A Mind is a Terrible Thing to Change: Confirmatory Bias in Financial Markets[*]

Sebastien Pouget[†], Julien Sauvagnat[‡] and Stephane Villeneuve[§]

September 5, 2016

[†]Toulouse School of Economics, University of Toulouse 1 Capitole (IAE-CRM-IDEI), 21 allee de Brienne, 31000 Toulouse, spouget@univ-tlse1.fr.

[‡]Bocconi University, CEPR and IGIER, 1 Via Roentgen, 20136 Milano, julien.sauvagnat@unibocconi.it.

[§]Toulouse School of Economics, University of Toulouse 1 Capitole (CRM-IDEI), 21 allee de Brienne, 31000 Toulouse, stephane.villeneuve@univ-tlse1.fr.

A Mind is a Terrible Thing to Change: Confirmatory Bias in Financial Markets

Abstract

This paper studies the impact of the confirmatory bias on financial markets. Building on Rabin and Schrag (1999), we propose a model in which some traders may ignore new evidence when it is inconsistent with their favorite hypothesis regarding the state of the world. The confirmatory bias provides a unified rationale for several existing stylized facts including excess volatility, excess volume and momentum. It also delivers novel predictions: at the individual level, traders' belief updating depends on the sign of past signals and previous beliefs, and, at the stock level, differences of opinion should be larger when past subsequent signals have different signs. Using data on US firms' earnings announcements and analysts' earnings forecasts from 1982 to 2014, we find strong empirical support for these predictions, suggesting that the confirmatory bias is at work in financial markets.

Keywords: financial markets, psychological biases, confirmatory bias, momentum, bubbles, trading strategies, differences of opinion, analysts' forecasts.

"A mind is a terrible thing to change... you believe stocks are going to outperform other assets, and all you can hear are warnings of the bloodbath to come in the bond and commodity markets. In short, your own mind acts like a compulsive yes-man who echoes whatever you want to believe." (Jason Zweig, in the *Wall Street Journal*, November 19, 2009)

The psychology literature defines the confirmatory bias as "the seeking or interpreting of evidence in ways that are partial to existing beliefs" (Nickerson, 1998).[1] This bias is intimately related to the dynamics of belief formation and thus appears particularly relevant in the context of trading and investing activities. To the best of our knowledge, the present paper is the first to study how the confirmatory bias affects beliefs and asset pricing in financial markets.

Building on Rabin and Schrag (1999), we propose a simple dynamic model of financial markets in which some traders are prone to the confirmatory bias: biased traders may ignore information when it is inconsistent with their prior views. In a framework with public information only, this bias creates differences of opinion between rational speculators and confirmatory-biased traders over the interpretation of public information.

These differences of opinion give rise to trading. Speculators take opposite positions with respect to biased traders and thus have a corrective impact on prices. Transaction costs however limit the effectiveness of corrective strategies causing the views of both speculators and biased traders to be incorporated into asset prices.

To understand how the confirmatory bias affects financial markets, consider that traders initially hold positive views about an asset's future cash flow. If subsequent information is

---

[1]In his book surveying biases in human reasoning, Jonathan Evans, a leading scholar in psychology, refers to the confirmatory bias as "the best known and most widely accepted notion of inferential error" (Evans, 1989). Starting with the seminal contributions of Lord, Ross, and Lepper (1979) and Darley and Gross (1983), the confirmatory bias has been extensively documented by psychologists. The prevalence of this bias has recently been confirmed by Hart et al. (2009) in a meta-analysis based on 67 articles and more than 8,000 individuals. See also the experiments on political stock markets by Forsythe, Nelson, Neumann, and Wright (1992) that display belief patterns consistent with the confirmatory bias.

also positive, then all traders interpret it correctly. However, if this information is negative, confirmatory-biased traders have a given probability to ignore the negative information.[2]

Our theoretical analysis shows that the confirmatory bias provides a unified explanation for several stylized facts, including excess volume (De Bondt and Thaler (1995)), excess volatility (Leroy and Porter (1981), Shiller (1981)), and momentum (Jegadeesh and Titman (1993)). Some behavioral finance theories have been proposed to account for these stylized facts (see for example Hirshleifer (2001) for a survey). Overconfidence may explain excess volume, excess volatility and even momentum when coupled with self-attribution bias (see Benos (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), and Odean (1998)). The representativeness heuristics may also rationalize momentum (see Barberis, Shleifer and Vishny (1998) and Rabin and Vayanos (2010)). As shown by Bouchaud, Krueger, Landier, and Thesmar (2016), sticky expectations can also be invoked to explain momentum, as well as other asset pricing phenomena such as the quality anomaly. Finally, Hong and Stein (2007) argue that gradual information flow and limited attention may explain excess volume and momentum.

To show the unique impact of the confirmatory bias, we then derive novel theoretical predictions. The core predictions of our model are that, at the individual level, i) after initial positive (respectively, negative) signals, some traders tend to ignore subsequent signals that are negative (respectively, positive), and ii) traders that have previously updated their beliefs upward (respectively, downward) are more likely to ignore subsequent signals that are negative (respectively, positive).

At the stock level, as conflicting signals opens opportunities for biased traders to misinterpret information, this creates differences of opinion. As a consequence, our model also predicts that when some traders suffer from the confirmatory bias, differences of opinion

---

[2]This is in line with the experimental psychology literature, e.g., Bodenhausen (1988), that shows that people are more likely to forget news that are inconsistent with their priors. All our results hold if we assume instead that biased traders misinterpret contradicting signals as supportive, as in Rabin and Schrag (1999).

should be larger when past subsequent signals have different signs.[3]

We next test these novel theoretical predictions using data on US firms' earnings announcements and analysts' earnings forecasts over the period 1982 to 2014. First, consistent with the basic prediction of the model, we find that analysts are less likely to revise their forecasts upward (respectively, downward) after good (respectively, bad) earnings surprises when one of the previous earnings surprises was negative (respectively, positive). Using previous analysts' forecast revisions (either upwards or downwards) as a proxy for their prior beliefs about an asset (either bearish or bullish), we also find consistent with the model that analysts are less likely to revise their forecasts upward (respectively, downward) after good (respectively, bad) earnings surprises, for earnings surprises that are of a different sign compared to their prior beliefs.

Finally, to test the last prediction, we follow the literature (see, e.g., Diether, Malloy, and Scherbina (2002)) and use dispersion in analysts' annual earnings forecasts as a proxy for differences of opinion. We find that forecasts dispersion is significantly larger when news revealed by earnings announcements over the previous quarters have different signs. Our findings hold when we consider changes in the sign of earnings surprises over either the previous two or previous three quarters.[4] These empirical results suggest that the confirmatory bias is at work in financial markets.

Overall, our paper's contribution is threefold. First, we propose a parsimonious and tractable model in which departure from perfect rationality is well-grounded in psychology

---

[3]Better understanding the formation of differences of opinion is important for several reasons. On the one hand, differences of opinion have been invoked to explain overvaluation in financial markets with short-selling restrictions (see Miller (1977)'s model and empirical tests by Chen, Hong and Stein (2002), Diether, Malloy, and Scherbina (2002), and Goetzmann and Massa (2005)). On the other hand, differences of opinion also seem to play an important role in volume formation (see Banerjee and Kremer (2010)), and, when they relate to the variance of returns, excess volatility (see Duchin and Levy (2010)).

[4]Our analyses control for numerous factors that are known to affect analysts' earnings forecasts and differences of opinion, including analyst coverage, market capitalization, stock returns volatility (a measure of uncertainty), book-to-market ratio, and return on asset. Finally, we obtain identical results when uncertainty is proxied by implied volatility as measured from option prices and when the sample is restricted to firms with fiscal year ending in December.

and driven by only one parameter: when the severity of the confirmatory bias is null, we are back to the perfectly rational benchmark. Second, we show that the confirmatory bias alone offers a unified rationale for several existing stylized facts, complementing previous explanations offered in the behavioral finance literature. Finally, we deliver and find support in the data for novel empirical predictions that follow from the confirmatory bias.

1. The model

We consider a pure exchange economy with one risky asset in fixed supply and a riskless asset in perfectly elastic supply. The riskless asset rate of return is normalized to zero. There are T dates of trading indexed by $t \in \{0, 1, 2, ..., T-1\}$. Consumption occurs only at date T when the risky asset distributes a cash flow v. For simplicity, we set $T = 4$.[5] The probability distribution of the random variable v depends on the *non observable* state of the economy H or L. We assume that, conditional on the state of nature $X \in \{H, L\}$, v has a Bernoulli distribution.[6] This distribution is given by:

$$\mathbb{P}(v = 1|X) = p_X = 1 - \mathbb{P}(v = 0|X) \text{ with } p_H > p_L. \tag{1.1}$$

There is a continuum of traders with mass normalized to one indexed by $j \in [0, 1]$. A trader pertains to one of two groups of agents. Speculators are represented by the subset $[\lambda, 1]$, and biased traders are represented by the subset $[0, \lambda]$. Therefore, the mass of specu-

---

[5]One could extend the model to include an arbitrary number of trading periods T, finite or infinite, and cash flow distribution and consumption could take place at various dates across time (for example, our set up with 4 dates of trading followed by the distribution of a cash flow could be repeated any number of times). This would not affect our results because the state of the world is not revealed in our model, even when a cash flow is distributed.

[6]We have also studied a more general model, presented in the Online Appendix, in which the cash flow is distributed according to a log-normal distribution: $v = e^d$, in which d is normally distributed with mean m_X and variance σ^2, and with $m_H = 1 = -m_L$. The signals are assumed independent and identically normally distributed with mean μ_X and variance σ^2, for simplicity. Our results hold in this continuous distribution model.

lators is $1 - \lambda$ while the mass of biased traders is λ. Each trader is endowed with one unit of the risky asset and no cash. Finally, information about the cash flow is generated by the observation of public signals s_t whose probability distribution depends on the state of nature. More precisely, we assume:

$$\mathbb{P}(s_t = 0) = 1 - \delta, \mathbb{P}(s_t = 1 | X = H) = \mathbb{P}(s_t = -1 | X = L) = \delta\theta, \text{ with } \theta > \frac{1}{2}. \quad (1.2)$$

Parameters δ and θ represent, respectively, the likelihood that the signal is informative and the level of informativeness. We introduce an uninformative signal, $s_t = 0$, so that biased traders who, as explained below, might neglect the informational content of news cannot realize that they are actually biased: in our model, it is also the case that rational traders do sometimes perceive that signals are uninformative.[7]

In order to focus on the informational aspects of financial markets, we assume that traders are risk neutral.[8] Absent market frictions, risk neutrality implies that traders would stand ready to exchange infinite amounts as long as prices do not equal their expectation of the asset value. This would prevent the existence of an equilibrium since the market would not clear. In order to avoid this phenomenon, we assume that traders incur an exogenous trading cost that is quadratic in the quantity traded and parameterized by $\frac{c}{2} > 0$.[9] The total cost of trading for trader j at date t with a demand d_t^j is thus equal to $\frac{c}{2}(d_t^j)^2$, for all $t < 4$. Trader j's objective at each date $t < 4$ is thus to maximize his expected wealth at date 4, conditional on information available at date t and on the fact that he will be able

---

[7]This statement is true as long as there is a finite number of signals. For the extreme case of an infinite number of signals (which is discussed in Section A.2 of the Online Appendix), biased traders, at infinity, will observe that they have received a proportion of uninformative signals that is higher than what their understanding of the world suggests. All our results hold if all signals are informative.

[8]This is in the spirit of Harris and Raviv (1993) and isolates our analysis from the influence of trading motives based on risk sharing or consumption smoothing.

[9]Alternatively, we could ensure existence of an equilibrium by assuming that traders can only trade up to a fixed amount of shares as in Abreu and Brunnermeier (2002). This different modeling framework would not affect our results.

to trade again in the future.[10] Trader j's wealth at date 4 is denoted by W^j and is equal to $\sum_{s=0}^{s=3} \left[d_s^j (v - P_s) - \frac{c}{2} (d_s^j)^2 \right] + v.$

The cost $\frac{c}{2}$ can be viewed as an explicit transaction cost traders have to pay to submit orders or as a proxy for the imperfect depth of financial markets. Imperfect market depth could be related to inventory or adverse selection risks borne by liquidity providers (see Madhavan (2000) and Biais, Glosten, and Spatt (2005) for surveys of the market microstructure literature dealing with those issues). The transaction cost creates limits to arbitrage and opens the scope for potential mispricings.[11]

In our model, differences of opinion emerge because all traders do not encode the public signal in the same way. On the one hand, speculators are perfectly rational in the sense that they are endowed with the actual probability model denoted by $\mathbb{P}$. On the other hand, each individual biased trader $i \in [0, \lambda]$ is endowed with a different probability model $\mathbb{P}^i$. When receiving a public signal s_t, biased trader i actually sees a signal σ_t^i.

To incorporate the confirmatory bias in our framework, we follow Rabin and Schrag (1999) and assume that biased traders misinterpret public information when it conflicts with their prior beliefs concerning the final cash flow.[12] More precisely, we consider that a confirmatory-biased trader treats some conflicting signals as uninformative and is otherwise

---

[10]Traders in our model are thus not myopic. However, as it will become clear below, our set up is such that traders optimally refrain from trading dynamically over time to take advantage of the price fluctuation across periods. Specifically, even if speculators anticipate momentum in the price over the subsequent period, they are not going to take a larger position in this period and then unwind in the next period. At equilibrium, they instead prefer speculating on the difference between the current asset price and the expected cash flow it delivers at date T. This feature of our model simplifies the analysis. We thank an anonymous referee for emphasizing this point.

[11]Alternative frameworks generating limits to arbitrage include noise trader risk as in De Long, Shleifer, Summers, and Waldmann (1990), short horizons as in Shleifer and Vishny (1997), and synchronization risk as in Abreu and Brunnermeier (2002). Barberis and Thaler (2001) survey this literature.

[12]Public signals constitute soft information in the sense that they open up a scope for misinterpretation. For example, public signals could represent the announcement of a change in corporate strategy or a report regarding the prospects of a firm's future products. On the contrary, we consider that the amount of cash flow distributed at date 4 is hard information in the sense that it cannot be misperceived by biased traders. Note that, when $\theta < 1$, the distributed cash flow, even if hard information, does not perfectly reveals the underlying state of the world: Equation (1.2) shows that, in this case, all public signals are possible in all states of the world.

completely Bayesian.[13]

To formally define the confirmatory bias, we have to specify belief dynamics regarding the state of the economy for each type of agents. We assume that the two states of nature are ex-ante equiprobable for each type of trader, that is $\mathbb{P}(X = H) = \mathbb{P}^i(X = H) = \frac{1}{2}$ for every i. We denote by $\mu_t = \mathbb{P}(X = H|s_1, \ldots, s_t)$, the beliefs of rational speculators (under the correct probability model $\mathbb{P}$) given the information they have received up to date t. To alleviate notation, we define the signal history up to date t as $h_t = \{s_1, \ldots, s_t\}$. We denote by $\mu_t^i = \mathbb{P}^i(X = H|\sigma_1^i, \ldots, \sigma_t^i)$, biased trader i's beliefs (under probability model $\mathbb{P}^i$) given the information he has received up to date t.

We consider that a signal s_t is inconsistent with prior beliefs μ_{t-1}^i if and only if s_t has a different sign than $\mu_{t-1}^i - \frac{1}{2}$, the difference between the conditional and the unconditional belief on the state of the economy. When biased trader i is bullish, that is, he believes that the good state of the economy is more likely, he might under-react to negative signals. Likewise, when biased trader i is bearish, he might under-react to positive signals.

We can now formally define how the confirmatory bias affects biased traders' perception of information. The following definition precisely characterizes biased traders' distortion of public signals.

Definition 1.1. *For $i \in [0, \lambda]$, the information perceived by biased trader i is $\sigma_1^i = s_1$ and,*

---

[13]In the empirical tests of our theory presented below, we use quarterly earnings announcements as a counterpart for the public signals in our model. Earnings surprises are not limited to the disclosure of quantitative information, but are also conveyed with value-relevant "soft" information. Market participants can predict future firms' outcomes by analyzing the "tone" of earnings press releases (David, Piger, and Sedor 2012), earnings conference calls (Matsumoto et al. 2011, Mayew and Venkatachalam, 2012) or news stories released around earnings announcements (Tetlock et al., 2008). Interestingly, Engelberg (2008) finds that the market incorporates soft information released during earnings announcements more slowly than hard information, which is consistent with the idea that soft information is more difficult to process and more likely to be subject to different interpretations (Petersen, 2008). We thank an anonymous referee for highlighting this point.

for $t \in \{2,3\}$, under the probability measure $\mathbb{P}$:

$$\sigma_t^i = s_t \mathbb{1}_{(\mu_{t-1}^i - \frac{1}{2})s_t \geq 0} + z_t^i s_t \mathbb{1}_{(\mu_{t-1}^i - \frac{1}{2})s_t < 0}, \tag{1.3}$$

where z_t^i is a random variable taking values in $\{0,1\}$ with probability distribution $\mathbb{P}(z_t^i = 0) = q$, and $\mathbb{1}_{(.)}$ is the indicator function that takes the value 1 if the condition is satisfied and 0 otherwise. We assume that the random variables $(z_t^i)_{i \in [0,\lambda]}$ are independent and that z_t^i is independent from z_s^i for $s \neq t$.

This definition indicates that biased traders may ignore a signal when this signal is inconsistent with their prior beliefs, i.e., biased traders might misperceive an informative signal as an uninformative one.[14] The timing of each date $t \in \{2,3\}$ is thus as follows: i) a signal is generated, ii) agents perceive a signal that might be different than the one generated, and iii) agents update their beliefs using Bayes' rule as if the perceived signal were the one truly generated.

Definition (1.1) deserves some comments. First, a biased trader may ignore, with probability q, a public signal that is inconsistent with his prior beliefs. The greater is q, the more severe the confirmatory bias is. Second, a proportion q of biased traders whose prior beliefs are inconsistent with the public signal misperceive the signal s_t. This is because the random variable z_t^i is independent across biased traders and because traders form a continuum. Third, we assume, for simplicity, that biased traders do not amplify signals that are consistent with prior beliefs. Fourth, at date 1, biased traders see the actual signal because they do not have formed an opinion yet. Their perception is thus not biased at this date.

---

[14]As shown in the Online Appendix, all our results hold if biased traders change the valence of signals that are inconsistent with their prior beliefs. This more extreme way of defining the confirmatory bias is used by Rabin and Schrag (1999) and allows our model to display belief polarization according to which, when some traders are optimistic and others are pessimistic, news can actually make beliefs diverge even more. It could also be interesting to study a form of confirmatory bias in which biased traders retroactively reinterpret past evidence in light of their new view of the world, which is referred to as an "epiphany" by Rabin and Schrag (1999). We leave open this avenue for further research that will require additional work in experimental psychology to provide guidance for the economic modeling.

We now turn to the analysis of the model.

2.   Equilibrium prices and rational benchmark

Our model features quadratic transaction costs. Individual asset demands are thus finite despite risk neutrality and there exists an equilibrium. Standard arguments, developed in the Appendix, show that prices in our financial market are weighted averages of speculators' and biased traders' beliefs given their respective information set. They are given in the following proposition where we denote by $\mathbb{P}_t$ the conditional probability $\mathbb{P}(.|h_t)$ and by $\mathbb{P}_t^i$ the probability $\mathbb{P}^i(.|(z^i h)_t)$, where $(z^i h)_t = \{s_1, z_2^i s_2, \ldots, z_t^i s_t\}$.

Proposition 2.1. *At each date $t \in \{0, 1, 2, 3\}$, the equilibrium price is given by:*

$$P_t = \int_0^\lambda \mathbb{P}_t^i(v = 1) \, di + (1 - \lambda)\mathbb{P}_t(v = 1)$$

$$\Leftrightarrow$$

$$P_t = p_L + (p_H - p_L)\, \bar{\mu}_t$$

where $\bar{\mu}_t = (1 - \lambda)\, \mu_t + \int_0^\lambda \mu_t^i di$ represents traders' average belief regarding the state of the world.[15]

The price is thus a linear function of traders' average belief. Let us describe explicitly the dynamics of equilibrium prices. Denote by $P_t(h_t)$ the equilibrium price that prevails after a history h_t of actual public signals. At dates $t \in \{0, 1\}$, all agents perceive correctly the public signal thus:

$$P_0 = P_1(0) = p_L + (p_H - p_L)\frac{1}{2},$$

---

[15]The beliefs of biased traders depend on the realizations of random variables z_i. In the paper, we are thus computing integrals and applying the law of large numbers for a continuum of random variables. As shown in Theorem 2 of Uhlig (1996), this is adequate in our context because z_i are square integrable and i.i.d. random variables.

$$P_1(1) = p_L + (p_H - p_L)\theta,$$

and

$$P_1(-1) = p_L + (p_H - p_L)(1 - \theta).$$

At date 2, the confirmatory bias may affect biased traders' perception: they are indeed optimistic if $s_1 = 1$ or pessimistic if $s_1 = -1$. Each biased agent has a probability q to ignore s_2 if it is inconsistent with their beliefs regarding the state of the world. Due to the assumption of independence of the variables z^i, the law of large numbers implies that there is a proportion q of biased agents that have an erroneous perception of the public signal. As a consequence, $P_2(s_1, s_2)$ can take seven values:

$$P_2(0,0) = P_0$$

$$P_2(1,0) = P_2(0,1) = P_1(1)$$

$$P_2(-1,0) = P_2(0,-1) = P_1(-1)$$

$$P_2(1,1) = p_L + (p_H - p_L)\theta(2,2),$$

$$P_2(1,-1) = p_L + (p_H - p_L)((1 - \lambda q)\theta(1,2) + \lambda q\theta(1,1)),$$

$$P_2(-1,1) = p_L + (p_H - p_L)((1 - \lambda q)\theta(1,2) + \lambda q\theta(0,1)),$$

$$P_2(-1,-1) = p_L + (p_H - p_L)\theta(0,2),$$

where $\theta(s,k) = \mathbb{P}(v = 1|s$ public signals among k are strictly positive, $k-s$ are strictly negative). The explicit expression for $\theta(s,k)$ is provided in the Appendix and is obtained by applying Bayes' rule for probability updating. Notice that the same function $\theta(s,k)$ applies both for rational and biased traders: both types of traders use Bayes' rule to update their beliefs. The difference in beliefs between these traders only comes from the fact that they may not perceive the same number of strictly positive or negative signals.

In our model, beliefs are symmetric around $\frac{1}{2}$. Without loss of generality, we thus focus

on paths h_t such that $s_1 = 1$.[16] On these paths, $P_3(s_1 = 1, s_2, s_3)$ can take six values:

$$P_3(1, 0, 0) = P_1(1)$$

$$P_3(1, 1, 0) = P_3(1, 0, 1) = P_2(1, 1)$$

$$P_3(1, 1, 1) = p_L + (p_H - p_L)\theta(3, 3),$$

$$P_3(1, -1, 0) = P_3(1, 0, -1) = P_2(1, -1),$$

$$P_3(1, 1, -1) = P_3(1, -1, 1) = p_L + (p_H - p_L)((1 - \lambda q)\theta(2, 3) + \lambda q\theta(2, 2)),$$

$$P_3(1, -1, -1) = p_L + (p_H - p_L)\left((1 - \lambda q)\theta(1, 3) + \lambda q(1 - q)\theta(1, 2) + \lambda q^2\theta(2, 3)\right).$$

Before analyzing the impact of the confirmatory bias on market outcomes, it is useful to study the benchmark case in which all traders are perfectly rational. Endogenous prices in this benchmark are indicated by a star. This benchmark is nested in our model and corresponds to the case in which $\lambda = 0$ or $q = 0$. In this case, we have $P_t^* = \mathbb{P}_t(v = 1) = p_L + (p_H - p_L)\mu_t$. Given the structure of the uncertainty in our model, it is straightforward to show that the following proposition holds.

Proposition 2.2. *When all traders are perfectly rational (that is, when $\lambda q = 0$), market outcomes are as follows:*

- *After the first public signal, expected prices are constant. In particular:*

$$\mathbb{E}\left(P_3^*|s_1 = 1\right) = \mathbb{E}\left(P_2^*|s_1 = 1\right) = P_1^*\left(s_1 = 1\right),$$

and

$$\mathbb{E}\left(P_3^*|s_1 = -1\right) = \mathbb{E}\left(P_2^*|s_1 = -1\right) = P_1^*\left(s_1 = -1\right).$$

[16]Note that when $s_1 = 0$, $P_3(s_1 = 0, s_2 = x, s_3 = y) = P_2(s_1 = x, s_2 = y)$ so that focusing on $s_1 = 1$ is indeed without loss of generality.

- *Expected returns are null. In particular:*

$$\mathbb{E}\left(P_2^* - P_1^* | P_1^* - P_0^*\right) = \mathbb{E}\left(P_3^* - P_2^* | P_2^* - P_1^*\right) = 0.$$

- *Volume is null at all dates. For all t:*

$$V_t^* = 0.$$

This proposition indicates that, when all traders are rational, there is no bubble. This is a consequence of the fact that prices are a martingale. It also shows that there is no momentum. It finally characterizes volume at date t, defined as the sum of the quantities purchased at price P_t: $V_t = \int_0^1 d_t^j \mathbb{1}_{\mathbb{E}_t^j(v) > P_t} dj$. In our model, there is no volume when all traders are rational. This is because the only motive for trading, related to differences of opinion, is in this case absent.[17]

3. Financial markets with the confirmatory bias

We now focus on the case in which some traders suffer from the confirmatory bias, i.e., on the case in which $\lambda q > 0$. The confirmatory bias then generates differences of opinion and systematic mistakes. We first offer an explanation for various phenomena that have been documented in actual financial markets: excess volume, excess volatility, momentum, and the occurrence of bubbles and crashes. We then highlight the distinctive impact of the confirmatory bias by deriving new theoretical predictions on how past signals relate to future differences of opinion and bias in beliefs updating.

We start with an example of price and trade formation to explain how beliefs, prices and

---

[17]Without transaction costs, traders would be indifferent between trading and not trading, and so the volume would be indeterminate.

volume evolve over time. For traders who suffer from the confirmatory bias, first impression matters: once they have formed an opinion, biased traders tend to misperceive the subsequent signals that contradict this prior opinion. To understand how the confirmatory bias influences financial markets, it is thus particularly interesting to focus on mixed history paths, i.e., paths that include signals with different signs.

3.1   An example of price and trade formation

On mixed history paths, the confirmatory bias affects beliefs. Consider for example the mixed history path $(s_1 = 1, s_2 = -1, s_3 = -1)$. After observing $s_1 = 1$, traders become optimistic: their beliefs at date 1 are higher than at date 0, $\mu_1 = \mu_1^i = \theta > \mu_0 = \frac{1}{2}$. Note that, at the first two dates, $t = \{0, 1\}$, there is no volume because all traders have the same beliefs.

Differences of opinion arise after the occurrence of the negative signal $s_2 = -1$. Rational speculators perceive correctly this information and their beliefs come back to their initial level, i.e., $\mu_2 = \mu_0 = \frac{1}{2}$. Biased traders on the other hand are split in two groups. One group, representing a proportion $\lambda(1-q)$ of traders, perceives the negative signal correctly and ends up with the correct belief $\mu_2^i = \mu_2 = \frac{1}{2}$. The other group, in proportion λq, perceives a null signal instead of the actual negative one. These biased traders thus maintain the same optimistic belief, $\mu_2^i = \theta(1,2) = \theta = \mu_1^i$. These differences of opinion generate trade.

Finally, after the announcement of the third public signal $s_3 = -1$, rational speculators' beliefs become pessimistic, $\mu_3 = 1 - \theta < \frac{1}{2}$. In this case, biased traders are split in three groups. Some biased traders, who correctly interpreted all public signals, hold the same pessimistic beliefs as rational speculators, $\mu_3^i = \theta(1,3) = 1 - \theta < \frac{1}{2}$. These traders are in proportion $\lambda(1-q)^2$. The second group includes biased traders who have misinterpreted one of the last two signals. They hold neutral beliefs, $\mu_3^i = \theta(1,2) = \frac{1}{2}$, and are in proportion $2\lambda q(1-q)$. The last group of biased traders includes the ones who have misinterpreted the

14

last two public signals. They hold a positive belief, $\mu_3^i = \theta(2, 3) = \theta$, and are in proportion λq^2. Again, these differences of opinion generate trading.

On history paths with identical signals, either when they are all positive or all negative, the confirmatory bias does not manifest itself. Prices are thus the same as in the perfectly rational benchmark. Moreover, in our model, the sign of price changes follows the sign of the signal that is received: if $s_t = 1$, $P_t - P_{t-1} > 0$, if $s_t = 0$, $P_t - P_{t-1} = 0$, and if $s_t = -1$, $P_t - P_{t-1} < 0$.

This example of price formation describes how the confirmatory bias influences financial markets. It is useful to better understand the logic of the general results provided in the next two subsections.

3.2   Explaining existing stylized facts on financial markets

The confirmatory bias provides an explanation for several empirically documented phenomena: excess volume and volatility, momentum, and the occurrence of bubbles and crashes. First, De Bondt and Thaler (1995, p. 392) indicate that "the high trading volume on organized exchanges is perhaps the single most embarrassing fact to the standard finance paradigm." Moreover, starting with Leroy and Porter (1981) and Shiller (1981), several contributions report that asset prices are excessively volatile. These two phenomena constitute long-standing puzzles, even if some behavioral explanations have already been proposed in the literature as we discuss below.

Jegadeesh and Titman (1993) document a momentum effect in the U.S. stock market: at quarterly, semi-annual or yearly horizons, stocks that have performed well in the past have a better performance than those that performed poorly. This result was confirmed in an international context by Rouwenhorst (1998). Momentum has also been uncovered in other financial markets, such as commodities and government bonds, for example by Asness, Moskowitz and Pedersen (2013).

Finally, several authors have argued that financial markets go through periods of bubbles and crashes. For example, Shiller (2000) indicates that stock price run-ups followed by bursts tend to occur after new technological discoveries that suggest a new "era" has begun. Xiong and Yu (2011) offer clean evidence of a bubble on the Chinese warrant market.

The next proposition indicates that our model can accommodate these various stylized facts.

Proposition 3.3. *When some traders are prone to the confirmatory bias (that is, when $\lambda q > 0$), market outcomes are as follows:*

- *There is excess volume at $t \in \{2,3\}$:*

$$V_t > V_t^* = 0.$$

- *There is excess volatility in asset prices at $t = 2$ and, for $\lambda q \geq \frac{1}{2}$ and $\delta \leq \frac{3}{5}$, at $t = 3$; in this case:*

$$Var(P_t) > Var(P_t^*).$$

- *There is momentum in asset prices at $t = 1$ and $t = 2$:*

$$\mathbb{E}(P_2 - P_1 | P_1 - P_0 > 0) > 0,$$

$$\mathbb{E}(P_3 - P_2 | P_2 - P_1 > 0) > 0.$$

- *A bubble forms after a good initial public signal:*

$$\mathbb{E}\left(P_3 | s_1 = 1\right) > \mathbb{E}\left(P_2 | s_1 = 1\right) > P_1\left(s_1 = 1\right) = P_1^*\left(s_1 = 1\right).$$

16

- *A crash occurs after a bad initial public signal:*

$$\mathbb{E}\left(P_3|s_1 = -1\right) < \mathbb{E}\left(P_2|s_1 = -1\right) < P_1\left(s_1 = 1\right) = P_1^*\left(s_1 = -1\right).$$

The intuitions for these results are as follows. Volume arises directly because of differences of opinion: biased and rational traders agree to disagree. Excess volatility may arise because, along mixed signal histories, prices are drifting away from the prices that would prevail if all traders were rational. Indeed, biased traders tend to be overoptimistic after observing a first positive signal and over-pessimistic after observing a first negative signal.

Referring to the example in the previous subsection, one can observe that, when a positive signal is followed by a negative one, the price is higher than in the rational benchmark, $P_2(1, -1) > P_2^*(1, -1) = \frac{1}{2}$. This is because, after observing a first positive signal, some biased traders ignore the second signal which is negative, end up over-optimistic, and push prices too high. By the same logic, we also have $P_2(-1, 1) < P_2^*(-1, 1) = \frac{1}{2}$, a case in which prices are too low after an initially negative signal. With confirmatory-biased traders, prices may thus be more scattered than in the rational benchmark, which explains the excessive volatility.

Momentum arises in our model because biased traders' beliefs at a given date affect their future misperception of public signals. In particular, when prices increase, it is more likely that biased traders are optimistic and therefore under-react to future bad signals. As a result future prices are expected to be higher than current prices, which creates the momentum effect.

In our model, a bubble forms after a first positive public signal is announced. This is because biased traders become optimistic and are expected to ignore some future negative signals. This phenomenon amplifies with time which explains why prices are expected to

17

grow. The same intuition applies when a negative signal initially hits the market.[18] [19]

Several behavioral models based on cognitive biases have been proposed to explain the stylized facts discussed above. Most of these models display differences of opinion. Overconfidence is one reason why people may display differences of opinion and agree to disagree. It has been prominently cited to explain excess volume (see for example, Odean (1998)) and excess volatility (see Daniel, Hirshleifer, and Subrahmanyam (1998)). Daniel, Hirshleifer, and Subrahmanyam (1998) further show that, coupled with the self-attribution bias, overconfidence may explain momentum.[20]Finally, Scheinkman and Xiong (2003) show that overconfidence can lead to speculative bubbles. The representativeness heuristics has also been invoked to rationalize momentum (see Barberis, Shleifer and Vishny (1998) and Rabin and Vayanos (2010)). Hong and Stein (2007) indicate that excess volume and momentum may be explained by gradual information flow or limited attention.

---

[18]Some authors define a crash as a price decline after a bubble bursts. We instead define a crash as a situation in which the price of an asset is lower than its fundamental value and is expected by rational traders to become even more depressed in the future. In our model, a bubble can be followed by a crash if a sequence of good news is followed by a sequence of bad news.

[19]Notice that, when T goes to infinity, there is a non-null probability that some biased traders end up being sure that the wrong state of the world has occurred. This is because, if some biased traders' beliefs become inconsistent with the realized state of the world (for example, if the state of the world is L but the first signal is positive), there is always a non-null set of biased traders who ignore all the signals that are inconsistent with their initial beliefs.

One can then wonder whether it is possible that biased traders' average beliefs may not always converge to the truth. It is easy to show that biased traders misperception survives in the long run when $q = 1$. When $q < 1$, one has to show that the mass of biased traders who misperceive all the signals inconsistent with their initial belief does not converge to zero as time goes to infinity. This interesting question is left for further research.

[20]Self-attribution and confirmatory biases are related but distinct psychological effects. Self-attribution bias had been studied by Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001). This bias comes into play through the interaction between private and public information. When agents are not aware of their ability to generate precise private information, if their private signal is confirmed by subsequent public signals, agents' perception of their own ability rationally rises. The bias expresses itself when public signals disconfirms the initial private signal. In this case, agents who suffer from the self-attribution bias reduce their perceived ability but not sufficiently compared to what they should if they were not biased.

The main difference between the self-attribution bias and the confirmatory bias is that they do not require the same type of information (private and public information for the self-attribution bias versus public information only for the confirmatory bias). As a result, in our model with public signals only, the self-attribution bias has no bite. Moreover, in our empirical analysis, we only use public news and thus our results cannot be explained by the self-attribution bias.

Our model offers a complementary explanation to these phenomena and has the following distinctive features: i) it provides a novel mechanism for differences of opinion, the confirmatory bias, which is theoretically and empirically well-grounded in the psychology literature (see for example the survey on the confirmatory bias by Hart et al. (2009)), ii) it is parsimonious in the sense that departures from perfect rationality are driven by only one parameter (when q, the bias severity, is null, we are indeed back to the perfectly rational benchmark), and iii) it offers novel empirical predictions that are derived and tested below.

3.3   Novel empirical predictions

In addition to providing a unified explanation for a variety of stylized facts, our model offers new theoretical predictions that are tested in the next section. We study how differences of opinion and beliefs updating are affected by the type of signals received. We now state our most novel theoretical results.

Proposition 3.4. *When some traders are prone to the confirmatory bias (that is, when $\lambda q > 0$),*

- *Beliefs dispersion is higher conditional on past signals having different signs:*

$$\mathbb{E}\left(Var[\nu_2|_{s_1 s_2 < 0}]\right) > \mathbb{E}\left(Var[\nu_2|_{s_1 s_2 > 0}]\right),$$

$$\mathbb{E}\left(Var[\nu_3|_{s_1 s_2 < 0 \ or \ s_2 s_3 < 0}]\right) > \mathbb{E}\left(Var[\nu_3|_{s_1 s_2 > 0 \ and \ s_2 s_3 > 0}]\right),$$

   where ν_t is the probability measure characterizing beliefs distribution at date t.

- *There are less beliefs revisions when signals have different signs. For example, after a positive signal, the proportion of upward beliefs revisions is lower if at least one of the*

19

previous signal was negative:

$$\int_0^1 \mathbb{P}(\mu_2^j - \mu_1^j > 0 | s_1 > 0, s_2 > 0)dj > \int_0^1 \mathbb{P}(\mu_2^j - \mu_1^j > 0 | s_1 < 0, s_2 > 0)dj$$

$$\int_0^1 \mathbb{P}(\mu_3^j - \mu_2^j > 0 | s_1 > 0, s_2 > 0, s_3 > 0)dj > \int_0^1 \mathbb{P}(\mu_3^j - \mu_2^j > 0 | s_1 < 0 \ or \ s_2 < 0, s_3 > 0)dj$$

- *A trader tends to revise his belief less often when signals have a sign that is different from the sign of his last belief revision. For example, after a positive signal, the probability that a trader revises upward his belief is larger if his last belief revision was positive than if it was negative:*

$$\int_0^1 \mathbb{P}(\mu_2^j - \mu_1^j > 0 | \mu_1^j - \mu_0^j > 0, s_2 > 0)dj > \int_0^1 \mathbb{P}(\mu_2^j - \mu_1^j > 0 | \mu_1^j - \mu_0^j < 0, s_2 > 0)dj$$

$$\int_0^1 \mathbb{P}(\mu_3^j - \mu_2^j > 0 | \mu_2^j - \mu_1^j > 0, s_3 > 0)dj > \int_0^1 \mathbb{P}(\mu_3^j - \mu_2^j > 0 | \mu_2^j - \mu_1^j < 0, s_3 > 0)dj$$

Beliefs distribution is defined for all $t \in \{0, 1, 2, 3\}$ as $\nu_t = \sum_{s \leq k \leq t} p_{s,k} D_{\theta(s,k)}$ where $D_{\theta(s,k)}$ is the Dirac measure centered on $\theta(s, k)$ and $p_{s,k}$ is the proportion of traders that have the belief $\theta(s, k)$ at time t. Note that the belief distribution depends on the public signals paths. For instance, consider date 2 and the path $\{s_1 = 1, s_2 = -1\}$. Along this path, the belief distribution charges the set $\{\theta(1, 2), \theta(2, 2)\}$ with $p_{1,2} = 1 - \lambda q$ and $p_{2,2} = \lambda q$.

Beliefs dispersion, computed as Var $[\nu_t]$, thus represents the variance of the beliefs taken across individual traders.[21] It is explicitly characterized in the Appendix. When all traders are rational, there is no dispersion in beliefs. The intuition for the first prediction in Proposition 3.4 is as follows. When signals with conflicting information accumulate, confirmatory-biased traders have plenty of occasions to misperceive public signals that contradict their

---

[21]Note that in Proposition 3.3, the variance of prices is computed as the dispersion of the average belief across histories while, in Proposition 3.4, belief dispersion refers to the variance of beliefs along a given history.

prior beliefs. This explains why differences of opinion are more pronounced in our model when there are changes in the sign of the public signals.

The two other predictions are related to the dynamics of beliefs. The second prediction builds on the fact that the influence of a signal on biased traders' beliefs depends on prior signals. For example, the impact of a positive signal at date 2 depends on the signal received at date 1. Indeed, if the first signal was positive, the positive signal at date 2 is perfectly perceived by biased traders. By contrast, if the first signal was negative, the positive signal at date 2 is discarded by a fraction of biased traders. As a result, if one measures the proportion of agents that revise upward their beliefs after receiving date 2 information, it is equal to one in the former case (when signals have the same sign), and it is less than one in the latter case (when signal have different signs).[22]

The third prediction focuses even more closely on individual traders' beliefs updating process. It indicates that it is more likely that a trader updates his beliefs upward after a good signal if he previously updated his beliefs upward rather than downward. The logic is that a trader who has previously updated downward his beliefs is more likely to be pessimistic and thus to ignore subsequent positive signals. A similar logic applies for bad signals.

These theoretical results are tested in the next section using data on US firms' earnings announcements and analysts' earnings forecasts.

4.   Empirical analysis

This section provides an empirical test of our three novel theoretical predictions: when some traders suffer from the confirmatory bias, i) traders are less likely to update their beliefs in the same direction as the latest signal, for signals that are of a different sign compared to the previous ones; ii) traders are less likely to update their beliefs in the same direction as the

---

[22]When they are all rational, the proportion of traders who revise their beliefs in the direction of the last signal is always equal to one.

latest signal, for signals that are of a different sign compared to their previous beliefs change; and iii) differences of opinion among traders are larger when past signals have different signs.

To construct a test of our theory, one must construct an empirical counterpart for the public signals in our model. For this, we build on earlier work (see for instance DellaVigna and Pollet (2009), and Hirshleifer, Lim and Teoh (2009)) and use quarterly earnings surprises as a proxy for the arrival of public news. Then, we follow the literature (e.g., Diether, Malloy, and Scherbina (2002)) and use the dispersion in analysts' earnings forecasts as a proxy for differences of opinion. Finally, we capture beliefs updating with analysts' revisions of annual earnings forecasts across quarterly earnings announcements.

We now present the sample formation, our empirical tests and the results.

4.1 Data and variables

The sample consists of all analyst-stock-year-quarter observations for which we have information on quarterly earnings announcements and analysts' earnings forecasts. We use quarterly earnings announcement data (and other firm-level accounting information) from the CRSP-Compustat merged database and data on analysts' annual earnings forecasts from the Institutional Brokers Estimates System (I/B/E/S) database (which are available from 1982 onwards).[23] When quarterly announcement dates are available in both the CRSP-Compustat merged database and in I/B/E/S and are different, we follow DellaVigna and Pollet (2009) and Hirshleifer, Lim and Teoh (2009) and take the earlier date. Our full sample covers the period 1982-2014. Earnings and forecasts are all split-adjusted. From CRSP, we obtain data on stock returns and closing stock prices. Following earlier work, we focus on companies' ordinary stocks - that is, companies with CRSP share codes of 10 or 11 - traded on NYSE, AMEX, and NASDAQ, and we exclude observations for which the stock

---

[23]As noted by Diether, Malloy and Scherbina (2002), there is a rounding error problem in the standard I/B/E/S "Detail History" data set. We thus use data on analysts' forecasts unadjusted for stock splits; we then scale analysts' forecasts by the CRSP cumulative adjustment factor.

price is less than \$5 dollars. Finally, OptionMetrics provides us with data on stocks' implied volatility (which are available only from 1996 onward).

Our measure of earnings surprises is the commonly used standardized unexpected earnings (SUE) variable (see for instance Bernard and Thomas (1989, 1990), Chan et al. (1996) and Sadka (2006)). SUE for stock i in quarter Q is defined as $(E_{i,Q} - E_{i,Q-4} - c_{i,Q})/\sigma_{i,Q}$, where $E_{i,Q}$ is the quarterly earnings per share in year-quarter Q, $E_{i,Q-4}$ the quarterly earnings four quarters ago, $c_{i,Q}$ and $\sigma_{i,Q}$ the average and standard deviation, respectively, of $E_{i,Q} - E_{i,Q-4}$ over the previous eight quarters.[24]

We construct our measures of forecasts revisions and forecasts dispersion using data from I/B/E/S. SAME-DIRECTION REV(Q) is a dummy that equals one if the analyst's annual earnings forecast revision, if any, made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as SUE(Q-1), the unexpected earnings in quarter Q-1.[25] In robustness tests, we check that our results hold when focusing separately on forecasts revisions following either positive and negative earnings surprises. For this, we construct the variable POS REV(Q) (respectively NEG REV(Q)), a dummy that equals one if the analyst's annual earnings forecast revision made between the announcement dates of the Q-1 and Q quarterly earnings is positive (respectively negative).

Forecasts dispersion is measured as the standard deviation of analysts' forecasts scaled by the prior year-end stock price to mitigate heteroskedasticity. DISP(Q) is the standard deviation multiplied by 100 of the most recent (annual) earnings forecast (for the current fiscal year) of each analyst covering the stock between the announcement date of the quarterly earnings in quarter Q-1 and the announcement date of the quarterly earnings in quarter Q

---

[24]We use earnings per share before extraordinary items from the CRSP-COMPUSTAT merged database (variable EPSPXQ), which are then split-adjusted.

[25]Individual revisions are computed as the difference between the last annual earnings forecast made between the announcement dates of the Q-1 and Q quarterly earnings and the last forecast, if any, made before the announcement date of the Q-1 quarterly earnings. Our results are identical when we construct SAME-DIRECTION REV(Q) based on the difference between the last annual earnings forecast made before the end date of two consecutive fiscal quarters (instead of before the announcement dates of two consecutive quarterly earnings).

normalized by the stock price at the end of the previous fiscal year.[26]

For every stock, we also construct COVERAGE, the number of analysts who covered the stock in the current fiscal year; LN(MKTCAP), the logarithm of the stock's total market capitalization (Compustat item CSHO × item PRCC_F) computed at the end of the current fiscal year; LNBM, the logarithm of the book-to-market ratio defined as in Fama and French (2008) in the current year; and ROA, firm return on assets, defined as operating income after depreciation (Compustat item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. Following Leahy and Whited (1996) and Bloom, Bond, and Van Reenen (2007), we use the standard deviation of daily stock returns for firm i in the current year, SIGMA, as our baseline measure of uncertainty.[27] We also use implied volatility from options prices in robustness checks, which provides us with a forward-looking measure of uncertainty (see for instance Bloom (2009), Kellogg (2014) and Stein and Stone (2013)).[28] As in Bloom (2009) and Stein and Stone (2013), we focus on the implied volatility of 30-day, at-the-money options. These are options that expire in 30 days for which the strike price is equal to the stock's forward price. Specifically, IMPLIEDVOL is defined as the average implied volatility of all 30-day, at-the-money stock i options issued in the current year.[29] Finally, to mitigate the effect of outliers, we windsorize all continuous variables at the first and ninety-ninth percentiles of their respective empirical distribution.

---

[26]The results are identical if DISP(Q) is measured using the latest earnings forecast of each analyst covering the stock between the end date of two consecutive fiscal quarters (instead of between the announcement dates of two consecutive quarterly earnings).

[27]Bloom, Bond, and Van Reenen (2007) show that firm-level stock returns volatility is significantly correlated with a range of alternative uncertainty proxies. Our results are identical when using monthly returns to construct stock-price volatility. As noted in Bloom, Bond, and Van Reenen (2007), using stock returns at a sufficiently high frequency has the advantage of reducing the sample variance. However, one might worry that daily observations are more subject to noise unrelated to fundamentals than monthly observations.

[28]The formula for the price of an option is based on Black and Scholes (1973)'s model. Given the price of an option, the price of the underlying, its time to maturity and strike price, and the riskless rate of interest, the formula can be inverted to yield the volatility implied by the option.

[29]The results are identical if we use implied volatility from at-the-money options 60 days or 90 days to expiration, or if we use implied volatility from either put or call options, separately.

[Table 1 here]

Table 1 presents summary statistics on the variables of interest. There are 996,129 analyst-stock-year-quarter observations (and 140,950 unique stock-year-quarters) in our sample. The mean and median analyst coverage of these stocks is 13 and 11 analysts. Forecast dispersion, DISP(Q), is right-skewed with mean (0.605) noticeably larger than the median (0.255). Finally, the average value of NEG REV(Q) is slightly above 0.5 which indicates that analysts are slightly more likely to revise their forecasts on annual earnings downward over the fiscal year.

4.2   Testing the predictions

Forecasts Revisions. We first study the effect of changes in the sign of earnings surprises on analysts' forecasts revisions. For this, we construct SUE SIGN CHANGE(Q-1,Q-2) (respectively, SUE SIGN CHANGE(Q-1,Q-2,Q-3) for robustness), a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters (respectively, three quarters).[30] The null hypothesis is that, if all agents are rational, they should all revise their forecasts in the direction of the latest signal, irrespective of the sign of any of the previous signals – i.e., the coefficient on the variable SUE SIGN CHANGE should be null. By contrast, as shown in the model, when some agents suffer from the confirmatory bias, changes in the sign of signals open opportunities for biased traders to ignore information contradicting their prior beliefs. As a result, we expect that some of them will not update their forecasts at the announcements of earnings news that are of a different sign compared to the previous ones.

---

[30]Our empirical analysis uses a specification at the quarterly frequency, but all our results hold if we run regressions on a yearly basis. In this case, the first three quarterly earnings announcements in a given fiscal year represent the signals s_1, s_2, and s_3.

To examine whether this prediction holds in the data, we use a linear probability model and estimate the following equation at the analyst-stock-year-quarter level:[31]

$$\text{SAME-DIRECTION REV}_{i,j,Q} = \alpha + \beta \text{ SUE SIGN CHANGE}_{i,(Q-1,Q-2)} + \gamma \text{ CONTROLS}_{i,j,Q} + \delta_Q + \eta_j + \varepsilon_{i,j,Q}$$

(4.4)

SAME-DIRECTION REV$_{i,j,Q}$ is a dummy that equals one if analyst j annual earnings forecast revision for stock i made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as the SUE announced at the end of quarter Q-1; δ_Q are year-quarter fixed effects; η_j are analyst fixed effects; and $\varepsilon_{i,j,Q}$ is the error term. All regressions also control for fiscal-quarter fixed effects. In some specifications, we also include industry and stock fixed effects.[32] Equation (4.4) also includes as control variables analyst coverage (COVERAGE), firm size (LN(MKTCAP)), book-to-market (LNBM), return on asset (ROA), and the volatility of stock returns (SIGMA). SIGMA is meant to control for firm uncertainty, which has been shown to be a key driver of analyst forecast behavior. To further mitigate the concern that SUE SIGN CHANGE might be simply reflecting uncertainty faced by analysts regarding firms fundamentals, we also use implied-volatility from equity options as an alternative (forward-looking) measure of uncertainty.

Finally, since forecast revisions are directly affected by the size of the earnings surprise, we include as an additional control in Equation (4.4) the absolute value of the latest SUE. In all regressions, standard errors are clustered at the stock level to account for serial correlation of the error term within the same stock.[33] The coefficient of interest, β, measures the average effect of past conflicting signals on the likelihood for analysts of revising their forecasts in the

---

[31]Discrete choice models such as logit and probit become unidentified when a large number of fixed effects is included. We therefore use below linear probability models that allow us to control in our strictest specifications for fiscal-quarter, year-quarter, stock, and analyst fixed effects.

[32]We construct the 48 Fama-French industry dummies using the firm's 4 digit SIC industry code and the conversion table in the Appendix of Fama and French (1997).

[33]The statistical significance of the coefficients of interest are stronger in all regressions presented in this paper if we report standard errors clustered at the analyst level.

direction of the latest signal, and should be negative if the confirmatory bias is at work.[34]

Heterogeneous behavior across analysts. To go deeper into the empirical analysis of the confirmatory bias, we also investigate differences across analysts in their individual forecast behavior. For this, we construct a dummy variable at the analyst-stock level which is meant to capture cases in which the announcement of earnings news differ from the analyst's prior beliefs about the prospect of the stock. We capture prior beliefs with past annual earnings forecasts revisions. Specifically, the dummy variable SIGN $\text{REV}_{j,Q-1} \neq$ SIGN $\text{SUE}_{i,Q-1}$ which equals one if the sign of analyst j's earnings revision made between the announcement dates of the earnings in quarter Q-2 and Q-1 is not the same as the sign of the SUE announced at the end of quarter Q-1. We can then study in our empirical tests whether two analysts that have different prior beliefs (as proxied by the sign of their prior earnings forecast revisions on a given stock) display differences in their reaction to future earnings announcements.

To examine whether this prediction holds in the data, we estimate the same Equation (4.4) as above, except that the independent variable of interest is in that case, SIGN $\text{REV}_{j,Q-1} \neq$ SIGN $\text{SUE}_{i,Q-1}$.[35] The null hypothesis is that, if all agents are rational, they should all revise their forecasts in the direction of the latest SUE, irrespective of their prior beliefs about the prospect of the stock– i.e., the coefficient on the variable SIGN $\text{REV}_{j,Q-1} \neq$ SIGN $\text{SUE}_{i,Q-1}$ in the regression discussed below should be null. By contrast, if some agents suffer from the confirmatory bias, we expect that agents for whom the announcement of earn-

---

[34]Earnings news announced in the midst of macroeconomic volatility are more likely to admit different interpretations of the news. We thus also expect our prediction that *traders are less likely to update their beliefs in the same direction as the latest signal for signals that are of a different sign compared to the previous ones* to be stronger for signals announced in the midst of macroeconomic volatility. To test this hypothesis, we augment Equation (1.1) by interacting the variables SUE SIGN CHANGE(Q-1,Q-2) and SUE SIGN CHANGE(Q-1,Q-2,Q-3) with a dummy, MACROVOL(HIGH), that equals one if macroeconomic volatility in quarter $Q-1$ lies in the highest tercile of its empirical distribution. As shown in Tables (A.1) and (A.2) of the Online Appendix, we do find strong empirical support for this additional prediction of our theory. We thank an anonymous referee for suggesting us this test.

[35]We also include as additional control a dummy indicating whether the revision of the analyst in year-quarter Q-1 has been positive.

ings news is inconsistent with their prior beliefs will be less likely to revise their forecasts in the direction of the latest SUE.

Beliefs dispersion. Finally, we look at the implications of the confirmation bias for beliefs dispersion at the stock level. Our theory predicts that a change in the sign of past signals has a positive effect on beliefs dispersion. The null hypothesis is that, if all agents are rational, this effect should be null: the sign of SUEs should not affect forecast dispersion because all agents should have the same interpretation of the signal. To examine whether this prediction holds in the data, we run the following equation at the stock-year-quarter level:

$$\text{DISP}_{i,Q} = \alpha + \beta\ \text{SUE SIGN CHANGE}_{i,(Q-1,Q-2)} + \gamma\ \text{CONTROLS}_{i,Q} + \delta_Q + \varepsilon_{i,Q} \quad (4.5)$$

$\text{DISP}_{i,Q}$ measures beliefs dispersion for stock i in year-quarter Q, δ_Q are year-quarter fixed effects, and $\varepsilon_{i,Q}$ is the error term. All regressions also control for fiscal-quarter fixed effects. In some specifications, we also include industry and stock fixed effects. Equation (4.5) includes the same control variables as in Equation (4.4), namely analyst coverage (COVERAGE), firm size (LN(MKTCAP)), book-to-market (LNBM), return on asset (ROA), the volatility of stock returns (SIGMA), and the absolute value of the latest SUE. The coefficient of interest, β, measures the effect of past conflicting signals on beliefs dispersion, and should be positive if the confirmatory bias is at work.

4.3 Results

Forecast revisions. We first look at how the signs of past signals affect analysts' forecast revision, and report the results in Tables 2 to 5. We first run the regression detailed in Equation (4.4). The variable of interest is SUE SIGN CHANGE(Q-1,Q-2) in columns [1] to [3], and SUE SIGN CHANGE(Q-1,Q-2,Q-3) in columns [4] to [6].

[Table 2 here]

We find support for our prediction that analysts are less likely to update their beliefs in the same direction as the latest signal, for signals of a different sign compared to the previous one(s): the first three columns of Table 2 indicate that, when the sign of the SUE in the last quarter differs from the sign of the SUE two quarters ago, the probability for a given analyst of revising her/his forecast in the direction of the last SUE is lower by around 2 percentage points. This effect is statistically significant at the 1 percent level, and corresponds to a drop of around 4% for the average analyst given that the sample mean of SAME-DIRECTION REV(Q) is 56.7%. This result is robust when we consider a change in the sign of SUEs over the previous three quarters (see the last three columns of Table 2).[36]

[Table 3 here]

We then check that these results are robust to controlling for implied volatility. As shown in Table 3, the estimates remain very similar to those in Table 2, and statistically significant at the 1 percent level across all specifications. Finally, we go one step further and check whether the prediction holds in the data when looking separately at positive and negative SUEs. For this, we split our sample into two subsamples depending on whether the latest SUE is positive or negative. In Table 4, we find that the probability for analysts to revise their forecasts upwards following a positive SUE in the previous quarter is significantly lower when the SUE two quarters ago (respectively two and/or three quarters ago) has a different

---

[36]The results are also robust to the way of computing the analysts' forecast revisions. In the paper, we compute the revisions only for analysts that issue a forecast both before and after the quarterly announcement of interest. However, we find very similar results when we compute revisions for all analysts that have covered the stock before the earnings announcement.

sign. The same results hold for negative revisions following negative SUEs, as shown in Table 5.

[Table 4 here]

[Table 5 here]

Heterogeneous behavior across analysts. We then investigate differences across analysts in their forecast revision behavior depending on whether earnings announcements conflict with their prior beliefs, and report the results in Tables 6 to 8. We run the same regression as in Equation (4.4), except that the variable of interest is now SIGN REV(Q-1) $\neq$ SIGN SUE(Q-1).

[Table 6 here]

We find support for our prediction that analysts are less likely to update their earnings forecasts in the same direction as the latest SUE, for SUEs of a different sign compared to their prior beliefs: as shown in Table 6, the probability of revising a forecast in the direction of the last SUE is lower by around 20 percentage points for an analyst with prior beliefs, as proxied by the sign of her/his forecast revision in the previous quarter, that conflict with the earnings surprise that have been announced. [37] This effect is statistically significant at the 1 percent level. The size of the effect is also significant and corresponds to a decrease in forecast revisions in the direction of the current SUE of around 35% for the average analyst.

---

[37]The results are virtually identical if we proxy for analysts' prior beliefs (either bearish or bullish) by the sign of analysts' last annual earnings forecast revision made in one of the two quarters before the announcement of the Q-1 earnings.

As shown in columns [4] to [6], this result is robust to controlling for implied volatility.[38] Again, we go one step further and check whether the prediction holds in the data when looking separately at positive and negative SUEs. As predicted by the model, we find in Table 7 that the probability of upwards revision following positive SUE is significantly lower for analysts that have revised their forecasts downwards in the previous quarter. Finally, the same results hold for negative revisions following negative SUEs, as shown in Table 8.

[Table 7 here]

[Table 8 here]

Belief dispersion. The regression results on forecast dispersion are reported in Tables 9 and 10.

[Table 9 here]

Table 9 presents estimates of different specifications of equation (4.5). Columns [2] and [5] include industry fixed effects whereas columns [3] and [6] include stock fixed effects. The coefficient on SUE SIGN CHANGE is always positive and statistically significant (at the one percent confidence level), indicating that changes in the sign of SUEs are associated with higher dispersion in analysts' forecasts, as predicted when some agents suffer from the confirmatory bias. The economic effect is also significant. For instance, in column [1], the coefficient on SUE SIGN CHANGE equals 0.048. Given that the sample mean of DISP(Q) equals 0.605, in this specification, a change in the sign of past SUEs is associated with an increase in belief dispersion for the average firm of around 8%.

[Table 10 here]

---

[38]As for the tests presented above, the results are also robust to the way of computing the analysts' forecast revisions.

The correlation between SUE SIGN CHANGE and SIGMA is negative and low (the correlation equals -0.008), which mitigates the concern that the coefficient on SUE SIGN CHANGE might simply reflect part of the effect of firm uncertainty on belief dispersion. Still, to mitigate further this concern, we use implied volatility from stock options as an alternative measure. Table 10 shows the results. Even though the size of the sample is reduced, the estimates remain remarkably stable, and statistically significant at the 1 percent level, across all specifications.

Overall, our empirical investigation provides new stylized facts on the link between the sign of public signals, on one hand, and analysts' forecast dispersion and revisions, on the other hand. We also find heterogeneous effects across analysts in their forecast behavior in the direction predicted by the model. These stylized facts are consistent with the confirmatory bias affecting financial analysts' perception of information.

5. Conclusion

This paper proposes and tests a theory of financial markets based on the premise that some traders are prone to the confirmatory bias. We model this cognitive bias by considering that biased traders tend to misperceive public signals that are inconsistent with their prior views. We show that, in the context of financial markets with public information only, this bias provides a rationale for various stylized facts including excess volume, excess volatility, and momentum.

The model also provides novel empirical predictions regarding both individual analysts' forecast behavior and differences of opinion. At the individual level, the confirmatory bias implies that first impression matters: i) after initial positive (respectively, negative) signals, some traders tend to ignore subsequent signals that are negative (respectively, positive), ii) traders that have previously updated their beliefs upward (respectively, downward) are more

likely to ignore subsequent signals that are negative (respectively, positive). At the stock level, as conflicting signals give ample room for biased traders to misinterpret information, the model predicts in presence of confirmatory-biased traders that differences of opinion should be larger when past signals exhibit changes in sign.

We test these novel predictions on U.S. stock market data over the period ranging from 1982 to 2014. For each stock and quarter, we measure dispersion in analysts' earnings forecasts (as a proxy for disagreement), and analysts' propensity to revise their forecasts (as a proxy for beliefs revisions) according to the sign of the latest earnings surprise. We then perform empirical tests of forecast revisions at the individual level, and find (after controlling for a variety of factors that affect analysts' forecasts), that analysts are less likely to update their beliefs in the same direction as the latest earnings surprise, i) following changes in the sign of earnings surprises over the previous quarters, and ii) for earnings surprises that are of a different sign compared to their previous beliefs. Moreover, we find at the stock level that differences of opinion across analysts are larger when past earnings surprises exhibit changes in sign. These results suggest that the confirmatory bias is at work in financial markets.

In future research, it would be interesting to estimate the parameters of our model and, in particular, the proportion of biased traders on financial markets and the magnitude of the confirmatory bias. This could be useful to evaluate the performance of a strategy designed to profit from the mistakes of confirmatory-biased traders.

Finally, our model could be used to study optimal corporate communication designed to create reversals in the signs of news. Indeed, firms that are confronted to financial markets populated by investors prone to the confirmatory bias might have an interest in appropriately choosing the timing of information releases. For example, firms in possession of both good and bad news might have an interest in first releasing the good news, hoping that the subsequent release of bad news might be ignored by some biased investors.

References

[1] Abreu D. and M. Brunnermeier, 2002, Synchronization risk and delayed arbitrage, Journal of Financial Economics, 66, pp. 341-360.

[2] Asness C., Moskowitz T., and L.H. Pedersen, 2013, Value and momentum everywhere, Journal of Finance, 68, 929-985.

[3] Banerjee S., and I. Kremer, 2010, Disagreement and learning: Dynamic patterns of trade, Journal of Finance, 65, 1269-1302.

[4] Barberis N. and R. Thaler, 2001, A survey of behavioral finance, Handbook of the Economics of Finance, G.M. Constantinides, M. Harris and R. Stulz, eds, Elsevier Science.

[5] Barberis N., A. Shleifer, and R. Vishny, 1998, A Model of investor sentiment, Journal of Financial Economics, 49, pp. 307-343.

[6] Benos, A.V., 1998, Aggressiveness and survival of overconfident traders, Journal of Financial Markets, 3-4, pp. 353-383.

[7] Bernard, V. L. and J. K. Thomas, 1989, Post-Earnings announcement drift and: Delayed price response or risk premium? Journal of Accounting Research, 27, pp.1-36.

[8] Bernard, V. L. and J. K. Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, Journal of Accounting and Economics, 13, pp. 305-340.

[9] Biais B., L. Glosten and C. Spatt, 2005, Market microstructure: a survey of microfondation, empirical results, and policy implications, Journal of Financial Markets, 8, pp. 217-264.

[10] Black, F., and M. S. Scholes, 1973, The Pricing of Options and Corporate Liabilities. Journal of Political Economy 81, pp. 637-654.

[11] Bloom, N., 2009, The Impact of Uncertainty Shocks, Econometrica, 77, pp. 623-685.

[12] Bloom, N., Bond, S., and J. Van Reenen, 2007, Uncertainty and Investment Dynamics, Review of Economic Studies, 74, pp. 391-415.

[13] Bodenhausen G., 1988, Stereotypic biases in social decision making and memory: testing process models of stereotype use, Journal of Personality and Social Psychology, 55, pp. 726-737.

[14] Bouchaud J.-P., P. Krueger, A. Landier, and D. Thesmar, 2016, Sticky Expectations and Stock Market Anomalies, HEC Paris Research Paper No. FIN-2016-1136.

[15] Chan, L. K. C., Jegadeesh N., and J. Lakonishok, 1996, Momentum Strategies, Journal of Finance 51, pp. 1681-1713.

[16] Chen, J. , Hong H., and J. Stein, 2002, Breadth of Ownership and Stock Returns, Journal of Financial Economics 66, pp. 171-205.

[17] Daniel K.,D. Hirshleifer, and A. Subrahmanyam, 1998, Investor psychology and securities market under- and overreactions, Journal of Finance, 53, pp. 1839-1885.

[18] Darley J. and P. Gross, 1983, A hypothesis-confirming bias in labeling effects, Journal of Personality and Social Psychology, 44, pp. 20-33.

[19] Davis, Angela K., Jeremy Piger, and Lisa M. Sedor, 2012, Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases, Contemporary Accounting Research, forthcoming.

[20] DeBondt W. and R.H. Thaler, 1995, Financial decision making in markets and firms, In Finance. Edited by R. Jarrow, V Maksimovic, and W. T. Ziemba. Elsevier-North Holland: Series of Handbooks in Operations Research and Management Science.

[21] DellaVigna, S., and J.M. Pollet, 2009, Investor Inattention and Friday Earnings Announcements, Journal of Finance, 64, pp. 709-749.

[22] De Long B., A. Shleifer, L. Summers, and R. Waldmann, 1990, Noise trader risk in financial markets, Journal of Political Economy, 98, pp. 703-738.

[23] Diether, K., C. Malloy, and A. Scherbina, 2002, Differences of opinion and the cross section of stock returns, Journal of Finance, 57, pp. 2113-2141.

[24] Duchin, R., and M. Levy, 2010, Disagreement, portfolio optimization, and excess volatility, Journal of Financial and Quantitative Analysis, 45, pp. 623-640.

[25] Engelberg, Joseph, 2008, Costly information processing: Evidence from earnings announcements, Working paper.

[26] Evans, J. St. B. T., 1989, Bias in human reasoning: causes and consequences. Hillsdale, NJ: Erlbaum.

[27] Fama, E., and K. French, 1997, Industry costs of equity, Journal of Financial Economics, 43, pp. 153-193.

[28] Fama E. and K. French, 2008, Dissecting anomalies, Journal of Finance, 63, pp. 1653-1678.

[29] Forsythe R., F. Nelson, G. Neumann, and J. Wright, 1992, Anatomy of an experimental political stock market, American Economic Review, 82, pp. 1142-1161.

[30] Gervais S. and T. Odean, 2001, Learning to be overconfident, Review of Financial Studies, 14, pp. 1-27.

[31] Goetzmann W. and M. Massa, 2005, Dispersion of opinion and stock returns, Journal of Financial Markets, 8, pp. 324-349.

[32] Harris M. and A. Raviv, 1993, Differences of opinion make a horse race, Review of Financial Studies, 6, pp. 473-506.

[33] Hart W., D. Albarracin, A. Eagly, I. Brechan, M. Lindberg, and L. Merrill, 2009, Feeling validated versus being correct: a meta-analysis of selective exposure to information, Psychological Bulletin, 135, pp. 555-588.

[34] Hirshleifer D., 2001, Investor Psychology and Asset Pricing, Journal of Finance, 56, pp. 1533-1597.

[35] Hirshleifer, D., Lim, S. S., and S.H. Teoh, 2009, Driven to Distraction: Extraneous Events and Underreaction to Earnings News, Journal of Finance, 64, pp. 2289-2325.

[36] Hong H. and J. Stein, 1999, A unified theory of underreaction, Momentum Trading and Overreaction in Asset Markets, Journal of Finance, 54, pp. 2143-2184.

[37] Jegadeesh N. and S. Titman, 1993, Returns to buying winners and selling losers: implications for stock market efficiency, Journal of Finance, 48, pp. 65-91.

[38] Kellogg, R., 2014, The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling, American Economic Review, 104, pp. 1698-1734.

[39] Lee, C.M.C and B. Swaminathan, 2000, Price momentum and trading volume, Journal of Finance, 55, pp. 2017-2069.

[40] Leroy S. and R. Porter, 1981, The present-value relation: tests based on implied variance bounds, Econometrica, 49, pp. 555-574.

[41] Lord C., L. Ross, and M. Lepper, 1979, Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence, Journal of Personality and Social Psychology, 37, pp. 2098-2109.

[42] Madhavan A., 2000, Market microstructure: a survey, Journal of Financial Markets, 3, pp. 205-258.

[43] Miller E., 1977, Risk, uncertainty, and divergence of opinion, Journal of Finance 32, pp. 1151-1168.

[44] Dawn Matsumoto, Maarten Pronk, and Erik Roelofsen (2011) What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. The Accounting Review: July 2011, Vol. 86, No. 4, pp. 1383-1414

[45] Nickerson, R.S., 1998, Confirmation bias: a ubiquitous phenomenon in many guises, Review of General Psychology, 2, pp. 175-220.

[46] Odean T., 1998, Volume, volatility, price, and profits when all traders are above average, Journal of Finance, 53, pp. 1887-1934.

[47] Mitchell A. Petersen, 2004, Information: Hard and Soft, Working Paper.

[48] Rabin M. and J. Schrag, 1999, First impressions matter: a model of confirmatory bias, Quarterly Journal of Economics, 114, pp. 37-82.

[49] Sadka, R., 2006, Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk, Journal of Financial Economics, 80, pp. 309-349.

[50] Scheinkman J. and W. Xiong, 2003, Overconfidence and speculative bubbles, Journal of Political Economy, 111, pp. 1183-1219.

[51] Stein, L.C.D, and E. Stone., 2013, The Effect of Uncertainty on Investment: Evidence from Equity Options, Working Paper.

[52] Rabin, M. and D. Vayanos, 2010, The gambler's and hot-hand fallacies: theory and applications, Review of Economic Studies, 77, pp. 730-778.

[53] Rouwenhorst, K.G., 1998, International momentum strategies, Journal of Finance, 53, pp. 267-284.

[54] Shiller R., 1981, The use of volatility measures in assessing market efficiency, Journal of Finance, 36, pp. 291-304.

[55] Shiller R., 2000, Irrational exuberance, Princeton University press, Princeton, New Jersey.

[56] Shleifer A. and R. Vishny, 1997, The limits of arbitrage, Journal of Finance, 52, pp. 35-55.

[57] Tetlock, P., M. Saar-Tsechansky, and S. Macskassy. 2008. More than words: Quantifying language to measure firms fundamentals. The Journal of Finance 63 (3): 1437-67.

[58] Uhlig H., 1996, A law of large numbers for large economies, Economic Theory, 8, pp. 41-50.

[59] Verardo M., 2009, Heterogeneous beliefs and momentum profits, Journal of Financial and Quantitative Analysis, 44, pp. 795-822.

[60] Xiong W. and J. Yu, 2011, The Chinese warrants bubble, American Economic Review, 101, pp. 2723-2753.

6. Tables

Table 1. SUMMARY STATISTICS

We consider a sample of stocks that have analysts' coverage in I/B/E/S over the period from January 1982 to December 2014 and non-missing information on quarterly earnings in the CRSP-Compustat merged database. SUE(Q) (respectively ABS SUE(Q)) are (respectively the absolute value of) standardized unexpected earnings in year-quarter Q; specifically SUE(Q) is defined as $(E_{i,Q} - E_{i,Q-4} - c_{i,Q})/\sigma_{i,Q}$, where $E_{i,Q}$ is the quarterly earnings per share in year-quarter Q, $E_{i,Q-4}$ the quarterly earnings four quarters ago, $c_{i,Q}$ and $\sigma_{i,Q}$ the average and standard deviation, respectively, of $E_{i,Q} - E_{i,Q-4}$ over the previous eight quarters. POS REV(Q) (respectively NEG REV(Q)) for a given analyst and a given stock is a dummy that equals one if the analyst covered the stock before the announcement date of the Q-1 quarterly earnings and has revised his/her annual earnings forecast upward (respectively downward) between the announcement dates of the Q-1 and Q quarterly earnings. SAME-DIRECTION REV(Q) is a dummy that equals one if the analyst's annual earnings forecast revision, if any, made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as SUE(Q-1), the unexpected earnings in quarter Q-1. Individual revisions are computed as the difference between the last annual earnings forecast made between the announcement dates of the Q-1 and Q quarterly earnings and the last forecast, if any, made before the announcement date of the Q-1 quarterly earnings. DISP(Q) is the standard deviation multiplied by 100 of the most recent (annual) earnings forecast (for the current fiscal year) of each analyst covering the stock between the announcement date of the quarterly earnings in quarter Q-1 and the announcement date of the quarterly earnings in quarter Q normalized by the stock price at the end of the previous fiscal year. POS SUE(Q-1) (respectively POS SUE(Q-1,Q-2)) is a dummy that equals one if SUE in quarter Q-1 (respectively in quarter Q-1 and/or quarter $Q-2$) is positive. SUE SIGN CHANGE(Q-1,Q-2) (respectively SUE SIGN CHANGE(Q-1,Q-2,Q-3)) is a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters (respectively over the previous three quarters). SIGN REV(Q-1) $\neq$ SIGN SUE(Q-1) is a dummy that equals one if the analyst's annual earnings forecast revision for a given stock, if any, made between the announcement dates of the Q-2 and Q-1 quarterly earnings has a different sign than SUE(Q-1). COVERAGE is the number of analysts who covered the stock in fiscal year t. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO $\times$ item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5.

	Obs	Mean	SD	10th	50th	90th
Analysts' Forecast Revision (at the analyst-stock-year-quarter level)						
POS REV(Q)	996129	0.454	0.498	0.000	0.000	1.000
NEG REV(Q)	996129	0.521	0.500	0.000	1.000	1.000
SAME-DIRECTION REV(Q)	996129	0.567	0.495	0.000	1.000	1.000
Analysts' Forecast Dispersion (at the stock-year-quarter level)						
DISP(Q)	140950	0.605	1.119	0.047	0.255	1.383
Quarterly Earnings Surprises (at the stock-year-quarter level)						
SUE(Q-1)	140950	-0.104	1.176	-1.402	-0.037	1.102
ABS SUE(Q-1)	140950	0.833	0.837	0.115	0.576	1.871
POS SUE(Q-2)	140950	0.488	0.500	0.000	0.000	1.000
POS SUE(Q-2,Q-3)	140950	0.632	0.482	0.000	1.000	1.000
SUE SIGN CHANGE(Q-1,Q-2,Q-3)	140950	0.439	0.496	0.000	0.000	1.000
SUE SIGN CHANGE(Q-1,Q-2)	140950	0.283	0.451	0.000	0.000	1.000
SIGN REV(Q-1) $\neq$ SIGN SUE(Q-1)	602209	0.406	0.491	0.000	0.000	1.000
Firm Characteristics (at the stock-year-quarter level)						
COVERAGE	140950	13.250	9.427	4.000	11.000	27.000
LN(MKTCAP)	140950	6.888	1.677	4.802	6.796	9.115
SIGMA	140950	2.654	1.315	1.301	2.355	4.384
IMPLIEDVOL	62830	0.453	0.183	0.246	0.419	0.704
LNBM	140950	-0.731	0.745	-1.713	-0.655	0.130
ROA	140950	0.031	0.032	0.004	0.031	0.065

Table 2. FORECAST REVISION

The dependent variable is a dummy that equals one if the analyst's annual earnings forecast revision for a given stock, if any, made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as SUE(Q-1), the unexpected earnings in quarter Q-1. Individual revisions are computed as the difference between the last annual earnings forecast made between the announcement dates of the Q-1 and Q quarterly earnings and the last annual earnings forecast, if any, made before the announcement date of the Q-1 quarterly earnings. SUE SIGN CHANGE(Q-1,Q-2) is a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters. SUE SIGN CHANGE(Q-1,Q-2,Q-3) is a dummy that equals one if there is one change in the sign of SUEs among the previous three quarters. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014.

| | REVISION IN THE SAME DIRECTION AS THE LATEST SUE | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
SUE SIGN CHANGE(Q-1,Q-2)	-0.021***	-0.021***	-0.018***			
	(0.003)	(0.003)	(0.003)			
SUE SIGN CHANGE(Q-1,Q-2,Q-3)				-0.018***	-0.018***	-0.017***
				(0.003)	(0.003)	(0.003)
ABS SUE(Q-1)	0.057***	0.057***	0.055***	0.057***	0.057***	0.054***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LN(MKTCAP)	-0.009***	-0.009***	0.005	-0.009***	-0.009***	0.005
	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.004)
COVERAGE	0.000	0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SIGMA	0.013***	0.014***	0.015***	0.013***	0.014***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LNBM	0.000	-0.000	-0.001	0.000	-0.000	-0.001
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
ROA	0.200***	0.207***	0.158*	0.200***	0.207***	0.159*
	(0.067)	(0.068)	(0.090)	(0.067)	(0.068)	(0.090)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	996129	996129	996129	996129	996129	996129
R^2	0.038	0.039	0.076	0.038	0.039	0.076

Table 3. FORECAST REVISION - ROBUSTNESS - CONTROLLING FOR IMPLIED VOLATILITY

This Table presents variants of the regressions in Table 2 in which we control for implied volatility. In columns [1] to [6], the dependent variable is a dummy that equals one if the analyst's annual earnings forecast revision for a given stock, if any, made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as SUE(Q-1), the unexpected earnings in quarter Q-1. Individual revisions are computed as the difference between the last annual earnings forecast made between the announcement dates of the Q-1 and Q quarterly earnings and the last annual earnings forecast, if any, made before the announcement date of the Q-1 quarterly earnings. SUE SIGN CHANGE(Q-1,Q-2) is a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters. SUE SIGN CHANGE(Q-1,Q-2,Q-3) is a dummy that equals one if there is one change in the sign of SUEs among the previous three quarters. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than \$5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1996 to December 2014.

	REVISION IN THE SAME DIRECTION AS THE LATEST SUE					
	[1]	[2]	[3]	[4]	[5]	[6]
SUE SIGN CHANGE(Q-1,Q-2)	-0.023***	-0.023***	-0.016***			
	(0.004)	(0.004)	(0.004)			
SUE SIGN CHANGE(Q-1,Q-2,Q-3)				-0.022***	-0.022***	-0.016***
				(0.004)	(0.004)	(0.004)
ABS SUE(Q-1)	0.048***	0.048***	0.048***	0.048***	0.048***	0.047***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LN(MKTCAP)	-0.007***	-0.008***	-0.015**	-0.007***	-0.008***	-0.015**
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
COVERAGE	0.000	0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
IMPLIEDVOL	0.120***	0.120***	0.139***	0.120***	0.120***	0.139***
	(0.020)	(0.020)	(0.026)	(0.020)	(0.020)	(0.026)
LNBM	-0.002	-0.003	-0.007	-0.002	-0.003	-0.007
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
ROA	0.243***	0.244***	0.106	0.244***	0.245***	0.108
	(0.091)	(0.090)	(0.120)	(0.091)	(0.090)	(0.120)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	517979	517979	517979	517979	517979	517979
R^2	0.041	0.042	0.083	0.041	0.042	0.083

Table 4. POSITIVE REVISION

This Table presents variants of the regressions in Table 2 in which we focus on analysts' positive revisions that follow positive earnings surprises. In columns [1] to [6], the dependent variable is for a given analyst and a given stock a dummy that equals one if the analyst covered the stock before the announcement date of the Q-1 quarterly earnings and has revised his/her annual earnings forecast upward between the announcement dates of the Q-1 and Q quarterly earnings. The sample is restricted to analyst-stock-year-quarter observations for which SUE(Q-1), the unexpected earnings in quarter Q-1, is positive. NEG SUE (Q-2) is a dummy that equals one if standardized unexpected earnings were negative two quarters ago. NEG SUE (Q-2,Q-3) is a dummy that equals one if standardized unexpected earnings were negative two and/or three quarters ago. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than \$5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014.

| | POSREV(Q) (Sample with SUE(Q-1)> 0 only) | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
NEG SUE(Q-2)	-0.017***	-0.018***	-0.019***			
	(0.004)	(0.004)	(0.004)			
NEG SUE(Q-2,Q-3)				-0.014***	-0.014***	-0.017***
				(0.004)	(0.004)	(0.004)
ABS SUE(Q-1)	0.048***	0.048***	0.047***	0.048***	0.048***	0.047***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LN(MKTCAP)	0.056***	0.057***	0.133***	0.056***	0.057***	0.133***
	(0.003)	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)
COVERAGE	-0.006***	-0.006***	-0.009***	-0.006***	-0.006***	-0.009***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
SIGMA	0.010***	0.010***	0.004	0.010***	0.010***	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LNBM	0.019***	0.019***	0.049***	0.019***	0.019***	0.049***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
ROA	3.190***	3.243***	4.029***	3.189***	3.242***	4.025***
	(0.104)	(0.106)	(0.153)	(0.105)	(0.106)	(0.153)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	476208	476208	476208	476208	476208	476208
R^2	0.137	0.138	0.208	0.136	0.138	0.208

Table 5. NEGATIVE REVISION

This Table presents variants of the regressions in Table 2 in which we focus on analysts' negative revisions that follow negative earnings surprises. In columns [1] to [6], the dependent variable is for a given analyst and a given stock a dummy that equals one if the analyst covered the stock before the announcement date of the Q-1 quarterly earnings and has revised his/her annual earnings forecast downward between the announcement dates of the Q-1 and Q quarterly earnings. The sample is restricted to analyst-stock-year-quarter observations for which SUE(Q-1), the unexpected earnings in quarter Q-1, is negative. POS SUE (Q-2) is a dummy that equals one if standardized unexpected earnings were positive two quarters ago. POS SUE (Q-2,Q-3) is a dummy that equals one if standardized unexpected earnings were positive two and/or three quarters ago. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014.

| | NEGREV(Q) (Sample with SUE(Q-1)< 0 only) | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
POS SUE(Q-2)	-0.025***	-0.025***	-0.026***			
	(0.004)	(0.004)	(0.004)			
POS SUE(Q-2,Q-3)				-0.023***	-0.023***	-0.024***
				(0.004)	(0.004)	(0.004)
ABS SUE(Q-1)	0.050***	0.050***	0.045***	0.050***	0.049***	0.045***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LN(MKTCAP)	-0.055***	-0.056***	-0.133***	-0.055***	-0.056***	-0.133***
	(0.002)	(0.002)	(0.006)	(0.002)	(0.002)	(0.006)
COVERAGE	0.005***	0.005***	0.007***	0.005***	0.005***	0.007***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
SIGMA	0.005**	0.005**	0.008**	0.005**	0.005**	0.008**
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
LNBM	-0.017***	-0.018***	-0.044***	-0.017***	-0.018***	-0.045***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
ROA	-2.369***	-2.383***	-2.961***	-2.365***	-2.379***	-2.955***
	(0.090)	(0.091)	(0.126)	(0.090)	(0.091)	(0.126)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	519921	519921	519921	519921	519921	519921
R^2	0.138	0.140	0.206	0.138	0.140	0.206

Table 6. FORECAST REVISION - HETEROGENEOUS BEHAVIOR ACROSS ANALYSTS

The dependent variable is a dummy that equals one if the analyst's annual earnings forecast revision for a given stock, if any, made between the announcement dates of the Q-1 and Q quarterly earnings has the same sign as SUE(Q-1), the unexpected earnings in quarter Q-1. Individual revisions are computed as the difference between the last annual earnings forecast made between the announcement dates of the Q-1 and Q quarterly earnings and the last annual earnings forecast, if any, made before the announcement date of the Q-1 quarterly earnings. SIGN REV(Q-1) ≠ SIGN SUE(Q-1) is a dummy that equals one if the analyst's annual earnings forecast revision for a given stock, if any, made between the announcement dates of the Q-2 and Q-1 quarterly earnings has a different sign than SUE(Q-1). POSREV(Q-1) is a dummy that equals one if the analyst has revised his/her annual earnings forecast upward between the announcement dates of the Q-2 and Q-1 quarterly earnings. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014 in columns [1] to [3], and from January 1996 to December 2014 in columns [4] to [6].

	REVISION IN THE SAME DIRECTION AS THE LATEST SUE					
	[1]	[2]	[3]	[4]	[5]	[6]
SIGN REV(Q-1) ≠ SIGN SUE(Q-1)	-0.220***	-0.219***	-0.198***	-0.198***	-0.198***	-0.175***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
POSREV(Q-1)	0.023***	0.024***	0.027***	0.027***	0.028***	0.031***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
ABS SUE(Q-1)	0.044***	0.044***	0.042***	0.038***	0.038***	0.038***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LN(MKTCAP)	-0.009***	-0.009***	-0.001	-0.008***	-0.008***	-0.014**
	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.006)
COVERAGE	0.000	0.000	-0.000	0.001	0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
SIGMA	0.010***	0.010***	0.012***			
	(0.002)	(0.002)	(0.002)			
IMPLIEDVOL				0.102***	0.102***	0.125***
				(0.020)	(0.020)	(0.027)
LNBM	0.000	-0.000	0.000	-0.000	-0.000	-0.006
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.006)
ROA	0.115*	0.118*	0.125	0.200**	0.196**	0.140
	(0.066)	(0.066)	(0.091)	(0.093)	(0.093)	(0.126)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	602187	602187	602187	334821	334821	334821
R^2	0.090	0.091	0.125	0.085	0.086	0.125

Table 7. POSITIVE REVISION - HETEROGENEOUS BEHAVIOR ACROSS ANALYSTS

This Table presents variants of the regressions in Table 6 in which we focus on analysts' positive revisions (in year-quarter Q) that follow positive earnings surprises. In columns [1] to [6], the dependent variable is for a given analyst and a given stock a dummy that equals one if the analyst covered the stock before the announcement date of the Q-1 quarterly earnings and has revised his/her annual earnings forecast upward between the announcement dates of the Q-1 and Q quarterly earnings. The sample is restricted to analyst-stock-year-quarter observations for which SUE(Q-1), the unexpected earnings in quarter Q-1, is positive. NEGREV(Q-1) is a dummy that equals one if the analyst has revised his/her annual earnings forecast downward between the announcement dates of the Q-2 and Q-1 quarterly earnings. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014 in columns [1] to [3], and from January 1996 to December 2014 in columns [4] to [6].

| | POSREV(Q) (Sample with SUE(Q-1)> 0 only) | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
NEGREV(Q-1)	-0.152***	-0.151***	-0.095***	-0.131***	-0.129***	-0.060***
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)
ABS SUE(Q-1)	0.040***	0.040***	0.041***	0.035***	0.035***	0.037***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)
LN(MKTCAP)	0.048***	0.048***	0.117***	0.048***	0.048***	0.136***
	(0.003)	(0.003)	(0.006)	(0.004)	(0.004)	(0.009)
COVERAGE	-0.005***	-0.005***	-0.008***	-0.004***	-0.004***	-0.007***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
SIGMA	0.008***	0.007***	0.007*			
	(0.003)	(0.003)	(0.004)			
IMPLIEDVOL				0.043	0.041	0.078*
				(0.029)	(0.029)	(0.044)
LNBM	0.019***	0.019***	0.052***	0.013**	0.012**	0.048***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.009)
ROA	2.898***	2.944***	3.864***	2.789***	2.812***	3.747***
	(0.104)	(0.105)	(0.163)	(0.126)	(0.128)	(0.224)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	282360	282360	282360	146731	146731	146731
R^2	0.177	0.179	0.245	0.174	0.175	0.258

Table 8. NEGATIVE REVISION - HETEROGENEOUS BEHAVIOR ACROSS ANALYSTS

This Table presents variants of the regressions in Table 6 in which we focus on analysts' negative revisions (in year-quarter Q) that follow negative earnings surprises. In columns [1] to [6], the dependent variable is for a given analyst and a given stock a dummy that equals one if the analyst covered the stock before the announcement date of the Q-1 quarterly earnings and has revised his/her annual earnings forecast downward between the announcement dates of the Q-1 and Q quarterly earnings. The sample is restricted to analyst-stock-year-quarter observations for which SUE(Q-1), the unexpected earnings in quarter Q-1, is negative. POSREV(Q-1) is a dummy that equals one if the analyst has revised his/her annual earnings forecast upward between the announcement dates of the Q-2 and Q-1 quarterly earnings. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects, fiscal-quarter fixed effects and analyst fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014 in columns [1] to [3], and from January 1996 to December 2014 in columns [4] to [6].

| | NEGREV(Q) (Sample with SUE(Q-1)< 0 only) | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
POSREV(Q-1)	-0.125***	-0.123***	-0.066***	-0.118***	-0.116***	-0.055***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
ABS SUE(Q-1)	0.042***	0.042***	0.039***	0.041***	0.041***	0.038***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
LN(MKTCAP)	-0.049***	-0.050***	-0.122***	-0.051***	-0.052***	-0.143***
	(0.002)	(0.002)	(0.006)	(0.003)	(0.003)	(0.008)
COVERAGE	0.004***	0.004***	0.007***	0.004***	0.004***	0.006***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
SIGMA	0.006**	0.006**	0.007*			
	(0.002)	(0.003)	(0.004)			
IMPLIEDVOL				0.053*	0.051*	0.040
				(0.027)	(0.027)	(0.040)
LNBM	-0.019***	-0.020***	-0.047***	-0.022***	-0.022***	-0.057***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)	(0.009)
ROA	-2.219***	-2.236***	-2.866***	-2.093***	-2.092***	-2.690***
	(0.096)	(0.096)	(0.140)	(0.131)	(0.132)	(0.196)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	319827	319827	319827	188090	188090	188090
R^2	0.170	0.171	0.234	0.166	0.168	0.237

Table 9. DISPERSION IN ANALYST FORECASTS

The dependent variable, DISP(Q), is the standard deviation multiplied by 100 of the most recent (annual) earnings forecast (for the current fiscal year) of each analyst covering the stock between the announcement date of the quarterly earnings in quarter Q-1 and the announcement date of the quarterly earnings in quarter Q normalized by the stock price at the end of the previous fiscal year. SUE SIGN CHANGE(Q-1,Q-2) is a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters. SUE SIGN CHANGE(Q-1,Q-2,Q-3) is a dummy that equals one if there is one change in the sign of SUEs among the previous three quarters. ABS SUE(Q-1) is the absolute value of standardized unexpected earnings in year-quarter Q-1. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO $\times$ item PRCC_F) at the end of the current fiscal year. SIGMA is the standard deviation of daily raw returns of the stock in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects and fiscal-quarter fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than \$5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1982 to December 2014.

	DISP(Q)					
	[1]	[2]	[3]	[4]	[5]	[6]
SUE SIGN CHANGE(Q-1,Q-2)	0.048***	0.045***	0.039***			
	(0.007)	(0.006)	(0.006)			
SUE SIGN CHANGE(Q-1,Q-2,Q-3)				0.051***	0.047***	0.040***
				(0.007)	(0.007)	(0.006)
ABS SUE(Q-1)	0.086***	0.082***	0.070***	0.087***	0.083***	0.071***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
LN(MKTCAP)	-0.022***	-0.030***	-0.201***	-0.022***	-0.030***	-0.201***
	(0.007)	(0.007)	(0.013)	(0.007)	(0.007)	(0.013)
COVERAGE	0.007***	0.006***	0.004***	0.007***	0.006***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SIGMA	0.274***	0.264***	0.220***	0.274***	0.264***	0.220***
	(0.010)	(0.010)	(0.012)	(0.010)	(0.010)	(0.012)
LNBM	0.304***	0.280***	0.306***	0.303***	0.279***	0.306***
	(0.012)	(0.012)	(0.014)	(0.012)	(0.012)	(0.014)
ROA	-4.483***	-4.560***	-3.038***	-4.476***	-4.556***	-3.036***
	(0.198)	(0.216)	(0.288)	(0.198)	(0.216)	(0.288)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	140950	140950	140950	140950	140950	140950
R^2	0.211	0.229	0.447	0.211	0.230	0.447

Table 10. DISPERSION IN ANALYST FORECASTS - ROBUSTNESS - CONTROLLING FOR IMPLIED VOLATILITY

This Table presents variants of the regressions in Table 9 in which we control for implied volatility. The dependent variable, DISP(Q), is the standard deviation multiplied by 100 of the most recent (annual) earnings forecast (for the current fiscal year) of each analyst covering the stock between the announcement date of the quarterly earnings in quarter Q-1 and the announcement date of the quarterly earnings in quarter Q normalized by the stock price at the end of the previous fiscal year. SUE SIGN CHANGE(Q-1,Q-2) is a dummy that equals one if there is a change in the sign of SUEs over the previous two quarters. SUE SIGN CHANGE(Q-1,Q-2,Q-3) is a dummy that equals one if there is one change in the sign of SUEs among the previous three quarters. COVERAGE is the number of analysts who covered the stock in the current fiscal year. LN(MKTCAP) is the logarithm of the stock total market capitalization (Compustat item CSHO × item PRCC_F) at the end of the current fiscal year. IMPLIEDVOL is the average of 30-day implied volatility of all at-the-money stock i options issued in the current fiscal year. LNBM is book-to-market defined as in Fama and French (2008) in the current year. Return on assets, ROA, is defined as operating income after depreciation (item OIBDP - item DP) over total assets (item AT) computed at the end of the current fiscal year. All regressions include year-quarter fixed effects and fiscal-quarter fixed effects. Columns [2] and [5] also include industry fixed effects. Columns [3] and [6] also include stock fixed effects. All continuous variables are windsorized at the first and ninety-ninth percentiles. We exclude all observations with stock price lower than $5. *, **, and *** denote significance at the 10%, 5%, and 1%, respectively. The sample period is from January 1996 to December 2014.

| | DISP(Q) | | | | | |
	[1]	[2]	[3]	[4]	[5]	[6]
SUE SIGN CHANGE(Q-1,Q-2)	0.039***	0.032***	0.025***			
	(0.009)	(0.009)	(0.007)			
SUE SIGN CHANGE(Q-1,Q-2,Q-3)				0.044***	0.036***	0.028***
				(0.009)	(0.009)	(0.008)
ABS SUE(Q-1)	0.062***	0.057***	0.044***	0.063***	0.059***	0.045***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)
LN(MKTCAP)	0.029***	0.018*	-0.163***	0.029***	0.018*	-0.164***
	(0.011)	(0.011)	(0.020)	(0.011)	(0.011)	(0.020)
COVERAGE	0.002*	0.001	0.001	0.002*	0.001	0.001
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
IMPLIEDVOL	2.030***	1.977***	1.698***	2.030***	1.977***	1.699***
	(0.112)	(0.118)	(0.154)	(0.112)	(0.118)	(0.154)
LNBM	0.253***	0.222***	0.256***	0.253***	0.222***	0.256***
	(0.016)	(0.015)	(0.019)	(0.016)	(0.015)	(0.019)
ROA	-3.322***	-3.284***	-1.971***	-3.319***	-3.283***	-1.971***
	(0.270)	(0.286)	(0.408)	(0.270)	(0.286)	(0.408)
Fiscal-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	No	Yes	No
Stock FE	No	No	Yes	No	No	Yes
Observations	62830	62830	62830	62830	62830	62830
R^2	0.202	0.226	0.484	0.202	0.226	0.484

7.   Appendix

7.1   Belief dynamics

To set the stage for the analysis of the evolution of prices and volumes, we start by studying belief dynamics.

Lemma 7.1. *For speculators, belief dynamics are given by:*

$$\mu_{t+1} = \mathbb{P}(X = H | s_1, \ldots, s_{t+1}) = f(\theta, \mu_t) \mathbb{1}_{\{s_{t+1}=1\}} + f(1 - \theta, \mu_t) \mathbb{1}_{\{s_{t+1}=-1\}} + \mu_t \mathbb{1}_{\{s_{t+1}=0\}},$$

and for biased traders $(i \in [0, \lambda])$,

$$\mu_{t+1}^i = \mathbb{P}^i(X = H | s_1, \ldots, s_{t+1}) = f(\theta, \mu_t^i) \mathbb{1}_{\{\sigma_{t+1}^i=1\}} + f(1 - \theta, \mu_t^i) \mathbb{1}_{\{\sigma_{t+1}^i=-1\}} + \mu_t \mathbb{1}_{\{\sigma_{t+1}^i=0\}},$$

where:

$$f(\theta, \mu) = \frac{\theta \mu}{\theta \mu + (1 - \theta)(1 - \mu)}.$$

Proof. Because the law of the rational beliefs μ_t under $\mathbb{P}$ is the same as the one of biased beliefs μ_t^i under the biased probability $\mathbb{P}^i$, it is enough to focus on belief dynamics for rational agents. Using Bayes' formula, we have:

$$
\begin{aligned}
\mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = 1) &= \frac{\mathbb{P}(X = H; s_{t+1} = 1 | s_1, \ldots, s_t)}{\mathbb{P}(s_{t+1} = 1 | s_1, \ldots, s_t)} \\
&= \frac{\mathbb{P}(s_{t+1} = 1 | X = H) \mathbb{P}(X = H | s_1, \ldots, s_t)}{\mathbb{P}(s_{t+1} = 1 | s_1, \ldots, s_t)} \\
&= \frac{\delta \theta \mu_t}{\delta \theta \mu_t + \delta(1 - \theta)(1 - \mu_t)} \\
&= \frac{\theta \mu_t}{\theta \mu_t + (1 - \theta)(1 - \mu_t)}.
\end{aligned}
$$

Likewise, Bayes' formula implies that:

$$\mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = -1) = \frac{(1-\theta)\mu_t}{(1-\theta)\mu_t + \theta(1-\mu_t)}$$

$$\mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = 0) = \mu_t.$$

We conclude by noting that:

$$\mu_{t+1} = \mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = 1)\mathbb{1}_{\{s_{t+1}=1\}}$$

$$+ \mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = -1)\mathbb{1}_{\{s_{t+1}=-1\}}$$

$$+ \mathbb{P}(X = H | s_1, \ldots, s_t, s_{t+1} = 0)\mathbb{1}_{\{s_{t+1}=0\}}.$$

$\diamond$

Note that for a belief history of length t, $h_t = (s_1, \ldots, s_t)$, the possible values for μ_t are:

$$\theta(s, k) = \frac{\theta^s(1-\theta)^{k-s}}{\theta^s(1-\theta)^{k-s} + \theta^{k-s}(1-\theta)^s}. \tag{7.6}$$

where s is the number of strictly positive signals with $s \leq k \leq t$, and $k - s$ is the number of strictly negative signals. This formula is similar to Rabin and Schrag (1999) for $k = t$: in Rabin and Schrag (1999), there is no uninformative signal.

We continue this subsection by giving some results concerning the probability distribution of the public signal. Clearly:

$$\mathbb{P}(s_1 = 1) = \frac{\delta\theta + \delta(1-\theta)}{2} = \frac{\delta}{2}.$$

Because of the symmetry of the model, the law of the belief process with a path starting with $s_1 = -1$ is obtained by changing θ in $1 - \theta$. The symmetry allows us to only focus on

52

belief paths with $s_1 = 1$. We have:

$$\begin{aligned}
\mathbb{P}(s_2 = 1 | s_1 = 1) &= \frac{\mathbb{P}(s_1 = 1; s_2 = 1)}{\mathbb{P}(s_1 = 1)} \\
&= \frac{\mathbb{P}(X = H)\mathbb{P}(s_1 = 1; s_2 = 1 | X = H) + \mathbb{P}(X = L)\mathbb{P}(s_1 = 1; s_2 = 1 | X = L)}{\mathbb{P}(s_1 = 1)} \\
&= \delta(\theta^2 + (1 - \theta)^2),
\end{aligned}$$

and:

$$\mathbb{P}(s_2 = -1 | s_1 = 1) = 2\delta\theta(1 - \theta),$$

and:

$$\mathbb{P}(s_2 = 0 | s_1 = 1) = 1 - \delta.$$

Likewise, it is straightforward to prove that:

$$\mathbb{P}(s_3 = 1 | s_1 = 1; s_2 = 1) = \frac{\delta(\theta^3 + (1 - \theta)^3)}{\theta^2 + (1 - \theta)^2},$$

$$\mathbb{P}(s_3 = -1 | s_1 = 1; s_2 = 1) = \frac{\delta\theta(1 - \theta)}{\theta^2 + (1 - \theta)^2},$$

and:

$$\mathbb{P}(s_3 = 0 | s_1 = 1; s_2 = 1) = 1 - \delta,$$

and:

$$\mathbb{P}(s_3 = 1 | s_1 = 1; s_2 = -1) = \mathbb{P}(s_3 = -1 | s_1 = 1; s_2 = -1) = \frac{\delta}{2}.$$

At date t, there exist 2^t paths for the public signal history. Let us denote by $h_t = (s_1, s_2, \ldots, s_t)$ a path and $(z^i h)_t = (s_1, z_2^i s_2, \ldots, z_t^i s_t)$ biased trader i's belief path. A speculator has a rational estimation of the future cash flow at date t that is given by $\mathbb{P}(v = 1 | h_t) = p_L + (p_H - p_L)\mu_t$. Biased trader i has an estimation of the future cash flow that is given by $\mathbb{P}^i(v = 1 | (z^i h)_t) = p_L + (p_H - p_L)\mu_t^i$.

7.2 Equilibrium prices and rational benchmark

Proof of Proposition 2.1. At date 4, the final wealth of trader j, Wj, is:

$$W^j = \sum_{t=0}^{t=3} \left[d_t^j (v - P_t) - \frac{c}{2} \left(d_t^j \right)^2 \right] + v,$$

with $j = \emptyset$ referring to an arbitrageur, and $j = i \in [0, \lambda]$ referring to a biased trader. Since traders only consume at the last date $t = 4$, their objective is to maximize their expected final wealth conditional on their information, anticipating that they may trade at future dates. To solve for the optimal demands, we proceed backward. Because traders are atomistic, they take prices as given. We start by solving the program of trader j at date 3:

$$\max_{d_3^j} \mathbb{E}_3^j \left(\sum_{t=0}^{t=2} \left[d_t^j (v - P_t) - \frac{c}{2} \left(d_t^j \right)^2 \right] + d_3^j (v - P_3) - \frac{c}{2} \left(d_3^j \right)^2 + v \right).$$

It is straightforward to check that the objective function is concave in d_3^j. The first order condition is thus necessary and sufficient to characterize the optimal demand at date 3:

$$d_3^j = \frac{\mathbb{E}_3^j (v) - P_3}{c}.$$

Applying the same analysis backward, we obtain that:

$$d_t^j = \frac{\mathbb{E}_t^j (v) - P_t}{c}, \quad \forall t \in \{0, 1, 2, 3\}.$$

At date $t \in \{0, 1, 2, 3\}$, the market clearing condition is given by:

$$\int_0^\lambda \frac{\mathbb{E}_t^i (v) - P_t}{c} di + \int_{(1-\lambda)}^1 \frac{\mathbb{E}_t (v) - P_t}{c} dj = 0$$

where 0 on the right-hand side corresponds to the fact that no new share is issued on

54

the market. The first pricing equation displayed in Proposition 2.1 derives from solving the market clearing condition for P_t and noting that $\mathbb{E}_t^j(v) = \mathbb{P}_t^j(v = 1)$.

The second pricing equation is obtained using the fact that $\mathbb{P}_t^j(v = 1) = p_L + (p_H - p_L)\mu_t^j$, with $j = \emptyset$ referring to an arbitrageur, and $j = i \in [0, \lambda]$ referring to a biased trader. $\diamond$

We now analyze the benchmark case in which all traders are perfectly rational. Endogenous prices in this benchmark are indicated by a star. This benchmark is nested in our model and corresponds to the case in which $\lambda = 0$ or $q = 0$. In this case, we have $P_t^* = \mathbb{P}_t(v = 1) = p_L + (p_H - p_L)\mu_t$. Given the structure of the uncertainty in our model, we have the following proposition.

Proof of Proposition 2.2. The proof of the result on expected returns relies on the martingale property of the belief process μ_t. The process μ_t is a martingale because:

$$
\begin{aligned}
\mathbb{E}(\mu_{t+1}|s_1, \ldots, s_t) &= \mathbb{E}(\mathbb{P}(X = H|s_1, \ldots, s_{t+1})|s_1, \ldots, s_t) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}_{X=H}|s_1, \ldots, s_{t+1})|s_1, \ldots, s_t) \\
&= \mathbb{E}(\mathbb{1}_{X=H}|s_1, \ldots, s_t) \\
&= \mu_t.
\end{aligned}
$$

The second equality derives from the definition of a probability and the third one from the law of iterated expectations.

To see that volume, defined by $V_t^* = \int_0^1 d_t^j \mathbb{1}_{\mathbb{E}_t^j(v) > P_t^*} dj$, is null at each date, recall that the price is $P_t^* = \mathbb{P}_t(v = 1) = \mathbb{E}_t^j(v)$, for all j and t.

$\diamond$

We now study how asset prices, returns, and volume are influenced by the fact that some traders are prone to the confirmatory bias. Statistical properties of equilibrium variables are evaluated based on the true probability measure $\mathbb{P}$ because we take the viewpoint of an

55

econometrician who would observe independent repetitions of the model.

7.3 Stylized facts

The proof of Proposition 3.3 is divided in various paragraphs in order to consider separately the different stylized facts. We assume that $\lambda q > 0$, i.e., that some traders suffer from the confirmatory bias.

7.3.1 Excess Volume

Confirmatory bias induces excess volume. Consider for example the mixed history path $(s_1 = 1, s_2 = -1)$. The equilibrium price along this path is:

$$
\begin{aligned}
P_2(1, -1) &= p_L + (p_H - p_L)(\lambda q \theta + (1 - \lambda q)\theta(1, 2)) \\
&= p_L + (p_H - p_L)(\lambda q \theta + \frac{1 - \lambda q}{2}).
\end{aligned}
$$

There is a proportion λq of agents (corresponding to the biased traders who did not perceive a signal at date 2 instead of the actual negative signal) who have a price estimation $p_L + (p_H - p_L)\theta$, which is higher than the equilibrium price. There is also a proportion $1 - \lambda q$ of agents (corresponding to the rational traders and to the biased traders who perceive correctly the signal at date 2) who have a price estimation $p_L + (p_H - p_L)\frac{1}{2}$, which is lower than the equilibrium price. Hence, volume equals:

$$
V_2(1, -1) = \int_0^1 d_2^j \mathbb{1}_{\mathbb{E}_2^j(v) > P_2} dj = \frac{(p_H - p_L)\lambda q(1 - \lambda q)}{c}(\theta - \frac{1}{2}) > 0 = V_2^*(1, -1).
$$

There are thus histories along which there is abnormally high volume compared to the situation in which all traders are rational.

7.3.2 Excess Volatility

To show that confirmatory bias may induce excess volatility, we start by showing that $\mathbb{E}(P_t) = \frac{p_H + p_L}{2}$, for $t = 1, 2, 3$. To see this, consider two opposite signal histories $h_t = \{s_1, s_2, \ldots, s_t\}$ and $-h_t = \{-s_1, -s_2, \ldots, -s_t\}$ where at least one signal $s_i \neq 0$. Consider also the path $h_t^0 = (0, \ldots, 0)$. We have $\sum_{h_t} (\mathbb{P}(h_t) + \mathbb{P}(-h_t)) + \mathbb{P}(h_t^0) = 1$. The law of signals s_t being symmetric, $\mathbb{P}(h_t) = \mathbb{P}(-h_t)$ and thus:

$$\sum_{h_t} \mathbb{P}(h_t) = \frac{1 - (1 - \delta)^t}{2},$$

because $\mathbb{P}(h_t^0) = (1-\delta)^t$.[39] Moreover, using Equation (A.3.7) and the fact that $f(\theta, \mu) + f(1 - \theta, 1 - \mu) = 1$, we can prove by induction that $\mu_t(h_t) + \mu_t(-h_t) = 1$ and $\mu_t^j(h_t) + \mu_t^j(-h_t) = 1$. As a consequence, the pricing formula yields $P_t(h_t) + P_t(-h_t) = p_H + p_L$.

$$
\begin{aligned}
\mathbb{E}(P_t) &= \sum_{h_t} (\mathbb{P}(h_t) P_t(h_t) + \mathbb{P}(-h_t) P_t(-h_t)) + \mathbb{P}(h_t^0) \frac{p_H + p_L}{2} \\
&= \sum_{h_t} \mathbb{P}(h_t) (P_t(h_t) + P_t(-h_t)) + \mathbb{P}(h_t^0) \frac{p_H + p_L}{2} \\
&= (p_H + p_L) \sum_{h_t} \mathbb{P}(h_t) + \mathbb{P}(h_t^0) \frac{p_H + p_L}{2} \\
&= \frac{p_H + p_L}{2}.
\end{aligned}
$$

We deduce from the price dynamics that:

$$P_2 = P_2^* + \varepsilon_2,$$

---

[39]In a more general setting in which signal informativeness is not the same for good and bad signals, the model would generate excess volatility only for some parameter values. However, all the other results do not crucially rely on the symmetric structure of the signals.

where ε_2 is the symmetric random variable given by:

$$\varepsilon_2 = (p_H - p_L)\lambda q(\theta - \frac{1}{2})(\mathbb{1}_{s_1=1;s_2=-1} - \mathbb{1}_{s_1=-1;s_2=1}).$$

Defining $m = \frac{p_H + p_L}{2}$, we have:

$$\mathrm{Var}(P_2) = \mathrm{Var}(P_2^*) + \mathrm{Var}(\varepsilon_2) + 2\mathbb{E}\left[(P_2^* - m)\varepsilon_2\right].$$

Note that $\mathbb{E}\left[(P_2^* - m)\varepsilon_2\right] = 0$. This is because, in all the potential signal histories $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$, there is one of the two factors in $(P_2^* - m)\varepsilon_2$ that is equal to 0. Indeed, we have: $P_2^*(1, -1) = P_2^*(-1, 1) = m$ and $\varepsilon_2(1, 1) = \varepsilon_2(-1, -1) = 0$. Given that $\mathrm{Var}(\varepsilon_2) > 0$, we conclude that:

$$\mathrm{Var}(P_2) > \mathrm{Var}(P_2^*).$$

There is thus always excess volatility for the price at date 2.

Regarding date $t = 3$, we have:

$$P_3 = P_3^* + \varepsilon_3,$$

with:

$$
\begin{aligned}
\varepsilon_3 &= \varepsilon_3(1,1,-1)(\mathbb{1}_{s_1=1;s_2=1;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=-1;s_3=1}) \\
&+ \varepsilon_3(1,-1,1)(\mathbb{1}_{s_1=1;s_2=-1;s_3=1} - \mathbb{1}_{s_1=-1;s_2=1;s_3=-1}) \\
&+ \varepsilon_3(1,-1,-1)(\mathbb{1}_{s_1=1;s_2=-1;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=1;s_3=1}) \\
&+ \varepsilon_3(1,-1,0)(\mathbb{1}_{s_1=1;s_2=-1;s_3=0} - \mathbb{1}_{s_1=-1;s_2=1;s_3=0}) \\
&+ \varepsilon_3(1,0,-1)(\mathbb{1}_{s_1=1;s_2=0;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=0;s_3=1}) \\
&+ \varepsilon_3(0,1,-1)(\mathbb{1}_{s_1=0;s_2=1;s_3=-1} - \mathbb{1}_{s_1=0;s_2=-1;s_3=1}) \\
&= (p_H - p_L)\lambda q(\theta(2,2) - \theta)(\mathbb{1}_{s_1=1;s_2=1;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=-1;s_3=1}) \\
&+ (p_H - p_L)\lambda q(\theta(2,2) - \theta)(\mathbb{1}_{s_1=1;s_2=-1;s_3=1} - \mathbb{1}_{s_1=-1;s_2=1;s_3=-1}) \\
&+ (p_H - p_L)\lambda q(2\theta - 1)(\mathbb{1}_{s_1=1;s_2=-1;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=1;s_3=1}) \\
&+ (p_H - p_L)\lambda q(\theta - \tfrac{1}{2})(\mathbb{1}_{s_1=1;s_2=-1;s_3=0} - \mathbb{1}_{s_1=-1;s_2=1;s_3=0}) \\
&+ (p_H - p_L)\lambda q(\theta - \tfrac{1}{2})(\mathbb{1}_{s_1=1;s_2=0;s_3=-1} - \mathbb{1}_{s_1=-1;s_2=0;s_3=1}) \\
&+ (p_H - p_L)\lambda q(\theta - \tfrac{1}{2})(\mathbb{1}_{s_1=0;s_2=1;s_3=-1} - \mathbb{1}_{s_1=0;s_2=-1;s_3=1}).
\end{aligned}
$$

The variance of the price at date 3 can be computed as:

$$
\mathrm{Var}(P_3) = \mathrm{Var}(P_3^*) + \mathrm{Var}(\varepsilon_3) + 2\mathbb{E}\left[(P_3^* - m)\varepsilon_3\right].
$$

We want to show that this variance is greater than $\mathrm{Var}(P_3^*)$.

We first compute $\mathrm{Var}(\varepsilon_3)$. Because $\mathbb{E}(\varepsilon_3) = 0$, we have $\mathrm{Var}(\varepsilon_3) = \mathbb{E}(\varepsilon_3^2)$. Moreover, we note that for every path h_3 containing only one change of sign in the public signal and no null signal, $\mathbb{P}(h_3) = \frac{\delta^3 \theta(1-\theta)}{2}$ and $P_3^*(h_3) - P_3^*(-h_3) = (p_h - p_L)(2\theta - 1)$. For path h_3 containing one change of sign and a null signal, $\mathbb{P}(h_3) = \delta^2(1-\delta)\theta(1-\theta)$ while $P_3^*(h_3) = m$. We thus

have:

$$\mathrm{Var}(\varepsilon_3) \;=\; (p_H - p_L)^2 (\lambda q)^2 \left[\frac{\delta^3 \theta(1-\theta)}{2} \left(4(\theta(2,2) - \theta)^2 + \right. \right.$$

$$\left. \left. +\;\; 2(2\theta - 1)^2 \right) + \delta^2(1-\delta)\theta(1-\theta)6(\theta - \frac{1}{2})^2 \right].$$

The same type of computations leads to:

$$\mathbb{E}[(P_3^* - m)\varepsilon_3] = (p_H - p_L)^2 \frac{\delta^3 \theta(1-\theta)}{2} (2\theta - 1)\lambda q \left(2(\theta(2,2) - \theta) - (2\theta - 1) \right).$$

Finally:

$$\mathrm{Var}(P_3) - \mathrm{Var}(P_3^*) \;=\; \mathrm{Var}(\varepsilon_3) + 2\mathbb{E}[(P_3^* - m)\varepsilon_3]$$

$$=\; (p_H - p_L)^2 \lambda q \delta^2 \theta(1-\theta) \left[2\delta\lambda q(\theta(2,2) - \theta)^2 + \delta\lambda q(2\theta - 1)^2 \right.$$

$$+\;\; (1-\delta)\frac{3}{2}\lambda q(2\theta - 1)^2 + 2\delta(2\theta - 1)(\theta(2,2) - \theta) - \delta(2\theta - 1)^2 \bigg]$$

$$=\; (p_H - p_L)^2 \lambda q \delta^2 \theta(1-\theta) \left[2\delta\lambda q(\theta(2,2) - \theta)^2 + 2\delta(2\theta - 1)(\theta(2,2) - \theta) \right.$$

$$+\;\; (\lambda q\delta + (1-\delta)\frac{3}{2}\lambda q - \delta)(2\theta - 1)^2 \bigg].$$

Obviously, there is no excess volatility when $\delta = 0$, i.e., when traders do not learn anything from public signals. The only potentially negative term in the above expression is $(\lambda q\delta + (1-\delta)\frac{3}{2}\lambda q - \delta)$. A sufficient condition for this term to be positive is that $\lambda q \geq \frac{1}{2}$ and $\delta \leq \frac{3}{5}$.

7.3.3 Momentum

Assume that $t = 1$. The random event $\{P_1 - P_0 > 0\}$ coincides with the event $\{s_1 = 1\}$. Therefore, we have to prove that $\mathbb{E}(P_2 - P_1 | s_1 = 1) > 0$. But,

$$
\begin{aligned}
\mathbb{E}(P_2 - P_1 | s_1 = 1) &= \mathbb{E}(P_2(1,1)\mathbb{1}_{s_2=1} + P_2(1,0)\mathbb{1}_{s_2=0} + P_2(1,-1)\mathbb{1}_{s_2=-1} | s_1 = 1) - ((p_H - p_L)\theta + p_L) \\
&= (p_H - p_L)[\theta(2,2)\mathbb{P}(s_2 = 1 | s_1 = 1) + \theta\mathbb{P}(s_2 = 0 | s_1 = 1) \\
&\quad + ((1 - \lambda q)\theta(1,2) + \lambda q\theta)\mathbb{P}(s_2 = -1 | s_1 = 1)) - \theta] \\
&= (p_H - p_L)\lambda q(\theta - \theta(1,2))\mathbb{P}(s_2 = -1 | s_1 = 1) > 0,
\end{aligned}
$$

where the last equality follows from the martingale property of the belief process μ_t. There is thus always momentum at date 1.

At $t = 2$, the return $P_2 - P_1$ is positive on the event $\{s_2 = 1\}$ only. Thus:

$$
\begin{aligned}
\mathbb{E}(P_3 - P_2 | P_2 - P_1 > 0) &= \mathbb{E}(P_3 - P_2 | s_1 = 1; s_2 = 1)\frac{\mathbb{P}(s_1 = 1; s_2 = 1)}{\mathbb{P}(s_2 = 1)} \\
&+ \mathbb{E}(P_3 - P_2 | s_1 = -1; s_2 = 1)\frac{\mathbb{P}(s_1 = -1; s_2 = 1)}{\mathbb{P}(s_2 = 1)} \\
&+ \mathbb{E}(P_3 - P_2 | s_1 = 0; s_2 = 1)\frac{\mathbb{P}(s_1 = 0; s_2 = 1)}{\mathbb{P}(s_2 = 1)}.
\end{aligned}
$$

Using the martingale property of the belief process, we get:

$$
\mathbb{E}(P_3 - P_2 | s_1 = 1; s_2 = 1) = (p_H - p_L)\lambda q(\theta(2,2) - \theta)\mathbb{P}(s_3 = -1 | s_1 = 1; s_2 = 1)
$$

$$
\mathbb{E}(P_3 - P_2 | s_1 = 0; s_2 = 1) = (p_H - p_L)\lambda q(\theta - \frac{1}{2})\mathbb{P}(s_3 = -1 | s_1 = 0; s_2 = 1)
$$

and:

$$
\begin{aligned}
\mathbb{E}(P_3 - P_2 | s_1 = -1; s_2 = 1) &= (p_H - p_L)\lambda q(1-\theta)\mathbb{P}(s_3 = 1 | s_1 = -1; s_2 = 1) \\
&+ (p_H - p_L)\lambda q(1-\theta)\mathbb{P}(s_3 = 0 | s_1 = -1; s_2 = 1) \\
&+ (p_H - p_L)\lambda q\theta(0,2)\mathbb{P}(s_3 = -1 | s_1 = -1; s_2 = 1) - \lambda q(1-\theta) \\
&= (p_H - p_L)\lambda q(\theta(0,2) - (1-\theta))\mathbb{P}(s_3 = -1 | s_1 = -1; s_2 = 1).
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
\mathbb{E}(P_3 - P_2 | P_2 - P_1 > 0) &= (p_H - p_L)\lambda q(\theta(2,2) - \theta)\frac{\mathbb{P}(s_1 = 1; s_2 = 1; s_3 = -1)}{\mathbb{P}(s_2 = 1)} \\
&+ (p_H - p_L)\lambda q(\theta(0,2) - (1-\theta))\frac{\mathbb{P}(s_1 = -1; s_2 = 1; s_3 = -1)}{\mathbb{P}(s_2 = 1)} \\
&+ (p_H - p_L)\lambda q(\theta - \frac{1}{2})\frac{\mathbb{P}(s_1 = 0; s_2 = 1; s_3 = -1)}{\mathbb{P}(s_2 = 1)} \\
&= (p_H - p_L)\lambda q(\theta - \frac{1}{2})\frac{\mathbb{P}(s_1 = 0; s_2 = 1; s_3 = -1)}{\mathbb{P}(s_2 = 1)} > 0,
\end{aligned}
$$

where the last equality uses $\theta(2,2) = 1 - \theta(0,2)$ and:

$$
\mathbb{P}(s_1 = 1; s_2 = 1; s_3 = -1) = \mathbb{P}(s_1 = -1; s_2 = 1; s_3 = -1).
$$

7.3.4   Bubble

In the section *Momentum*, we have proved that: $\mathbb{E}(P_2 - P_1 | s_1 = 1) > 0$ which is equivalent to $\mathbb{E}(P_2 | s_1 = 1) \geq \mathbb{E}(P_1 | s_1 = 1) = p_L + (p_H - p_L)\theta$.

Moreover, we have:

$$
\begin{aligned}
\mathbb{E}(P_3 - P_2|s_1 = 1) \;=\;\; & \mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = 1)\mathbb{P}(s_2 = 1|s_1 = 1) \\
+\;\; & \mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = 0)\mathbb{P}(s_2 = 0|s_1 = 1) \\
+\;\; & \mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = -1)\mathbb{P}(s_2 = -1|s_1 = 1).
\end{aligned}
$$

We have already proved that:

$$
\mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = 1) > 0.
$$

And we also have:

$$
\mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = -1) = (p_H - p_L)\lambda q^2 \left(\theta - (1-\theta)\right)\mathbb{P}(s_3 = -1|s_1 = 1; s_2 = -1) > 0.
$$

and

$$
\mathbb{E}(P_3 - P_2|s_1 = 1, s_2 = 0) = (p_H - p_L)\lambda q \left(\theta - \frac{1}{2}\right)\mathbb{P}(s_3 = -1|s_1 = 1; s_2 = 0) > 0.
$$

Therefore,

$$
\mathbb{E}(P_3|s_1 = 1) \geq \mathbb{E}(P_2|s_1 = 1).
$$

The proof for the case of a crash follows the same logic.

7.4 Novel empirical predictions

7.4.1 Beliefs dispersion

Let us denote by ν_t traders' beliefs distribution. For each history h_t, $\nu_t(h_t)$ is a probability measure on the finite space $\theta(s, k)$, for $s \in \{0, 1, \ldots, k\}$, representing traders' belief distri-

bution function after observing s strictly positive signals out of k non-null signals. Various fractions of traders have beliefs that take value in $\theta(s, k)$, for $s \in \{0, 1, \ldots, k\}$. We can thus compute the average and the variance of the belief distribution, across traders at a given date. We will define beliefs dispersion for a given date as the variance of each probability measure $\nu_t(h_t)$ and the conditional beliefs dispersion as the conditional variance of each probability measure.

Let us write precisely traders' beliefs distribution for $t \in \{0, 1, 2, 3\}$. Clearly,

$$\nu_0 = D_{\frac{1}{2}},$$

where D stands for the Dirac measure. Because all traders have the same perception of date 1 signal, the distribution of biased traders' beliefs at date 1 is:

$$\nu_1 = \frac{\delta}{2}(D_\theta + D_{1-\theta}) + (1 - \delta)D_{\frac{1}{2}}.$$

In particular, $\nu_1(1) = D_\theta$ and $\nu_1(-1) = D_{1-\theta}$ and therefore Var $\nu_1(1) = $ Var $\nu_1(-1) = 0$. The average belief dispersion is thus $\mathbb{E}(\text{Var } \nu_1) = 0$.

From date 2 onward, beliefs dispersion depends on the public signal's path. More precisely, we have:

$$
\begin{aligned}
\nu_2 \;=\; & \mathbb{P}(s_1 = 1; s_2 = 1)\nu_2(1; 1) + \mathbb{P}(s_1 = 1; s_2 = -1)\nu_2(1; -1) \\
& + \; \mathbb{P}(s_1 = -1; s_2 = 1)\nu_2(-1; 1) + \mathbb{P}(s_1 = -1; s_2 = -1)\nu_2(-1; -1) \\
& + \; \mathbb{P}(s_1 = 0; s_2 = 1)\nu_2(0; 1) + \mathbb{P}(s_1 = 1; s_2 = 0)\nu_2(1; 0) \\
& + \; \mathbb{P}(s_1 = 0; s_2 = -1)\nu_2(0; -1)\mathbb{P}(s_1 = -1; s_2 = 0)\nu_2(-1; 0)
\end{aligned}
$$

where:

$$\nu_2(1; 1) = D_{\theta(2,2)}; \nu_2(1; -1) = qD_{\theta(2,2)} + (1 - \lambda q)D_{\theta(1,2)};$$

64

$$\nu_2(-1; -1) = D_{\theta(0,2)}; \nu_2(-1; 1) = qD_{\theta(0,2)} + (1 - \lambda q)D_{\theta(1,2)}.$$

At date 3, we have for every path h_3 a dispersion measure $\nu_3(h_3)$. For instance:

$$\nu_3(1; 1; 1) = D_{\theta(3,3)}; \nu_3(1; 1; -1) = qD_{\theta(3,3)} + (1 - \lambda q)D_{\theta(2,3)};$$

$$\nu_3(1; -1; 1) = qD_{\theta(3,3)} + (1-\lambda q)D_{\theta(2,3)}; \nu_3(1; -1; -1) = q^2 D_{\theta(3,3)} + 2q(1-q)D_{\theta(2,3)} + ((1-\lambda) + \lambda(1-q)^2)D_{\theta(1,3)}.$$

We first prove that:

$$\mathbb{E}(\mathrm{Var}(\nu_2)|s_1 s_2 > 0) < \mathbb{E}(\mathrm{Var}(\nu_2)|s_1 s_2 < 0).$$

We observe that:

$$\mathbb{E}(\mathrm{Var}(\nu_2)|s_1 s_2 > 0) = \frac{\mathrm{Var}(\nu_2)(1, 1)\mathbb{P}(s_1 = 1, s_2 = 1) + \mathrm{Var}(\nu_2)(-1, -1)\mathbb{P}(s_1 = -1, s_2 = -1)}{\mathbb{P}(s_1 s_2 > 0)}$$

$$= 0,$$

because $\mathrm{Var}(\nu_2)(1, 1) = \mathrm{Var}(\nu_2)(-1, -1) = 0$. On the other hand,

$$\mathbb{E}(\mathrm{Var}(\nu_2)|s_1 s_2 < 0) = \frac{\mathrm{Var}(\nu_2)(1, -1)\mathbb{P}(s_1 = 1, s_2 = -1) + \mathrm{Var}(\nu_2)(-1, 1)\mathbb{P}(s_1 = -1, s_2 = 1)}{\mathbb{P}(s_1 s_2 < 0)}$$

$$= \frac{\mathrm{Var}(\nu_2)(1, -1) + \mathrm{Var}(\nu_2)(-1, 1)}{2}$$

$$> 0,$$

which completes the proof.

We now prove that: $\mathbb{E}\left(\mathrm{Var}\left[\nu_3|_{s_1 s_2 < 0 \text{ or } s_2 s_3 < 0}\right]\right) > \mathbb{E}\left(\mathrm{Var}\left[\nu_3|_{s_1 s_2 > 0 \text{ and } s_2 s_3 > 0}\right]\right).$ First, as before, remark that:

$$\mathrm{Var}\left[\nu_3|_{s_1 s_2 > 0 \text{ and } s_2 s_3 > 0}\right] = 0.$$

Notice that without of generality, we may assume $s_1 = 1$,

$$
\begin{aligned}
\mathbb{E}(\mathrm{Var}(\nu_3)|s_1 s_2 < 0 \text{ or } s_2 s_3 < 0) &= \mathbb{E}(\mathrm{Var}(\nu_3)|s_1 = 1, s_2 = 1, s_3 = -1)\frac{\mathbb{P}(s_1 = 1, s_2 = 1, s_3 = -1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)} \\
&+ \mathbb{E}(\mathrm{Var}(\nu_3)|s_1 = 1, s_2 = -1, s_3 = -1)\frac{\mathbb{P}(s_1 = 1, s_2 = -1, s_3 = -1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)} \\
&+ \mathbb{E}(\mathrm{Var}(\nu_3)|s_1 = 1, s_2 = -1, s_3 = 1)\frac{\mathbb{P}(s_1 = 1, s_2 = -1, s_3 = 1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)} \\
&= \mathrm{Var}(\nu_3)(1, 1, -1)\frac{\mathbb{P}(s_1 = 1, s_2 = 1, s_3 = -1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)} \\
&+ \mathrm{Var}(\nu_3)(1, -1, -1)\frac{\mathbb{P}(s_1 = 1, s_2 = -1, s_3 = -1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)} \\
&\mathrm{Var}(\nu_3)(1, -1, 1)\frac{\mathbb{P}(s_1 = 1, s_2 = -1, s_3 = 1)}{\mathbb{P}(s_1 s_2 < 0 \text{ or } s_2 s_3 < 0)}
\end{aligned}
$$

To conclude, it suffices to observe that the last three terms of the right-hand side are strictly positive.

7.4.2 Beliefs revision

Let us focus on date 2. We will compute the probability of upward belief's revision when past public signals have different signs, and show that it is lower than the probability of upward belief's revision when past public signals have the same signs.

For every agent j, the probability of upward belief's revision when past public signals have the same signs equals one because all agents have perceived correctly the public signals, i.e.

$$
\mathbb{P}\left(\mu_2^j - \mu_1^j > 0 | s_1 > 0, s_2 > 0\right) = 1.
$$

Along a path where public signals have different signs, for instance $s_1 < 0$ and $s_2 > 0$, belief's revision depends on the nature of the agent. Rational agents, and a proportion $\lambda(1 - q)$ of biased traders, perceive correctly the public signals and therefore revise upward their belief after a good signal at date 2. However, there is a proportion λq of biased traders that does

not perceive the public signal at date 2. Therefore:

$$\mathbb{P}\left(\mu_2^j - \mu_1^j > 0 | s_1 < 0, s_2 > 0\right) = \lambda(1 - q) + 1 - \lambda = 1 - \lambda q < 1.$$

The same line of reasoning applies at date 3. We have:

$$\mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 > 0, s_2 > 0, s_3 > 0\right) = 1.$$

Moreover,

$$\mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 < 0 \text{ or } s_2 < 0, s_3 > 0\right) =$$

$$= \mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 < 0, s_2 = 0, s_3 > 0\right) \frac{\mathbb{P}(s_1 < 0, s_2 = 0, s_3 > 0)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)}$$

$$+ \mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 < 0, s_2 > 0, s_3 > 0\right) \frac{\mathbb{P}(s_1 < 0, s_2 > 0, s_3 > 0)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)}$$

$$+ \mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 = 0, s_2 < 0, s_3 > 0\right) \frac{\mathbb{P}(s_1 = 0, s_2 < 0, s_3 > 0)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)}$$

$$+ \mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 > 0, s_2 < 0, s_3 > 0\right) \frac{\mathbb{P}(s_1 > 0, s_2 < 0, s_3 > 0)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)}$$

$$+ \mathbb{P}\left(\mu_3^j - \mu_2^j > 0 | s_1 < 0, s_2 < 0, s_3 > 0\right) \frac{\mathbb{P}(s_1 < 0, s_2 < 0, s_3 > 0)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)}.$$

We define:

$$p_1 = \mathbb{P}(s_1 < 0, s_2 = 0, s_3 > 0) = (1 - \delta)\delta^2\theta(1 - \theta),$$

$$p_2 = \mathbb{P}(s_1 = 0, s_2 < 0, s_3 > 0) = (1 - \delta)\delta^2\theta(1 - \theta),$$

$$p_3 = \mathbb{P}(s_1 < 0, s_2 > 0, s_3 > 0) = \frac{\delta^3}{2}\theta(1 - \theta),$$

$$p_4 = \mathbb{P}(s_1 > 0, s_2 < 0, s_3 > 0) = \frac{\delta^3}{2}\theta(1 - \theta),$$

$$p_5 = \mathbb{P}(s_1 < 0, s_2 < 0, s_3 > 0) = \frac{\delta^3}{2}\theta(1 - \theta).$$

We thus obtain:

$$
\begin{aligned}
\mathbb{P}\left(\mu_3^j - \mu_2^j > 0 \,|\, s_1 < 0 \text{ or } s_2 < 0, s_3 > 0\right) &= 2(1 - \lambda q)\frac{(1 - \delta)\delta^2\theta(1 - \theta)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)} \\
&+ (1 - \lambda q)\frac{\frac{\delta^3}{2}\theta(1 - \theta)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)} \\
&+ \frac{\frac{\delta^3}{2}\theta(1 - \theta)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)} \\
&+ (1 - \lambda q^2)\frac{\frac{\delta^3}{2}\theta(1 - \theta)}{\mathbb{P}(s_1 < 0 \text{ or } s_2 < 0, s_3 > 0)} \\
&= (1 - \lambda q)\frac{p_1}{\sum_{i=1}^{5} p_i} + (1 - \lambda q)\frac{p_2}{\sum_{i=1}^{5} p_i} + (1 - \lambda q)\frac{p_3}{\sum_{i=1}^{5} p_i} \\
&+ \frac{p_4}{\sum_{i=1}^{5} p_i} + (1 - \lambda q^2)\frac{p_5}{\sum_{i=1}^{5} p_i} \\
&< \frac{p_1}{\sum_{i=1}^{5} p_i} + \frac{p_2}{\sum_{i=1}^{5} p_i} + \frac{p_3}{\sum_{i=1}^{5} p_i} + \frac{p_4}{\sum_{i=1}^{5} p_i} + \frac{p_5}{\sum_{i=1}^{5} p_i} \\
&= 1.
\end{aligned}
$$

7.4.3 Individual beliefs updating

The proof of the last prediction in Proposition 3.4 derives directly from the previous proof after remarking that, if a belief change is strictly different from zero, it has the same sign as the signal that generated this change. For example, if one observes that $\mu_2^j - \mu_1^j > 0$, for a given j, then we have $s_2 > 0$.