
	(0.013)
$size_{i,t-4}$	0.441^{***}
	(0.029)
$\Delta BNH_{i,t-4}$	-0.067^{***}
	(0.024)
$D_{i,F} \times \Delta BNH_{i,t-4}$	0.066^{***}
	(0.024)
$D_{i,13F} \times \Delta BNH_{i,t-4}$	0.077^{***}
	(0.024)
$D_{i,F} \times D_{i,13F} \times \Delta BNH_{i,t-4}$	-0.078^{***}
	(0.025)
$D_{i,CG} \times \Delta edgar_{i,t-4}$	0.0003^{**}
	(0.0001)
Observations	25391
Sargan Test	2.3036
A-B m-statistic (1st)	-11.1240^{***}
A-B m-statistic (2nd)	0.4296

As a value of 0.0003 seems very small, I also run a regression using standardized variables of $\Delta edgar_{i,t-4}$ and $\Delta BNH_{i,t-4}$ which I denote as $Z^{\Delta edgar}_{i,t-4}$ and $Z^{\Delta BNH}_{i,t-4}$. As these results, shown in table 2.10, indicate, an increase in one standard deviation of $\Delta edgar_{i,t-4}$ (23.2984), is related to a bias increase of 0.5% for stocks in the Consumer Goods sector.

Table 2.10: Results Using Standardized Variables

Optimism $y_{i,t}$ is the dependent variable. In the difference equation, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $Z^{\Delta BNH}_{i,t-4}$ are $Z^{\Delta BNH}_{i,t-6}$ and $Z^{\Delta BNH}_{i,t-7}$; for $Z^{\Delta edgar}_{i,t-4}$ are $\Delta edgar_{i,t-6}$ and $\Delta edgar_{i,t-7}$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

$y_{i,t-1}$	0.496^{***}
	(0.034)
$roa_{i,t-5}$	0.289^{***}
	(0.053)
$return_{i,t-4}$	-0.134^{***}
	(0.025)
$sd_{i,t-4}$	-0.039^{***}
	(0.013)
$size_{i,t-4}$	0.441^{***}
	(0.029)
$Z^{\Delta BNH}_{i,t-4}$	-0.029^{***}
	(0.010)
$Z^{\Delta edgar}_{i,t-4}$	0.002
	(0.001)
$D_{i,F} \times Z^{\Delta BNH}_{i,t-4}$	0.029^{***}
	(0.010)
$D_{i,13F} \times Z^{\Delta BNH}_{i,t-4}$	0.034^{***}
	(0.010)
$D_{i,F} \times D_{i,13F} \times Z^{\Delta BNH}_{i,t-4}$	-0.034^{***}
	(0.011)
$D_{i,CG} \times Z^{\Delta edgar}_{i,t-4}$	0.005^{*}
	(0.003)
Obs.	25391
Sargan	3.5234
m-stat. (1st)	-11.0857^{***}
m-stat. (2nd)	0.4205

2.4.1 Only Banks Instead of All Financials

Till now, I have argued that financial firms provide more information than firms in other sectors. While all publicly traded companies disclose detailed information on quarterly financial statements (Form 10-Q), audited annual financial performance (Form 10-K), and unscheduled events at a firm that could be of importance to the shareholders such as the hiring of a new director (Form 8-K), many financial firms such as banks, report additional hard information about measures of available capital v.s risk-weighted credit exposures, asset quality (loan's quality), management's ability to ensure the safe operation and sensitivity to particular risk exposures. Nevertheless, within financials, there is heterogeneity on the type and quantity of information that they release. Hedge funds or close ended funds, to give two examples, provide much less information than commercial banks to the general public.

In my sample of financials there are 832 firms. In this section, I show the results of selecting those firms in the financial sector that are banks, and using only these to capture those stocks for which investors can better establish the quality of the reports. In order to do so, I use a list of 5185 banks that appear at usbanklocations.com which provides information on banks in the United States. I carry out a web scraping of the web page and select the 38 banks of my original sample of financials that are in the list of banks in the US.

Figure 2.4: Banks in the Sample of Financials

Bank of America	Morgan Stanley	Charles Schwab	State Street	Progressive
Key	Bank of the Ozarks	FNB	South State	Green Dot
BancorpSouth Bank	First Merchants	Trustmark	Renasant	Banner
CenterState Bank	First Ban	Univest	Opus Bank	Cadence Ban
Mercantile Bank	First Financial	Access National	1st Source	Hingham Institution for Savings
Citizens & Northern	Northrim Ban	Paragon Commercial	ACNB	BankFinancial
LCNB	Community Financial	Summit State Bank	Citizens First	Central Federal
Goldman Sachs Group Inc	Elmira Savings Bank	Public Storage		

I calculate a dummy $D_{i,Banks}$ that takes the value of one for firms that are exclusively banks and estimate a dynamic panel model whose results I show in table 2.11, indicating that forecast bias is reduced in 36.3% for those stocks issued by banks in the holdings of institutional investors, when information acquisition increases in one standard deviation.

Table 2.11: Results Using Banks

Optimism $y_{i,t}$ is the dependent variable. In the difference equation, the valid instruments for $y_{i,t-1}$ are $y_{i,t-2}$ and $y_{i,t-3}$; for $Z^{\Delta BNH}_{i,t-4}$ are $Z^{\Delta BNH}_{i,t-6}$ and $Z^{\Delta BNH}_{i,t-7}$; for $Z^{\Delta edgar}_{i,t-4}$ are $\Delta edgar_{i,t-6}$ and $\Delta edgar_{i,t-7}$. *Note:* *p<0.1; **p<0.05; ***p<0.01.

$y_{i,t-1}$	0.495^{***}
	(0.034)
$roa_{i,t-5}$	0.287^{***}
	(0.053)
$return_{i,t-4}$	-0.135^{***}
	(0.025)
$sd_{i,t-4}$	-0.040^{***}
	(0.013)
$size_{i,t-4}$	0.441^{***}
	(0.029)
$Z^{\Delta BNH}_{i,t-4}$	-0.024^{***}
	(0.008)
$Z^{\Delta edgar}_{i,t-4}$	0.005
	(0.004)
$D_{i,Banks} \times Z^{\Delta BNH}_{i,t-4}$	0.025
	(0.019)
$D_{i,13F} \times Z^{\Delta BNH}_{i,t-4}$	0.026^{***}
	(0.008)
$D_{i,Banks} \times D_{i,13F} \times Z^{\Delta BNH}_{i,t-4}$	-0.365^{**}
	(0.153)
$D_{i,CG} \times Z^{\Delta edgar}_{i,t-4}$	0.005^{*}
	(0.003)
$D_{t,2012} \times Z^{\Delta edgar}_{i,t-4}$	-0.003
	(0.004)
Obs.	25391
Sargan	3.2372
m-stat. (1st)	-11.0775^{***}
m-stat. (2nd)	0.38574

2.5 Conclusions

In this paper I estimate the effects of investors' information gathering on analyst forecast bias, and find evidence in favor of the hypothesis that investors' information acquisition is positively related to analysts' optimism when the potential for trading is larger, and negatively related to optimism when investors are more likely to identify inflated forecasts. These results suggest that information acquisition is not only correlated to analysts' optimism but also that its effect does not work trivially and solely in one direction but it activates two different incentives in analysts' decisions.

These results highlight the importance of information acquisition: being more informed in the stock market may increase or decrease analyst forecast bias which deteriorates or improves the decision-making process of naive investors. These also help explaining why there are analysts who are systematically biased in the market: even with widespread access to relevant information, not all investors have the same potential to identify inflated forecasts and to penalize analysts who, driven by their trading incentives, incur in an opportunistic behavior. In the words of Lipman (1991): "knowing a fact does not mean that one knows all the logical implications of that fact...the agent can fail to recognize the appropriate action because this requires him to process his information." People, including investors, can be persuaded to pay for even blatantly useless predictions, as we may irrationally act, for instance, to feel in "control" over a random situation, to avoid the regret of not acquiring the forecast if it turns out to be "correct", or to avoid blaming ourselves if the decision outcome goes wrong (Powdthavee and Riyanto, 2015).

These results also highlight the importance of an unanswered question: what drives more and less sophisticated investors to follow a set of stocks and avoid others. The answer is not trivial since it may be related to information access, technology usage in the processing of information and and cognitive efforts. For instance, before deciding how much information is optimal to acquire (i.e. how much effort to make) on a stock, an investor must decide on which stock to gather information, and then the investor may first compute or understand the implications of following a stock instead of another. If finding the best set of stocks to follow is costly, then the investor may first construct a decision procedure in order to decide which stocks to follow, which involves trading off the benefits to improving the choice with the costs of improving the decision-making. But if the construction of this procedure is costly and there are several alternative procedures, then the investor may first also compute an algorithm to decide which procedure to apply, and so on. This is the infinite regress problem of bounded rationality well explained by Lipman (1991) and is related to the idea that "[c]ognitive resources should be allocated just like other scarce resources" (Gabaix and Leibson, 2005).

The difference between more sophisticated and less sophisticated investors may also be related to government regulation. Thus, for further research it would also be worth studying the effects of regulation on analysts behavior. Regulatory agencies have addressed analysts' conflicts of interests by regulating information issuers, with rules such as the Securities Act of 1933 and the Reg FD of 2000 which have the purpose of increasing the amount of information that investors can access directly from firms, or rules that affect analysts compensations such as the Global Research Analyst Settlement and the Sarbanes-Oxley

Act of 2002. An interesting question is whether laws that focus on investors i.e. financial information consumers, would improve the quality of the information issued by sell-side analysts. For instance, Choi (2000) proposes limiting the investments of the less informed investors to passive index mutual funds, while requiring from other investors that they make their transactions through brokerage firms with research departments certified as highly respected.

Appendix

The summary statistics on the list of section 13(f) securities ($D_{i,13F}$) are in table 2.12. The list of section 13(f) securities of 2017Q4 available in pdf format from the website of the U.S Securities and Exchange Commission, counts with more than 17,000 individual securities from which I identify the 504 stocks in my sample that are included in this report and represent 32% of the observations in my sample.

The summary statistics for the complete set of variables are in table 2.12. Consistent with the literature on analyst optimism, both the median and the mean of forecast bias in my sample are positive, and the absolute value of the minimum is smaller than the maximum.

Table 2.12: Summary Statistics.

$\Delta eps_{i,t}$ is calculated as $\frac{EPS_{i,t}-EPS_{i,t-1}}{P_{i,t-1}}$; $roa_{i,t}$ is $\frac{Earnings_{i,t}}{Assets_{i,t}}$; $size_{i,t}$ is the log of market capitalization; $return_{i,t}$ are stock returns; $sd_{i,t}$ is the quarterly standard deviation from daily returns.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
$optim_{i,t}$	-2.07728	-0.16093	0.04082	0.09685	0.28658	4.05181
$\Delta eps_{i,t}$	-2.0763320	-0.0048225	0.0003587	-0.0008504	0.0055807	0.7644606
$roa_{i,t}$	-5.38368	0.00691	0.03721	0.01321	0.07864	1.53539
$size_{i,t}$	1.829	6.637	7.726	7.788	8.874	13.494
$return_{i,t}$	-0.93347	-0.07138	0.02868	0.03469	0.12961	7.77441
$sd_{i,t}$	0.003848	0.013670	0.018893	0.022091	0.026988	0.236444
$D_{i,F}$	0.0000	0.0000	0.0000	0.2249	0.0000	1.0000
$D_{i,T}$	0.0000	0.0000	0.0000	0.1145	0.0000	1.0000
$D_{i,CG}$	0.0000	0.0000	0.0000	0.08849	0.0000	1.0000
$D_{i,13F}$	0.0000	0.0000	0.0000	0.3195	1.0000	1.0000
$\Delta BNH_{i,t}$	-4.0000	-0.1667	0.0000	0.0344	0.1788	4.0000
$\Delta edgar_{i,t}$	-1718.5275	-0.6598	0.0000	-0.0011	0.9507	1504.1848

Author's calculations.

References

Abarbanell, J. S. (1991). Do analysts' earnings forecasts incorporate information in prior stock price changes?. *Journal of Accounting and Economics*, 14(2), 147-165.

Abarbanell, J. S. and Bernard, V. L. (1992). Tests of analysts' overreaction/underreaction to earnings information as an Explanation for Anomalous Stock Price Behavior. *The Journal of Finance*, 47(3), 1181 - 1207.

Abarbanell, J. and Lehavy, R. (2003). Biased Forecasts or Biased Earnings? The Role of Reported Earnings in Explaining Apparent Bias and Over/Underreaction in Analysts' Earnings Forecasts. *Journal of Accounting and Economics*, 36, 105 - 146.

Ali, A., Klein, A. and Rosenfeld J. (1992). Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS. *Accounting Review*, 67(1), 183 - 198.

Atiase, R. K. (1985). Predisclosure Information, Firm Capitalization, and Security Price Behavior Around Earnings Announcements. *Journal of Accounting Research*, 21-36.

Asquith, P., M. B. Mikhail and A. S. Au. (2005). Information Content of Equity Analyst Reports. *Journal of Financial Economics*, 75(2), 245 - 82.

Banerjee, S., Breon-Drish, B., and Engelberg, J. (2020). Discussion of "disclosure processing costs, investors' information choice, and equity market outcomes: A review". Journal of Accounting and Economics, 101337.

Beaver, W. H. (2002). Perspectives on Recent Capital Market Research. *The Accounting Review*, 77(2), 453 - 474.

Ben-Rephael, A., Da, Z. and Israelsen, R. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *Review of Financial Studies*, 30(9), 3009-3047.

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

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, 11(2-3), 255-274.

Boehmer, E. and Kelley, E. K. (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies*, 22(9), 3563-3594.

Bonini, S., Zanetti, L., Bianchini, R. and Salvi, A. (2010). Target price accuracy in equity research. *Journal of Business Finance & Accounting*, 37(9-10), 1177-1217.

Chang, J. W. and Choi, H. M. (2017). Analyst Optimism and Incentives under Market Uncertainty. *Financial Review*, 52(3), 307-345.

Bradshaw, M. T., Brown, L. D., and Huang, K. (2013). Do sell-side analysts exhibit differential target price forecasting ability?. Review of Accounting Studies, 18(4), 930-955.

Choi, S. (2000). Regulating Investors Not Issuers: A Market-Based Proposal. *California Law Review*, 88(2), 279-334.

Cowen, A., Groysberg, B. and Healy, P. (2006). Which Types of Analyst Firms Are More Optimistic? *Journal of Accounting and Economics*, 41, 119 - 146.

Da, Z., J. Engelberg, and P. Gao. (2011). In search of attention. *Journal of Finance*, 66, 1461-99.

Da, Z., Hong, K. P. and Lee, S. (2016). What drives target price forecasts and their investment value?. *Journal of Business Finance & Accounting*, 43(3-4), 487-510.

Dahiya, S., Iannotta, G. and Navone, M. (2017). Firm opacity lies in the eye of the beholder. *Financial Management*, 46(3), 553-592.

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

Das, S., Guo, R. J. and Zhang, H. (2006). Analysts' selective coverage and subsequent performance of newly public firms. *The Journal of Finance*, 61(3), 1159-1185.

DeBondt, W. F. M. and Thaler, R. (1990). Do Security Analysts Overreact? *The American Economic Review*, 80(2), 52 - 27.

Ely, J. C., and Välimäki, J. (2003). Bad reputation. The Quarterly Journal of Economics, 118(3), 785-814.

Engelberg, J. E. and Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, 66(1), 67-97

Easley, D., Kiefer, N. M., O'hara, M., and Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. The Journal of Finance, 51(4), 1405-1436.

Easterwood, J. C. and Nutt, S. R. (1999). Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or SystematicOptimism? *The Journal of Finance*, 54(5), 1777 - 1797.

Fama, E. (1965). Random Walks in Stock Market Prices. Financial Analysts Journal, 51(1), 75-80

Fang, L. and Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5), 2023-2052.

Fang, L., and Yasuda, A. (2009). The effectiveness of reputation as a disciplinary mechanism in sell-side research. The Review of Financial Studies, 22(9), 3735-3777.

Fisher, I. (1930). The theory of interest: as determined by impatience to spend income and opportunity to invest it. Augustusm Kelly Publishers, Clifton.

Fischer, P. E. and Stocken, P. C. (2010). Analyst Information Acquisition and Communication. *The Accounting Review*, 85(6), 1985 - 2009.

Gabaix, X., and Laibson, D. (2005). Bounded rationality and directed cognition. Harvard University.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. American Economic Review, 96(4), 1043-1068.

Galperti, S. (2019). Persuasion: The art of changing worldviews. American Economic Review, 109(3), 996-1031.

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

Givoly, D., Hayn, C. and Lehavy, R. (2009). The Quality of Analysts' Cash Flow Forecasts. *The Accounting Review*, 84(6), 877 - 1911.

Goldsmith-Pinkham, Paul and Hirtle, Beverly and Lucca, David O., Parsing the Content of Bank Supervision (2016-03-01). FRB of NY Staff Report No. 770. Available at SSRN: https://ssrn.com/abstract=2750315

Gunther, J. W., and Moore, R. R. (2003). Loss underreporting and the auditing role of bank exams. Journal of Financial Intermediation, 12(2), 153-177.

Hayes, R. M. (1998). The impact of trading commission incentives on analysts' stock coverage decisions and earnings forecasts. *Journal of Accounting Research*, 36(2), 299-320.

Hayes, R. M. and Levine, C. B. (2000). An approach to adjusting analysts' consensus forecasts for selection bias. *Contemporary Accounting Research*, 17(1), 61-83.

Hilary, G. and Hsu, C. (2013). Analyst forecast consistency. *The Journal of Finance*, 68(1), 271-297.

Hillert, A., Jacobs, H. and Müller, S. (2014). Media makes momentum. *The Review of Financial Studies*, 27(12), 3467-3501.

Hong, H., Lim, T. and Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *The Journal of Finance*, 55(1), 265-295.

Hugon, A., and Muslu, V. (2010). Market demand for conservative analysts. Journal of Accounting and Economics, 50(1), 42-57.

Hugonnier, J., and Morellec, E. (2017). Bank capital, liquid reserves, and insolvency risk. Journal of Financial Economics, 125(2), 266-285.

Jackson, A. R. (2005). Trade Generation, Reputation, and Sell-Side Analysts. *The Journal of Finance*, 60(2), 673 - 717.

Jullien, B., and Park, I. U. (2014). New, like new, or very good? Reputation and credibility. The Review of Economic Studies, 81(4), 1543-1574.

Kaniel, R. and Parham, R. (2017). WSJ Category Kings–The impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics*, 123(2), 337-356.

Kartik, N., Ottaviani, M. and Squintani, F. (2007). Credulity, Lies, and Costly Talk. *Journal of Economic Theory*, 134, 93 - 116.

Lim, T. (2001). Rationality and analysts' forecast bias. *The Journal of Finance*, 56(1), 369-385.

Lin, J. T., Bumcrot, C., Mottola, G., and Walsh, G. (2019). Investors in the United States: A report of national financial capability study. FINRA: Investor Education Foundation.

Lipman, B. L. (1991). How to decide how to decide how to...: Modeling limited rationality. Econometrica: Journal of the Econometric Society, 1105-1125.

Loughran, T. and McDonald, B. (2017). The use of EDGAR filings by investors. *Journal of Behavioral Finance*, 18(2), 231-248.

Ljungqvist, A., Marston, F., Starks, L. T., Wei, K. D. and Yan, H. (2007). Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics*, 85(2), 420-456.

Malmendier, U. and Shanthikumar, D. (2014). Do security analysts speak in two tongues?. *The Review of Financial Studies*, 27(5), 1287-1322.

Mahmud, S., Scarth, N., Shah, N. and Molloy, P. (2016). The Future of Equity Research. Bloomberg Professional Services, available at:
https://www.bloomberg.com/professional/blog/future-equity-research/

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

Mikhail, M. B., Walther, B. R. and Willis, R. H. (2007). When security analysts talk, who listens?. *The Accounting Review*, 82(5), 1227-1253.

Mikhail, M. B., Walther, B. R. and Willis, R. H. (1999). Does Forecast Accuracy Matter to Security Analysts? *The Accounting Review*, 74(2), 185-200.

Millimet, D. L. and Tchernis. (2013). Estimation of treatment effects without an exclusion restriction: with an application to the analysis of the school breakfast program. *Journal of Applied Econometrics*, 28(6), 982-1017.

Niehaus, G. and Zhang, D. (2010). The Impact of Sell-Side Analyst Research Coverage on an Affiliated Broker's Market Share of Trading Volume. *Journal of Banking & Finance*, 34, 776 - 787.

Ottaviani, M., and Sørensen, P. N. (2006). Reputational cheap talk. The Rand journal of economics, 37(1), 155-175.

Oppel Jr, R. A. (2002). Wall Street analysts faulted on Enron. *The New York Times* (Feb. 28).

Peress, J. (2008). Media coverage and investors' attention to earnings announcements, Working Paper. INSEAD.

Pinto, J. E., Robinson, T. R. and Stowe, J. D. (2019). Equity valuation: A survey of professional practice. Review of Financial Economics, 37(2), 219-233.

Powdthavee, N., and Riyanto, Y. E. (2015). Would you pay for transparently useless advice? A test of boundaries of beliefs in the folly of predictions. Review of Economics and Statistics, 97(2), 257-272.

Ramnath, S., Rock, S. and Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting*, 24(1), 34-75.

Solomon, D. H., Soltes, E. and Sosyura, D. (2014). Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, 113(1), 53-72.

Soltes, E. (2009). News dissemination and the impact of the business press. The University of Chicago.

Stickel, S. E. (1992). Reputation and performance among security analysts. The Journal of Finance, 47(5), 1811-1836.

Tetlock, P. C., Saar-Tsechansky, M. and Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3), 1437-1467.

Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica: Journal of the Econometric Society*, 1163-1181.

Womack, K. L. (1996). Do Brokerage Analysts' Recommendations Have Investment Value? *The Journal of Finance*, 51(1), 137 - 167.

Zhang, Y. J. and Yao, T. (2016). Interpreting the movement of oil prices: driven by fundamentals or bubbles?. *Economic Modelling*, 55, 226-240.

Chapter 3

Investors' Information Choice

Abstract: I estimate a demand model for online services of financial data, from a random parameters or mixed logit model, using a sample with searches at Bloomberg Terminals and at the EDGAR system. My preliminary results suggest that the substitution investors make of financial information providers, are affected by the subscription prices, investors' expectations on stock returns, and investors' income.

Key words: random parameters, open access services, subscription providers, market shares.

3.1 Introduction

Investors need to get informed in order to take good investment decisions, and the idea that having many providers of information, improves the decision making process of investors, is intuitive. Nevertheless, as acquiring data is costly and attention is a scarce cognitive resource (Peng and Xiong, 2006), investors also need to take decisions on which source will provide them the needed information. The current massive use of the internet offers to investors a large number of sources which include no-fee open access providers such as Yahoo Finance and EDGAR, that co-exist with the services of subscription providers such as Bloomberg, Refinitiv Eikon, FactSet and Capital IQ. These sources are not perfect substitutes: although the sets of information they offer have common elements, some may present it in a different manner. In this research, I empirically study the drivers of investors' demand for financial information services.

In previous studies, little attention has been devoted to analyzing investors' choices over information sources. While the study of product demand is extensive in the academic literature, understanding investors substitution patterns over information providers has its own issues. First, financial information about stock markets is useful for investors as long as they have the financial capability to buy financial assets, and therefore, even with open-access information providers, income is important to understand information acquisition activities. Second, financial information is not consumed for its own sake but it is acquired with the purpose of making a decision on another product: a financial asset. As the acquisition of financial information is expected to help in the decision making process of asset allocation, it requires the investor to process it, making it relevant to evaluate the potential effects of investor's personal characteristics, such as financial literacy and willingness to take risks, on changes in choices over information providers.

I estimate a demand model for financial information services, using data on investor demographics, stocks and search activities. In particular, I use survey results at the investor level from the FINRA Foundation, returns and fundamental variables of stocks belonging to the CRSP index, daily non-robot EDGAR downloads and the News Heat index of Bloomberg searches. I estimate the parameters of a random parameters or mixed logit model, controlling for unobserved service characteristics as well as for unobserved investor characteristics, as proposed in Nevo (2001). As the service price is endogenous, I use, as instruments, (lagged) Google's SVI of the expression "Bloomberg Terminal" as well as sales revenues of Microsoft, arguing that the popularity of Bloomberg Terminals is related to Bloomberg's capacity to charge higher prices, and based on the facts that the Terminals are not available for MAC operating systems and many functionalities involve an intensive use of Excel. In addition, as the demand for information may influence the expectations investors have on stock returns, I use as instruments a set of variables that I argue, are not affected by choices on information services, such as firm return on assets or earnings per share volatility, among others. I find that subscription prices as well as expected returns are important to explain financial information demand, and that income have an influence on investor's marginal valuation of expected returns when deciding over sources of information. This research is, to the best of my knowledge, the first attempt to understand the substitution patterns between sources of information.

While, as discussed in section 3.7, I can improve my estimates, my current results show that there are investor characteristics that induce investors to change their information providers, and that some investors perceive, to some extent, service informational advantages. This is important because new open-access sources of financial information may affect the viability for existent subscription providers, who in turn may react in order to stay in the market, and thus, investors' decisions on services affect the overall informational environment of the stock market and investors well being. For instance, a subscription provider may reduce its effort to deliver that information also available on open access providers or may increase its effort to present this same information in a more friendly manner or in higher detail.

There are eight sections in this document including the introduction. In section two I present the literature related to financial information demand. In section three I show my sample, describing the data related to information acquisition activities, as well as that related to investor demographics. In sections four and five I explain my model of investor's choices over sources of financial information, the definition of the outside option and identification issues. In section six I show my results, and in sections seven and eight, I discuss my results and conclude.

3.2 Literature Review

The relevance of financial information demand is well known in the theoretical literature. For instance, in a model where acquiring information is costly, Grossman and Stiglitz (1980) show that the degree in which stock prices convey information from the informed to the uninformed investors, depends on the number of individuals who get informed. At the same time, the availability of many information sources require investors to allocate their attention, which is a limited resource. From the experimental and theoretical literature, it is clear that information search is a costly cognitive operation, for which individuals must also make attention allocation decisions (Gabaix, et al., 2006; Sadler, 2021). Recent empirical literature in finance (e.g. Hu, et al., 2021) recognize that investors have limited attention and expose that the improvement of the informational environment for investors, that comes from having a greater number of information channels is not straightforward, since having a great number of them creates a poverty of attention, and thus, there is a need to make an optimal allocation of attention, across the overabundant sources that might consume it (Da, Engelberg and Gao, 2011). The implications of attention allocation decisions for stock markets has been studied in the theoretical literature, as in Peng and Xiong (2006), who model a representative investor who must allocate his attention among types of information, namely, market information, sector information and firm information. In my research, I empirically study investors' choice of information providers.

The demand for information is likely related to expectations on stock returns. For instance, Willinger (1989) models an investor who solves a problem of asset allocation in which he must also decide to pay (or not) the cost of acquiring information. As the contents of the new information are uncertain ex-ante, information itself is risky. In Willinger's (1989) model, information increases the expected return of the investor's portfolio and thus is valuable. Moreover, Vlastakis and Markellos (2012) empirically show that Google's search index on company names, increases for periods of higher stock returns. Information demand also likely depends on some investor characteristics such as financial literacy and the willingness to take risks. Willinger (1989), finds that the expected value of information (first defined in LaValle, 1968) and risk aversion are related. Lipman (1991) posits that allocation decisions (here choosing information providers) vary across individuals, even if we know our preferences and our feasible sets, since we may not know, all the logical implications of a same fact. For example, it is very likely that one does not know many theorems of set theory, even when these are logically implied by the axioms one knows. A more interesting example in the context of this paper, can be found in the results of the FINRA Foundation's survey of 2015, in which around 16% of those who had investments in stocks, gave an answer different to "More than $102" to the following question: "Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?."

Although, to the best of my knowledge, the empirical literature on investors' choices of financial information sources is scarce, there is a considerable amount of literature studying the characteristics of the users of a single open-access provider of information. Da, Engelberg and Gao (2011) use the Search Volume Index (SVI) of Google and find that increments in SVI, are related to trading of less sophisticated retail investors. Also, Behrendt, Peter and Zimmermann (2020) show that patterns of retail (collective) trading, are related to Wikipedia searches for firm information. Loughran and McDonald (2017), find that the most requested EDGAR filings of non-robot investors, tend to be about popular companies such as Facebook, and argue that retail investors would not be using robots to search information at the SEC.gov website. Chi and Shanthikumar (2018) find that retail trading (buying and selling) is related to searches for 10-K and 10-Q filings at EDGAR and Asthana, Balsam and Sankaraguruswamy (2004) show that when firms filed the form 10-K on EDGAR for the first time there was an increase in the volume of small trades whereas there is not an effect for large investors. At the same time, subscription providers are more likely to be used by institutional investors. For instance, Ben-Rephael, Da, and Israelsen (2017) propose a measure of institutional investor attention based upon searches and reading activity at Bloomberg terminals. The authors document that, as of August 26, 2016, around 80% of Blommberg Terminal users worked in financial industries including banking, asset management, and institutional financial services, with 32% of the job titles being portfolio managers or traders, 19% presidents or directors, and 17% being analysts, including buy-side and sell-side ones. This is in line with the research that, following the intuition, assume that Bloomberg is a source of information for more sophisticated or institutional investors, which include Da, Engelberg and Gao (2011) and Wang (2020).

3.3 Data and Preliminary Evidence

This paper uses four main data sets. First, it uses data on search activities at two different financial information services, namely, EDGAR and Bloomberg. Second, it uses data on investor characteristics, or demographics, from the Financial Industry Regulatory Authority (FINRA) Foundation, and in particular, it uses the results of the National Financial Capability Study of 2012 and 2015. Third, data on stocks belonging to the CRSP index (3,611) are used in this paper. I collect historical stock prices (from which I calculate stock returns), earnings per share, return on assets, market capitalization, number of analyst recommendations, target prices (from which I calculate bias volatility), the sectors to which they belong and whether the stock is in the list of section 13(f) securities of the SEC. Finally, this research uses historical data on Google's SVI for the expression "Bloomberg Terminal", as well as Microsoft's sales revenues. I provide more details in the next subsections.

3.3.1 Financial Information Services

For firms included in the Center for Research in Security Prices (CRSP) stock index from the second quarter of 2010 to the fourth quarter of 2015, I observe daily data on the News Heat - Daily Max Readership index of Bloomberg. The index is constructed by Bloomberg based upon the "number of times each article is read by its users, as well as the number of times users search for news for a specific stock" (Ben-Rephael, Da and Israelsen, 2017) and takes higher values for higher levels of readers activity going from 0 to 4. Additionally, I observe the daily non-robot searches of EDGAR filings through SEC.gov. The Division of Economic and Risk Analysis (DERA) constructed the EDGAR log file data set containing statistics on user access to the SEC.gov website and is "intended to provide insight into the usage of publicly accessible EDGAR company filings".

Among other variables, the EDGAR log file data, includes the IP addresses, dates, Central Index Key

(CIK) codes and whether the user self-identified as a crawler. In the SEC website[1] there are 2,880 zip folders in the EDGAR log file data set for the period 2010 - 2017, one for each day, and each folder contains a "README" file documenting the variables, and a *csv* file with the statistics on SEC.gov website traffic, which can include more than ten million entries (e.g. the file of January first of 2015 contains 15'682,916 rows and 15 columns). This allows me to count the number different non-crawler IP addresses, grouped by CIK codes, which I use to estimate the average number of investors that follow one stock in a day, with the purpose of having a definition of a potential market. In particular in this research, I filter out crawlers and count the number of IPs grouped by stocks, using the EDGAR log files from 2015-01-01 to 2015-02-05.

It is important to note that, as I am comparing EDGAR users with Bloomberg Terminal users, who are humans, filtering out crawlers from EDGAR users is important. Therefore, I also use a filtered version of the data for my whole sample period (2010 - 2015), from The Software Repository for Accounting and Finance[2] of the University of Notre Dame, which provides an extension of the information used in Loughran and McDonald (2017), and counts with data from 2006 to 2015. This data have been filtered to eliminate damaged files[3], irrelevant entries, or those with missing CIK, accession number, IP, or date. Also, robot downloads (those with more than 49 downloads from a single IP within a single day or self identified as a web crawler), or with a server code larger or equal than 300, or records of traffic on the index page of a set of documents (e.g. index.htm) have been filtered out. The information set I use in this paper includes daily data, from 2010 to 2015, of the number of (exclusively) non-robot downloads for each stock, identified with the CIK number, with their respective dates. Filtering robot activity, is an advantage of using data on EDGAR searches, since this filtering is not possible to carry out when analyzing data related to traffic on other free sources of information, such that available at Google Trends or Wiki-media. Moreover, as the EDGAR database provides only hard financial information, the records of information acquisition at EDGAR are records more narrowed to investors' activities, when compared to, for example, Google Trends (Search Volume Index) which registers Google usage by potentially anybody (investors and non-investors).

As Bloomberg provides an Index of user activity, but not the number of users or number of searches, these data is not directly comparable to EDGAR downloads. I count the number of different stocks belonging to the CRSP, searched each day at each source of information and compare information acquisition activities between these two sources.

[1] https://www.sec.gov/dera/data/edgar-log-file-data-set.html
[2] https://sraf.nd.edu/
[3] All files from 2005-09-24 to 2006-05-10 were labeled by the SEC as "lost or damaged" (Loughran and McDonald, 2017).

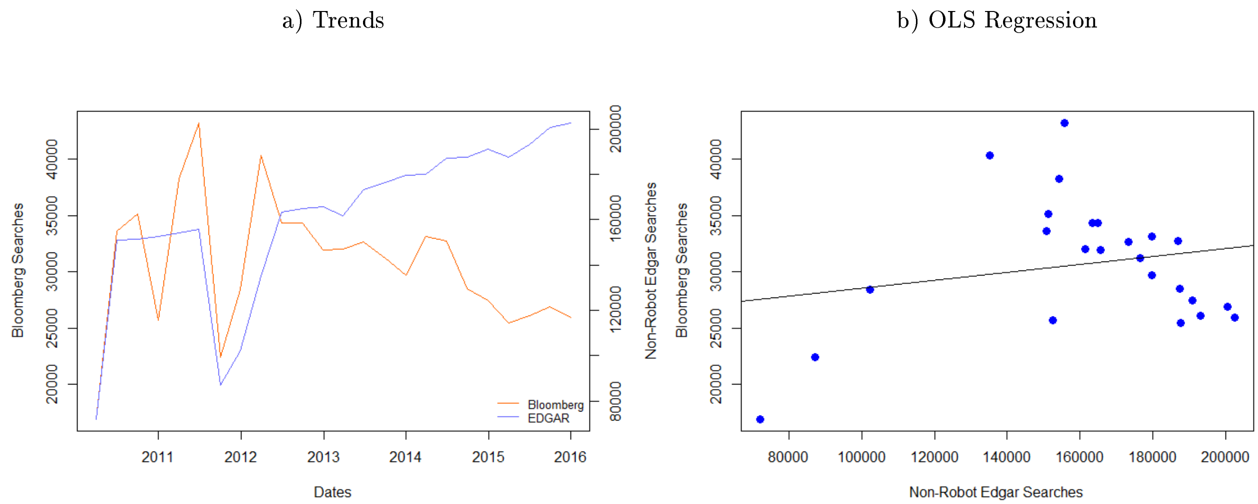


Figure 3.1: **Trends and Linear Regression of Information Acquisition Activities.** In panel a), the orange line represents, the quarterly sum of the daily number of stocks in the CRSP with a positive News Heat Index of Bloomberg, and the blue line represents the quarterly sum of the daily number of stocks in the CRSP with EDGAR searches. Panel b) shows the linear regression line between the two series of panel a).

Table 3.1: Summary Statistic on Search Activity at Each Source. Quarterly Number of Stocks in the CRSP Between 2010-03-31 and 2015-12-31.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
EDGAR	72,025	152,246	165,350	161,492	187,148	202,540
Bloomberg Terminals	16,952	26,679	31,575	30,684	33,811	43,181

In a first look at the data, from the trends of information acquisition activities at EDGAR and Bloomberg shown in panel *a*) of figure 3.1, we cannot get a clear idea of the substitution patterns. In the figure, the orange line represents, the quarterly sum of the daily number of stocks in the CRSP with a positive News Heat Index of Bloomberg, and the blue line represents the quarterly sum of the daily number of stocks in the CRSP with EDGAR searches. While the two series seem to follow opposite directions from around 2013 onward, they seem to move in a similar direction between 2010 and 2013, and this unclear pattern produces a non significant OLS slope estimate of 0.035 between the two series (panel b)).

I also collect data on service prices. EDGAR services are free of charge to the user and, as Bloomberg does not publicize its prices, I use a news media report containing historical subscription (two-year contract) prices per terminal from 2001 to 2013 of a single client in the US, which I project to 2015, using the average growth rate in price. I use an inflation-adjusted series, so that price changes reflect changes with respect to 2009 prices of Bloomberg subscriptions.

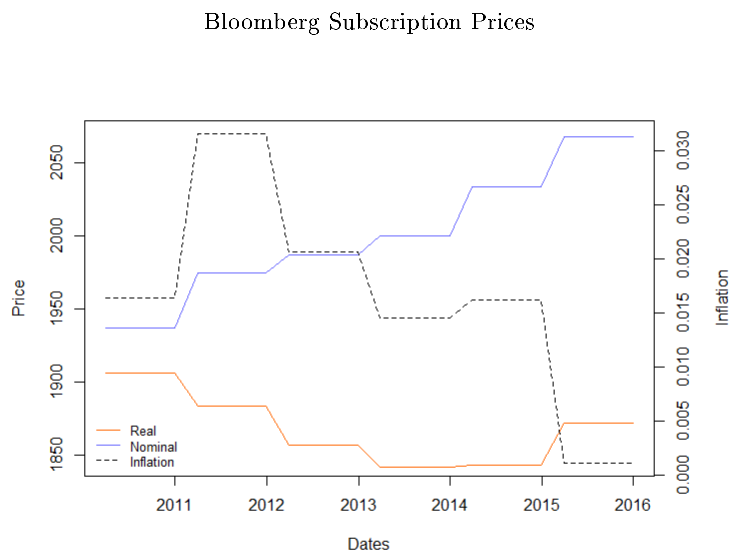


Figure 3.2: **Bloomberg prices 2010 - 2015.** The blue line represents two-year contract subscription prices per terminal; the orange line represents prices adjusted for inflation based on 2009 prices; the dashed line represents the annual inflation rate in the United States from fred.stlouisfed.org.

Interestingly, while the nominal service price has increased, the inflation-adjusted price of Bloomberg shows a slight decline. I assign a price of zero to EDGAR services and the inflation-adjusted price to Bloomberg services.

3.3.2 Demographics

I use the results of the National Financial Capability Study of 2012 and 2015 carried out by the Financial Industry Regulatory Authority (FINRA) Foundation in order to collect data on demographics. These are national self-report studies of the financial capability of American adults, which consists of a state-by-state online survey of 25,509 and 27,564 American adults in 2012 and 2015 respectively. In particular, I gather data on those individuals who responded "Yes" to the following question: "Not including retirement accounts, do you [does your household] have any investments in stocks, bonds, mutual funds, or other securities?" (variable $B14$[4]). From this population, I keep the answers on the variables of sex, ranges of annual income, willingness to take risks when investing, and the recorded answers to the following savings hypothetical problem: "Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?" (variable $M6$). I use the answers to this question as a proxy for financial literacy, which provides an idea of the ability of investors to understand the implications of the information available to them.

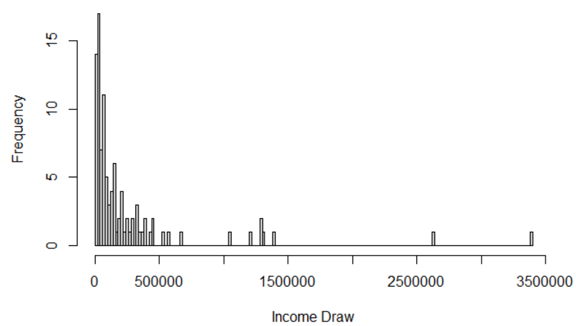
[4]This is the only question related to investments in stocks.

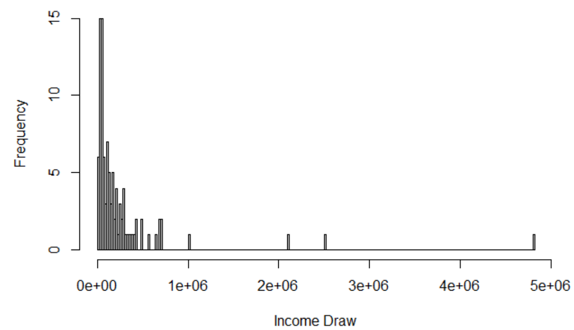
Table 3.2: Summary Statistic on Observations for Demographic Draws.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
				2012		
"Male"	0.00	0.00	1.00	0.51	1.00	1.00
Income Levels	12,750	62,500	87,500	277,371	125,000	1'479,052
"Very Willing"	0.000	0.000	0.000	0.055	0.000	1.000
"More than $102"	0.00	1.00	1.00	0.85	1.00	1.00
				2015		
"Male"	0.00	0.00	1.00	0.54	1.00	1.00
Income Levels	12,750	62,500	87,500	253,809	125,000	1'479,052
"Very Willing"	0.000	0.000	0.000	0.081	0.000	1.000
"More than $102"	0.00	1.00	1.00	0.84	1.00	1.00

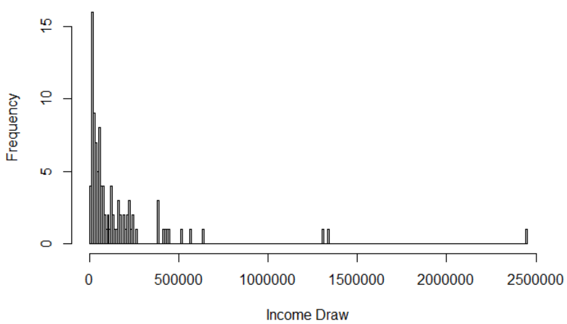
These variables are categorized with integers that are different from zero and one in many cases. For income, I transform the categories into levels of income (see table 3.2) using the average of the range for each category, where the wages of the the top 0.1% earners of the US population (2'808,104 USD) is the upper bound of the last category. With these income observations from the FINRA's survey, I calculate the logarithm of income levels, as well as the mean and standard deviation of the log-income for 2012 and 2015, parameters that I use to create a 100 draws of income from a log-normal distribution for each quarter, as shown in figure 3.3 and table 3.3. I denote the logarithm of income draws (which follows a normal distribution) as $Income$. For the other variables, I use dummies to obtain draws for my estimation, i.e. a dummy that takes the value of one for male, a dummy that equals one for "Very Willing", and a dummy that equals one for "More than $102." I use the average of each variable to generate 100 draws for each quarter from a binomial distribution, draws which I denote as $Male$, *Risk Lover* and *Literacy*.

a) Log-Normally Distributed Income Draws of Period 1b) Log-Normally Distributed Income Draws of Period 2





c) Log-Normally Distributed Income Draws of Period 23d) Log-Normally Distributed Income Draws of Period 24



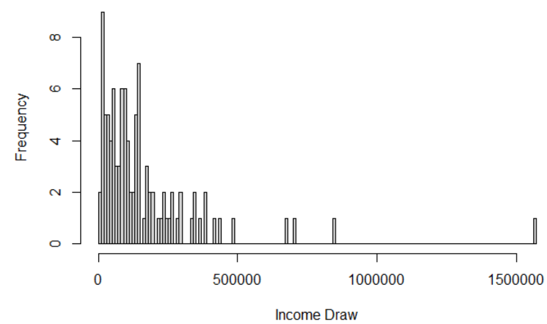


Figure 3.3: **Histograms of Log-Normally Distributed Income Draws.** Draws are obtained from a log-normal distribution, using the parameters estimated from survey observations.

Table 3.3: Summary Statistics on Log-Normally Distributed Income Draws of Four Quarters.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Period (market) 1	3,646	33,108	88,369	258,519	243,558	3'382,801
Period (market) 2	4,991	43,871	115,774	266,533	266,656	4'808,749
Period (market) 23	4,528	24,810	61,677	166,644	174,420	2'449,473
Period (market) 24	3,781	50,593	103,022	164,175	195,186	1'565,939

3.3.3 Data with Variation Across Alternatives

I calculate the cross-sectional average of stock returns, which I use to estimate the expected returns of the stocks sought through service j at quarter t, denoted $return_{jt}$. I interpret $return_{jt}$ as the expected attainable stock return from, or the expectations investors have on returns when, gathering financial information at provider j. Notice that the relationship between stock returns and information acquisition works in two directions: acquiring information on a set of stocks may influence their average returns, and having expectations on stock returns may have an effect on decisions of information acquisition.

I also calculate the cross-sectional average of the number of recommendations (in logs) issued by financial analysts, on stocks with activity at provider j, which I denote rec_{jt}. The number of stock recommendations in the literature is used as a proxy for analyst coverage (see e.g. Niehaus and Zhang, 2010). Furthermore, for each information provider, I calculate for each quarter, the average of the characteristics of the stocks sought at that source. I have in my sample, the data on firm size (log of market capitalization), return on assets (accounting variable), the volatility of earnings changes (accounting variable), stocks in the portfolios of institutional investors (list of section 13(f) securities of the SEC), and non-exhaustive sector dummies. I denote the average of these characteristics as $size_{jt}$, roa_{jt}, Δeps^{v}_{jt}, and D^{13F}_{ij}, and denote sectors with superscripts on the dummies such as D^{ID}_{jt} (industrials sector), D^{CG}_{jt} (consumer goods sector), D^{Tech}_{jt} (technology sector), D^{Fin}_{jt} (financials sectors), among others. In addition, my sample contains the projected stock prices from financial analysts from which I calculate the variance of forecast bias, i.e., of the difference between projected stock prices and realized stock prices, which captures the difficulty to evaluate the prospects of stocks (Hilary and Hsu's, 2013), and denote its average for the set of stocks sought at service j as $bvol_{jt}$.

Table 3.4: Summary Statistic on Alternative Specific Variables.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
$return_{jt}$	46	0.035	0.096	−0.222	−0.004	0.103	0.187
$F13_{ij}$	46	0.242	0.036	0.198	0.220	0.254	0.338
rec_{jt}	46	2.435	0.338	2.038	2.167	2.681	3.583
Δeps^{v}_{jt}	46	0.058	0.024	0.002	0.043	0.072	0.108
roa_{jt}	46	0.012	0.040	−0.063	−0.015	0.040	0.109
$size_{jt}$	46	8.167	0.889	6.966	7.487	8.851	11.396
$bvol_{jt}$	46	0.711	0.308	0.328	0.538	0.776	2.234
D^{CG}_{jt}	46	0.101	0.028	0.065	0.080	0.118	0.162
D^{CS}_{jt}	46	0.155	0.055	0.000	0.114	0.191	0.275
D^{MT}_{jt}	46	0.039	0.008	0.000	0.038	0.043	0.049
D^{Fin}_{jt}	46	0.175	0.055	0.056	0.131	0.221	0.266
D^{HC}_{jt}	46	0.133	0.035	0.000	0.116	0.154	0.188
D^{ID}_{jt}	46	0.170	0.022	0.111	0.156	0.186	0.204
D^{Tech}_{jt}	46	0.127	0.084	0.072	0.104	0.123	0.667
D^{TE}_{jt}	46	0.011	0.004	0.000	0.009	0.012	0.034
D^{UT}_{jt}	46	0.023	0.011	0.000	0.012	0.033	0.040

3.4 Empirical Strategy

I follow Berry, Levinsohn and Pakes (1995) and Nevo (2001). Two information providers offer their services in each of T quarters, and each market contains numerous investors (consumers of information) $i \in I$. In the estimation below, a market will be defined as a quarter. All investors within a given information market, face the same services with the same attributes so that the attributes of the information services may vary over markets but not over investors within each market. Let J_t be the number of options available to each investor in market $t \in T$. Variables of stock performance or characteristics related to service j in market t (attributes), such as the stock return attainable from seeking information at provider j in market t, are denoted by the K-dimensional row vector $\boldsymbol{x}_{jt}$; besides exogenous variables, this vector includes the endogenous stock return and service price. The unobserved attributes are denoted collectively as ξ_{jt} which, in this setup, represents the common utility that investors obtain from the unobserved attributes of service j in market t. I specify ξ_{jt} as the sum of two components: the mean valuation of the service characteristics that I (the researcher, as opposed to the investor) do not observe, ξ_j, and a quarter specific deviation from this mean, $\Delta\xi_{jt}$. I control for ξ_j, by including brand-specific dummy variables in the regressions. Market-specific components are included in $\Delta\xi_{jt}$ and are left as error terms. The utility that investor i in market t obtains from information provider j depends on observed and unobserved variables of stock performance (attributes) attainable from using the service. I assume that utility takes the form

$$U_{ijt} = \delta_{jt} + \boldsymbol{x}_{jt}\tilde{\boldsymbol{\beta}}_i + \varepsilon_{ijt} \tag{1}$$

where $\delta_{jt} = \boldsymbol{x}_{jt}\bar{\beta} + \xi_{jt}$, $\bar{\boldsymbol{\beta}} = (\bar{\beta}_1, ..., \bar{\beta}_k)'$ are parameters that are the same for all investors, $\tilde{\boldsymbol{\beta}}_i = (\tilde{\beta}_{i1}, ..., \tilde{\beta}_{ik})'$ are parameters that vary across investors, and ε_{ijt} is i.i.d. Type I extreme value with zero mean. Also,

$$\tilde{\boldsymbol{\beta}}_i = \Pi D_i + \Sigma \boldsymbol{v}_i \tag{2}$$

with $\boldsymbol{v}_i \sim N(0, I_K)$ (normal distribution of preferences over the K characteristics), where D_i is a $d \times 1$ vector of observed demographic variables, Π is a $K \times d$ matrix of coefficients that measure how the taste characteristics vary with demographics, and Σ is a scaling matrix so that $\mathbb{E}(v_{i,k}^2) = 1$. I assume investors observe all the service characteristics and take them into consideration when making decisions. Investors may decide not to purchase any of the services in my sample and, as usual, I normalize the mean utility of the outside option to zero.

3.4.1 Market Size and the Outside Option

In order to calculate market size, I use the aggregate number of stocks that belong to the CRSP, searched daily, during a quarter at each source of financial information. Subsequently, I multiply the aggregate number of stocks, by the average of daily counts of non-crawler IP addresses that access information on a stock (1,296). In order to obtain this estimate, I filter out crawlers and count the number of different IP addresses, grouped by CIK codes, that sought information at EDGAR using the 36 EDGAR log files in the period 2015-01-01 - 2015-02-05. Unfortunately, I do not have data on the daily number of Bloomberg terminal users who searched information at the stock or the aggregate level. Thus, I am assuming that those who access Bloomberg Terminals can also access EDGAR files, and then, this is a potential number for Bloomberg.

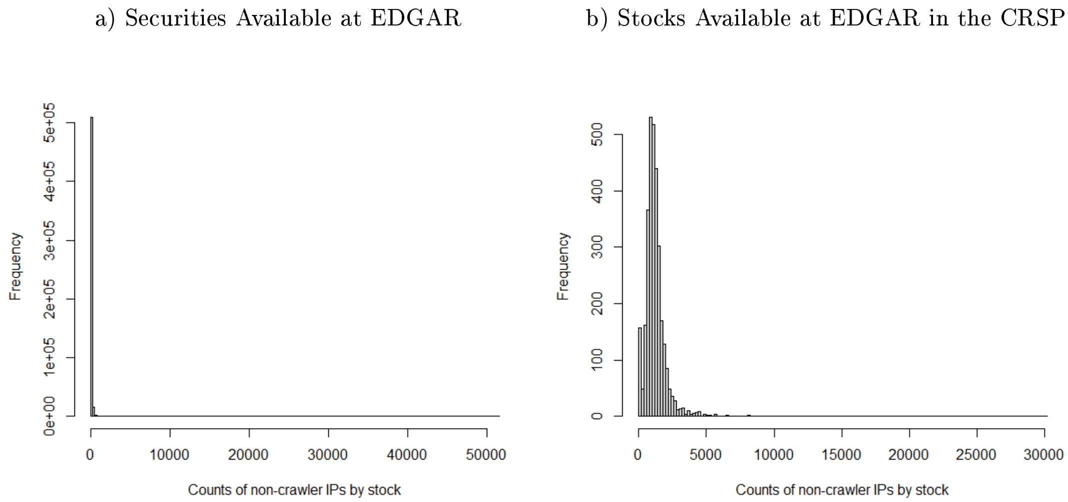


Figure 3.4: **Histograms of the number of different non-crawler IPs accessing data on a given stock, between 2015-01-01 and 2015-02-05.** In panel a), the data includes all securities available at EDGAR. Panel b) shows data on CRSP stocks only.

Table 3.5: Summary Statistic on Non-Crawler IP Counts by Stock, between 2015-01-01 and 2015-02-05.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Securities Available at EDGAR	1	4	10	41.73	22	51,406
Stocks Available at EDGAR in the CRSP	2	821	1,125	1,296	1,477	30,131

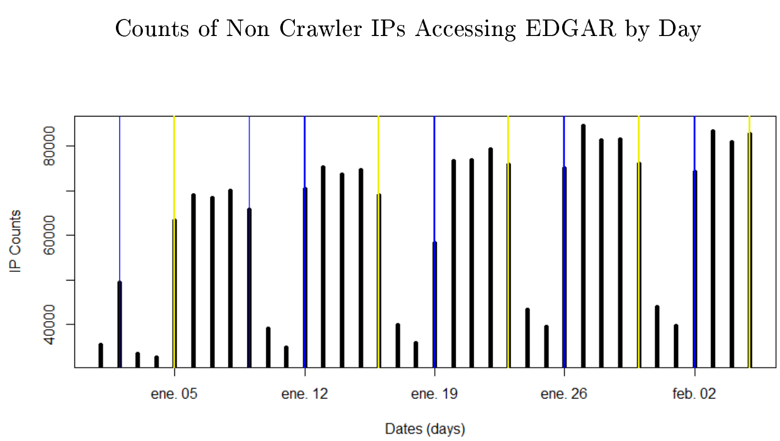


Figure 3.5: **Number of different non-crawler IPs accessing data at EDGAR by day, from 2015-01-01 to 2015-02-05.** Yellow lines represent Mondays and blue lines represent Fridays.

Around 32% of the US population have investments in stocks, bonds, mutual funds, or other securities outside of retirement accounts, percentage that has changed little over the years (34% in 2009, 32% in 2012, 30% in 2015 and 32% 2018), and from these, 74% own individual stocks in 2015 as well as in 2018 (see Lin et al., 2019a, b). There are sources of financial information alternative to internet-based sources such as radio and TV programs, printed news, interpersonal sources, workplace-based sources and personal financial advisors, and many investors do not acquire data often. For instance, Loibl and Hira (2009) analyze survey results of investors in the US, and identify that 30% of investors in their sample, are "reluctant" who reported to "seldom" or "never" gather financial information, whereas only 11% reported to obtain information "often" or "very often" from online sources. With 32% $\times$ 74% $\times$ 11% as the portion of the population (327'677,163) interested in online stock data, there are 8'535,334.7 online potential searches in a day, and 537'726,086 in a quarter. The definition of market shares I use is then the following,

$$Market\ Share_{jt} = \frac{Stocks_{jt} \times I\bar{P}s}{Potential\ Searches} \tag{3}$$

where $Stocks_{jt}$ is the aggregate daily number of stocks of the CRSP, searched at source j during quarter t, $I\bar{P}s = 1,296$ is the average count of non-crawler IP addresses that access information on a stock and $Potential\ Searches = 537'726,086$ is the number of potential searches in a quarter as defined before. The market share of the outside option is

$$1 - \left[\frac{Stocks_{jt} \times I\bar{P}s}{Potential\ Searches}\right] \tag{4}$$

Among the paid information providers in 2018, Bloomberg's market share was around 33%[5], Refinitiv Eikon's (formerly Reuters) was around 23%, Fact-Set's around 4% and Capital IQ's around 6%. Including other not observed sources of financial information, such as, radio and TV programs, printed news and other internet-based sources, it is sensible to observe that the outside option reaches a market share as high as 73,6% in the third quarter of 2011 (see table 3.6).

[5]See Wall Street Prep's article *Bloomberg vs. Capital IQ vs. FactSet vs. Thomson Reuters Eikon* and The Financial Times' article *Bloomberg and Reuters lose data share to smaller rivals*

Table 3.6: Market Shares In Number of Searches

Market shares are defined as the aggregate number of stocks that belong to the CRSP, searched in a quarter at each source of financial information, times the average count of non-crawler IP addresses that access information on a stock, divided by the population interested in online stock data.

Date	Product id	Shares	Date	Product id	Shares
2010-06-30	1	0.081	2013-06-28	1	0.079
2010-06-30	2	0.363	2013-06-28	2	0.418
2010-09-30	1	0.085	2013-09-30	1	0.075
2010-09-30	2	0.365	2013-09-30	2	0.426
2010-12-31	1	0.062	2013-12-31	1	0.072
2010-12-31	2	0.368	2013-12-31	2	0.433
2011-03-31	1	0.092	2014-03-31	1	0.080
2011-03-31	2	0.372	2014-03-31	2	0.433
2011-06-30	1	0.104	2014-06-30	1	0.079
2011-06-30	2	0.376	2014-06-30	2	0.451
2011-09-30	1	0.054	2014-09-30	1	0.069
2011-09-30	2	0.210	2014-09-30	2	0.452
2011-12-30	1	0.068	2014-12-31	1	0.066
2011-12-30	2	0.246	2014-12-31	2	0.460
2012-03-30	1	0.097	2015-03-31	1	0.061
2012-03-30	2	0.326	2015-03-31	2	0.452
2012-06-29	1	0.083	2015-06-30	1	0.063
2012-06-29	2	0.394	2015-06-30	2	0.466
2012-09-28	1	0.083	2015-09-30	1	0.065
2012-09-28	2	0.398	2015-09-30	2	0.484
2012-12-31	1	0.077	2015-12-31	1	0.062
2012-12-31	2	0.399	2015-12-31	2	0.488
2013-03-28	1	0.077			
2013-03-28	2	0.389			

3.4.2 Predicted Market Shares

Investors are assumed to use the service that gives the highest utility. This implicitly defines the set of unobserved variables that lead to the choice of information provider j:

$$A_{jt}(\boldsymbol{x}_t, \boldsymbol{\delta_t}, \Sigma, \Pi) = \{(D_i, \boldsymbol{v}_i, \boldsymbol{\varepsilon}_{it}) | U_{ijt} \geq U_{ilt} \quad \forall l = 0, 1, 2\}$$

where $\boldsymbol{\delta_t} = (\delta_{1t}, \delta_{2t})'$. Assuming ties occur with zero probability, the market share s_{jt} of the jth provider as a function of the mean utility levels of all the $J+1$ services, given the parameters, is

$$s_{jt}(\boldsymbol{x}_t, \boldsymbol{\delta_t}, \Sigma, \Pi) = \int_{A_{jt}} dP^*(\boldsymbol{\varepsilon}) dP^*(\boldsymbol{v}) dP^*(D) \tag{5}$$

where $P^*(\cdot)$ denotes population distribution functions. With the specified distributions for ε_{ijt} and $\boldsymbol{v_i}$, the market shares are

$$s_{jt} = \int \left[\frac{e^{(\boldsymbol{x}_{jt}\bar{\boldsymbol{\beta}} + \xi_j + \Delta\xi_{jt}) + (\boldsymbol{x}_j \tilde{\boldsymbol{\beta}}_i)}}{1 + \sum_j e^{(\boldsymbol{x}_{jt}\bar{\boldsymbol{\beta}} + \xi_j + \Delta\xi_{jt}) + (\boldsymbol{x}_j \tilde{\boldsymbol{\beta}}_i)}} \right] dP^*(\boldsymbol{v}) dP^*(D) \tag{6}$$

In my specification, from all the service characteristics that I include, namely $return_{jt}$, $price_{jt}$ and $F13_{jt}$, only the coefficient on $return_{jt}$ will be random. I estimate the parameters of the model by following the algorithm used by Berry, Levinsohn, and Pakes (1995), without a supply model as in Nevo (2001).

3.5 Identification

I control for the mean valuation of the service characteristics that I do not observe, ξ_j, by including a brand dummy $D_{jt}^{provider}$ that takes the value of one for Bloomberg. Furthermore, I address endogeneity problems of my regressors as follows. Service price affects information demand but prices are also a consequence of changes in demand, making this variable endogenous. I instrument subscription prices using its four-quarters lag as well as the first and second lags of Google's search volume index (SVI) on the expression "Bloomberg Terminal." Choi and Varian (2012) show that Google Trends queries help describing real variables such as travel destination planning and automobile sales. I argue that the popularity of Bloomberg Terminals is related to Bloomberg's capacity to charge higher prices for their services. Furthermore, I use Microsoft sales revenues as a proxy for Bloomberg service prices, since the Terminals are not available for MAC platforms and many functionalities involve an intensive use of Excel.

a) Google SVI on "Bloomberg Terminal" b) Microsoft Sales Revenues (Adjusted for Inflation)

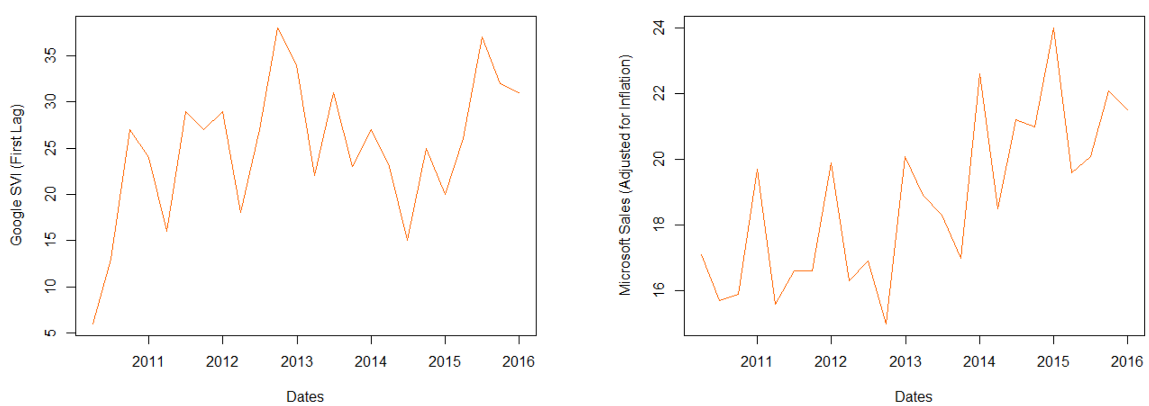


Figure 3.6: **Service Price Instruments.** In panel a), I show the quarterly series of the lagged values of Google's SVI on the expression "Bloomberg Terminal." Panel b) shows the contemporaneous quarterly values of Microsoft sales adjusted for inflation.

In addition, as the demand for financial information may influence the expectations investors have on stock returns, I instrument $return_{jt}$ with a set of plausibly exogenous variables. In particular, I assume that decisions on information sources at a given quarter, do not affect the values of the average firm size

($size_{jt}$) or the average analyst coverage on stocks with activity at each provider (rec_{jt}), and I expect these variables to affect decisions on information acquisition only through their effects on expectations on stock returns.

I also include among instruments for $return_{jt}$, the mean of return on assets (roa_{jt}) which I argue, is an accounting variable not directly affected by decisions on information acquisition at a given quarter, and also the variance of earnings per share changes (Δeps^{v}_{jt}). Furthermore, I include the variance of the difference between projected stock prices and realized stock prices ($bvol_{jt}$), which captures the difficulty to evaluate the prospects of stocks. Among the instruments are also included, the average of (non-exhaustive) sector dummies of Consumer Goods (D^{CG}_{jt}), Consumer Services (D^{CS}_{jt}), Materials (D^{MT}_{jt}), Health Care (D^{HC}_{jt}), Industrials (D^{ID}_{jt}), Technology (D^{Tech}_{jt}), Financials (D^{Fin}_{jt}), Telecomunications (D^{TE}_{jt}) and Utilities (D^{UT}_{jt}) which are exogenous. I argue that these variables affect the choice of information provider, only through their effect on stock return expectations. Additionally to the contemporaneous variables, I include the first lag of roa_{jt}, Δeps^{v}_{jt}, $bvol_{jt}$, rec_{jt} and $size_{jt}$.

Table 3.7: Linear Correlation With Expected Returns

	Estimate	p-value
$return_{jt}$	1	0
rec_{jt}	0.117	0.44
$rec_{j,t-1}$	0.024	0.87
$size_{jt}$	0.181	0.23
$size_{j,t-1}$	-0.056	0.71
roa_{jt}	0.115	0.44
$roa_{j,t-1}$	0.091	0.55
Δeps^{v}_{jt}	-0.099	0.51
$\Delta eps^{v}_{j,t-1}$	-0.005	0.97
$bvol_{jt}$	-0.023	0.88
$bvol_{j,t-1}$	0.173	0.25

With the purpose of analyzing the instruments, I estimate the conditional first-stage F statistic of Angrist and Pischke (2009). I regress service prices on instruments, and in turn regress the fitted values, on exogenous variables (brand and 13(f) list dummies) and on the fitted values, of the regression of expected returns on the set of instruments. Subsequently, I regress the last residuals on instruments and estimate the F-statistic to have an idea of the goodness of instruments for price. In order to analyze the instruments for the return expectations, I carry out a similar exercise in which the first regression contains expected returns as the dependent variable. The results of the estimations of the first regressions on instruments are in table 3.8. As a rule of thumb, values larger than 10 of Angrist and Pischke's (2009) F-statistic, are a good indicator that instruments are not weak, which is the case for the regressions of service prices and expected returns.

Table 3.8: Linear Regression of Residuals on Instruments.
Variables svi_{t-1} and $msft_t$ are Google's SVI of the expression "Bloomberg Terminal" and Microsoft Sales respectively. Prices and sales adjusted for annual inflation starting from 2010. Note: *p<0.1; **p<0.05; ***p<0.01.

	Dependent variable:	
	$service\ price_{jt}$	$return_{jt}$
	(1)	(2)
$price_{jt-4}$	0.994***	0.0001
	(0.016)	(0.0001)
svi_{jt-1}	−0.397	0.001
	(0.423)	(0.003)
svi_{jt-2}	−0.899**	−0.003
	(0.365)	(0.002)
$msft_{jt}$	1.693	0.005
	(1.219)	(0.008)
roa_{jt}	190.573	−0.622
	(272.451)	(1.734)
roa_{jt-1}	−30.621	−0.291
	(206.526)	(1.314)
$bvol_{jt}$	24.660	−0.138
	(16.908)	(0.108)
$bvol_{jt-1}$	0.984	0.020
	(7.233)	(0.046)
$size_{jt}$	20.503	0.605***
	(15.644)	(0.100)
$size_{jt-1}$	13.848	−0.502***
	(14.635)	(0.093)
rec_{jt}	19.431	−1.395***
	(49.741)	(0.317)
rec_{jt-1}	−46.109	1.443***
	(47.224)	(0.301)
Δeps^v_{jt}	−599.608***	−1.289
	(193.182)	(1.229)
Δeps^v_{jt-1}	20.453	0.561
	(221.537)	(1.410)
D^{CG}_{jt}	−816.908	0.179
	(493.782)	(3.142)
D^{CS}_{jt}	−794.252*	2.016
	(383.213)	(2.439)
D^{MT}_{jt}	157.202	3.522
	(511.184)	(3.253)
D^{Fin}_{jt}	−282.980	−0.755
	(336.189)	(2.140)
D^{HC}_{jt}	−446.065	2.869
	(402.882)	(2.564)
D^{ID}_{jt}	−662.113**	2.140
	(308.803)	(1.965)
D^{Tech}_{jt}	−693.811**	1.267
	(302.983)	(1.928)
D^{TE}_{jt}	−2,270.632*	5.802
	(1,118.614)	(7.119)
D^{UT}_{jt}	1.689	0.038
	(628.467)	(4.000)
Constant	335.217	−2.227
	(316.974)	(2.017)
Observations	46	46
Residual Std. Error (df = 22)	8.990	0.057
F Statistic (df = 23; 22)	21,539.810	4.529
Angrist-Pischke F Statistic	12.770	130.014

3.6 Results

I estimate a full random coefficients model in order to control for unobserved heterogeneity, using several specifications. The results are in table 3.9. The statistically negative estimate on $price_{jt}$ suggests that lower prices induce a higher demand for financial information services, as one would expect from the standard economic intuition. Moreover, the statistically positive estimate on $return_{jt}$, suggests that higher expectations on the stock returns attainable when searching at a source of information, induce a higher demand for the services of this provider. The statistically negative estimate on $A8_i \times return_{jt}$ suggest that the marginal valuation of expected returns decreases with higher income levels.

Table 3.9: Results: Full Random Coefficients Model.

Instruments for $return_{jt}$ are roa_{jt}, Δeps^v_{jt}, $bvol_{jt}$, rec_{jt}, $size_{jt}$ and their one-quarter lags, as well as $D^{CG}_{jt}, D^{CS}_{jt}, D^{MT}_{jt}, D^{HC}_{jt}, D^{ID}_{jt}, D^{Tech}_{jt}, D^{Fin}_{jt}, D^{TE}_{jt}$ and D^{UT}_{jt}. Bloomberg service price is instrumented with the one-quarter lag and the two-quarter lag of Google's SVI of the expression "Bloomberg Terminal", as well as with a four-quarters lag of price and the contemporaneous value of Microsoft's sales, both adjusted for annual inflation starting from 2010. Codes in parenthesis are the labels in the survey, used to obtain the corresponding draws. **Computational Details:** market shares are integrated with modified latin hypercube sampling draws (MLHS) and 100 demographic draws; method for standard errors: heteroskedastic.

	(1)	(2)	(3)	(4)
(Intercept)	1.90***	1.900***	2.000***	1.900***
	(0.520)	(0.610)	(0.600)	(0.570)
$return_{jt}$	103.00*	103.000*	102.000*	103.000*
	(54.000)	(55.000)	(54.000)	(59.000)
$price_{jt}$	-0.008**	-0.008**	-0.008**	-0.008**
	(0.004)	(0.004)	(0.004)	(0.004)
$F13_{jt}$	-8.50***	-8.400***	-8.600***	-8.200***
	(2.100)	(2.600)	(2.600)	(2.400)
$D^{provider}_{jt}$	13.00*	13.000*	13.000*	13.000*
	(6.600)	(6.700)	(6.900)	(6.600)
Demographics (Random Coefficients)				
(Unobserved) $\sigma \times return_{jt}$	15.0**	15.0**	15.0**	15.0*
	(7.200)	(7.500)	(7.200)	(7.900)
(A8) $Income_i \times return_{jt}$	-9.0*	-9.0*	-9.0*	-9.0*
	(4.700)	(5.200)	(5.300)	(5.300)
(M6) $Literacy_i \times return_{jt}$	–	-0.400	-0.400	-0.400
	–	(20.000)	(20.000)	(17.000)
(A3) $Male_i \times return_{jt}$	–	–	2.0	2.0
	–	–	(10.000)	(12.000)
(J2) $Risklover_i \times return_{jt}$	–	–	–	-13.0
	–	–	–	(39.000)

3.6.1 Robustness Checks

To calculate market sizes, in the previous section I used the average, of daily counts of non-crawler IP addresses that access information on a stock, with information from 36 EDGAR log files, as well as the number of stocks searched each period. Yet, the average number of times that unique visitors, access the system during a quarter, per stock, could be not the same for EDGAR and Bloomberg. 12.7% of respondents of the Investor Survey of 2015[6], which is a separate follow-up survey of investors from the 2015

[6] Available only for 2015 and 2018.

NFCS, reported to have used paid online services during the past 12 months (13.38% in 2018 reported to have used paid subscription services). With a market share of 33% among subscription providers, and 537'726,086 online potential searches in a quarter, Bloomberg's market size is 22'536,100.26. Meanwhile, the number of non-robot searches at EDGAR was 18'462,680 in the fourth quarter of 2015, for which, in terms of visits, Bloomberg market size is 1.22 times the EDGAR size. In this subsection, I show the results of multiplying Bloomberg's unique visitors per stock, by 1.22 (see market shares in table 3.11 of the appendix). The estimates shown in table 3.10, remain statistically positive for $return_{jt}$ and statistically negative for $price_{jt}$. Moreover, the estimates on the interactions between income and return are significant in all specifications.

Table 3.10: Results: Full Random Coefficients Model.

Instruments for $return_{jt}$ are roa_{jt}, Δeps^{v}_{jt}, $bvol_{jt}$, rec_{jt}, $size_{jt}$ and their one-quarter lags, as well as $D^{CG}_{jt}, D^{CS}_{jt}, D^{MT}_{jt}, D^{HC}_{jt}, D^{ID}_{jt}, D^{Tech}_{jt}, D^{Fin}_{jt}, D^{TE}_{jt}$ and D^{UT}_{jt}. Bloomberg service price is instrumented with the one-quarter lag and the two-quarter lag of Google's SVI of the expression "Bloomberg Terminal", as well as with the contemporaneous value of Microsoft's sales, both adjusted for annual inflation starting from 2010. Codes in parenthesis are the labels in the survey, used to obtain the corresponding draws. **Computational Details:** market shares are integrated with modified latin hypercube sampling draws (MLHS) and 100 demographic draws; method for standard errors: heteroskedastic.

	(1)	(2)	(3)	(4)
(Intercept)	1.303***	1.216***	1.318***	1.377***
	(0.301)	(0.311)	(0.302)	(0.325)
$return_{jt}$	102.446**	104.618**	102.021*	102.417*
	(44.960)	(47.700)	(61.520)	(53.320)
$price_{jt}$	-0.005*	-0.006**	-0.005*	-0.006**
	(0.003)	(0.003)	(0.003)	(0.003)
$price_{jt-4}$	0.002	0.004*	0.002)	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
$F13_{jt}$	-5.549***	-5.125***	-5.636***	-5.797***
	(1.246)	(1.289)	(1.282)	(1.283)
$D^{provider}_{jt}$	3.100	3.223	3.614	4.705
	(4.368)	(4.324)	(4.599)	(4.710)
Demographics (Random Coefficients)				
(Unobserved) $\sigma \times return_{jt}$	5	5	5	5
	(11.589)	(11.613)	(13.415)	(13.689)
(A8) $Income_i \times return_{jt}$	-9**	-9**	-9*	-9*
	(3.952)	(4.546)	(5.287)	(4.663)
(M6) $Literacy_i \times return_{jt}$	–	-3	-3	-3
	–	(11.409)	(14.254)	(15.435)
(A3) $Male_i \times return_{jt}$	–	–	6	6
	–	–	(17.301)	(16.245)
(J2) $Risk\ lover_i \times return_{jt}$	–	–	–	-13
	–	–	–	(17.773)

3.7 Discussion of Results and Limitations

These results were obtained from data on the daily number of stocks, sought by investors through each source of information, and then transformed in number of investors by using the number of IP addresses per stock in EDGAR. My data does not capture all differences of information acquisition between sources, since up to now I have not included the number of user accounts of Bloomberg terminals in my sample. Ideally, the variation in information acquisition between providers in the data would be a consequence of observing the number of users with activity at each source of information, and the type of security that was of each user's interest, so that I could associate investor choice of information provider directly to his portfolio characteristics.

In order to capture differences of information acquisition between sources in a more direct manner, I will look for data on Bloomberg users. Bloomberg counts with identity objects that contain information on, e.g. the last time the user logged in or the functions users are accessing, information that is used by Bloomberg's sales force through the function UUID (not available for final users). Incorporating this data in my sample will allow me to compare search activity at each source in a direct manner and to capture all differences in information acquisition. Furthermore, I will estimate market shares from media attention, as information providers that receive more attention should count with higher search activity. In particular, I will count the number of times Bloomberg and EDGAR appear in the crowd-sourced content service Seeking Alpha, which has over 16,000 contributors who share finance news and investing ideas, covering around 8,000 tickers in an archive of over 1 million articles as of 2021. Web-scrapping *seekingalpha.com* and counting the number of times contributors make a reference to a source of information, has important advantages over downloading data from the more traditional source of media attention Google Trends. First, a great amount of Google's records for searches on the words "EDGAR" or "Bloomberg" is not related to financial information acquisition as many Google users are not investors interested in the prospect of firms, but are non investors curious about these platforms and their workings. Second, the number of references per period of time is more intuitive than the Search Volume Index which takes values from 0 to 100, where a value of 50 is vaguely described as "half of popularity" by Google.

Also, my data on the costs of using each provider, included historical information on only one Bloomberg client, but, besides the fact that there are several clients around the world, the cost of acquiring data varies with the price of technology even for open access sources. Including data on more clients of Bloomberg as well as data on technology, will offer me better inputs to estimate the effects of prices on financial information acquisition.

Two important counterfactual questions can be answered with this research. The first one, related to the fact that open access sources of information also provide hard financial data (besides news), consists in understanding whether the market will eventually consider information on accounting variables, provided by subscription services, as completely substitutable, or in question form: would searches of financial statements at Bloomberg be completely absorbed by EDAGAR in the event that Bloomberg disappears from the market? Would be there investors who stop being market participants if the subscription provider with the highest market share disappears? The second one, is related to the fact that many investors, use EDGAR not only because it contains relevant financial information but also because it does not require a subscription. What would happen to search activities if EDGAR became a subscription provider and its access price increases?

3.8 Conclusions

The widespread access to internet, offers to investors a large set of choice alternatives of financial information providers, which comprehends subscription services as well as no-fee open access providers. In this paper, I present empirical evidence that investors substitute information sources according to subscription prices, as well as to stock return expectations and investor's income.

Understanding investors' choices over sources is important for information vendors as well as for investors' welfare. Standard economic theory tells us that stock prices and stock returns convey information from the informed investors to the uninformed, in a degree that depends on the number of individuals who are informed. While having greater amounts of information delivered through many providers improves the quantity and diversity of information available to investors, it comes with its new own difficulties for decision making: cognitive resources and attention are scarce resources that must be allocated. This research contributes to the literature on information acquisition in financial markets, presenting empirical evidence that help to understand decisions over financial data vendors.

3.9 Appendix

Table 3.11: Market Shares for Robustness Check

EDGAR market size is defined as the aggregate of the daily number of stocks that belong to the CRSP, searched in a quarter, times the average count of non-crawler IP addresses that access information on a stock, For Bloomberg, the average count of IP addresses is estimated using that of EDGAR, times 1.22. Market shares equal market sizes divided by the population interested in online stock data.

Date	Product id	Shares	Date	Product id	Shares
2010-06-30	1	0.099	2013-06-28	1	0.096
2010-06-30	2	0.363	2013-06-28	2	0.418
2010-09-30	1	0.103	2013-09-30	1	0.092
2010-09-30	2	0.365	2013-09-30	2	0.426
2010-12-31	1	0.076	2013-12-31	1	0.087
2010-12-31	2	0.368	2013-12-31	2	0.433
2011-03-31	1	0.113	2014-03-31	1	0.097
2011-03-31	2	0.372	2014-03-31	2	0.433
2011-06-30	1	0.127	2014-06-30	1	0.096
2011-06-30	2	0.376	2014-06-30	2	0.451
2011-09-30	1	0.066	2014-09-30	1	0.084
2011-09-30	2	0.210	2014-09-30	2	0.452
2011-12-30	1	0.084	2014-12-31	1	0.081
2011-12-30	2	0.246	2014-12-31	2	0.460
2012-03-30	1	0.119	2015-03-31	1	0.075
2012-03-30	2	0.326	2015-03-31	2	0.452
2012-06-29	1	0.101	2015-06-30	1	0.077
2012-06-29	2	0.394	2015-06-30	2	0.466
2012-09-28	1	0.101	2015-09-30	1	0.079
2012-09-28	2	0.398	2015-09-30	2	0.484
2012-12-31	1	0.094	2015-12-31	1	0.076
2012-12-31	2	0.399	2015-12-31	2	0.488
2013-03-28	1	0.094			
2013-03-28	2	0.389			

3.10 References

Angrist, J., and Pischke, J.S., (2009). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton.

Asthana, S., Balsam, S., and Sankaraguruswamy, S. (2004). Differential response of small versus large investors to 10-K filings on EDGAR. *The Accounting Review*, 79(3), 571-589.

Baum, J. A., Bowers, A., and Mohanram, P. (2016). Mutual forbearance and competition among security analysts. *Management Science*, 62(6), 1610-1631.

Behrendt, S., Peter, F. J., and Zimmermann, D. J. (2020). An encyclopedia for stock markets? Wikipedia searches and stock returns. *International Review of Financial Analysis*, 101563.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.

Chi, S. and Shanthikumar, D. M. (2018). Do Retail Investors Use SEC Filings? Evidence from EDGAR Search (October 25, 2018). Available at SSRN

Choi, H., and Varian, H. (2012). Predicting the present with Google Trends. *Economic record*, 88, 2-9.

Coeurdacier, N., and Rey, H. (2013). Home bias in open economy financial macroeconomics. *Journal of Economic Literature*, 51(1), 63-115.

Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5), 1461-1499.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96(4), 1043-1068.

Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, 70(3), 393-408.

Hilary, G., and Hsu, C. (2013). Analyst forecast consistency. *the Journal of Finance*, 68(1), 271-297.

Hu, Y., Li, X., Goodell, J. W., and Shen, D. (2021). Investor attention shocks and stock co-movement: substitution or reinforcement?. *International Review of Financial Analysis*, 73, 101617.

LaValle, I. H. (1968). On cash equivalents and information evaluation in decisions under uncertainty Part I: Basic theory. *Journal of the American Statistical Association*, 63(321), 252-276.

Loughran, T. and McDonald, B. (2017). The use of EDGAR filings by investors. *Journal of Behavioral Finance*, 18(2), 231-248.

Loibl, C., and Hira, T. K. (2009). Investor information search. *Journal of economic psychology*, 30(1), 24-41.

Lipman, B. L. (1991). How to decide how to decide how to...: Modeling limited rationality. *Econometrica: Journal of the Econometric Society*, 1105-1125.

Lin, J., Bumcrot, C., Ulicny, T., Mottola, G., Walsh, G., Ganem, R., Kieffer, C. and Lusardi, A. (2019a). The State of US Financial Capability: The 2018 National Financial Capability Study. FINRA Investor Education Foundation.

Lin, J. T., Bumcrot, C., Mottola, G., and Walsh, G. (2019b). Investors in the United States: A report of national financial capability study. FINRA: Investor Education Foundation.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3), 483-510.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2), 307-342.

Niehaus, G., and Zhang, D. (2010). The impact of sell-side analyst research coverage on an affiliated broker's market share of trading volume. *Journal of Banking & Finance*, 34(4), 776-787.

Peng, L. and Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3), 563-602.

Sadler, E. (2021). A Practical Guide to Updating Beliefs From Contradictory Evidence. *Econometrica*, 89(1), 415-436.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2), 461.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, 20(4), 518-529.

Vlastakis, N., and Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, 36(6), 1808-1821.

Wang, Pingle. (2020). Demand for Information and Stock Returns: Evidence from EDGAR. Available at SSRN

Willinger, M. (1989). Risk aversion and the value of information. *Journal of Risk and Insurance*, 320-328.