Working Paper Series n° 720

March, 2012

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

Sébastien Pouget and Stéphane Villeneuve

Institut d'Economie Industrielle (IDEI) – Manufacture des Tabacs
Aile Jean-Jacques Laffont – 21, allée de Brienne – 31000 TOULOUSE – FRANCE
Tél. + 33(0)5 61 12 85 89 – Fax + 33(0)5 61 12 86 37 – **www.idei.fr** – contact@cict.fr

INSTITUT D'ECONOMIE INDUSTRIELLE

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

Sebastien Pouget[†]and Stephane Villeneuve [‡]

March 7, 2012

*We are grateful to Nicholas Barberis, Bruno Biais, Sylvain Friederich, Ron Kaniel, Hyun-woo Lee, Thomas Mariotti, Richard Payne, Matthew Rabin, Jean-Charles Rochet, Hersh Shefrin, Dimitri Vayanos, Wei Xiong, seminar participants at Bristol University and Toulouse University, and participants to the 2010 American Finance Association meeting and the fall 2011 NBER Behavioral Finance meeting for providing useful comments. A previous version of this paper circulated under the title "Price formation with confirmation bias". All errors are ours. Financial support from the Agence Nationale de la Recherche (ANR-09-BLAN-0358-01) and from the Chaire SCORE at IDEI-R is gratefully acknowledged.

[†]University of Toulouse (TSE-IAE-CRM-IDEI), 21 Alle de Brienne, 31 000 Toulouse, spouget@univ-tlse1.fr

[‡]University of Toulouse (TSE-CRM-IDEI), 21 alle de Brienne, 31000 Toulouse and Europlace Institute of Finance, stephane.villeneuve@univ-tlse1.fr.

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

Abstract

This paper proposes a dynamic model of financial markets where some investors are prone to the confirmation bias. Following insights from the psychological literature, these agents are assumed to amplify signals that are consistent with their prior views. In a model with public information only, this assumption provides a rationale for the volume-based price momentum documented by Lee and Swaminathan (2000). Our results are also consistent with a variety of other empirically documented phenomena such as bubbles, crashes, reversals and excess price volatility and volume. Novel empirical predictions are derived: i) return continuation should be stronger when biased traders' beliefs are more extreme, and ii) return continuation should be stronger after an increase in trading volume. The implications of our model for short-term quantitative investments are twofold: i) optimal trading strategies involve riding bubbles, and that ii) contrarian trading can be optimal in some market circumstances.

Keywords: financial markets, psychological biases, confirmation bias, momentum, reversal, bubbles, trading strategies

"A mind is a terrible thing to change. You decide gold is a good bet to hedge against inflation, and suddenly the news seems to be teeming with signs of a falling dollar and rising prices down the road. Or 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. Psychologists call this mental gremlin the "confirmation bias"." (Jason Zweig, in the *Wall Street Journal*, November 19, 2009)

Information processing is a major aspect of trading in financial markets. Investors need to learn about the fundamental characteristics of the assets that are traded. They also need to forecast what other investors will think of these fundamentals in order to better estimate future asset prices. Economic theory traditionally assumes that investors have perfect information perception capabilities, that is, investors have unlimited and unbiased cognition. In contrast, the psychological literature on judgment and decision-making suggests that some individuals are prone to cognitive biases and make systematic mistakes when processing information (see, for example, Kahneman, Slovic, and Tversky (1982) and Gilovich, Griffin, and Kahneman (2002)).

Models of behavioral finance have analyzed the financial consequences of various psychological biases. Overconfidence is studied by Kyle and Wang (1997), Odean (1998), Benos (1998), Scheinkman and Xiong (2003), and Dumas, Kurshev, and Uppal (2005). Self-attribution bias is examined by Daniel, Hirshleifer, and Subrahmanyam (1998), Gervais and Odean (2001), and Choi and Lou (2010), gambler's fallacy by Rabin and Vayanos (2010), and representativeness and conservatism by Barberis, Shleifer, and Vishny (1998). To complement this literature, the present paper incorporates the confirmation bias in a model of financial markets.

Confirmation bias is defined by Nickerson (1998) as "the seeking or interpreting of evidence in ways that are partial to existing beliefs". Jonathan Evans indicates that "confirmation bias is perhaps the best known and most widely accepted notion of inferential error to come out of the literature on human reasoning" (Evans (1989), p. 41). The prevalence of the confirmation bias has recently been quantitatively established in a meta-analysis by Hart et al. (2009).

In light of the evidence offered by psychologists (such as Bodenhausen

(1998)), our model assumes that a trader prone to the confirmation bias amplifies information that is consistent with his or her prior views. In our framework, this bias creates differences of opinion between rational arbitrageurs and biased traders over the interpretation of public information. These differences of opinion in turn motivate trading. Arbitrageurs take opposite positions with respect to biased traders and thus have a stabilizing impact on prices. Transaction costs however limit the effectiveness of arbitrage strategies causing the views of both arbitrageurs and biased traders to be incorporated into asset prices.[1]

Our results are related to the properties of mispricings and to short-term quantitative trading strategies. To the best of our knowledge, our model is the first to rationalize the volume-based momentum documented by Lee and Swaminathan (2000): in our model, a price trend associated with a high volume induces more momentum than a price trend associated with a low volume. This is because a high volume indicates a large difference of opinion (and thus a large mistake in biased traders' beliefs). In turn, the higher the mistake is, the longer biased traders stick with their opinion, thereby causing price momentum.

We also show that the presence of investors prone to the confirmation bias creates bubbles and crashes, and induces short-term momentum, long-term reversals, and excess price volatility and volume. Bubbles occur when traders hold optimistic beliefs.[2] This is because positive beliefs increase the

[1] We base our model on exogenous transaction costs that reflect commissions or market impact. We make this modeling choice for simplicity only. Transaction costs can have a significant pricing impact (see, for example, Lesmond, Schill and Zhou (2004) for a study of the transaction costs involved in momentum trading), but other factors can also limit the willingness to trade of arbitrageurs and would also support our results. As analyzed by Wurgler and Zhuravskaya (2002), assets' fundamental risk diminishes the attractiveness of arbitrage when perfect substitutes do not exist. Noise trader risk coupled with short investment horizons as modeled by De Long, Shleifer, Summers, and Waldmann (1990) and by Shleifer and Vishny (1997) are other factors preventing arbitrageurs from fully exploiting unsophisticated investors' erroneous sentiment. Finally, Abreu and Brunnermeier (2002)'s synchronization risk is another type of risk borne by arbitrageurs that reduces their willingness to fight against mispricings.

[2] We use the terms optimism and pessimism to reflect the fact that traders' beliefs are respectively above or below their initial expectation of the fundamental value of the asset. Our paper is therefore different from the analysis of Jouini and Napp (2007), among others, that considers optimism or pessimism as a personal and stable psychological trait that leads agents to systematically over- or under- estimate the probability of successes of

likelihood that biased traders will in the future exaggerate positive signals. During a bubble, prices are above perfectly rational levels and, from a rational arbitrageur's perspective, the mispricing is expected to worsen in the future. The same logic applies to crashes as they relate to pessimism. Momentum and reversals are a consequence of the fact that when biased traders' beliefs shift from, for example, optimistic to pessimistic, these traders switch from amplifying positive news to amplifying negative news. These autocorrelations in price changes are not driven by fundamental factors since public information is modeled as identically and independently distributed signals. Excess volatility derives from the fact that biased traders sometimes overreact to information. Excess variance is however not an arte-fact of the overreaction embedded in our definition of confirmation bias. Indeed, there is also excess variance, for a large range of parameter values, when biased traders *underreact* to information that is *inconsistent* with their prior beliefs. Excess volume arises naturally in our framework because the confirmation bias creates differences of opinion. Finally, we find that, in the short run (i.e., as long as the cash flow has not been distributed), biased traders are expected to have higher capital gains than rational arbitrageurs. This is in line with the result of De Long et al. (1990) in which irrational traders' expected profits can be higher than rational traders' ones. Since rational arbitrageurs have a long horizon and only care about their consumption after the final cash-flow has been distributed, this inferior interim performance is not a problem for them. We show below that the situation would be different if rational arbitrageurs had a short horizon.

These asset pricing patterns are in line with the results of various empirical studies. At quarterly, semi-annual or yearly horizons, Jegadeesh and Titman (1993) document a momentum effect: assets 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). DeBondt and Thaler (1985) and Fama and French (1988) document a reversal effect: at a three- or five-year horizon, past losers tend to outperform past winners. Starting with Leroy and Porter (1981) and Shiller (1981), several contributions report that asset prices are excessively volatile. 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 fi-

their investments.

nance paradigm". The present paper offers a parsimonious model based on investors' misperception of public information that is consistent with these stylized facts.[3]

Novel empirical predictions of our model are two-fold. On the one hand, when biased traders' beliefs are more extreme, future price changes should be less positively autocorrelated than when biased traders' beliefs are neutral. Indeed, when biased traders are extremely optimistic (respectively, pessimistic), they almost surely amplify positive (respectively, negative) news. This in turn implies that, when biased traders' beliefs are extreme, there is almost no positive autocorrelation in short-term price changes. Our model thus predicts an inverted U-shaped relationship between the strength of biased traders' beliefs and the future autocorrelation in price changes: autocorrelation should be maximal when biased traders are neutral and should be minimal when biased traders are extremely pessimistic or optimistic. To identify the case in which biased traders beliefs are extreme, one can use past volume: a price trend associated with a high volume indicates high disagreement of biased traders with rational arbitrageurs, and is associated with more extreme beliefs than low volume price trends. In the categorization of Lee and Swaminathan (2000), one would expect lower autocorrelation in future price changes for large volume winners and losers than for the other assets.

On the other hand, price continuations should be more pronounced when past price trends are associated with an increasing trend in volume. This is because such an increase in volume signals that biased traders' beliefs become more and more extreme. These predictions are not present in other behavioral models of momentum (see for example, Barberis, Shleifer and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), Hong and Stein (1999), and Rabin and Vayanos (2010)) and could thus be helpful to disentangle the sources of the momentum effect observed in stock markets.

To analyze the implications of our model for quantitative investments, we derive the optimal trading strategy of a short-term rational trader referred

[3]The model is parsimonious in the sense that it includes the perfectly rational benchmark as a special case and that departures from perfect rationality are driven by only one parameter. This parameter is denoted by q and measures the severity of the confirmation bias. When $q = 1$, the model features perfectly rational asset prices and trading strategies. In contrast, when $q > 1$, asset prices and trading strategies are influenced by the confirmation bias.

to as a hedge fund. The short horizon can be interpreted as coming from a liquidity constraint and implies that the hedge fund does not have the same incentives as the rational arbitrageurs who can hold positions until asset's cash flow is distributed.[4] We indeed show that the hedge fund has an incentive to "ride" bubbles despite its knowledge of an over-valuation. This is because it correctly anticipates that mispricings will worsen in the future. This is consistent with hedge funds' behavior during the technology bubble as described by Brunnermeier and Nagel (2004) or with Hoare's (a London-based bank) trading behavior during the 1720's South Sea bubble as reported by Temin and Voth (2003). We also show that the hedge fund should short-sell during crashes despite its knowledge of an under-valuation.

We finally show that, while the hedge fund in our model uses a positive-feedback strategy more often than a contrarian strategy, both of these trading strategies can be optimal depending on market circumstances: when biased traders are optimistic, a hedge fund should adopt a positive-feedback strategy after positive public announcements as well as after strongly negative announcements. The fund should buy after price increases and sell after large price drops.[5] On the contrary, after mildly negative public announcements, a hedge fund should adopt a contrarian strategy, that is, the fund should buy after moderate price drops. When biased traders are pessimistic, the optimal strategy of the hedge fund follows the inverse logic. Overall, the important input that drives this optimal trading strategy is biased trader's beliefs. We provide an explicit characterization of the optimal strategy, referred to as confirmation-based strategy, and compare it to momentum and technical analysis strategies.

The rest of this paper is organized as follows. Next section briefly re-

[4]Because rational arbitrageurs are constrained by transaction costs, their presence does not completely eliminate the anomalies generated by biased traders. This leaves some profit opportunities for an additional rational trader such as the hedge fund in our model. We assume that the hedge fund is atomistic so it does not influence prices or eliminate the anomalies. On the top of transaction costs, the hedge fund is constrained by a short-term horizon; Its profitability even if positive will thus be lower than the one of rational arbitrageurs.

[5]Strong public announcements are defined as signals that are extreme enough to shift biased traders' beliefs from optimistic to pessimistic or inversely from pessimistic to optimistic. By opposition, an announcement is characterized as mild when it does not change the direction of biased traders' beliefs. These definitions will be more formally laid out in Section 4.

views the literatures on the confirmation bias and on theoretical behavioral fiannce. Section 2 describes our model. Our asset pricing results are presented in Section 3. Section 4 offers implications for quantitative trading strategies. Section 5 concludes. Proofs are in the Appendix.

1 Literature

1.1 On the confirmation bias

Central to our model is the assumption that some traders are prone to the confirmation bias. This bias has been extensively documented in the psychological literature following the seminal contribution of Lord, Ross, and Lepper (1979). They study how people react to new information regarding a social dispute, namely capital punishment. Confronted with mixed evidence regarding the deterrent effectiveness of this type of punishment, both initial proponents and opponents to capital punishment tend to rate the evidence that favors their initial view as more convincing and reliable. Furthermore, the mixed evidence induces a polarization in people's beliefs, a phenomenon at odd with Bayes' rule prediction that beliefs should move in the same direction after seeing the same information.

The results of Lord et al. (1979) could be due to a systematic tendency to only consider (or give credit) to particular types of information. Darley and Gross (1983) address this issue and show that confirmation bias is indeed at work. They experimentally manipulate prior beliefs of participants in a task that consisted in the evaluation of the academic ability of a fourth-grade student. They show that subjects who are given a good (bad) initial opinion of the student end up with an even more positive (negative) evaluation after receiving mixed evidence of her ability in various academic tests.

Confirmation bias is related to two psychological mechanisms. On the one hand, *meaning change* reflects the tendency to change the valence of information in order to make it support prior beliefs. On the other hand, *biased assimilation* refers to the tendency to give more attention and more weight on the decision to information that are in line with prior beliefs. Bodenhausen (1988) evaluates the relative importance of these two mechanisms in the context of a judicial trial experiment. Subjects are initially given a neutral or a bad opinion about a defendant. They then are given various pieces of information, some containing evidence in favor of the defense, others containing evidence in favor of the prosecution. Bodenhausen (1988) shows that subjects with an initially bad opinion of the defendant tend to interpret correctly the meaning of the evidence but better remember the evidence that favors the prosecution. Bodenhausen (1988) interprets these results as favoring the biased attribution mechanism.[6]

Hart et al. (2009) offer more systematic evidence regarding the prevalence of the confirmation bias. They perform a meta-analysis of 91 studies that include almost 8,000 subjects. They find that people are two times more likely to pay attention to an information that is in line rather than not in line with their prior beliefs. They further show that the confirmation bias is stronger when the information received are of good quality, when the decision maker has a high commitment towards his prior beliefs (for example, because they are closely related to his or her personal values), and when the decision maker is less confident that his or her prior beliefs are correct. Finally, individuals who are prone to the confirmation bias appear to be more closed-minded than others.

The economics literature has recently recognized the importance of the confirmation bias. Rabin and Schrag (1999) study the consequences of the confirmation bias for beliefs' formation. They show that confirmation bias can lead to overconfidence and can prevent adequate learning of the realized state of nature even after receiving a infinite amount of information. [7] Evidence of such judgmental bias has been documented in an economic setting by Forsythe, Nelson, Neumann, and Wright (1992). They analyze data from the Iowa Political Stock Market organized during the 1988 U.S. presidential election campaign. They report that, after the third debate, supporters of one candidate were more likely to think that their preferred candidate won the debate. They also show that these beliefs influenced trading behavior. In the present paper, we propose a model that formally incorporates the

[6]Other contributions to the study of the confirmation bias include Plous (1991), Dougherty, Turban, and Callender (1994), and Dave and Wolfe (2003). Additional references are offered by Nickerson (1998) and by Rabin and Schrag (1999).

[7]Some papers offer rational theories of attitude polarization. In Suen (2004), attitude polarization is due to decision makers relying on advisors. In Kondor (2011)'s model of financial markets, it is due to investors' differential trading horizons. Confirmation bias can also be related to the need to avoid dissonance between past choices and current beliefs (see Yariv (2005) for a model of cognitive dissonance).

confirmation bias and studies its influence on trading strategies and on the price formation process in financial markets.

1.2 On behavioral finance models

Our model unifies two mostly separate streams of literature about financial markets, one on differences of opinion and the other on cognitive biases. The first literature on differences of opinion has generated a lot of predictions regarding the relationship between volume and asset returns. For example, Harris and Raviv (1993) predict that price changes and volume are positively correlated. When there are short-sales constraints, Miller (1977) predicts that high disagreement lead to lower expected returns and Hong and Stein (2003) to negatively skewed returns. In a dynamic setting, Scheinkman and Xiong (2003) show that differences of opinion lead to a speculative motive for trade, and to a positive link between trading volume and price volatility.

Behavioral finance models based on cognitive biases offer more insights about the time-series properties of asset prices, in particular about the momentum effect documented by Jegadeesh and Titman (1993) for the US stock market or more recently by Asness, Moskowitz, and Pedersen (2009) for a variety of international asset classes. In Barberis, Shleifer, and Vishny (1998), conservatism induces traders to underreact to isolated news, and representativeness heuristic induces them to overreact to a series of news. When associated with limited learning capabilities, these two effects generate excessive price continuation. In Rabin and Vayanos (2010), investors are subject to the gambler's fallacy and expect outcomes in random series to exhibit systematic reversals. Rabin and Vayanos (2010) show that the gambler's fallacy can lead to the hot-hand fallacy: when investment performance can be due to luck or skill, investors over-infer that their asset manager is good after observing a streak of above average performance. Thus, the gambler's fallacy coupled with delegated portfolio management can explain momentum.

The model that is closest to ours is offered by Daniel, Hirshleifer, and Subrahmanyam (1998). In their model, traders observe private information and then receive a public signal. Traders are assumed to suffer from the self-attribution bias: if the realization of the public signal is in line with private information, traders' confidence in the accuracy of private information rises, whereas if it is not in line, confidence falls less than it should. In this context with public and private information, Daniel et al. (1998) show that self-attribution generates momentum. In addition to being based on a different psychological bias, our model differs from theirs in several dimensions. First, in our model, it is possible (even if not as frequent as it should be) that biased traders change their mind and overreact to different signals compared to the ones they were overreacting to initially. For example, if biased traders are initially optimistic, they will start by overreacting to positive information. But if a very negative information reaches the market such that traders become pessimistic, these traders will now overreact to negative information. Second, our model matches the empirical evidence on the volume-based price momentum offered by Lee and Swaminathan (2000): price continuation is stronger after a high past volume than after a low past volume. Third, our rationalization of the momentum effect does not rely on the access to private information. We use a framework with differences of opinions la Harris and Raviv (1993) while they consider an asymmetric information setting la Grossman and Stiglitz (1980). In their setting, it is the addition of private information and the self-attribution bias that creates momentum.

Another type of model of the momentum effect is based on the interaction between different types of traders. In Hong and Stein (1999), news-watchers and trend-followers are trading with each others. Information is assumed to diffuse gradually within the population of news-watchers, thereby generating initial price underreaction and predictability. This predictability is exploited by trend-followers. Because they are assumed to be using only simple strategies, trend-followers generate price overreaction. The interaction between heterogenous agents is thus inducing price momentum. [8]

Most of the models described above also predict price reversals and excess volatility but they cannot rationalize the volume-based price momentum. The various explanations, including ours, of the momentum effect and of the other price patterns are not mutually exclusive. In this paper, we provide several new empirical predictions, in particular about the relationship between volume and price patterns. These predictions could be useful to evaluate the relative importance of the confirmation bias relative to other

[8] Some articles such as Johnson (2002), Biais, Bossaerts, and Spatt (2010), and Vayanos and Woolley (2010) offer a rational explanation for the momentum effect but do not address the volume-based price momentum.

cognitive biases as an explanation for the behavior of financial markets.

2 Model

We consider a pure exchange economy with one risky asset and a riskless asset whose rate of return is normalized to zero. There are $T+1$ periods of trading indexed by $t \in \{0, 1, ..., T\}$. Consumption occurs only at period $T+1$ when the asset's cash flow $\tilde{v_A}$ is distributed. We consider that $\tilde{v_A} = m + \sum_{t=1}^{t=T} s_t + \varepsilon_A$ where m is a constant that represents prior beliefs, s_t is a public signal announced before trading occurs at period t, and ε_A is a random noise realized at date $T+1$. Note that no signal reaches the market before trading at date 0.[9] For simplicity, we assume that there exists a probability measure $\mathbb{P}_A$ under which we have $(s_1, ..., s_T, \varepsilon_A) \overset{law}{=} \mathcal{N}\left(\mathbf{0}, \left(\Sigma^2 \mathrm{Id}_{T+1}\right)\right)$, where $\mathbf{0}$ is the null vector of dimension $T+1$ and Id_{T+1} is the identity matrix of dimension $T+1$. The number of shares of the risky asset is normalized to one. There is a continuum of traders with mass normalized to one. A trader pertains to one of two groups of agents. Arbitrageurs (denoted by A) are in proportion $1-\lambda$, and biased traders (denoted by B) are in proportion λ. Traders are endowed with one unit of the risky asset and no cash.

In order to focus on the informational aspects of financial markets, we assume that traders are risk neutral.[10] Absent market frictions, risk neutrality implies that traders would stand ready to submit infinite demands as long as asset 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 $\frac{c}{2} > 0$.[11] 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. This cost can be viewed as an explicit transaction cost traders have to pay to submit orders or as a proxy

[9] This assumption is made for simplicity and has no influence on the results.

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

[11] 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.

for the imperfect depth of financial markets.[12] This transaction cost creates limits to arbitrage and opens the scope for potential mispricings.[13]

In our model, differences of opinion emerge because all traders do not encode the public signal in the same way. On the one hand, arbitrageurs are perfectly rational in the sense that they are endowed with the actual probability model $\mathbb{P}_A$. On the other hand, biased traders are endowed with a different probability model $\mathbb{P}_B$. When receiving public signals s_t, biased traders actually see σ_t, for every t. Under the probability model $\mathbb{P}_B$, the asset's cash flow is: $\tilde{v}_B = m + \sum_{t=1}^{t=T} \sigma_t + \varepsilon_B$. To incorporate the fact that biased traders believe that they correctly perceive signals s_1,s_2,...,s_T when they in fact observe σ_1,σ_2,...,σ_T, we consider that, under $\mathbb{P}_B$, we have $(\sigma_1, ..., \sigma_T, \varepsilon_B) \overset{law}{=} \mathcal{N}\left(\mathbf{0}, \left(\Sigma^2 \mathrm{Id}_{T+1}\right)\right)$.[14]

To incorporate the confirmation bias in our framework, we assume that biased traders amplify public information when it is consistent with their prior beliefs concerning the final cash flow.[15].

We denote by $\mu_{t-1} = \mathbb{E}^B\left(\tilde{v}_B | \sigma_1, \ldots, \sigma_{t-1}\right)$, the beliefs of biased traders (under probability model $\mathbb{P}_B$) given the information they have received up to date $t-1$. We denote by $\mu^A_{t-1} = \mathbb{E}^A\left(\tilde{v}_A | s_1, \ldots, s_{t-1}\right)$, the beliefs of rational arbitrageurs (under the correct probability model $\mathbb{P}_A$) given the information they have received up to date $t-1$. We consider that a signal s_t is consistent with prior beliefs μ_{t-1} if and only if s_t has the same sign as $\mu_{t-1} - m$, the difference between the conditional and the unconditional

[12]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.

[13]Alternative modeling 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.

[14]This implies that, before receiving the first public signal, the biased traders have a correct understanding of the statistical model underlying the financial market. As stated above, their bias derives from their improper perception of information. Except from that, biased traders maximize their expected utility, update their beliefs using Bayes' rule, and have rational expectations.

[15]The Appendix shows that our results qualitatively hold when biased traders attenuate information which is inconsistent with their prior view. In this case, excess volatility also arise for a large set of parameter values

expectation of the asset cash flow. When biased traders are bullish, that is, if they believe that the expected cash flow is above the unconditional mean m, they over-react to positive signals. Likewise, when traders are bearish, they over-react to negative signals.[16] We can now formally define how the confirmation bias affects biased traders' perception of public information.

Definition 2.1 *Under the appropriate probability measure $\mathbb{P}_A$, the information perceived by biased traders is:*

$$\sigma_t = q s_t \mathbb{1}_{(\mu_{t-1}-m)s_t>0} + s_t \mathbb{1}_{(\mu_{t-1}-m)s_t\leq 0} \tag{2.1}$$

in which $q \geq 1$ measures the severity of the confirmation bias, and $\mathbb{1}_{(.)}$ is the indicator function that takes the value 1 if the condition is satisfied and 0 otherwise.

In our model, the departure from perfect rationality is parameterized by q. If $q = 1$, we are in the perfect rationality case: biased traders correctly perceive the public signal, that is, $\sigma_t = s_t$ for every t. However, when $q > 1$, biased traders are prone to the confirmation bias: they amplify information when it is consistent with their prior view. The following lemma shows that the confirmation bias as modeled here induces biased traders to keep an opinion for too long.

Lemma 2.1 *When $q > 1$, under the appropriate probability measure $\mathbb{P}_A$, biased traders' beliefs are too persistent compared to what rationality would prescribe:*

$$\mathbb{P}^A\left(\mu_{t+x} - m > 0 | \mu_t - m = n > 0\right) > \mathbb{P}^A\left(\mu^A_{t+x} - m > 0 | \mu^A_t - m = n > 0\right) \tag{2.2}$$

in which x is an integer with $2 \leq x \leq T - t$.

[16] This definition is a reminiscence of category-like thinking as modeled for example by Mullainathan (2000). We consider here that biased traders classify an asset into two categories, good assets are expected to deliver more cash flows than initially expected while bad asset are expected to deliver less. We depart from Mullainathan (2000) because we assume that the category to which an asset belongs influences information perception as shown by the literature on psychology. One could generalize our approach to consider more than two categories of assets.

The asset's payoff distributed at date $T+1$ corresponds to the realization of $\tilde{v}_A$ for both arbitrageurs and biased traders because the probability measure $\mathbb{P}_A$ represents the truth. Reality thus strikes at date $T+1$ when the asset distributes cash, a variable that can hardly be misinterpreted. The probability measure $\mathbb{P}_B$ is an abstraction that is used to model biased traders' beliefs. To complete the set up of our model, we endow biased traders with rational expectations. This implies that they are not surprised when observing transaction prices that differ from their conditional expectation of the asset cash flow. We thus posit the following definition.

Definition 2.2 *Under $\mathbb{P}_B$:*

$$s_t = \frac{1}{q}\sigma_t \mathbb{1}_{(\mu_{t-1}-m)\sigma_t>0} + \sigma_t \mathbb{1}_{(\mu_{t-1}-m)\sigma_t\leq 0}$$

This definition indicates that, from biased traders' perspective, arbitrageurs attenuate public information when it is consistent with biased traders' prior views. Let us point out that the set of information generated by the observation of the signals s coincides with the set of information generated by the biased signals σ. In other words, there is no asymmetry of information in our setting. For any randon variables $\tilde{x}$ and $\tilde{y}$, we use $\mathbb{E}_t^A(\tilde{x})$ and $\mathbb{E}_t^B(\tilde{y})$ to denote $\mathbb{E}^A\left(\tilde{x}|s_1,\ldots,s_t\right)$ and $\mathbb{E}^B\left(\tilde{y}|\sigma_1,\ldots,\sigma_t\right)$, respectively.

To summarize, the situation in our model is as follows: there are two types of traders who receive the same public signals. The two types of traders however do not perceive this information in the same way: biased traders amplify signals that are consistent with their prior views. All traders then use Bayes' rule to update their beliefs. The two types of traders know that not all the market participants perceive information in the same manner but both of them believe that their interpretation is the correct one.[17] This

[17]We could consider a model where traders assign a non-zero probability to the event that their probabilistic measure is incorrect and engage in bayesian learning to appropriately choose the measure that is the most plausible. However, since in our setting the asset pays off a cash flow only at the last period, and since agents including biased traders have rational expectations, none of the information that traders observe (that is, market prices) would reduce the likelihood that their probabilistic measure is the correct one. Alternatively, if the asset was distributing several cash flows over time, biased agents could learn the true probabilistic measure. Our model can thus be viewed as focusing on the period during which biased traders have not yet learned what the true probability model is.

creates differences of opinion across traders offering them a rationale to trade. The next section characterizes the trading outcomes of this financial market. We focus on the case in which $T = 3$ because it is rich enough to deliver our main results.[18]

3 Equilibrium Prices, Returns and Volumes

Our model features quadratic transaction costs. As a result, asset demands, d_t^A for arbitrageurs and d_t^B for biased traders, are finite despite risk neutrality and there exists an equilibrium. Standard arguments show that prices in our financial market are weighted averages of arbitrageurs and biased traders' beliefs given their respective information set. They are given in the following proposition.

Proposition 3.1 *At each date t:*

- *The price is:*

$$\begin{aligned} P_t &= (1-\lambda)\,\mathbb{E}_t^A(\tilde{v}_A) + \lambda \mathbb{E}_t^B(\tilde{v}_B) \\ &= m + (1-\lambda)\sum_{i=1}^{i=t} s_i + \lambda \sum_{i=1}^{i=t} \sigma_i. \end{aligned}$$

- *The volume, defined as $(1-\lambda)\,|d_t^A|$ or $\lambda |d_t^B|$, is:*

$$\begin{aligned} V_t &= \frac{\lambda(1-\lambda)}{c} |\mathbb{E}_t^A(\tilde{v}_A) - \mathbb{E}_t^B(\tilde{v}_B)| \\ &= \frac{\lambda(1-\lambda)}{c} |\sum_{i=1}^{i=t}(s_i - \sigma_i)|. \end{aligned}$$

For $t > 1$, the confirmation bias affects biased traders' conditional beliefs, $\mathbb{E}_t^B(\tilde{v}_B)$, and thus in turn it impacts equilibrium variables. To analyze this impact, it is useful to analyze the benchmark case in which all traders are perfectly rational. Endogenous variables 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 $q = 1$. In this case, we have $P_t^* = \mathbb{E}_t^A(\tilde{v}_A) = m + \sum_{i=1}^{i=t} s_i$. Given the structure of the uncertainty in our model, it is straightforward to show that the following proposition holds.

[18] The appendix shows that our results extend to the general case in which $T > 3$.

Proposition 3.2 *When all traders are perfectly rational (that is, when $\lambda = 0$ or $q = 1$), market outcomes are as follows:*

- *The volume is null:*

$$V_t^* = 0, \text{ for all } t$$

- *Expected returns are null. In particular:*

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

- *There is no autocorrelation in short-term price changes:*

$$Cov^A \left(P_{t+k}^* - P_t^*, P_t^* - P_{t-k}^*\right) = 0, \text{ for } t > 0, 1 \leq k \leq \min(t, 3-t)$$

- *There is no autocorrelation in long-term price changes:*

$$Cov^A \left(P_t^* - P_0^*, \tilde{v}_A - P_t^*\right) = 0, \text{ fort} > 0$$

- *The variance of prices reflects the arrival of fundamental information:*

$$Var^A \left(P_t^*\right) = t\Sigma^2, \text{ for all } t$$

When there are no biased traders, prices are equal to the present value of the asset's cash flow conditional on the available information. Public signals are the only factor influencing asset prices. Because we assume that public signals are independent and identically distributed, the variance of prices is just the sum of the public signals' variances. There is no autocorrelation in returns. Finally, trading volume is null because, in the benchmark case, there are no differences of opinion.

We now study how asset prices, returns, and volumes are influenced by the fact that some traders are prone to the confirmation bias. [19] Statistical properties of equilibrium variables are evaluated based on the true probability measure P_A because we take the viewpoint of an econometrician who would observe independent repetitions of our model.

[19] As indicated above, the Appendix shows that our results qualitatively hold when biased traders attenuate information which is inconsistent with their prior view.

3.1 Volume-based price momentum

We start by showing that our model provides a rational for the volume-based price momentum documented by Lee and Swaminathan (2000). They show that price continuation is higher for stocks with past price trends that were associated with high volume. We also show that our model is consistent with the evidence offered by Jegadeesh and Titman (1993) on the momentum effect. The next proposition establishes these results.

Proposition 3.3 *When some traders are prone to the confirmation bias (that is, when $\lambda > 0$ and $q > 1$), market prices are such that:*

- *Future expected returns are higher after higher past price changes, that is:*
$$\frac{\partial \mathbb{E}^A \left(P_3 - P_2 | P_2 - P_1\right)}{\partial \left(P_2 - P_1\right)} > 0.$$
- *Price continuation is stronger after higher past volume, that is:*
$$Cov^A \left(V_2, \left|\mathbb{E}^A \left(P_3 - P_2 | P_2 - P_1\right)\right|\right) > 0.$$

Momentum in our model derives from the fact that a positive (respectively, negative) price change makes it more likely that biased traders are optimistic (respectively, pessimistic) and will thus overreact to future positive (respectively, negative) news. In addition to this effect, the volume-based momentum is driven by the fact that a high past absolute price change increases the expected size of the disagreement between biased traders and rational arbitrageurs, leading to a high past trading volume. Together, these two elements indicate that, in our model, the difference in future expected return between winning and losing assets is higher for assets with a high rather than a low past volume. This is in line with the empirical findings of Lee and Swaminathan (2000).

3.2 Mispricings

We now explicitly derive the best estimate a rational agent can make about future prices (and thus biased traders' beliefs) at date 2 and 3 conditional on the information at date 1. We have the following proposition.

Proposition 3.4 *When some traders are prone to the confirmation bias (that is, when $\lambda > 0$ and $q > 1$), market prices are such that:*

- *When $\mu_1 - m = s_1$ is positive, there is a bubble in the sense that*

$$\mathbb{E}_1^A (P_3) > \mathbb{E}_1^A (P_2) > m + s_1 = \mathbb{E}_1^A (P_2^*) = \mathbb{E}_1^A (P_3^*)$$

- *Alternatively, when $\mu_1 - m = s_1$ is negative, there is a crash in the sense that*

$$\mathbb{E}_1^A (P_3) < \mathbb{E}_1^A (P_2) < m + s_1$$

Proposition 3.4 shows that the presence of biased traders creates mispricings: asset prices differ from the fundamental value as measured by the sum of the public signals. This proposition is a reminiscence of the fact that, as underlined by Rabin and Schrag (1999), first impression matters for agents who are prone to the confirmation bias. When initial news (s_1) make them optimistic (that is, push their beliefs above m), biased traders exacerbate future good news thus pushing prices above fundamental levels. On the contrary, if initial news make them pessimistic, biased traders subsequently push prices below fundamental levels. At initial dates $t \in \{0, 1\}$, biased traders have not formed an opinion yet so the confirmation bias has no bite and prices equal fundamental values: $P_0 = P_0^*$ and $P_1 = P_1^*$.

Furthermore, Proposition 3.4 shows that mispricings are expected to worsen instead of being corrected. We interpret this phenomenon as a price bubble or crash depending on the direction of biased traders' beliefs at date 1. Suppose, for example, that public information at date 1 is positive, $s_1 > 0$. After observing this signal, a rational agent expects the price at date 2 to be higher than the fundamental value and to drift further away from the rational benchmark from date 2 to date 3. This is because biased traders' beliefs are too persistent: for a given optimistic belief (that is, a belief above the initial belief m), biased traders' are likely to be even more optimistic in the future. After a negative public signal at date 1, the logic is the same and a crash develops: a rational agent expects the price at date 2 to be lower than the fundamental value and to drift further downward from date 2 to date 3. As explained above, rational arbitrageurs' trades partly correct mispricings but do not completely eliminate them because of transaction costs.

3.3 Excess volatility

We now study how the confirmation bias affects price volatility. The next proposition compares the volatility of prices when some traders are biased to the volatility of prices in the perfectly rational benchmark.

Proposition 3.5 *At date $t \in \{2, 3\}$, market prices exhibit excess variance:*

$$Var(P_t) = Var(P_t^*) + \Sigma^2 \Theta_t(\lambda, q).$$

The appendix provides the exact expressions for $\Theta_t(\lambda, q)$, and shows that $\Theta_t(0, .) = \Theta_t(., 1) = 0$ and $\Theta_t(\lambda, q) > 0$ if $\lambda > 0$ and $q > 1$.

Proposition 6.5 indicates that excess variance arises because of the confirmation bias. The reason for this result is twofold. On the one hand, biased traders sometimes over-react to public signals. This naturally increases the variance of prices. On the other hand, biased traders put more weight on future positive signals when being optimistic and on future negative signals when being pessimistic. This positive correlation between current views and future misperceptions exacerbates price volatility. The Appendix shows that all our results hold when biased traders underreact to information that is inconsistent with their prior views. In this case also, excess variance arises in our model, even if only for a large set of parameters' values. This indicates that, in our model, excess variance is not an arte-fact of the overreaction assumption.

3.4 Volume

In our model, volume is positively related to the magnitude of differences of opinion. When there are no biased traders (that is, $\lambda = 0$ or $q = 1$), volume is null: $V_t^* = 0$, for all t. Also, since at initial dates $t \in \{0, 1\}$, biased traders have not formed an opinion yet, volume is always null: $V_0 = V_1 = 0$. The next proposition shows that, once biased traders have formed an opinion, volume arises, is positively autocorrelated, and is larger when contemporaneous absolute returns are larger.

Proposition 3.6 *When some traders are prone to the confirmation bias (that is, $\lambda > 0$ and $q > 1$):*

- *Volume is positively autocorrelated:*

$$Cov^A(V_2, V_3) \geq 0$$

- *At $t \in \{2, 3\}$, there is a positive correlation between the trading volume and the magnitude of contemporaneous returns:*

$$Cov^A(V_t, |P_t - P_{t-1}|) \geq 0.$$

Once biased traders have formed an opinion, excess volume arises at $t \in \{2, 3\}$ and tend to be higher when large price change occur. This is not a mechanic result. Indeed, once a difference of opinion has emerged, trading occurs even if there is no news ($s_t = 0$) so that the price remains unchanged ($P_t = P_{t-1}$). This is due to the convex nature of transaction costs: instead of trading a large quantity once, traders prefer to split their orders and trade repeatedly small quantities. This implies that trading is observed even in the absence of any new information arrival. Proposition 3.6 also indicates that there is a positive correlation between returns' magnitude and volume. This is due to the fact that both the magnitude of price changes and trading volume are positively affected by biased traders' beliefs.

3.5 Reversal effect

Our model also displays a reversal effect. When some traders are prone to the confirmation bias (that is, when $\lambda > 0$ and $q > 1$), after a price increase, the price is expected to decrease when the actual cash-flow of the asset is distributed, that is, when reality strikes. We have the following proposition:

Proposition 3.7 *Expected long-term returns, as measured by $\mathbb{E}^A(v_A - P_2 | P_2 - P_1)$, are negative (respectively, positive) after a positive (respectively, negative) past return.*

For example, after a price increase, it is more likely that prices are above the fundamental value. As a result, when the cash-flow v_A is realized, reality strikes thereby revealing the overly optimistic market valuation. This result also holds when biased traders underreact to information that is inconsistent with their prior views for a set of parameters' values identical to the excess volatility case.

3.6 Trading performance

This subsection studies how the trading performance of rational arbitrageurs compares to the one of biased traders. Performance is measured by the expected capital gains.[20] At dates 0 and 1, because all traders have identical endowments and beliefs, there is no trade and thus no capital gains. At date 2, different positions are established by rational arbitrageurs and biased traders. Because they misperceive the public information, biased traders' performance is always lower than rational arbitrageurs' one when the information revealed at date 4 is taken into account, date at which reality strikes and the cash-flow is distributed. The following proposition characterizes the expected capital gains realized from date 2 to date 3, defined as $\mathbb{E}^A\left(\left(1+d_2^X\right)(P_3-P_2)\right)$, with $X \in \{A, B\}$.

Proposition 3.8 *Expected capital gains from date 2 to date 3 are strictly positive for biased traders and strictly negative for rational arbitrageurs.*

In our model, not only rational arbitrageurs have lower expected capital gains than biased traders, but also they incur expected losses as long as the final dividend is not distributed. This is because biased traders' misperception worsens overtime in expectation. Prices are thus expected to drift further and further away from fundamentals, thereby impacting negatively the performance of the rational arbitrageurs. These arbitrageurs only fare better than biased traders when the final dividend of the asset is distributed. At this date, biased traders end up with a negative expected wealth while rational arbitrageurs have a positive one. Proposition 3.8 indicates that rational arbitrageurs' stabilizing behavior is damageable for their performance in the short-run. It is however beneficial in the long-run when the actual cash flow of the asset is distributed. The short-run underperformance is not an issue for rational arbitrageurs in our model because they only care about their expected final wealth. They have a long horizon. The next section studies the trading strategy of a rational agent with a short horizon.

[20]Focusing on expected capital gains abstracts from the impact of transaction costs. Because traders are assumed risk neutral, we do not need to control for risk.

3.7 New empirical predictions

This subsection proposes new empirical predictions based on the autocorrelation in short-term price changes. In the perfectly rational benchmark, there is no autocorrelation in returns because prices are only influenced by independent public signals. The following proposition shows that positive autocorrelation arises in our model because of the confirmation bias.

Proposition 3.9

When some traders are prone to the confirmation bias (that is, $\lambda > 0$ and $q > 1$), price changes exhibit positive autocorrelation, that is,

$$Cov^A (P_1 - P_0, P_2 - P_1) \geq 0.$$

$$Cov_1^A (P_2 - P_1, P_3 - P_2) \geq 0.$$

The autocorrelation is weaker when biased traders' beliefs are more extreme, that is, $Cov_1^A (P_2 - P_1, P_3 - P_2 | \mu_1 - m)$ decreases with $|\mu_1 - m|$.

This autocorrelation however does not imply that a rational observer can predict whether the next return will be positive or negative. Rather, a rational agent predicts that, when biased traders are optimistic, the size of positive price changes is on average larger than the one of negative price changes. Likewise, when biased traders are pessimistic, the size of negative price changes is on average larger than the one of positive price changes.

Proposition 3.9 states that autocorrelation varies with the strength of biased traders' beliefs. The intuition for this result is as follows. Consider for example that biased traders are optimistic. When they are extremely optimistic, it is very unlikely that biased traders will ever become pessimistic in the future. As a result, they will almost always overreact to positive signals and correctly interpret negative ones. Because signals are independent, biased traders' changes in beliefs are uncorrelated. On the contrary, when biased traders' beliefs are less extreme, the next signal can affect their views on the asset's cash flow: they may become pessimistic if the next signal is sufficiently negative. This shift in beliefs is then going to affect the future information distortions. There is thus a positive correlation between the public signal and the future signal distortions. This correlation is maximal when biased traders are neutral because, in this case, their state of mind (and thus future information distortions) are extremely sensitive to public information: a positive (negative) signal turns them into optimistic (pessimistic) which implies that they will over-react to future positive (negative) news.

3.8 Discussion

Our model is, to the best of our knowledge, the first one that is in line with Lee and Swaminathan (2000)'s empirical observation of the volume-based momentum. This is made possible by the fact that our model is a differences of opinion model (and thus can generate volume predictions) in which the source of the disagreement is explicitly modeled (which gives rise to peculiar dynamics in price and volume). In addition to this, our model provides a unified explanation for a variety of empirically documented phenomena and delivers novel empirical predictions.

First, price momentum arises in our model. This is in line with the empirical analysis of Jegadeesh and Titman (1993). In addition, confirmation bias creates autocorrelation in price changes. A novel empirical prediction is that this autocorrelation decreases with the strength of biased traders' beliefs. Along with the volume-based momentum result, this prediction differentiates our theory from other behavioral finance models of momentum such as Barberis, Shleifer, and Vishny (1998), Daniel, Hirshleifer, and Subrahmanyam (1998), or Hong and Stein (1999).

Second, after good initial news, our model shows that a bubble is expected to form. Favorable initial news induce traders to be optimistic concerning the future cash-flow to be distributed by the asset. Biased traders then distort subsequent information and, in expectation, become more and more optimistic. Prices are set above fundamental value and are expected to drift further up. This expected price run-up goes together with an increasing expected volume: after good initial news, a rational agent would not only expect prices to inflate above fundamental values but also trading volume to increase as prices rise. Our model offers a joint characterization of prices and volumes in bubble episodes and thus complements the theory of bubbles based on overconfidence and differences of opinion developped by Scheinkman and Xiong (2003), and Hong, Scheinkman, Xiong (2006). The bubble in our model is expected to crash only when a strike of bad news is strong enough to make biased traders pessimistic. If a bad signal is announced but does not turn biased traders optimism into pessimism, the price of the asset goes down but the bubble is still expected to keep on inflating afterward.

Third, biased traders have positive expected capital gains in the short run and negative ones in the long run. Long run negative returns are consistent with the empirical results of Odean (1999). Moreover, the fact that biased traders trade a lot and incur losses is consistent with the empirical results of Barber and Odean (2000). Our model suggests that confirmation bias can complement overconfidence as an explanation for these phenomena. Rational arbitrageurs are expected to accumulate losses except at the final date at which the dividend is distributed. The performance of rational arbitrageurs is better than the one of biased traders only when reality strikes, that is, when the cash flow is distributed. This result is related to the analysis of De Long, Shleifer, Summers, and Waldmann (1990) in which some irrational investors create risk on the market and end up, in some cases, with a higher expected wealth than rational agents because they pocket in a higher fraction of the risk premium. In our model, in the short-run, biased traders have a better performance than rational arbitrageurs because they pocket in profits from a momentum effect that they themselves create. However, this is not a problem for rational arbitrageurs since they have a long-term horizon: they want to maximize their expected final wealth. If a rational agent had a short-term horizon though, he would adopt a different strategy. This issue is the focus of the next section.

Finally, our model features excessive price volatility and price reversal. These results are consistent with the empirical evidence documented by Shiller (1981), and Mankiw, Romer, and Shapiro (1985) on volatility and De Bondt and Thaler (1985) on reversal. In our model, these phenomena derive from two sources. On the one hand, they are related to the fact that traders who are prone to the confirmation bias may amplify some of the public signals. On the other hand, even when biased traders underreact rather than overreact, these phenomena arise because of the positive correlation between past news and future distortions. Both effects translate into higher price volatility and reversal because of the price pressure biased traders exert on the market.

4 Short-term investments and the Confirmation Bias

Some investors in financial markets are subject to liquidity constraints that prevent them from holding positions for an extended period of time (see the analysis of Shleifer and Vishny (1997)). We thus investigate how an investor with a short horizon would behave when confronted with traders who suffer from the conrmation bias. To do so, we introduce in the model an additional risk neutral rational trader, referred to as a hedge fund, with a negligible mass and a one-period horizon. The negligible mass implies that the hedge fund trading behavior does not affect prices. As a consequence, the pricing results derived above are still valid. The short horizon implies that, at each date, the hedge fund's objective is to maximize next period's return.[21] The question that we aim to address here is whether the fund uses a contrarian or a positive-feedback strategy, whether the fund is fighting or riding bubbles and crashes.

Proposition 4.10 below shows how to set up the optimal short-term trading strategy, that is, the strategy that exploits mispricings when they arise.

Proposition 4.10 *For $t \in \{1, 2\}$, the demand of a short-term trader is:*

$$d_t = \frac{\lambda(q-1)\Sigma}{c\sqrt{2\pi}} \left[2\mathbb{1}_{(\mu_t > m} - 1\right].$$

Proposition 4.10 highlights the fact that one of the crucial dimensions of a short-term strategy in our model is to track the evolution of biased traders' beliefs. Before the final trading date, if biased traders are optimistic (pessimistic), that is $\mu_t - m > 0$ ($\mu_t - m < 0$), the hedge fund establishes a long (short) position because it rationally anticipates that potential price increases (drops) will on average be larger than drops (increases). The hedge fund rides the bubble or the crash. Such a behavior is in sharp contrast with arbitrageurs' behavior in our model. These arbitrageurs indeed sell when the price is above the rational fundamental value and buy when it is below. This is because they hold their inventories until the asset's cash flow is distributed.

[21] All of the following results hold as long as the horizon of the hedge fund does not include the date at which the asset's cash flow is distributed. If it were the case, its behavior would be similar to the one of rational arbitrageurs.

At the final trading date, the hedge fund trades against the mispricing and is thus in line with arbitrageurs' strategy. In addition to this, it is easy to show from Proposition 4.10 that the trading activity and the profit of the hedge fund increase with the intensity of the confirmation bias, and with the proportion of biased traders. These parameters can be estimated thanks to maximum likelihood techniques. The likelihood function is provided in the appendix. Based on Proposition 4.10, we now characterize the type of strategy used by the fund.

Proposition 4.11 *For $t \in \{1, 2\}$, a short-term trader engages in:*

- Positive-feedback trading when the public signal is positive ($s_t > 0$) and biased traders are optimistic ($\mu_t - m > 0$), or when the public signal is negative ($s_t < 0$) and biased traders are pessimistic ($\mu_t - m < 0$);

- Contrarian trading when the public signal is positive ($s_t > 0$) and biased traders are pessimistic ($\mu_t - m < 0$), or when the public signal is negative ($s_t < 0$) and biased traders are optimistic ($\mu_t - m > 0$);

- A short-term trader adopts a positive-feedback trading strategy more often than a contrarian strategy: for a given μ_{t-1}, the probability that the fund uses a positive-feedback strategy equals $\frac{1}{2} + \Phi(\frac{-|\mu_{t-1} - m|}{\Sigma})$ which is greater than $\frac{1}{2}$.

Proposition 4.11 shows that a short-term trader is more likely to follow a positive-feedback strategy and that there are circumstances in which contrarian trading is optimal. These circumstances correspond to cases in which the public signal is inconsistent with biased traders' beliefs but is not strong enough to change their views. For example, a contrarian strategy is optimal when biased traders are optimistic ($\mu_{t-1} - m > 0$) and the public signal is mildly negative ($-(\mu_{t-1} - m) < s_t < 0$).

To illustrate the optimal strategy of a short-term trader, consider that biased traders are optimistic, that is, $\mu_{t-1} - m > 0$ (the logic described here also applies when they are pessimistic). The most recent return is $P_t - P_{t-1} = (1 - \lambda)s_t + \lambda\sigma_t$ which is positive (negative) if s_t is positive (negative). When s_t is positive, biased traders' beliefs remain positive ($\mu_{t-1} - m + \sigma_t > 0$) in which case the fund is buying ($d_t > 0$): upon receiving a good news at date t, the market price increases and the fund engages in positive-feedback trading. When s_t is negative, there are two cases. First, if $s_t < -(\mu_{t-1} - m)$, biased traders' beliefs become negative in which case the fund is selling ($d_t < 0$).

Again, this case corresponds to positive-feedback trading since, after a very bad news at date t, the price declines and the fund is going short. Second, if $-(\mu_{t-1} - m) < s_t < 0$, biased traders' beliefs remain positive despite the bad news in which case the fund is buying ($d_t > 0$). In this case, contrarian trading occurs: the fund is buying after a price decline.

This example can be useful to highlight the difference between a strategy based on the confirmation bias and some popular strategies such as momentum and technical analysis strategies. We define a momentum strategy as a strategy according to which a fund buys after a price increase (that is, if s_t is positive) and sells after a price decline (that is, if s_t is negative). Such a momentum strategy would generate positive expected profits but these profits would be lower than the one of the optimal strategy described above. This is because a momentum strategy does not take into account the cases in which information arrivals do not change the valence of biased traders' beliefs. Consider for example that biased traders are optimistic. A negative information arrives that does not affect biased traders' optimism. In this case, the momentum strategy would issue a selling signal while the optimal strategy issues a buying signal to reflect the fact that biased traders are still likely to overreact to positive news.

We now compare the confirmation-based strategy to technical analysis strategies. We define a technical analysis strategy as a strategy that prescribes buying (selling) as soon as the price path crosses a particular resistance (support). In our model, the level m constitutes a natural candidate for a support or a resistance. Such a technical analysis strategy imperfectly captures the change in biased traders' beliefs. Consider for example, that the price is above the level m but that biased traders are pessimistic. The technical analysis strategy would issue a buying signal while the optimal strategy would issue a selling signal. In addition, the technical analysis strategy would not take into account the autocorrelation in price changes as effectively as the optimal strategy.[22] In a sense, the optimal strategy based on the confirmation bias can be viewed as a combination of the momentum

[22] Another type of technical analysis consists in buying when the price is above a given support and selling when it is below a given resistance. Again, a natural candidate in our model for the support and the resistance would be the level m. This strategy would benefit from the autocorrelation in prices but only imperfectly because it does not perfectly track biased traders' beliefs

and the technical analysis strategies. Because it is based on a correct understanding of price formation in our model, this strategy will outperform the two others by construction.[23] It would be interesting to test whether this is also the case on actual data.

5 Conclusion

This paper proposes a theory of price formation based on the premise that some traders are prone to the confirmation bias. We model this cognitive bias by considering that people prone to the confirmation bias tend to amplify signals that are consistent with their prior views. We show that, in the context of financial markets, this bias provides a rationale for the volume-based price momentum documented by Lee and Swaminathan (2000): the return differential between winning and losing assets is higher for high past volume than for low past volume assets. In addition, we show that excess volatility, momentum and reversals arise in our model. Also, when biased traders are optimistic (pessimistic), the market experiences a bubble (crash) in the sense that prices are above (below) fundamentals, and that the mispricing is expected to worsen in the future. Finally, a new prediction is that autocorrelation in price changes is negatively related to the strengths of biased traders' beliefs: when biased traders' beliefs are more extreme (either more optimistic or more pessimistic), there is less autocorrelation in price changes. This prediction could be the basis for an empirical test of our model. One way to proceed could be to compare price continuation and future price change autocorrelation after different volume episodes: a price trend associated with a decreasing volume indicates that biased traders believes are becoming more in line with fundamentals. This predicts a lower price continuation but a higher autocorrelation in price changes.

We highlight the implications of our model for quantitative investments by deriving the optimal strategy of a short-term trader. We analytically indicate how to structure a portfolio in order to profit from biased traders' misperception. In particular, we show that most of the time, a short-term

[23]If we assume that the hedge fund invests one monetary unit at each date (either long or short), the expect profit of a confirmation-based strategy is $c/2$ which is greater than $(c(2\mathbb{1}_{\mu_t>m}-1)-\frac{c}{2})\mathbb{1}_{s_t>0}+(-c(2\mathbb{1}_{\mu_t>m}-1)-\frac{c}{2})\mathbb{1}_{s_t<0}$, the expected profit of a momentum strategy.

trader adopts a positive feedback strategy. We also characterize market conditions in which a contrarian trading strategy is optimal. These conditions correspond to cases in which the public signal is inconsistent with biased traders' beliefs but is not strong enough to change their views. For example, a contrarian strategy is optimal when biased traders are optimistic and the public signal is mildly negative (that is, negative but not enough to turn biased traders from optimistic to pessimistic). Overall, our model highlights the fact that, to successfully benefit from short-term strategies, traders should track the evolution of biased traders' beliefs that are at the origin of the momentum effect and of the other mispricings. Our model provides an explicit characterization of these biased traders' beliefs and their evolution and may thus be helpful in structuring dynamic trading strategies bound to profit from the momentum effect.

In future work, it would be interesting to estimate the parameters of our model (in particular the proportion of biased traders on the market, , and the magnitude of the bias, q) in order to evaluate the abnormal returns generated by a strategy based on the confirmation bias. Also such estimation could be useful to test the model against the fully rational benchmark and against the overconfidence case where traders overreact to every signal. Finally, our model could be used to study optimal corporate communication. Indeed, firms that are confronted to financial market populated by investors prone to the confirmation bias might have an interest in appropriately choosing the timing and the strength of information releases.

6 Appendix

This appendix provides the analytical expressions that lead to the results stated in our various propositions.

Proof of Lemma 2.1 Let us consider $x = 2$. We will prove that

$$\mathbb{P}^A\left(\mu_{t+2} - m > 0 | \mu_t - m = n > 0\right) > \mathbb{P}^A\left(\mu^A_{t+2} - m > 0 | \mu^A_t - m = n > 0\right),$$

or equivalently for any $n > 0$,

$$\mathbb{P}^A\left(n + \sigma_{t+1} + \sigma_{t+2} > 0\right) > \mathbb{P}^A\left(n + s_{t+1} + s_{t+2} > 0\right).$$

From the equality,

$$n+\sigma_{t+1}+\sigma_{t+2} = n+s_{t+1}+s_{t+2}+(q-1)s_{t+1}\mathbb{1}_{s_{t+1}>0}+(q-1)s_{t+1}\mathbb{1}_{(n+\sigma_{t=1})s_{t+2}>0},$$

we deduce that the set $\{n+s_{t+1}+s_{t+2} > 0\} \subset \{n+\sigma_{t+1}+\sigma_{t+2} > 0\}$ because $n+\sigma_{t+1}+\sigma_{t+2} \geq n+s_{t+1}+s_{t+2}$ almost surely. To show that the inequality is strict, let us consider a path where $s_{t+1} > 0$ and $-(n + qs_{t+1}) < s_{t+2} < -(n + s_{t+1})$.

Proof of Proposition 3.1

At date $T+1$, the final wealth of trader j, W_j, is:

$$W_j = \sum_{t=0}^{t=T}\left[d^j_t\left(v_X - P_t\right) - \frac{c}{2}\left(d^j_t\right)^2\right] + v_X,$$

with X=A if j is an arbitrageur and X=B if j is a biased trader. Since traders only consume at the last date $T+1$, their objective is to maximize their expected final wealth conditional on their information. To solve for the optimal demands, we start by solving the program of trader j at period T:

$$\max_{d^j_T}\mathbb{E}^X_T\left(\sum_{t=0}^{t=T-1}\left[d^j_t\left(v_X - P_t\right) - \frac{c}{2}\left(d^j_t\right)^2\right] + d^j_T\left(v_X - P_T\right) - \frac{c}{2}\left(d^j_T\right)^2 + v_X\right),$$

subject to the budget constraint of trader j at date T.

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

$$d_T^j = \frac{\mathbb{E}_T^X(v) - P_T}{c}.$$

Proceeding backward, we obtain that:

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

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

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

where 0 corresponds to the fact that no new share is being issued on the market. From this market clearing condition, we derive the following pricing equation:

$$\begin{aligned} P_t &= (1-\lambda)\,\mathbb{E}_t^A(\tilde{v}_A) + \lambda \mathbb{E}_t^B(\tilde{v}_B) \\ &= m + (1-\lambda)\sum_{i=1}^{i=t} s_i + \lambda \sum_{i=1}^{i=t} \sigma_i. \end{aligned} \tag{6.3}$$

Given these results, deriving the expression for the volume is straightforward.

Proof of Proposition 3.2

This proposition derives from the fact that, in the rational benchmark case, prices constitute a random walk. In this case, the volume is null because all agents hold the same belief regarding the value of the future cash flow.

Proof of Proposition 3.3

Momentum is related to the fact that:

$$\mathbb{E}^A(P_3 - P_2 | P_2 - P_1) = \lambda(q-1)\frac{\Sigma}{\sqrt{2\pi}}(2\Phi(P_2 - P_1)).$$

$\Phi(.)$ represents the cumulative distribution function of a standardized normal random variable. It is straightforward to see that this expression increases with $P_2 - P_1$.

Volume-based price momentum derives from the fact that:

$$Cov^A\left(V_2, \left|\mathbb{E}^A\left(P_3 - P_2 | P_2 - P_1\right)\right|\right) = \frac{\lambda^2 \left(1-\lambda\right)\left(q-1\right)^2}{c} \frac{\Sigma}{\sqrt{2\pi}} \mathbb{E}^A\left(\mathbb{1}_{s_1>0}|s_2|\left(1 - 2\Phi\left(-ks_2\right)\right)\right) > 0.$$

where $k = \lambda(q-1) + 1$.

Proof of Proposition 3.4
The mispricing at date t$\in \{2, 3\}$ *is* :

$$P_t - P_t^* = \lambda \sum_{i=1}^{i=t} s_i \left((q-1)\mathbb{1}_{(\mu_{i-1}-m)s_i>0}\right).$$

When $\mu_1 - m = s_1$ is positive, there is a bubble in the sense that:

$$\mathbb{E}_1^A\left(P_3\right) = \mathbb{E}_1^A\left(P_2\right) + \lambda \frac{\Sigma}{\sqrt{2\pi}}(q-1)(1-\Phi(-\frac{s_1}{\Sigma})) > \mathbb{E}_1^A\left(P_2\right) = m + s_1 + \lambda \frac{\Sigma}{\sqrt{2\pi}}(q-1) > m + s_1.$$

When $\mu_1 - m = s_1$ is negative, there is a crash in the sense that

$$\mathbb{E}_1^A\left(P_3\right) = \mathbb{E}_1^A\left(P_2\right) - \lambda \frac{\Sigma}{\sqrt{2\pi}}(q-1)\Phi(-\frac{s_1}{\Sigma}) < \mathbb{E}_1^A\left(P_2\right) = m + s_1 - \lambda \frac{\Sigma}{\sqrt{2\pi}}(q-1) < m + s_1.$$

In order to prove the propositions relative to excess variance, volume-based momentum and reversal effect, we need to prove technical results concerning the distribution function of the biased signal.

Result .1 *For every t, the random variable $\mu_t - m$ has a symmetric distribution.*

Proof of Result .1 We will make a proof by induction. The result is obvious for $t = 1$ because $\mu_1 - m = s_1$.
Let us assume that $\mu_{t-1} - m$ has a symmetric distribution. Let us first show that σ_t has a symmetric distribution. Take $x > 0$, we have

$$\mathbb{P}(\sigma_t \leq x) = \mathbb{P}(\mu_{t-1} - m > 0)\mathbb{P}(s_3 \leq \frac{x}{q}) + \mathbb{P}(\mu_{t-1} - m < 0)\mathbb{P}(s_3 \leq x).$$

On the other hand,

$$\mathbb{P}(\sigma_t \geq -x) = \mathbb{P}(\mu_{t-1} - m > 0)\mathbb{P}(s_3 \geq -x) + \mathbb{P}(\mu_{t-1} - m < 0)\mathbb{P}(s_3 \geq -\frac{x}{q}).$$

We conclude using that both s_3 and $\mu_{t-1} - m$ have a symmetric distribution.
Now, by denoting F_{t-1} the distribution function of $\mu_{t-1} - m$,

$$\begin{aligned}
\mathbb{P}(\mu_t - m \leq x) &= \mathbb{P}(\mu_{t-1} - m + \sigma_t \leq x) \\
&= \int_{\mathbb{R}} \mathbb{P}(\sigma_t \leq x - y) F_{t-1}(dy) \\
&= \int_{\mathbb{R}} \mathbb{P}(\sigma_t \geq y - x) F_{t-1}(dy) \text{ since } \sigma_t \text{ has a symmetric distribution} \\
&= \int_{\mathbb{R}} \mathbb{P}(\sigma_t \geq -z - x) F_{t-1}(dz) \text{ since } \mu_{t-1} - m \text{ has a symmetric distribution} \\
&= \mathbb{P}(\mu_t - m \geq -x)
\end{aligned}$$

Result .2 *The second order moments of the signals are*

1. $$Var(\sigma_2) = Var(\sigma_3) = (1 + q^2)\Sigma^2.$$

2. $$cov^A(\sigma_2, s_1) = (q - 1)\frac{\Sigma^2}{\pi}.$$

3. $$cov^A(\sigma_2, s_2) = (q + 1)\frac{\Sigma^2}{2}.$$

4. $$cov^A(\sigma_3, s_1) = (q - 1)\frac{\Sigma^2}{\sqrt{\pi}}.$$

5. $$cov^A(\sigma_3, s_2) = (q - 1)\frac{\Sigma^2}{\pi}.$$

6. $$cov^A(\sigma_3, s_2) = (q - 1)\frac{\Sigma^2}{\pi}\left(q - 1 + \frac{1}{\sqrt{2}}\right).$$

Proof of Result .2 Using the independence of the signals s_i and Result .1 the first three equalities are straightforward. We focus on the last three.

$$\begin{aligned}
\mathrm{cov}^A(\sigma_3, s_1) &= (q-1)\mathbb{E}^A(s_1 s_3 \mathbb{1}_{(s_1+\sigma_2)s_3 \geq 0}) \\
&= (q-1)\frac{\Sigma}{\sqrt{\pi}}\mathbb{E}^A(s_1(\mathbb{1}_{s_1+\sigma_2\geq 0} - \mathbb{1}_{s_1+\sigma_2\leq 0})) \\
&= 2(q-1)\frac{\Sigma}{\sqrt{\pi}}\mathbb{E}^A(s_1(\mathbb{1}_{s_1+\sigma_2\geq 0}))
\end{aligned}$$

Using,

$$\mathbb{P}_1^A(s_1+\sigma_2 \geq 0) = \mathcal{N}\left(\frac{s_1}{\Sigma}\right),$$

we get

$$\mathrm{cov}^A(\sigma_3, s_1) = 2(q-1)\frac{\Sigma^2}{\sqrt{2\pi}}\mathbb{E}^A\left(s_1\mathcal{N}\left(\frac{s_1}{\Sigma}\right)\right).$$

By integration by parts, we have $\mathbb{E}^A\left(s_1\mathcal{N}\left(\frac{s_1}{\Sigma}\right)\right) = \frac{\Sigma}{\sqrt{2}}$ and thus the result. On the same manner, we have

$$\mathrm{cov}^A(\sigma_3, s_2) = 2(q-1)\frac{\Sigma}{\sqrt{\pi}}\mathbb{E}^A(s_2(\mathbb{1}_{s_1+\sigma_2\geq 0})).$$

Using $\mathbb{E}_1^A(s_2(\mathbb{1}_{s_1+\sigma_2\geq 0})) = \Sigma e^{-\frac{s_1^2}{2\Sigma^2}}$ and $\mathbb{E}^A(e^{-\frac{s_1^2}{2\Sigma^2}}) = \frac{1}{\sqrt{2}}$, we get the result. On the other hand, using Result .1

$$\mathrm{cov}_A(\sigma_2, \sigma_3) = \mathbb{E}^A(\sigma_2\sigma_3).$$

We have,

$$\begin{aligned}
\mathbb{E}^A(\sigma_3\sigma_2) &= \mathbb{E}^A(\mathbb{E}_2^A(\sigma_3)\sigma_2) \\
&= (q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(\sigma_2(2\mathbb{1}_{\{\mu_2\geq m\}} - 1)) \\
&= 2(q-1)\frac{\Sigma}{\sqrt{2\pi}}\left(\mathbb{E}^A(\sigma_2(\mathbb{1}_{\{s_1+\sigma_2\geq 0\}}\right).
\end{aligned}$$

But,

$$\begin{aligned}
\mathbb{E}^A(\sigma_2\mathbb{1}_{\{s_1+\sigma_2\geq 0\}}) &= (q-1)\mathbb{E}^A(s_2\mathbb{1}_{\{s_1 s_2\geq 0\}}\mathbb{1}_{\{s_1+\sigma_2\geq 0\}}) + \mathbb{E}^A\left(s_2\mathbb{1}_{\{s_1+\sigma_2\geq 0\}}\right) \\
&= (q-1)\mathbb{E}^A(s_2\mathbb{1}_{\{s_1\geq 0\}}\mathbb{1}_{\{s_2\geq 0\}}) + \mathbb{E}^A\left(s_2\mathbb{1}_{\{s_1+\sigma_2\geq 0\}}\right) \\
&= (q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{P}_A(s_1>0) + \frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(e^{-\frac{(s_1)^2}{2\Sigma^2}}) \\
&= \frac{1}{2}(q-1)\frac{\Sigma}{\sqrt{2\pi}} + \frac{\Sigma}{2\sqrt{\pi}}.
\end{aligned}$$

Proof of Proposition 6.5
Using Result .2, we obtain

$$Var^A(P_2) = Var^A(P_2^*) + \Sigma^2 \Theta_2(\lambda, q).$$

where

$$\Theta_2(\lambda, q) = \lambda^2 \frac{(q-1)^2}{2} + 2\lambda \left(\frac{q-1}{\pi} + \frac{q-1}{2} \right) > 0.$$

Likewise, the variance of P_3 is:

$$Var^A(P_3) = Var^A(P_3^*) + \Sigma^2 \Theta_3(\lambda, q),$$

where

$$\begin{aligned} \Theta_3(\lambda, q) \;\; &= \;\; \frac{\lambda^2}{2} (q-1)^2 + \frac{4\lambda^2(q-1)}{2\pi} \left(\frac{q-1}{2} + \sqrt{2} \right) \\ &- \;\; \frac{4\lambda^2(q-1)}{2\pi} \sqrt{2} + 2\lambda(q-1) + \frac{4\lambda(q-1)}{2\pi} \sqrt{2} \\ &+ \;\; \frac{4\lambda^2(q-1)}{2\pi} \sqrt{2} + \frac{2\lambda(q-1)}{\pi} > 0. \end{aligned}$$

Our model exhibits excess variance both at dates 2 and 3.

When biased traders underreact to information, the biased signal is modeled as

$$\sigma_t = s_t \mathbb{1}_{(\mu_{t-1}-m)s_t > 0} + r s_t \mathbb{1}_{(\mu_{t-1}-m)s_t \leq 0} \tag{6.4}$$

with $r \leq 1$. All the previous formula remains valid with $1 - r$ in the place of $q - 1$. Therefore, excess variance arises when confirmation bias implies underreaction of biased traders.

Proof of Proposition 3.6
Volume at date 2 is:

$$V_2 = \frac{\lambda (1-\lambda)(q-1)}{c} \mathbb{1}_{s_1 s_2 > 0} |s_2|.$$

Volume at date 3 is:

$$V_3 = \frac{\lambda (1-\lambda)(q-1)}{c} \left| \mathbb{1}_{s_1 s_2 > 0} s_2 + \mathbb{1}_{(\mu_2 - m)s_3 > 0} s_3 \right|.$$

Because $V_2 = 0$ if $s_1 s_2 \leq 0$ then we will compute volume at date 3 when $s_1 s_2 \geq 0$. Hence,

$$\frac{\lambda (1-\lambda)(q-1)}{c} \left| \mathbb{1}_{s_1>0} \mathbb{1}_{s_2>0} (s_2 + \mathbb{1}_{s_3>0} s_3) - \mathbb{1}_{s_1<0} \mathbb{1}_{s_2<0} (s_2 + \mathbb{1}_{s_3<0} s_3) \right|.$$

As a result,

$$Cov^A(V_2, V_3) = \frac{\lambda^2 (1-\lambda)^2 (q-1)^2}{c^2} \left(\frac{\Sigma^2}{2} + \frac{1}{2} \frac{\Sigma^2}{2\pi} \right) > 0.$$

On the other hand,

$$Cov^A (V_2, |P_2 - P_1|) = \frac{\lambda (1-\lambda)(q-1)}{c} \left((\lambda q + (1-\lambda)) \left(\frac{\Sigma^2}{2} - \frac{\Sigma^2}{4\pi} \right) - \frac{\Sigma^2}{4\pi} \right) > 0.$$

The same computations apply at date 3. We conclude that volume is positively correlated with the magnitude of price changes.

Proof of Proposition 3.7

Expected long-term returns conditional on past performance are:

$$\mathbb{E}^A (v_A - P_2 | P_2 - P_1) = -\frac{\lambda(q-1)}{2(\lambda(q-1)+1)} (P_2 - P_1).$$

These returns are negative after a price increase and positive after a price decrease.

Proof of Proposition 3.8

Expected capital gains from date 2 to date 3 for a biased trader are:

$$\mathbb{E}^A \left(\left(1 + d_2^B\right) (P_3 - P_2) \right) = \frac{(1-\lambda)}{c} (q-1)^2 \frac{\Sigma^2}{\pi} > 0.$$

Expected capital gains from date 2 to date 3 for a rational arbitrageurs are:

$$\mathbb{E}^A \left(\left(1 + d_2^A\right) (P_3 - P_2) \right) - \frac{\lambda}{c} (q-1)^2 \frac{\Sigma^2}{\pi} < 0.$$

Proof of Proposition 3.9

Our new predictions regards various autocorrelations in price changes provided below.

$$Cov^A (P_1 - P_0, P_2 - P_1) = \lambda (q-1) \frac{\Sigma^2}{\pi} > 0.$$

Also, we have:

$$\begin{aligned}\mathrm{cov}_A(P_3 - P_2, P_2 - P_1) &= \mathrm{cov}_A(\lambda\sigma_3 + (1-\lambda)s_3, \lambda\sigma_2 + (1-\lambda)s_2) \\ &= \lambda^2 \mathrm{cov}_A(\sigma_2, \sigma_3) + \lambda(1-\lambda)\mathrm{cov}_A(s_2, \sigma_3)\end{aligned}$$

which is positive according to Result .2.
Furthermore, conditionally on the information set at time 1, we prove in the same manner that

$$\mathrm{cov}_1^A(P_3 - P_2, P_2 - P_1) = f(|s_1|),$$

where f is defined on $(0, +\infty)$ by

$$f(x) = \lambda^2(q-1)\frac{\Sigma^2}{\pi}e^{-\frac{x^2}{2\Sigma^2}} + \lambda(1-\lambda)\frac{\Sigma^2}{2\pi}\left((q-1) + (q-1)^2\mathcal{N}\left(\frac{-x}{\Sigma}\right)\right).$$

Obviously, f is a decreasing function. Moreover, $\lim_{x\to+\infty} f(x) = 0$ and f attains its maximum at $x = 0$.

Case in which $T > 3$

This part of the Appendix extends our model to the case in which $T > 3$ and in which biased traders only amplify public signals that are consistent with their prior views ($q > 1$).

In order to understand how future prices are related to overreacting biased traders' current beliefs, we introduce an operator $\mathcal{M}$ acting on bounded function defined as:

$$(\mathcal{M}f)(\mu_t) = \mathbb{E}_t^A\left[f(\mu_{t+1})\right],$$

with

$$\begin{aligned}(\mathcal{M}f)(x) &= \left[\int_0^\infty f(x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x>m} \\ &+ \left[\int_0^\infty f(x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x<m}.\end{aligned}$$

Applied to the function $f_m(x) = \mathbb{1}_{\{x>m\}}$, the function

$$(\mathcal{M}f_m)(x) = \mathbb{P}_A(\mu_{t+1} > m|\mu_t = x)$$

corresponds to a rational agent expectation of biased traders' future beliefs given that the current beliefs of these biased traders equal x.

According to bounded dominated convergence, the belief operator maps the bounded function in the set of continuous functions on the open set $]-\infty, 0[\cup]0, +\infty[$. Moreover, it has several properties that will govern the behavior of equilibrium prices. These properties are summarized in the following lemma.

Lemma 6.2 *We have,*

- *If f is increasing then $(\mathcal{M}f)$ is an increasing function.*
- *Let f be a bounded function satisfying $f(x) + f(-x) = 1$ for every positive real x. Then, we have for every positive real x*

$$(\mathcal{M}f(x)) + (\mathcal{M}f)(-x) = 1.$$

A nice consequence of Lemma 6.2 is that $(Mf)(x) \geq \frac{1}{2}$ for $x > m$. Applied to the function $f_m(x)$ defined earlier, this consequence indicates that, if biased traders are currently optimistic, it is more likely that they will also be optimistic in the future (and vice versa when they are pessimistic). This persistence in biased traders' beliefs is at the core of our results on the properties of prices in our model.

Proof of Lemma 6.2

- Let f be an increasing function. By definition, we have

$$\begin{aligned}
(\mathcal{M}f)(x) &= \mathbb{E}^A\left(f(\mu_{t+1}) | \mu_t = x\right) \\
&= \mathbb{E}_t^A\left(f(x + \sigma_{t+1})\right) \\
&= \mathbb{E}^A\left(f(x + q s_{t+1} \mathbb{1}_{(x-m)s_{t+1} \geq 0} + s_{t+1} \mathbb{1}_{(x-m)s_{t+1} \leq 0})\right).
\end{aligned}$$

Let us consider the function $\theta_y(x) = x + qy\mathbb{1}_{(x-m)y \geq 0} + y\mathbb{1}_{(x-m)y \leq 0}$. We will prove that the function θ_y is increasing for every real numbers y which will imply that the function $(\mathcal{M}f)$ is increasing. To see this, we first assume that $y \geq 0$. We have,

$$\theta_y(x) = (x + y)\mathbb{1}_{\{x<m\}} + (x + qy)\mathbb{1}_{\{x>m\}}.$$

Therefore, θ_y is an increasing function if its jump at m is positive. But the jump at m is $(q-1)y$ which is positive since $q>1$ and $y>0$.
If we assume $y \leq 0$, we have

$$\theta_y(x) = (x+qy)\mathbb{1}_{\{x<m\}} + (x+y)\mathbb{1}_{\{x>m\}}.$$

The jump at m is $(1-q)y$ which is positive for $y \leq 0$.

- Take $x > m$ and a bounded function f such that $f(x) + f(-x) = 1$. We have

$$\begin{aligned}
& (\mathcal{M})f(x) + (\mathcal{M})f(-x) \\
= \; & (\mathcal{M}f)(x) + \left[\int_0^\infty f(-x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(-x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x<m} \\
+ \; & \left[\int_0^\infty f(-x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(-x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x>m} \\
= \; & (\mathcal{M})f(x) + \left[\int_{-\infty}^0 f(-x-qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_0^\infty f(-x-y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x<m} \\
+ \; & \left[\int_{-\infty}^0 f(-x-y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_0^\infty f(-x-qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x>m} \\
= \; & \left[\int_0^\infty f(x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x>m} \\
+ \; & \left[\int_0^\infty f(x+y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_{-\infty}^0 f(x+qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x<m} \\
+ \; & \left[\int_{-\infty}^0 f(-x-qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_0^\infty f(-x-y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x<m} \\
+ \; & \left[\int_{-\infty}^0 f(-x-y)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}} + \int_0^\infty f(-x-qy)e^{-\frac{y^2}{2\Sigma^2}}\frac{dy}{\sqrt{2\pi\Sigma^2}}\right]\mathbb{1}_{x>m} \\
= \; & 1.
\end{aligned}$$

□

Before proceeding to prove that there is excess variance in our model, we state a technical lemma claiming that overreacting biased traders' beliefs at time t are unconditionally positively correlated with the true signal s_t.

Lemma 6.3 *Let f be a bounded increasing function. We have for every t,*

$$\mathbb{E}^A(s_t f(\mu_t)) \geq 0.$$

Proof of Lemma 6.3: Let us denote F_t the distribution function of μ_t. We have

$$\begin{aligned}
\mathbb{E}^A(s_t f(\mu_t)) &= \mathbb{E}^A(s_t f(\mu_{t-1}+\sigma_t)) \\
&= \int_{\mathbb{R}} \mathbb{E}^A\left(s_t \mathbb{1}_{\{(x-m)s_t\geq 0\}} f(x+qs_t) + s_t \mathbb{1}_{\{(x-m)s_t\leq 0\}} f(x+s_t)\right) dF_{t-1}(x) \\
&= \int_{\mathbb{R}} \mathbb{E}^A\left(s_t \mathbb{1}_{\{(x-m)s_t\geq 0\}} (f(x+qs_t) - f(x+s_t))\right) dF_{t-1}(x).
\end{aligned}$$

We conclude by noting that the function $s \to s(f(x+qs) - f(x+s))$ is positive for any increasing function f. $\square$

Now, we come back to the proof of the excess variance in our model. To do this, we express P_t in terms of P_t^* using the definition of the biased signals σ_t,

$$\begin{aligned}
P_t &= P_t^* + \lambda \sum_{i=1}^{t} (\sigma_i - s_i) \\
&= P_t^* + \lambda(q-1) \sum_{i=1}^{t} s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i\geq 0\}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\text{(6.5)} \qquad \mathrm{var}_A(P_t) &= \mathrm{var}_A(P_t^*) + \\
&+ 2\lambda(q-1)\mathrm{cov}_A\left(P_t^*, \sum_{i=1}^{t} s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i\geq 0\}}\right) \\
&+ \lambda^2(q-1)^2 \mathrm{var}_A\left(\sum_{i=1}^{t} s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i\geq 0\}}\right).
\end{aligned}$$

We focus on the second term and show that it is positive. Because $\mathbb{E}_A(P_t^*) =$

0, we have

$$\begin{aligned}
\mathrm{cov}_A\left(P_t^*, \sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) &= \mathbb{E}^A\left[\left(\sum_{i=1}^t s_i\right)\left(\sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right)\right] \\
&= \sum_{i=1}^t \mathbb{E}^A\left(s_i^2 \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) \\
&+ 2\sum_{1\leq i<j\leq t} \mathbb{E}^A\left(s_i s_j \mathbb{1}_{\{(\mu_{j-1}-m)s_j \geq 0\}}\right).
\end{aligned}$$

Now, consider $i < j$. We have,

$$\begin{aligned}
\mathbb{E}^A\left(s_i s_j \mathbb{1}_{\{(\mu_{j-1}-m)s_j \geq 0\}}\right) &= \mathbb{E}^A\left(s_i \mathbb{E}_i^A(s_j \mathbb{1}_{\{(\mu_{j-1}-m)s_j \geq 0\}})\right) \\
&= 2(q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(s_i \mathbb{P}_i(\mu_{j-1} > m)).
\end{aligned}$$

Let us remind that $f_m(x) = \mathbb{1}_{\{x>m\}}$. Note that f_m is a bounded increasing function. We have

$$\begin{aligned}
\mathbb{P}_i(\mu_{j-1} > m) &= \mathbb{E}_i^A(f_m(\mu_{j-1})) \\
&= (\mathcal{M}^{j-i-1} f_m)(\mu_i).
\end{aligned}$$

Finally, we obtain

$$\begin{aligned}
\text{(6.6)} \qquad \mathrm{cov}_A\left(P_t^*, \sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) &= \sum_{i=1}^t \mathbb{E}^A\left(s_i^2 \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) \\
+ \ 4(q-1)\frac{\Sigma}{\sqrt{2\pi}}\sum_{1\leq i<j\leq t} \mathbb{E}^A(s_i(\mathcal{M}^{j-i-1} f_m)(\mu_i)).
\end{aligned}$$

We conclude by applying Lemma 6.3 since $\mathcal{M}^{j-i-1} f_m$ is a bounded increasing function.

In order to show the momentum effect, we focus on the case $k = 1$ without loss of generality. Let us take $t \in \{0, \ldots, T\}$, we have

$$\begin{aligned}
\mathrm{cov}_A(P_{t+1} - P_t, P_t - P_{t-1}) &= \mathrm{cov}_A(\lambda\sigma_{t+1} + (1-\lambda)s_{t+1}, \lambda\sigma_t + (1-\lambda)s_t) \\
&= \lambda^2 \mathrm{cov}_A(\sigma_{t+1}, \sigma_t) + \lambda(1-\lambda)\mathrm{cov}_A(\sigma_{t+1}, s_t).
\end{aligned}$$

We will prove that the two terms of the right-hand side are positive. Indeed, we have

$$\begin{aligned}
\mathrm{cov}_A\left(\sigma_{t+1}, s_t\right) &= \mathbb{E}^A\left(\sigma_{t+1} s_t\right) \\
&= q\mathbb{E}^A\left(s_t s_{t+1} 1\!\!1_{\{(\mu_t - m)s_{t+1} \geq 0\}}\right) + \mathbb{E}^A\left(s_t s_{t+1} 1\!\!1_{\{(\mu_t - m)s_{t+1} < 0\}}\right) \\
&= 2(q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_t \geq m\}}\right).
\end{aligned}$$

Now

$$\begin{aligned}
\mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_t \geq m\}}\right) &= \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_{t-1} + \sigma_t \geq m\}}\right) \\
&= \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_{t-1} + q s_t \geq m\}} 1\!\!1_{\{(\mu_{t-1} - m)s_t \geq 0\}}\right) + \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_{t-1} + s_t \geq m\}} 1\!\!1_{\{(\mu_{t-1} - m)s_t \leq 0\}}\right) \\
&= \mathbb{E}^A\left(s_t 1\!\!1_{s_t \geq 0\}} 1\!\!1_{\{\mu_{t-1} \geq 0\}}\right) + \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_{t-1} + s_t \geq m\}} 1\!\!1_{\{\mu_{t-1} \leq m\}}\right) \\
&+ \mathbb{E}^A\left(s_t 1\!\!1_{\{m - \mu_{t-1} < s_t < 0\}} 1\!\!1_{\{\mu_{t-1} \geq m\}}\right) \\
&= \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_{t-1} + s_t \geq m\}}\right) \\
&= \frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(e^{-\frac{(\mu_{t-1} - m)^2}{2\Sigma^2}}) \\
&\geq 0.
\end{aligned}$$

Thus, $\mathrm{cov}_A\left(\sigma_{t+1}, s_t\right) \geq 0$. On the other hand,

$$\mathrm{cov}_A\left(\sigma_{t+1}, \sigma_t\right) = \mathbb{E}^A(\sigma_{t+1}\sigma_t) - \mathbb{E}^A(\sigma_{t+1})\mathbb{E}^A(\sigma_t).$$

We have, for every t,

$$\begin{aligned}
\mathbb{E}^A(\sigma_t) &= (q-1)\mathbb{E}^A(s_t 1\!\!1_{\{(\mu_{t-1} - m)s_t \geq 0\}}) \\
&= (q-1)\frac{\Sigma}{\sqrt{2\pi}}(2\mathbb{P}_A(\mu_{t-1} > m) - 1).
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}^A(\sigma_{t+1}\sigma_t) &= \mathbb{E}^A(\mathbb{E}_t^A(\sigma_{t+1})\sigma_t) \\
&= (q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(\sigma_t(2 1\!\!1_{\{\mu_t \geq m\}} - 1))
\end{aligned}$$

But,

$$\begin{aligned}
\mathbb{E}^A(\sigma_t 1\!\!1_{\{\mu_t \geq m\}}) &= (q-1)\mathbb{E}^A(s_t 1\!\!1_{\{(\mu_{t-1} - m)s_t \geq 0\}} 1\!\!1_{\{\mu_t \geq m\}}) + \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_t \geq m\}}\right) \\
&= (q-1)\mathbb{E}^A(s_t 1\!\!1_{\{\mu_{t-1} \geq m\}} 1\!\!1_{\{s_t \geq 0\}}) + \mathbb{E}^A\left(s_t 1\!\!1_{\{\mu_t \geq m\}}\right) \\
&= (q-1)\frac{\Sigma}{\sqrt{2\pi}}\mathbb{P}_A(\mu_{t-1} > m) + \frac{\Sigma}{\sqrt{2\pi}}\mathbb{E}^A(e^{-\frac{(\mu_{t-1} - m)^2}{2\Sigma^2}}).
\end{aligned}$$

Finally,

$$\text{cov}_A(\sigma_t, \sigma_{t+1}) = (q-1)\frac{\Sigma^2}{2\pi}\left(2\mathbb{E}(e^{-\frac{(\mu_{t-1}-m)^2}{2\Sigma^2}}) + (q-1)\left(1 - (2\mathbb{P}_A(\mu_{t-1} > m) - 1)(2\mathbb{P}_A(\mu_t > m) - 1)\right)\right).$$

The latter quantity is always positive since

$$|(2\mathbb{P}(\mu_{t-1} > m) - 1)(2\mathbb{P}(\mu_t > m) - 1)| \leq 1.$$

Furthermore, conditionally to the information set at time $t-1$, computations above give that

$$\text{cov}^A_{t-1}(P_{t+1} - P_t, P_t - P_{t-1}) = f(|\mu_{t-1} - m|),$$

where f is defined on $(0, +\infty)$ by

$$f(x) = \lambda(q-1)\frac{\Sigma^2}{\pi}e^{-\frac{x^2}{2\Sigma^2}} + \lambda^2(q-1)^2\frac{\Sigma^2}{\pi}\mathcal{N}\left(\frac{-x}{\Sigma}\right).$$

Obviously, f is a decreasing function. Moreover, $\lim\limits_{x\to+\infty} f(x) = 0$ and f attains its maximum at $x = 0$.

We finally focus on the price reversal effect as measured by the number

$$\text{cov}_A(P_t - P_0, \tilde{v}_A - P_t) = \text{cov}_A(P_t, \tilde{v}_A - P_t).$$

According to equation (6.5), we have

$$\begin{aligned}
\text{var}_A(P_t) \quad &= \quad \text{var}_A(P_t^*) + 2\lambda(q-1)\text{cov}_A\left(P_t^*, \sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) \\
&+ \quad \lambda^2(q-1)^2\text{var}_A\left(\sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) \\
&\geq \quad \text{var}_A(P_t^*) + \lambda(q-1)\text{cov}_A\left(P_t^*, \sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right) \\
&= \quad \text{cov}_A\left(\tilde{v}_A, P_t\right)
\end{aligned}$$

because $\text{cov}_A\left(P_t^*, \sum_{i=1}^t s_i \mathbb{1}_{\{(\mu_{i-1}-m)s_i \geq 0\}}\right)$ is positive according to Equation (6.6).

To derive the results on mispricings, we start by recalling their analytical expression, as measured by the difference $P_t - P_t^*$. We have, for every t,

$$\begin{aligned} P_t - P_t^* &= \lambda \sum_{i=1}^{t} (\sigma_i - s_i) \\ &= \lambda(q-1) \sum_{i=1}^{t} s_i \mathbb{1}_{s_i(\mu_{i-1}-m)>0} \end{aligned}$$

Therefore, for $0 \leq k \leq l$,

$$P_{t+l} - P_{t+l}^* = P_{t+k} - P_{t+k}^* + \lambda(q-1) \sum_{i=k+1}^{l} s_{t+i} \mathbb{1}_{\{s_{t+i}(\mu_{t+i-1}-m)>0\}}$$

Taking expectations conditionally on the information available at time t, we obtain

$$\mathbb{E}_t^A \left(P_{t+l} - P_{t+l}^*\right) = \mathbb{E}_t^A \left(P_{t+k} - P_{t+k}^*\right) + \lambda(q-1) \frac{\Sigma}{\sqrt{2\pi}} \sum_{i=k+1}^{l} \left(2\mathbb{P}_t(\mu_{t+i-1} > m) - 1\right).$$

Now,

$$\mathbb{P}_t^A(\mu_{t+i-1} > m) = (\mathcal{M}^{i-1} f_m)(\mu_t),$$

with as usual $f_m(x) = \mathbb{1}_{x \geq m}$. According to Lemma 6.2, we have that $(\mathcal{M}^{i-1} f_m)(\mu_t) \geq \frac{1}{2}$ if and only if $\mu_t > m$. This shows that, for example, when biased traders are optimistic, there is a bubble: prices are too high and are expected to increase further.

We now analyze the investment strategy of a short-term trader (referred to as a hedge fund). The trading strategy of a one-period horizon hedge fund is

$$d_t = \frac{1}{c} \left(\mathbb{E}_t^A (P_{t+1} - P_t)\right).$$

Because,

$$P_{t+1} = P_t + (1-\lambda) s_{t+1} + \lambda \sigma_{t+1},$$

we get

$$
\begin{aligned}
d_t &= \frac{\lambda}{c}\mathbb{E}_t^A(\sigma_{t+1}) \\
&= \frac{\lambda(q-1)\Sigma}{c}\mathbb{E}_t^A(s_{t+1}\mathbb{1}_{(\mu_t-m)s_{t+1}\geq 0} \\
&= \frac{\lambda(q-1)\Sigma}{c\sqrt{2\pi}}\left[2\mathbb{1}_{(\mu_t>m} - 1\right]
\end{aligned}
$$

which implies that $d_t > 0$ if $\mu_t - m > 0$. The maximal profit of a short term trader is thus

$$
\Pi_t^* = \frac{c}{2}\left(\frac{\lambda(q-1)\Sigma}{c\sqrt{2\pi}}\right)^2.
$$

In order to study the behavior of a short term trader, we assume without loss of generality that we are at time t and $\mu_t > m$. We will compute the probability of a feedback (contrarian) trading.

$$
\begin{aligned}
\mathbb{P}_t^A(\text{feedback trading}) &= \mathbb{P}_t^A(s_{t+1}d_{t+1} \geq 0) \\
&= \mathbb{P}_t^A(s_{t+1} \geq 0 \cap d_{t+1} \geq 0) + \mathbb{P}_t^A(s_{t+1} \leq 0 \cap d_{t+1} \leq 0) \\
&= \mathbb{P}_t^A(s_{t+1} \geq 0 \cap \mu_{t+1} \geq m) + \mathbb{P}_t^A(s_{t+1} \leq 0 \cap \mu_{t+1} \leq m) \\
&= \mathbb{P}_t^A(s_{t+1} \geq 0) + \mathbb{P}_t^A(s_{t+1} \leq -(\mu_t - m)) \\
&= \frac{1}{2} + \mathcal{N}\left(\frac{-(\mu_t - m)}{\Sigma}\right)
\end{aligned}
$$

- If the public signal $s_{t+1} \geq 0$, the variation of price $P_{t+1} - P_t = \lambda s_{t+1} + (1-\lambda)qs_{t+1}$ is positive and the biased traders'belief $\mu_{t+1} = \mu_t + qs_{t+1} > 0$ which implies that $d_{t+1} > 0$. The short term trader engages in feeback trading.

- If the public signal $s_{t+1} \leq -\mu_t$, the variation of price $P_{t+1} - P_t = s_{t+1}$ is negative and the biased traders' belief $\mu_{t+1} = \mu_t + s_{t+1} < 0$ which implies that $d_{t+1} < 0$. The short term trader engages in feeback trading.

- If the public signal $-\mu_t \leq s_{t+1} \leq 0$, the variation of price $P_{t+1} - P_t = s_{t+1}$ is negative but the biased traders' belief $\mu_{t+1} = \mu_t + s_{t+1} >$

0 which implies that $d_{t+1} > 0$. The short term trader engages in contrarian trading.

References

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

[2] 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.

[3] Barberis N., A. Shleifer, and R. Vishny, 1998, A Model of Investor Sentiment, Journal of Financial Economics, 49, pp. 307-343.

[4] 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.

[5] 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.

[6] Brunnermeier M. and S. Nagel, 2004, Hedge funds and the technology bubble, Journal of Finance, 59, pp. 2013-2040.

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

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

[9] Dave C. and K. Wolfe, 2003, On confirmation bias and deviations from bayesian updating, Working Paper, Pittsburgh University.

[10] DeBondt W. and R. Thaler, 1985, Does the stock market overreact?, Journal of Finance, 40, pp. 793-805.

[11] 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.

[12] Dougherty T., D. Turban, and J. Callender, 1994, Confirming first impressions in the employment interview: A field study of interviewer behavior, Journal of Applied Psychology, 79, pp. 659-665.

[13] Dumas B., A. Kurshev, and R. Uppal, 2005, What can rational investors do about excessive volatility and sentiment fluctuations?, Working Paper, INSEAD and LBS.

[14] Fama E. and K. French, 1988, Permanent and temporary components of stock prices, Journal of Political Economy, 96, pp. 246-273.

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

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

[17] Gilovich T., D. Griffin, and D. Kahneman, 2002, Heuristics and biases: The psychology of intuitive judgment, Cambrige University Press, New York.

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

[19] Hong H. and J. Stein, 1999, A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets, Journal of Finance.

[20] 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.

[21] Jouini E. and C. Napp, 2007, Consensus Consumer and Intertemporal Asset Pricing with Heterogeneous Beliefs, Review of Economic Studies, 74, pp. 1149-1174.

[22] Kahneman D., P. Slovic and A. Tversky, eds, 1982, Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press.

[23] Kyle A. and A. Wang, 1997, Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?, Journal of Finance, 52, pp. 2073-2090.

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

[25] Lesmond D., M. Schill, and C. Zhou, 2004, The illusory nature of momentum profits, Journal of Financial Economics, 71, pp. 349-380.

[26] 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.

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

[28] Mankiw G., D. Romer, and M. Shapiro, 1985, An unbiased reexamination of stock market volatility, Journal of Finance, 40, pp. 677-687.

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

[30] Plous S., 1991, Biases in the assimilation of technological breakdowns: Do accidents make us safer?, Journal of Applied Social Psychology, 21, pp. 1058-1082.

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

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

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

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

[35] Temin P. and H.-J. Voth, 2003, Riding the south sea bubble, Working Paper, Universitat Pompeu Fabra.

[36] Wurgler J. and E. Zhuravskaya, 2002, Does arbitrage flattens demand curves for stocks?, Journal of Business, 75, pp. 583-608.