GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 388

Path-Dependent Behavior with Asymmetric Information about Traders' Types

Alessia Testa*

* University of Bonn

July 2012

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Path-Dependent Behavior with Asymmetric Information about Traders' Types*

Alessia Testa†

University of Bonn

July 29, 2012

Abstract

We define path-dependency as the generic phenomenon according to which agents take an action regardless of their private information. Path-dependency can be of two types contingent on whether agents act with the crowd (herding) or against the crowd (contrarianism). We consider a quote-driven market where traders can in some cases observe whether their predecessors were informed, although they cannot observe their private information, while in other cases they are left with the uncertainty that their predecessors acted purely for liquidity motives. In this setting we recover herding and contrarianism and we find that better-informed markets (i.e. where informed traders receive high precision signals) can generate path-dependent behavior more easily than poorly informed ones. Moreover, we illustrate how a market dominated by herding features a price that is more informative of the asset value than the price of a market where traders always follow their signal. We also discuss how contrarianism has the exact opposite effect by decreasing price informativeness. (JEL D82, D83, G14)

*I would like to thank Amil Dasgupta, Ian Jewitt, Pete Kyle, Meg Meyer, Sujoy Mukerji, Han Ozsoylev, Bruno Strulovici and Dezsö Szalay for their useful comments and suggestions. I also benefited from conversations with Godfrey Keller, Andrea Patacconi, Robert Ritz and Chris Wallace. All mistakes are mine.

†Contact: alessia@uni-bonn.de.

1 Introduction

We study the failure of information aggregation in financial markets where traders show up sequentially in front of a competitive and risk neutral market maker to trade the only asset in the economy. By acting upon the reception of private information about the value of the asset, traders release this information to the public, allowing the market maker and, in general, the economy to learn about its true value.

The types of failure of information aggregation we focus on are herd behavior, contrarianism and informational cascades. We group herding and contrarianism under the label of *path-dependent behavior*, and we describe this as the phenomenon by which an individual takes an action independently of his private information. While engaging in herd behavior an individual disregards his information to follow the crowd, whereas when engaging in contrarian behavior an individual disregards his private information to go against the crowd. An informational cascade takes place when everybody in the economy engages in path-dependent behavior, whether this is herding or contrarianism. In the spirit of Chamley [3], we trace the difference between path-dependency and informational cascades in the fact that the former occurs at an individual level, while there is still the possibility that some agent's private information induces him to take a different action. This possibility is not realized, but the very fact that it is not realized yields information that is incorporated in the social learning: learning slows down but does not stop completely. On the other hand, in an informational cascade all the individuals in the economy are engaging in path-dependency, for the public belief dominates the private belief of every individual. As a consequence, actions do not disclose any new private information, beliefs do not change and, if everything else remains the same, learning stops forever. It follows that, in general, an informational cascade implies path-dependency, but the contrary does not need to be true.

The seminal paper by Bikhchandani, Hirshleifer and Welch [2] (BHW henceforth) seems to suggest that informational cascades are more likely to occur when each individual's private information is very accurate. However, in another seminal paper, Avery and Zemsky [1] (AZ henceforth) study herding in financial markets and conclude that when the quality of the private information is poor enough one can build trading histories generating herd behavior. We would like to cast some light on the role that the quality of private information plays in achieving herding in financial markets, and on whether there is something special about financial markets that overturns the early intuition found in BHW. We start by reviewing their results and by illustrating some of the latest findings in the literature.

BHW challenges the efficiency outcome of market equilibrium when individuals act sequentially and the price for taking an action is fixed ex ante. In their model, the first person who does not follow his signal to follow the crowd triggers something irreversible and learning stops completely. Agents differ only by their private information, so once their actions become independent of it, they become identical. It follows that whenever a particular individual takes an action regardless of his private information, this would be the case for any other individual in the economy: once someone herds, anybody else would herd in his place and an informational cascade is generated.

The key feature of models with a fixed price of adoption is that, eventually, individual beliefs reach some threshold above (below) which no private information can take the expected value of adoption given those beliefs below (above) the price ever again. In financial markets, however, prices respond to the trading activity. AZ recover the efficiency result of asset markets with sequential trading, where prices are set via a bid-ask mechanism. They show that informational cascades can never occur as long as market prices are flexible and, in general, that herding cannot occur unless asymmetric information of a particular kind, for instance in the form of *event uncertainty*, is introduced. Under event uncertainty, not only there is uncertainty about the value of the asset (which they call *value uncertainty*), but also the asset could have changed its value in an one-off event (a shock) in the course of trading; traders are informed on whether a shock has occurred, whereas the market maker can only learn this by observing the trading behavior of the market participants. Event uncertainty makes traders respond to the informational event faster than the market maker, who assigns too much of the trading to the activity of noise traders. Hence, event uncertainty slows the price adjustment in the short run, creating some form of inflexibility and allowing for herd behavior to occur. AZ find that, for a given trading history, herding occurs more easily in poorly informed markets.

In a recent paper, Park and Sabourian [5] (PS henceforth) generalize AZ's private information structure. Within a state space of at least three elements, they investigate the impact of different signal configurations on path-dependent behavior. In particular, they define U-shaped signals as those signals moving probability mass from moderate to extreme states; correspondingly, they call signals moving probability mass from extreme to moderate states hill-shaped. PS show that with a state space of at least three elements, U-shaped signals with positive or negative bias (i.e. giving more weight to the good or bad state, respectively) are necessary and (almost) sufficient to generate herd behavior, while biased hill-shaped signals are necessary and (almost) sufficient to generate contrarianism. For an intuition of their result, consider, as an example, a trader who receives a U-shaped signal with a negative bias. This signal assigns more weight to the extreme states and, due to the negative bias, it induces a trader who has observed nothing but his signal to sell. Suppose that, before placing his order, the trader in question observes a trading history that convinces him that the negative state is almost impossible. In light of this new information, the negative bias has no bite and the signal transfers almost all of the probability mass from the intermediate state to the positive one, inducing the trader to buy following the crowd.

The financial market we consider is represented by a trading room populated by a countable number of agents who are randomly selected one at the time to trade one unit of an asset with a perfectly competitive and risk neutral market maker. Traders can be of two 'types': with probability $(1-\lambda)$ they are of type I, in which case they are informed with probability one; or, with probability λ, they are type II traders, in which case they can either be informed with probability $(1-\mu)$, or noise traders with probability μ. Informed traders receive informative private signals conditional on which they update their beliefs before trading, whereas noise traders do not receive any private information and buy and sell with equal probability. Traders can observe each others' types, namely they can tell whether someone trading at some time t is a type I or a lottery trader; however, in the case a trader is of type I, his private signal cannot be observed, and in case he is a lottery trader it cannot be said whether he is informed or not, let alone, in case he is informed, his private information. The market maker cannot observe either types or signals.

This set-up can be seen as capturing the concept of neighborhood such as, for instance, the trading floor of a financial institution, where traders know the types but not the specific information of their clients, and share this information and the corresponding order flow, but not necessarily the identity of the clients, with the rest of the trading floor. The latter serves the purpose of gathering together the order flow, with traders acting on behalf of their clients. With the concept of the trading room, we are modelling the clients directly as traders who can observe each others' types. The two types of clients we have in mind are hedge funds trading primarily upon the reception of by private information, and investment funds whose trades can be driven by private information, but also by the issuance and redemption of shares. The latter reason for trading does not reveal any private information and it is completely unrelated to the market's trading history. A trader revealing that he has been executing the order of a type I (hedge) fund can be in the interest of his client: by moving the market in the direction of his trade, the client is more likely to profit from a capital gain.

The existence of herd and contrarian behavior is driven by the asymmetry of information between the traders and the market maker, who sets prices via a bid-ask mechanism à la Glosten and Milgrom [4] (GM henceforth). In particular, in face of a long enough realization of type I buys, traders update their beliefs as if they had full information about the signal realizations; on the other hand, the market maker has to consider the eventuality that the trades he observes are generated by noise. This causes his pricing to be rigid enough for traders to find it advantageous to buy regardless of their signal. Contrarianism is the consequence of type II trading activity. The probability of noise conditional on observing a type II trader is higher than the overall fraction of noise traders: this causes the price to react 'too much' to the trading activity and causes traders to ignore their private information to go against the market.

We prove the existence of trading histories such that herd and contrarian behaviors occur with positive probability. We find that, for any trading history, herding and contrarianism occur more easily in better-informed markets, i.e. markets with high-precision signals, than in poorly-informed markets. This is because a very precise signal exacerbates the informational asymmetry between the traders and the market maker. Once a trader acts regardless of his signal, the rest of the traders do not update their beliefs anymore whereas the market maker is learning that either herding or contrarianism is taking place. This continues until the market maker realigns prices to reflect the information present in the market.

The analysis suggests that allowing traders to observe each other's types when the share of type I traders is high is beneficial to the informativeness of the price: herd behavior causes the price to realign with the valuation of the fully informed type I traders and it suppresses the opinion of the (mistaken) minority. Correspondingly, allowing traders to observe each other's types when the fraction of type II traders is high generates contrarianism. Type II trades introduce too much noise in the eyes of the traders, and they do so in a non-neutral way but in the direction of the correct state of the world. For instance, when $V = 1$ traders attribute too many of the buy orders to noise originating contrarian selling to the detriment of price informativeness.

We also reconcile our results with AZ. Indeed, our information structure is similar to the one in AZ's event uncertainty in that it is formed by three levels: a first level is the public information that is common knowledge among the traders and the market maker and it is represented by the trading history; a second level can be thought of as *public information among traders*, which is common knowledge only among traders and consists, on top of the trading history, of the types of the traders associated with past trades; as a third level there is the traders' private information. The difference in the information structure between this model and AZ's is that, in the latter, an exogenous shock is needed to create this intermediate level of information, whereas in the present paper this is obtained by allowing traders to know something more about each other. As in AZ, this intermediate level of information creates price rigidity and allows for herd behavior to occur. As the information structure is equivalent to the one in AZ, similar notions of herding and contrarianism are used.

Changing the nature of the intermediate level in the information structure is not just a way of re-interpreting the source of uncertainty and informational asymmetry. In fact, this has an important implication, namely a gain in tractability which allows for a smaller state space and that enables us to state the conditions for herding and contrarianism in terms of exogenous variables only. This has deep consequences for our comparative statics. In fact, we find that herding and contrarianism occurs with positive probability after some trading history when the signal precision is *above* some threshold. AZ, though, construct trading histories such that herd behavior occurs at a certain time t if the signal's precision is *below* some threshold. This apparent contradiction disappears if we notice that the threshold below which AZ's market herds depends on the market participants' posterior beliefs prior to herding, whereas our cut-off level is only a function of exogenous variables, namely the level of noise trading, and of the type-composition of the market.

AZ do not link the formation of the posterior beliefs to the quality of information. They simply state that, once those beliefs come into place, it is easier to herd for low values of the precision. To fix ideas, consider the case of herd buying; herd buying occurs when a trader buys regardless of his signal, in particular with a low signal. For this to happen, after observing the trading history, the traders' valuation of the asset prior receiving a low signal must be high enough, and the market maker valuation prior receiving a buy order must be low enough, in relative terms, so that even after the reception of a low signal and of a buy order, respectively, the gap between the updated valuation and the ask price is preserved. As long as the traders' valuation of the asset is higher than the market maker's valuation, there exists a threshold below which herding occurs: the lower the signal precision the less sensitive the decrease in valuation upon the reception of a low signal by the traders, and the smaller the increase in the ask price by the market maker. Hence, once the asymmetry of information has generated some gap in valuations, a lower signal works better in order to preserve this gap. Our result does not contradict all this as, in addition, it has something to say about how the gap in the posterior beliefs was formed in the first place in relation to the quality of information. We find conditions on the signal precision and on the level of noise under which the gap in valuations can be built through the trading history starting from $t = 0$ and such that herd buying occurs.

We face a trade-off: if on the one hand it is still true that once a gap in beliefs is in place a lower signal precision works better to preserve it, on the other hand, the higher the signal precision the stronger the asymmetry of information between traders and market maker, the faster their valuations diverge. The trade-off is resolved in favor of a high signal precision, for any level of noise and type-composition of the market. From the perspective of a single period's trading, assuming that the valuations have already accumulated a gap in the 'right' direction, there is no discordance between AZ's result and ours, although ours identifies conditions such that the 'right' gap is built through the trading history.

Finally, we can interpret our result within PS's general signal structure: for instance, the signal configurations that generate herd buying (selling) more easily are those that are the fastest in moving probability mass towards the good (bad) state.

The paper is organized as follows: Section 2 presents the structure of the model and addresses the problem of the market maker's pricing rule in presence of herd behavior; Section 3 studies the case where the possibility of herd and contrarian behavior occurs for the first time in the trading history and establishes its existence; Section 4 reconciles our results with those in AZ and Section 5 concludes by discussing the informational properties of the price in a market where herding and contrarianism take place compared to a market where everybody always follows his private information.

2 The Model

There is a finite set of risk neutral players $N = \{1, 2, ...\}$ who show up randomly and anonymously at the post of a perfectly competitive and risk neutral market maker in order to trade one unit of the only asset available in the economy, asset that can take value $V \in \{0, 1\}$. Trading happens sequentially, i.e. only one trader at the time can show up in front of the market maker; moreover, agents can only trade once in their lifetime. Time is discrete, $t \in \{1, 2, ..., T\}$, where T is the time when the asset is liquidated and the capital gain (loss) is realized. At each point in time, the market maker posts a bid B_t and an ask A_t price at which he commits to trade.

Agents, if called to trade, decide whether to buy, sell or not trade at all given those prices. We indicate a generic action/trade at t with $a_t \in \{\text{buy, sell, no trade}\}$, while the realized price at time t is denoted by V_t^m.

Agents can be of two types: they are either type I traders with probability $(1-\lambda)$, or they are type II traders. Type I traders are informed with probability one and receive an informative signal σ about the value of the asset, while type II traders are either informed with probability $(1-\mu)$ or noise traders with probability μ. Both types' draws and the draws determining whether a type II trader is informed or not are independent and independent from each other. Noise traders trade for liquidity reasons and they are assumed to buy and sell with equal probability[1].

Signals can be either high (H) or low (L) and, conditional on V, they are independent. The probability that a signal reveals the true state is $p > \frac{1}{2}$, i.e. $\Pr\{\sigma = H \mid V = 1\} = \Pr\{\sigma = L \mid V = 0\} = p$, where the initial common prior is $\pi_0 = \Pr\{V = 1\} = \frac{1}{2}$. We take the convention according to which informed traders receive their signal only at the moment in which they are called to trade. This is without loss of generality and it is going to make the exposition simpler, as it implies that, before being called to trade, traders share the same valuation for the asset.

Each period the selected trader is assigned a type which can be observed by the other traders but not by the market maker. It follows that if a trader is observed to be of type I, he is automatically recognized to be informed, although his signal cannot be observed by anybody else but himself. If the selected trader is observed to be of type II, his fellow traders cannot distinguish whether he is informed or a noise trader. The market maker does not know either the type or the signal of the trader: he just receives the trading order and executes it, while the other traders observe the realized price. Therefore, traders not only can observe the trading history that has unfolded up to the point when they are called to trade, but they are also given a little bit more information compared to the market maker, namely they know the type of the trader active at each point in time.

[1]The case where noise traders do not trade is omitted, as by its own definition a noise trader needs to trade to be such.

The different possible transactions in every period together with all the possible bid and ask prices form the space $\Omega = \{\text{buy, sell, no trade}\} \times [0,1]^2$, which is identical for all $t \in \{1, 2, ..., T\}$, so that the space of all possible trading sequences is $\mathcal{H} = \prod_{t=1}^{T} \Omega_t$, where $\Omega_t = \Omega$. Call $\mathcal{F}$ the algebra on $\mathcal{H}$, and $\{\mathcal{F}_t\}$ the corresponding filtration.

Each trader's information is composed of three parts: the trading history, a vector of types of those who traded before him, and his private signal. Formally, indicate with τ_t a random variable that takes value 1 if the trader at t is of type I and 0 if the trader at t is of type II. Call $\mathcal{N} = \Pi_{t=1}^{T} \{1, 0\}_t$ the history of types and $\{\mathcal{T}_t\}$ the corresponding filtration. Then, each trader's information structure at time t is represented by the filtration

$$\{\mathcal{I}_t^i\} = \{\mathcal{F}_t, \mathcal{T}_t, \sigma^i\},$$

where agent i's private information consists in his private signal σ^i.

We use the convention of indicating with π_{t-1} the generic posterior probability generated by trading at $t-1$ and that is carried over to period t as a prior, so we use the phrase 'posterior at $t-1$' and 'prior at t' interchangeably to mean π_{t-1}.

Call π_t^i the posterior probability that, at time t, agent i assigns to the event that the true value of the asset is 1, where $\sigma^i = \varnothing$ if i is not trading a t:

$$\begin{aligned} \pi_t^i &= \Pr\left(V = 1 \mid a_t, \mathcal{I}_t^i\right) \\ &= \frac{\Pr\left(a_t \mid V = 1, \mathcal{F}_t, \mathcal{T}_t, \sigma^i\right) \pi_{t-1}^i}{\Pr\left(a_t \mid V = 1, \mathcal{F}_t, \mathcal{T}_t, \sigma^i\right) \pi_{t-1}^i + \Pr\left(a_t \mid V = 0, \mathcal{F}_t, \mathcal{T}_t, \sigma^i\right) \left(1 - \pi_{t-1}^i\right)}, \end{aligned} \tag{1}$$

Correspondingly, the traders' valuation of the asset is

$$\begin{aligned} V_t^i &= E\left[V \mid a_t, \mathcal{I}_t^i\right] = E\left[V \mid \mathcal{I}_{t+1}^i\right] \\ &= 1 \times \pi_t^i + 0 \times \left(1 - \pi_t^i\right) = \pi_t^i \end{aligned}$$

It is understood that the optimal decision for an informed trader i called to trade at time t

is

$$\text{buy if} \quad V_t^i > A_t, \tag{2}$$

$$\text{sell if} \quad V_t^i < B_t. \tag{3}$$

The market maker is Bayesian and updates his beliefs given the trading history. Call π_t^m the probability that, at time t, the market maker assigns to the event that $V = 1$ given the trading history:

$$\begin{aligned} \pi_t^m &= \Pr\left(V = 1 \mid a_t, \mathcal{F}_t\right) \\ &= \frac{\Pr\left(a_t \mid \mathcal{F}_t, V = 1\right) \pi_{t-1}^m}{\Pr\left(a_t \mid \mathcal{F}_t, V = 1\right) \pi_{t-1}^m + \Pr\left(a_t \mid \mathcal{F}_t, V = 0\right) \left(1 - \pi_{t-1}^m\right)}. \end{aligned} \tag{4}$$

Correspondingly, the market maker's valuation of the asset is:

$$\begin{aligned} V_t^m &= E\left[V \mid a_t, \mathcal{F}_t\right] = E\left[V \mid \mathcal{F}_{t+1}\right] \\ &= \Pr\left[V = 1 \mid \mathcal{F}_{t+1}\right] = \pi_t^m. \end{aligned} \tag{5}$$

In setting the price at the beginning of time t, the market maker does not know whether he will be facing a buy or a sell order. Conditional on a buy or a sell, he will post an ask and a bid price respectively so that the zero profit condition is satisfied:

$$A_t = E\left[V \mid \mathcal{F}_t, a_t = \text{buy}\right] = E\left[V \mid \mathcal{F}_t, V_t^i > A_t\right] \tag{6}$$

$$B_t = E\left[V \mid \mathcal{F}_t, a_t = \text{sell}\right] = E\left[V \mid \mathcal{F}_t, V_t^i < B_t\right]. \tag{7}$$

If at time t the trading order is a buy (sell), then we will refer to A_t (B_t) as the realized price and we will indicate it with V_t^m, the valuation of the asset by the market maker at the beginning of time $t + 1$.

Definition 1 ***(Equilibrium).*** *An equilibrium consists of a system of individual trading strategies as in (2) and (3), and a system of prices satisfying (6) and (7).*

Typically, in models where generic agents or traders are the only actors in the economy (e.g. [2]), the public information is common knowledge and the private signal constitutes the only private information. In an economy where traders are not alone because of the presence of a market maker, the distinction between private and public information is not that stark. In the model presented here there are three levels of knowledge: the public information $\mathcal{F}_t$, which is common knowledge among the traders and the market maker; the information contained in $(\mathcal{F}_t, \mathcal{T}_t)$, which is common knowledge only among traders; finally, the private signal σ_i, which we will refer to as private information. As pointed out in the introduction, this information structure parallels AZ's one under event uncertainty. For this reason, we adopt the same definition of herding. Roughly, an agent is herding if he disregards his private signal to trade in the direction of the market. Furthermore, an agent engages in contrarian behavior if he disregards his private signal to trade against the market. The definitions are given abstracting from bid and ask prices.

Definition 2 ***(Herding - Contrarianism).*** *A trader with signal σ^i engages in herd behavior at time t if he buys when $V_0^i(\sigma^i) < V_0^m < V_t^m$ or if he sells when $V_0^i(\sigma^i) > V_0^m > V_t^m$; and buying (or selling) is strictly preferred to other actions.*

A trader with signal σ^i engages in contrarian behavior if he buys when $V_0^i(\sigma^i) < V_0^m$ and $V_t^m < V_0^m$, and if he sells when $V_0^i(\sigma^i) > V_0^m$ and $V_t^m > V_0^m$; and buying buying (or selling) is strictly preferred to other actions.

As AZ point out, for herd buying to occur three things need to happen: without observing any trading history, the trader sells at $t = 0$; the history of trades must be positive and, despite of the increase in price the trader must be willing to buy after having observed the trading history. Herding can be interpreted as a situation where the price has not moved as much as the trader's valuation after observing a positive trading history. Correspondingly, for contrarian buying to occur, the trader needs to be wanting to sell at $t = 0$, and to buy after observing a trading history that leads to a decrease in the price. Contrarianism is the consequence of the price reacting too much to the trading history compared to the traders' valuation. In general, *path-dependent behavior.* is triggered when traders who would have followed their signal at $t = 0$ observe the unfolding of specific trading history paths and change in their behavior.

2.1 Traders' Updating and Market Maker's Pricing Rule

Traders' Updating As we use the convention according to which informed traders receive a signal only when called to trade, before their turn comes all of the traders have the same valuation of the asset at each point in time, for there is no informational asymmetry among them. In particular, they all agree on whether the conditions for a path-dependent behavior have arised. Consider the case of herding (the same reasoning holds for the case of contrarian behavior). Since traders differ only by their signal, once one of them herds, all of them would be herding. It follows that a type I trader who is observed herding at t does not release any information to the other traders, so that $V_t^i = V_{t-1}^i$ for every i. Similarly, if the conditions for herding are in place and a type II trader is observed, his actions are uninformative as well because he is either a herding informed trader or a noise trader.

Whenever the traders observe a type I trade and there is no possibility for herd behavior, then it is as if they could observe each other's signals. In case they receive a high signal as well as if they observe a type I buy in absence of herding, traders update their valuation according to:

$$V_t^i = E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i = H\right] = \frac{pV_t^i}{pV_t^i + (1-p)\left(1 - V_t^i\right)}. \tag{8}$$

We can conclude that what matters for the traders' valuation is the number of non-path-dependent buys and sells. Indicate with h_t and l_t the numbers of high and low signals implicitly 'observed' (through the trades of type I traders) up to time t, and with b_t^i and s_t^i the number of type II buys and sells observed up to time t. Then, for every i,

$$V_t^i = \frac{p^{h_t - l_t}\left[\frac{\mu}{2} + (1-\mu)\,p\right]^{b_t^i - s_t^i}}{p^{h_t - l_t}\left[\frac{\mu}{2} + (1-\mu)\,p\right]^{b_t^i - s_t^i} + (1-p)^{h_t - l_t}\left[\frac{\mu}{2} + (1-\mu)\,(1-p)\right]^{b_t^i - s_t^i}}. \tag{9}$$

Market Maker's Pricing Rule We have already established that the market maker fixes bid and ask prices given the history of trades and the eventuality he will face a sell or a buy order respectively, as formalized by (4). In general, the market maker sets the prices conditional on the traders' strategies. Without the possibility of path dependent behavior, the strategy of an informed trader is to buy upon the reception of a high signal and to sell upon the reception of a low signal. If there are the conditions for herding, though, an informed trader buys regardless of his signal.

We say that the market maker is 'naive' if he sets bid and ask prices assuming that every informed trader follows his signal even when the conditions for path-dependent behavior are in place. Correspondingly, we will indicate the bid and ask prices of a naive market maker at time t with B_t^{naive} and A_t^{naive} respectively. These prices satisfy:

$$\begin{aligned} A_t^{naive} &= E\left[V \mid \mathcal{F}_t, V_t^i\left(\mathcal{F}_t, \mathcal{T}_t, \sigma^i = H\right) > A_t^{naive}\right], \\ B_t^{naive} &= E\left[V \mid \mathcal{F}_t, V_t^i\left(\mathcal{F}_t, \mathcal{T}_t, \sigma^i = L\right) > B_t^{naive}\right] \ . \end{aligned}$$

On the other hand, a sophisticated market maker is aware of the possibility of herding or contrarianism. He first checks whether by fixing $B_t = B_t^{naive}$ and $A_t = A_t^{naive}$ traders are going to engage in path-dependent behavior at those prices and, if this is not the case, those are the prices he actually posts. However, if he finds that at those prices there is room for herding or contrarianism, he revises them to take into account this eventuality. In equilibrium, the ask and the bid price of a sophisticated market maker satisfy

$$\begin{aligned} A_t &= E\left[V \mid \mathcal{F}_t, a_t = \text{buy}\right] = E\left[V \mid \mathcal{F}_t, V_t^i\left(\mathcal{F}_t, \mathcal{T}_t, \sigma^i\right) > A_t\right], \\ B_t &= E\left[V \mid \mathcal{F}_t, a_t = \text{sell}\right] = E\left[V \mid \mathcal{F}_t, V_t^i\left(\mathcal{F}_t, \mathcal{T}_t, \sigma^i\right) < B_t\right], \end{aligned}$$

where the first equality follows from the assumption of perfect competition and the second equality from the assumption of rational expectations.

In what follows we are going to focus on the ask price: we will first compute the price posted by the market maker under the assumption of perfect competition, and then we will check that this is in fact a rational expectation price.

Example 1. To fix ideas, consider the very simple case where $t = 2$ and $\mathcal{F}_3 = \{(B, V_1^m), (B, V_2^m)\}$. Suppose that up until $t = 3$ no possibility of path-dependency arised so that the prices of a naive and sophisticated market maker coincide, namely, their valuation is equal to

$$V_2^m = E[V \mid \mathcal{F}_3] = \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right]^2}{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right]^2 + \left[\frac{\lambda\mu}{2} + (1-\lambda\mu)(1-p)\right]^2}.$$

The trading history $\mathcal{F}_3$ is compatible with four type-signal histories, $\left\{\mathcal{G}_3^j\right\}_{j=1}^4$:

$$\overbrace{(B, B)}^{\mathcal{F}_3} \rightarrow \underbrace{\begin{cases} ((\tau = 1, B), (\tau = 1, B)) & \leftarrow \mathcal{G}_3^1 \\ ((\tau = 1, B), (\tau = 0, B)) & \leftarrow \mathcal{G}_3^2 \\ ((\tau = 0, B), (\tau = 1, B)) & \leftarrow \mathcal{G}_3^3 \\ ((\tau = 0, B), (\tau = 0, B)) & \leftarrow \mathcal{G}_3^4 \end{cases}}_{\left\{\mathcal{G}_3^j\right\}_{j=1}^4}$$

Indicate with $E\left[V \mid \mathcal{G}_3^j\right]$ the evaluation of a trader who has observed $\mathcal{G}_3^j$. Then,

$$\begin{aligned}
E\left[V \mid \mathcal{G}_3^1\right] &= \frac{p^2}{\left[p^2 + (1-p)^2\right]} = V_2^i, \\
E\left[V \mid \mathcal{G}_3^2\right] &= E\left[V \mid \mathcal{G}_3^3\right] = \frac{p\left[\frac{\mu}{2} + (1-\mu)\, p\right]}{p\left[\frac{\mu}{2} + (1-\mu)\, p\right] + (1-p)\left[\frac{\mu}{2} + (1-\mu)(1-p)\right]}, \\
E\left[V \mid \mathcal{G}_3^4\right] &= \frac{\left[\frac{\mu}{2} + (1-\mu)\, p\right]^2}{\left[\frac{\mu}{2} + (1-\mu)\, p\right]^2 + \left[\frac{\mu}{2} + (1-\mu)(1-p)\right]^2}.
\end{aligned}$$

Suppose that p, λ and μ are such that, if $\mathcal{G}_3^1$ is realized and the market maker sets $A_3 = A_3^{naive}$, an informed agent selected to trade at $t = 3$ is going to herd, i.e. a trader with a low signal is going to buy. Then, in face of a buy at $t = 3$, a sophisticated market maker needs to consider the following scenarios:

$$
(B,B,B) \rightarrow \begin{cases} & \nearrow \quad L \\ \mathcal{G}_3^1 & \rightarrow \quad H \\ & \searrow \quad noise \\ \mathcal{G}_3^2, \mathcal{G}_3^3, \mathcal{G}_3^4 & \begin{matrix} \nearrow & H \\ \searrow & noise \end{matrix} \end{cases}
$$

For every type-signal path compatible with $\mathcal{F}_3$, not only the market maker needs to account for the possibility that the buy order comes from an informed trader with a high signal or a noise trader, but also he needs to take into consideration that on $\mathcal{G}_3^1$ an informed trader with a low signal is going to buy as well. □

We are now going to generalize what illustrated in the previous example to the case where path-dependent behavior occurs for the first time at a generic t. Set

$$
\Sigma = \{(\tau = 0, B), (\tau = 0, S), (\tau = 1, B), (\tau = 1, S)\}
$$

to be the type-signal space and $\Sigma^T = \prod_{t=1}^T \Sigma_t$, where $\Sigma_t = \Sigma$ for every t, the set of all possible T-dimensional vectors of type-signal sequences. Call $\mathcal{G}$ the algebra on Σ^T and $\{\mathcal{G}_t\}$ its generic filtration.

At time t, without the possibility of path-dependent behavior, the trading history $\mathcal{F}_t$ is compatible with 2^t type-signal histories $\{\mathcal{G}_t^j\}_{j=1}^{2^t}$ as shown below, where we indicate with $E\left[V \mid \mathcal{G}_t^j\right]$ the valuation of a trader who 'saw' $\mathcal{G}_t^j$ at time t.

$$
\mathcal{F}_t \quad \rightarrow \quad \begin{cases} \mathcal{G}_t^1 & \rightarrow E\left[V \mid \mathcal{G}_t^1\right] \\ \mathcal{G}_t^2 & \rightarrow E\left[V \mid \mathcal{G}_t^2\right] \\ \vdots & \qquad \vdots \\ \mathcal{G}_t^{2^t} & \rightarrow E\left[V \mid \mathcal{G}_t^{2^t}\right] \end{cases}
$$

The traders' valuations along the $\mathcal{G}_t^j$ paths depend on the number and the type of buys and sells in absence of herd behavior as specified in (9). In particular, the higher the number of high signals implicitly observed or the number of type II sells, the bigger the gap $E\left[V \mid \mathcal{G}_t^j\right] - E\left[V \mid \mathcal{F}_t\right]$ and the easier for path-dependent buying to occur. For any trading history, in absence of past herding or contrarian possibilities, there is a unique $\mathcal{G}_t^{j^*}$ such that $E\left[V \mid \mathcal{G}_t^{j^*}\right] = \max_j E\left[V \mid \mathcal{G}_t^j\right]$, which is associated with the path where all buys are type I buys and all sells are type II sells. If path-dependent buying occurs for the first time at t, it will happen on the path $\mathcal{G}_t^{j^*}$: if $b_t > s_t$ we observe herd buying, while if $b_t < s_t$ we observe contrarian buying. It follows that, in deriving the ask price for the case of first time path-dependent buying, we only need to consider one possible type-signal path leading to it and the associated value for the traders 'on' that path, value that we indicate with $V_{t-1}^B = E\left[V \mid \mathcal{G}_t^{j^*}\right]$.

Define the measure $\eta_t : \left\{\mathcal{G}_t^j\right\}_j \rightarrow [0,1]$ to be the probability of the type-signal history that leads to the type-signal set $\mathcal{G}_t^j$, given that we reached $\mathcal{F}_t$. The probability of a single type-signal history vector it is simply the probability of r successes in t Bernoulli trials, where r is the number of Type I trades in the vector:

$$\Pr\left(\mathcal{G}_t^j \mid \mathcal{F}_t\right) = \binom{t}{r} \left[\Pr\left(\tau = 0\right)\right]^{t-r} \left[\Pr\left(\tau = 1\right)\right]^{r}.$$

Define, at each time t, the probability of the type-signal history corresponding to the maximal valuation V_{t-1}^B as

$$\eta_t^B = \Pr\left(\mathcal{G}_t^{j^*} \mid \mathcal{F}_t\right) = \binom{t}{b_t} \left[\Pr\left(\tau = 0\right)\right]^{t-b_t} \left[\Pr\left(\tau = 1\right)\right]^{b_t},$$

where $\left[\Pr\left(\tau = 0\right)\right] = \lambda$ and $\left[\Pr\left(\tau = 1\right)\right] = 1 - \lambda$.

Proposition 1 *Consider a trading history $\mathcal{F}_t$ such that, given p, λ and μ, no path-dependent behavior could have occurred until time t. Moreover, suppose that, on the maximal type-signal history $\mathcal{G}_t^{j^*}$ compatible with $\mathcal{F}_t$, an informed trader buys regardless of his signal. Then, market* *maker's expected value of the asset given a buy order at t is given by:*

$$E\left[V \mid \mathcal{F}_t, a_t = buy\right] \tag{10}$$
$$= \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\,p\right] V_{t-1}^m + \eta_t^B (1-\lambda\mu)(1-p)\, V_{t-1}^B}{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[pV_{t-1}^m + (1-p)\left(1 - V_{t-1}^m\right)\right] + \eta_t^B (1-\lambda\mu)\left[(1-p)\, V_{t-1}^B + p\left(1 - V_{t-1}^B\right)\right]}.$$

Proof. See Appendix A.1. ■

The market maker updates in a 'normal' way along all the type-signal paths, attributing a buy (sell) order either to a high (low) signal or to noise, while on the path leading to path-dependent buying he also updates as if the trader received a low signal, $(1-\lambda\mu)(1-p)\, V_{t-1}^B$, times the probability of that path given the trading history.

By construction, (10) satisfies the zero profit condition of the market maker. The next proposition states that, by setting $A_t = E\left[V \mid \mathcal{F}_t, a_t = \text{buy}\right]$, the market maker cannot prevent herding or contrarianism from happening, even if he accounts for this possibility in his pricing rule. This makes A_t the competitive rational expectation equilibrium price.

Proposition 2 *Consider a trading history $\mathcal{F}_t$ and a type-signal history $\mathcal{G}_t^j$ compatible with it such that $V_t^i\left(\mathcal{G}_t^j, \sigma^i = L\right) > A_t^{Naive}$. Then,*

$$V_t^i\left(\mathcal{G}_t^j, \sigma^i = L\right) > E\left[V \mid \mathcal{F}_t, a_t = buy\right].$$

The converse is also true. Hence, $A_t = E\left[V \mid \mathcal{F}_t, a_t = buy\right]$ is a rational expectation equilibrium price.

Proof. See Appendix A.2. ■

One might be tempted to think that the previous result is driven by the fact that the price of a sophisticated market maker needs to be lower than the price of a naive market maker because the former is accounting for both a high and a low signal driving the buy order on $\mathcal{G}_t^{j^*}$. This last intuition is true, but it is only part of the story: a buy order makes $\mathcal{G}_t^{j^*}$ and its associated high prior V_{t-1}^B more likely in the eyes of the market maker, inducing an overall increase of his valuation. In general we always have that $A_t > A_t^{naive}$.

Proposition 3 *Consider a trading history $\mathcal{F}_t$ and a type-signal history $\mathcal{G}_t^j$ compatible with it such that $E\left[V \mid \mathcal{G}_t^j, \sigma^i = L\right] > A_t^{Naive}$. Then $A_t > A_t^{naive}$. The converse is also true.*

Proof. See Appendix A.3. ■

We have established that $E\left[V \mid \mathcal{G}_t^{j^*}, \sigma^i = L\right] > A_t > A_t^{naive}$. This is easily interpreted if we observe that, by moving from the highest of the valuations to the lowest, we are 'losing' something either at the level of information or at the level of rationality. In fact, if $E\left[V \mid \mathcal{G}_t^{j^*}, \sigma^i = L\right]$ is the valuation of a fully informed and fully rational agent, A_t is the valuation of a partially informed and fully rational agent, to conclude with A_t^{naive}, which is the valuation of a partially informed and partially rational agent.

One consequence of these results is that, at every time period t, investigating the conditions for path-dependent behavior under the pricing rule of a naive market maker is equivalent to studying the same market under a sophisticated market maker. This does not mean that using the pricing rule of a market maker who is *always* naive is equivalent to using the pricing rule of a market maker who is not. The simplification is used to check conditions at a specific t, not as the market maker's pricing rule over time.

3 Herding and Contrarian Behavior

In this section we study the case of path-dependent behavior occurring for the first time at t, so when we talk about informed trading up to the herding/contrarian time t, we are assuming that those traders are following their signal. Typically, in the literature, the reasoning to prove the existence of herd behavior goes as follows: in finite time any trading history has positive probability; noise trading can always generate trading histories compatible with herding, hence the existence of herding. In this model, type-signal realizations are as important for herding as the trading history they are compatible with. This implies that we will not be able to appeal to noise trading generating any possible trading history in order to prove the existence of herding. In fact, noise can generate trading histories whose compatible type-signal histories meet the conditions for herding, but the very fact of assuming noise to generate those histories rules out herding. We can conclude that noise can generate trading histories such that the market maker attaches a positive probability to the eventuality of herding at some time t, but if the trading history at t were in fact caused by noise, then herding would not be possible. However, consistency requires that if the market maker deems some trading history compatible with herding, then there must exist a type-signal path leading to it. On the other hand, noise can lead to contrarianism, as during the realization of type II trades the price moves too much for the amount of information actually present in the market.

The proof of the existence of path-dependent behavior is achieved in two steps: first, in Theorem 1, we establish conditions on p, λ and μ, involving both the trading history and the compatible type-signal realizations, under which herd and contrarian buying occur with probability one; then, in Theorem 2, we prove that trading histories satisfying the conditions for path-dependency exist with positive probability.

Herd buying is a consequence of the delay, due to the presence of noise, with which the market maker reacts to a positive history of type I buys. However, traders' advantage in observing the nature of trading is not per se sufficient to generate herd behavior. Consider $1 - \lambda\mu$ vis à vis $(h_t + l_t)/t$ as the level of imbalance between theoretical informed trading and observed informed trading in the history up to t. Not all observed informed trades exceeding their theoretical frequency have an impact in causing price rigidity. In fact, if $|h_t - l_t|$ is steadily low, although there might be a big discrepancy between observed and theoretical frequencies of informed trading, type I buys and sells balance out, and the price does not need to adjust too much (on average) so that price rigidity has no bite. If $|h_t - l_t|$ gets high, then the imbalance starts to matter, and a high theoretical noise will cause price rigidity. Everything else equal, herd buying occurs when a relative high number of type I high signals is realized. Correspondingly, contrarian buying is a consequence of the excess of informational content the market maker assigns to a string of type II sells. This is once again due to the possibility for the traders to observe their colleagues' types: the share of informed type II traders is only $1 - \mu$, while the market maker believes that any trade has a probability $1 - \lambda\mu \geq 1 - \mu$ to be informed. Similarly we have observed for type I trades, what matters is the magnitude of the difference $|b_t^i - s_t^i|$ and not the absolute value of type II sells.

In Theorem 1 we find that it is easier for a well-informed market (i.e. a market where the quality p of information is high) to engage in path-dependent behavior than it is for a poorly-informed market. Intuitively, for any level of noise and type composition, the higher the signal's precision the bigger the effect of the asymmetry of information between the traders and the market maker. This is because the difference in informational content of a type I trade (correspondingly, the difference in the scarcity of information of a type II trade) between traders and market maker is much higher when the signal is very precise. It follows that a long enough sequence of type I buys or type II sells is going to make the difference in valuations $V_{t-1}^i - V_{t-1}^m$ increase more the higher the precision.

There is a trade off between building up, between periods 0 and $t-1$, a gap $V_{t-1}^i - V_{t-1}^m$ big enough for herd or contrarian buying to occur at t and, given prior valuations V_{t-1}^i and V_{t-1}^m having the best conditions for path-dependent buying to occur at t. We have already noticed that the gap grows bigger the higher the precision. At the same time, though, we would like valuations not to react too much to the new information at time t: we would like V_{t-1}^i not to decrease too much upon the reception of a low signal, and V_{t-1}^m not to increase too much upon observing a buy, so that the gap stays positive and an agent receiving a low signal at t buys. The situation just described, where $V_{t-1}^i - V_{t-1}^m$ is preserved at t, is easier to achieve the lower the precision of the signal. The following theorem states that, in any circumstance, the trade off is at the advantage of a high precision signal: it is easier for a high precision signal to accumulate a big enough gap $V_{t-1}^i - V_{t-1}^m$ such that at time t it will not be reverted by a low realization of the signal and a buy order, than to build the gap itself with a low precision signal. By 'easier' we mean that it takes a shorter sequence of type I buys or type II sells, and that it is more likely that any given trading history $\mathcal{F}_t$ will result in path-dependent behavior. It also means that, for a long enough positive (negative) history of type I (II) trades, herd (contrarian) buying occurs earlier with a high precision signal.

To ease notation, indicate with L the log-likelihood ratio between $V = 1$ and $V = 0$ for a trader upon the (implicit) observation of a high signal, with L^μ the log-likelihood ratio between

$V=1$ and $V=0$ for a trader upon the observation of a type II buy, and with L^{λ} the log-likelihood ratio between $V=1$ and $V=0$ for the market maker upon the reception of a buy order. Formally,

$$\begin{aligned} L &= \frac{\Pr\left(V=1 \mid a_t=\text{buy},\ \tau_t=1\right)}{\Pr\left(V=0 \mid a_t=\text{buy},\ \tau_t=1\right)}, \\ L^{\mu} &= \frac{\Pr\left(V=1 \mid a_t=\text{buy},\ \tau_t=0\right)}{\Pr\left(V=0 \mid a_t=\text{buy},\ \tau_t=0\right)}, \\ L^{\lambda} &= \frac{\Pr\left(V=1 \mid a_t=\text{buy}\right)}{\Pr\left(V=0 \mid a_t=\text{buy}\right)}. \end{aligned}$$

It is easy to check that $L \geq L^{\lambda} \geq L^{\mu}$ for every p, λ and μ.

Theorem 1 ***(Conditions for Path-Dependent Behavior)*** *Consider a trading history $\mathcal{F}_t$ such that no path-dependent behavior could have occurred before t and such that*

$$h_{t-1}-l_{t-1} > \frac{L+L^{\lambda}}{L-L^{\lambda}} + \left(b^{i}_{t-1}-s^{i}_{t-1}\right)\frac{L^{\lambda}-L^{\mu}}{L-L^{\lambda}} \tag{11}$$

Then,

1. *If $h_{t-1}-l_{t-1} < \frac{2}{\lambda\mu}-1+\left(b^{i}_{t-1}-s^{i}_{t-1}\right)\left(\frac{1}{\lambda}-1\right)$ there exists a cutoff level $\frac{1}{2} < p^{*}(\lambda,\mu) < 1$ such that for $1 > p > p^{*}(\lambda,\mu)$ path-dependent buying occurs at t;*
2. *If $h_{t-1}-l_{t-1} \geq \frac{2}{\lambda\mu}-1+\left(b^{i}_{t-1}-s^{i}_{t-1}\right)\left(\frac{1}{\lambda}-1\right)$ path-dependent buying occurs at t for every value of $p \in \left(\frac{1}{2},1\right)$.*

Moreover, we have that $\frac{\partial p^{*}}{\partial \lambda} < 0$ if and only if $b^{m}_{t} > s^{m}_{t}$, whereas $\frac{\partial p^{*}}{\partial \mu} < 0$ for any $\mathcal{F}_t$ leading to path-dependent behavior.

Proof. See Appendix A.4. ■

As already mentioned, a high p has the effect of increasing the informational asymmetry between the traders and the market maker. In fact, as $L \geq L^{\lambda}$, a higher p increases, in presence of noise, the informational content of type I trades more to the traders than to the market maker: coeteris paribus, a smaller h_t is needed to generate herd buying. In a similar fashion, since $L^{\lambda} \geq L^{\mu}$, a higher p increases the informational content of a type II trade more to the market maker than to the traders and, coeteris paribus, a smaller s_t is needed to generate contrarian buying.

When $p = 1$, the signal perfectly reveals the value of the asset, which also implies that signals are all the same. Conditional on $V = 1$, the gap in valuations between the traders and the market maker is always (weakly) positive, and informed traders all act alike not because they are herding but because they all receive the same, perfectly revealing information.

When $p = \frac{1}{2}$ the signal is completely uninformative and traders do not learn anything from it or by observing the other market participants' behavior. The price remains equal to π_0 and since traders are indifferent between buying and selling, we can assume that they simply follow their signal. However, no learning takes place.

The overall effect of λ is to make the market maker's pricing mechanism less reactive to the trading history. Price stickiness is beneficial in order to generate herding, as the gap $\left|V_{t-1}^{i} - V_{t-1}^{m}\right|$ is caused mainly by type I trade, but it makes the occurrence of contrarianism more difficult, as the gap increases following a preponderance of type II trades. Hence, if the condition for path-dependent buying (11) is consistent with herding, which requires $b_t^m > s_t^m$, then we observe that the cut-off p^* above which a trader buys with a low signal decreases with λ: even lower levels of the signal precision are compatible with herd behavior. At the same time, if path dependent buying occurs when $b_t^m < s_t^m$, an increase in λ causes the minimum level of precision p^* above which contrarianism takes place to increase.

The effect of noise trading is always beneficial to the occurrence of path-dependency: in case of herding, it dampens the market maker's price adjustment to type I trading, while in case of contrarianism it makes the price too reactive to type II trading. It follows that, for any given trading history satisfying condition (11) an increase in the mass of noise traders increases the set of values for the signal precision leading to path dependency.

Theorem 2 ***(Existence of Path-Dependent Behavior)*** *For* $p \in \left\{\frac{1}{2}, 1\right\}$ *path-dependent behavior does not occur for any trading history* $\mathcal{F}_t$ *and any* $\lambda, \mu \in [0,1]$. *For* $p \in \left(\frac{1}{2}, 1\right)$ *and for every* $1 > \lambda, \mu > 0$, *price paths with path-dependent behavior occur with positive probability. In* *particular, path-dependency occurs with positive probability for the first time at t whenever $\mathcal{F}_t$ is such that*

$$b_{t-1}^m \geq \frac{L + L^\lambda}{L - L^\lambda} - s_{t-1}^m \frac{L^\lambda - L^\mu}{L - L^\lambda}. \tag{12}$$

Proof. See appendix A.5. ■

For an intuition of Theorem 2, consider the simple case in which the market opens with a string of b_{t-1}^m type I buys in the first $t-1$ periods. Then $h_{t-1} = b_{t-1}^m$ and if at t an individual with a low signal is called to trade, he will herd at t for given values of p and μ if $b_{t-1}^m \geq \left[L^\lambda + L\right] / \left[L - L^\lambda\right]$. Suppose, instead, that the market opens with a string of s_{t-1}^m type II sells in the first $t-1$ periods. Then, an individual with a low signal who is called to trade at t engages in contrarian buying if $s_{t-1}^m \geq \left[L + L^\lambda\right] / \left[L^\lambda - L^\mu\right]$.

Indicate with $\lceil x \rceil$ the smallest integer greater than x. The minimum amount of time that is needed for the possibility of herd buying to realize given some values of p and μ is equal to

$$t_h^* = \left\lceil \frac{2L}{L - L^\lambda} \right\rceil,$$

whereas, the minimum amount of time that is needed for the possibility of contrarian buying to realize is equal to

$$t_c^* = \left\lceil \frac{2L^\lambda + L - L^\mu}{L^\lambda - L^\mu} \right\rceil.$$

It is easy to show that there exists a threshold $\lambda^* (p, \mu)$ such that $t_h^* < (>)\, t_c^*$ when $\lambda > (<)\, \lambda^*$. This is clearly related to the fact that $\frac{\partial p^*}{\partial \lambda} < 0$ if and only if $b_t^m > s_t^m$ and it is a complementary result: for any level of p and for any *given* trading history, herding occurs faster when prices are stickier, while contrarianism occurs faster if, coeteris paribus, prices move 'too much'. Concurrently, a low λ facilitates the price alignment to the traders' valuation following a period of type I trades which slows down the possibility of herding. In both cases, $\frac{\partial t_h^*}{\partial p} < 0$ and $\frac{\partial t_c^*}{\partial p} < 0$: both herding and contrarian behavior are more likely to happen earlier for higher levels of the signal precision.

4 Reconciling with Avery and Zemsky

In this section we are going to reconcile the results of Theorem 1 with those already found in the literature. In particular, we are going to show that, if we take prior evaluations V_{t-1}^{i} and V_{t-1}^{m} as given, herding at time t is easier to achieve for low levels of the signal's precision as stated in [1]. In this respect, we isolate herding from any other path dependent phenomena, and we make the assumption that the only path-dependent behavior arising at t at the ask price A_t is herd buying. Moreover, in line with AZ's results, we look for conditions for herding to occur at a generic time t regardless of whether other path-dependent behaviors could have occurred already in the trading history. In order to do this, we need to make sure we can still use (10) as the correct ask price. In fact, when the possibility of herding arises for the first time, the herding type-signal path and hence the traders' valuation attached to it is unique. However, when the possibility of herding has already occurred in the past, new herding paths might not be unique and more than one valuation along those paths might be taken into account when forming the ask price. The next result establishes that, although herding paths are not unique, they must share the same valuation of the asset, allowing us to use the same formulation as in (10) for a generic ask price at t when the only path-dependent behavior we need to worry about is herding.

To fix ideas, consider again the simple case of $\mathcal{F}_3 = \{B, B\}$ leading to the possibility of herd buying at $t = 3$. In fact, if the correct type-signal path is $\mathcal{G}_3^1$, traders do not update their beliefs at $t = 3$, as the agent trading is either herding or he is a noise trader. If there is still an opportunity for herd buying at time four, the traders' prior valuation at $t = 4$ is the same as in the previous period. If the correct type-signal path is not $\mathcal{G}_3^1$, then traders on path $\mathcal{G}_3^1$ and those emanating from paths $\mathcal{G}_3^2$ and $\mathcal{G}_3^3$ who observed a type I buy at $t = 3$ will have the same prior at the beginning of $t = 4$. With this logic in mind, we can state the next proposition.

Proposition 4 *If there is the possibility that agents with prior $V_{t-1}^{i}$are herding at t on $\mathcal{G}_t^j$ for some j, then all the herding agents have prior V_{t-1}^{i} for every j.*

Proof. *See Appendix A.6.* ■

Proposition 4 shows that, at the beginning of $t+1$, following the possibility of herd buying at t on $\mathcal{G}_t^j$, the traders' valuation on the paths emanating from $\mathcal{G}_t^j$ continues to be V_{t-1}^i: whether they observed a type I or a type II trader at t they would not update their prior, as all informed traders would be herding. Moreover, traders on the paths where the number of type I buys is one less than in $\mathcal{G}_t^j$ update their valuation upon observing an informed buy t. Then, at the beginning of $t+1$, this valuation is aligned to the one of traders coming from $\mathcal{G}_t^j$, and which is equal to V_{t-1}^i. Hence, V_{t-1}^i remains the highest traders' prior valuation at the beginning of time $t+1$. In fact, V_{t-1}^i remains the highest traders' prior valuation throughout all the periods where there is the possibility of herd buying, whereas both V_{t-1}^m and the probability η_t of being on the type-signal path leading to herding increase. It follows that the formula used for the ask price for the case of first time herding is in fact a general formula for the ask price at any period, the only things changing over time being V_{t-1}^m and η_t, with $V_t^i = V_{t-1}^i$ whenever $\eta_{t-1} > 0$.

Having just clarified the way the market maker updates his beliefs given the possible type-signal scenarios compatible with the trading history, it comes natural to explain here why informational cascades cannot occur in this market. An informational cascade means that nobody updates anymore, not even the market maker. The latter does not update when he thinks that the probability of an action is independent of the state of the world. For instance, in a cascade the market maker might believe that all the informed traders are buying and all sell orders come from noise traders. For this to occur, the market maker needs to believe that on all the type-signal paths informed traders are doing the same thing. However, there will always be paths were a buy does not mean herding and hence a sell does not mean noise. For this reason an informational cascade cannot happen.

Using (10), we can state and prove the following theorem, which delivers conditions for herd buying stated, analogously to AZ, as a function of posterior beliefs.

Theorem 3 *Suppose that the trading history at t is such that $V_{t-1}^i > V_{t-1}^m$. Then there exist a cut off level $\overline{p}\left(\lambda, \mu, V_{t-1}^i, V_{t-1}^m\right)$ such that traders engage in herd buying if and only if $p < \overline{p}\left(\lambda, \mu, V_{t-1}^i, V_{t-1}^m\right)$.*

Proof. See Appendix A.7. ■

The first thing to notice is that we have a cut off level $\overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_{t}^{m}\right)$ *below* which herd buying occurs. This is because, given the priors V_{t-1}^{i} and V_{t-1}^{m} such that $V_{t-1}^{i} > V_{t-1}^{m}$, a low precision of the signal decreases trader's valuation less upon the reception of a low signal and make herd buying easier. However, for any $\mathcal{F}_t$ such that $V_{t-1}^{i} > V_{t-1}^{m}$, the difference $V_{t-1}^{i} - V_{t-1}^{m}$ is larger the higher the precision p. Consider, in fact, the extreme case where $p = \frac{1}{2}$: the precision is the lowest possible, however, $V_{t-1}^{i} = V_{t-1}^{m}$ for every λ, μ and any $\mathcal{F}_t$. When $p = \frac{1}{2}$ neither the market maker nor the traders change their beliefs over time and herding is impossible (no action is strictly preferred to any other action).

Consider the case where herding occurred for the first time at t. Then $V_t^i = V_{t-1}^i$, while the market maker's prior valuation at $t+1$ has increased to $V_t^m = A_t$. Proposition 2 ensures that $V_{t-1}^{i} > V_t^m$, and that at $t+1$ we just need to check whether $A_{t+1}^{naive} < E\left[V \mid \mathcal{G}_{t+1}^{j}, \sigma^i = L\right]$. The proof of the theorem shows that this condition is equivalent to

$$\begin{aligned}\Delta_{t+1}(p) &= p^2\left(V_{t-1}^{i} - V_t^m\right)(1-\lambda\mu) + \left(V_{t-1}^{i} - V_{t-1}^{i}V_t^m\right)\left(1 - \frac{\lambda\mu}{2}\right) \\ &\quad + p\frac{\lambda\mu}{2}\left(V_{t-1}^{i} - V_t^m\right) - 2p\left(V_{t-1}^{i} - V_t^m V_{t-1}^{i}\right)\left(1 - \frac{\lambda\mu}{2}\right) > 0. \end{aligned} \tag{13}$$

The theorem guarantees that there exists a cut off $\overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_t^m\right)$ such that, when $V_{t-1}^{i} > V_t^m$, $\Delta_{t+1}(p) > 0$ for $p < \overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_t^m\right)$. Moreover, it is easy to check that when $V_{t-1}^{i} = V_t^m$, $\overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_t^m\right) < \frac{1}{2}$, so condition (13) can never be satisfied when $p = \frac{1}{2}$. If $\overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_t^m\right) < p^*(\lambda, \mu)$ herding stops with t. If $\overline{p}\left(\lambda, \mu, V_{t-1}^{i}, V_t^m\right) > p > p^*(\lambda, \mu)$ then herding continues at $t+1$, and at the beginning of $t+2$ the priors will be $V_{t+1}^m = A_{t+1}$ and $V_{t+1}^i = V_{t-1}^i$. In general, as $\frac{\partial \overline{p}}{\partial V_t^m} < 0$, at each herding period the cut off level of precision below which herding continues decreases.

5 Discussion

Although we have derived all our results for the case of path-dependent buying, it is clear that the case of path-dependent selling follows symmetrically. In particular, the equivalent of condition (11) for path-dependent selling is

$$l_{t-1} - h_{t-1} > \frac{L + L^{\lambda}}{L - L^{\lambda}} + \left(s_{t-1}^{i} - b_{t-1}^{i}\right) \frac{L^{\lambda} - L^{\mu}}{L - L^{\lambda}},$$

while the minimum waiting time for the first possibility of herd and contrarian selling are still equal to t_h^* and t_c^* respectively. Moreover, whether the first instance of herd selling occurs earlier or later than contrarian selling depends on the same threshold $\lambda^*(p, \mu)$ found for the case of buying.

We would like to know how the market price behaves in the long run. In particular, we are interested in the informational properties of the price in our market, where the private information about types leads to path dependency, compared to a market where traders always follow their signal. In this respect, consider a market which is identical to the one we have been studying so far with the exception that traders cannot observe each other's types: this is simply a market à la Glosten and Milgrom, where the intermediate layer of information which is common knowledge only among traders is shut down. Indeed, in a GM market traders always follow their signal. We would like to know whether the price in our market is a better or a worse predictor of the true value of the asset compared to a market à la GM.

For this purpose, indicate with $Var\left(V \mid V_t^{GM}\right) = \pi_t^{GM}\left(1 - \pi_t^{GM}\right)$ the variance of the value of the asset after having observed the realized price in a market where traders have no information about types, where π_t^{GM} is the probability that the true state is $V = 1$ at t in such a market. We would like to compare this with $Var\left(V \mid V_t^{m}\right) = \pi_t^{m}\left(1 - \pi_t^{m}\right)$. It is clear then that the price is more informative the higher (or lower) is π_t^m.

Assuming $V = 1$, $\pi_t^m > \pi_t^{GM}$ if and only if the occurrence of path-dependent buying is more frequent than the occurrence of path-dependent selling. This for two reasons: the first reason lies in the fact that path-dependent buying 'hides' low signals, whereas it does not alter the traders' behavior compared to a GM market when a high signal is realized; moreover, as we have seen in Proposition 3, the price is more volatile in periods of path-dependent trading which reduces[2] the conditional volatility of V even further.

When $V = 1$, the incidence of herd buying is naturally higher than the incidence of herd selling, whereas the incidence of contrarian selling is higher than the incidence of contrarian buying. Then, to make path-dependent buying more likely than path-dependent selling, the share of type II trades λ, which is responsible for contrarianism, needs to be 'small enough'[3].

We can (roughly and perhaps not surprisingly) conclude that $Var\left(V \mid V_t^m\right)$ is smaller than $Var\left(V \mid V_t^{GM}\right)$ when the share of type I traders is high: then, being able to observe traders' types is very informative and it leads to herding. By suppressing the opinion of the minority, given that the majority is correct in the limit, herding pushes the market closer to the true value of the asset. When the market is mostly composed of type II traders, the conditional variance of the state given the realized price is lower in a GM-type of market. In fact, the ability to observe types induces traders to attribute too much noise to the type II trading activity in a non-neutral way. For instance, when $V = 1$ more high signals than low signals moving the type II trading activity are thought to likely be noise, which in turns leads to more frequent contrarian selling and the hiding of high signals.

A Appendix

A.1 Proof of Proposition 1

Suppose that at t the first possibility of herd/contrarian buying arises on $\mathcal{G}_t^{j^*}$, where $j^* \in \arg\max_j E^i\left[V \mid \mathcal{G}_t^j\right]$ and $\left\{\mathcal{G}_t^j\right\}_j$ is the family of type-signal histories at time t compatible with $\mathcal{F}_t$. We know that j^* is unique, namely that the argmax is a singleton, Indicate with V_{t-1}^i the valuation of the asset by a trader i who has observed $\mathcal{G}_t^{j^*}$. We can write the market maker's valuation of the asset as the weighted sum of the value of the asset along the paths $\left\{\mathcal{G}_t^j\right\}_j$ with weights given by the probability of each of the paths given the trading history.

[2] Vives ([6], p.131) proves that $Var\left(V_t^m\right) = Var\left(V\right) - Var\left(V \mid V_t^m\right)$.

[3] Notice that we can have $t_h < t_c$ and at the same time have contrarianism more likely than herding: t_h and t_c indicate how fast herding and contrarianism can occur, not how likely they are.

We can then rewrite V_{t-1}^m as

$$\begin{aligned} V_{t-1}^m &= E[V \mid \mathcal{F}_t] = \Pr(V = 1 \mid \mathcal{F}_t) = \frac{\Pr(\mathcal{F}_t \mid V=1)\Pr(V=1)}{\Pr(\mathcal{F}_t)} \\ &= \frac{\sum_i \Pr\left(\mathcal{G}_t^i \mid V=1\right)\Pr(V=1)}{\Pr(\mathcal{F}_t)} \\ &= \frac{\sum_i \frac{\Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right)}{\Pr(V=1)}\Pr(V=1)}{\Pr(\mathcal{F}_3)} = \frac{\sum_i \Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right)}{\Pr(\mathcal{F}_t)} \end{aligned} \tag{14}$$

Using the fact that $\Pr\left(\mathcal{F}_t \mid \mathcal{G}_t^i\right) = 1$,

$$\Pr\left(\mathcal{G}_t^i \mid \mathcal{F}_t\right)\Pr(\mathcal{F}_t) = \Pr\left(\mathcal{F}_t \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right) = \Pr\left(\mathcal{G}_t^i\right).$$

It follows that

$$\frac{\Pr\left(\mathcal{G}_t^i\right)}{\Pr(\mathcal{F}_t)} = \Pr\left(\mathcal{G}_t^i \mid \mathcal{F}_t\right) \tag{15}$$

and we can re-write (14) as

$$V_{t-1}^m = \sum_i \Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i \mid \mathcal{F}_t\right). \tag{16}$$

To ease notation, we use $a_t = B$ to indicate a buy. Using the law of conditional expectations and Bayes' rule we can write (4) as

$$\begin{aligned} \Pr(V=1 \mid a_t = B, \mathcal{F}_t) &= \frac{\Pr(\mathcal{F}_t, a_t = B, V=1)}{\Pr(a_t = B, \mathcal{F}_t)}, \\ &= \frac{\Pr(a_t = B, \mathcal{F}_t, V=1)}{\Pr(a_t = B, \mathcal{F}_t, V=1) + \Pr(a_t = B, \mathcal{F}_t, V=0)}, \\ &= \frac{\Pr(\mathcal{F}_t \mid a_t = B, V=1)\Pr(a_t = B, V=1)}{\Pr(\mathcal{F}_t \mid a_t = B, V=1)\Pr(a_t = B, V=1) + \Pr(\mathcal{F}_t \mid a_t = B, V=0)\Pr(a_t, V=0)}. \end{aligned}$$

Since $\left\{\mathcal{G}_t^i\right\}_i$ 'partitions' $\mathcal{F}_t$,

$$\begin{aligned} &\Pr(V=1 \mid a_t = B, \mathcal{F}_t) \\ &= \frac{\sum_{i=1}^{2^t} \Pr\left(\mathcal{G}_t^i \mid a_t = B, V=1\right)\Pr(a_t = B, V=1)}{\sum_{i=1}^{2^t}\left[\Pr\left(\mathcal{G}_t^i \mid a_t = B, V=1\right)\Pr(a_t = B, V=1) + \Pr\left(\mathcal{G}_t^i \mid a_t = B, V=0\right)\Pr(a_t = B, V=0)\right]}, \\ &= \frac{\sum_{i=1}^{2^t} \Pr\left(a_t = B \mid V=1, \mathcal{G}_t^i\right)\Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right)}{\sum_{i=1}^{2^t}\left[\Pr\left(a_t = B \mid V=1, \mathcal{G}_t^i\right)\Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right) + \Pr\left(a_t = B \mid V=0, \mathcal{G}_t^i\right)\Pr\left(V=0 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\right)\right]}. \end{aligned}$$

Dividing both the numerator and the denominator by $\Pr(\mathcal{F}_t)$ and using (14), the numerator can be written as

$$\sum_{i=1}^{2^t} \Pr\left(a_t = B \mid V=1, \mathcal{G}_t^i\right)\Pr\left(V=1 \mid \mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i \mid \mathcal{F}_t\right). \tag{17}$$

On $\mathcal{G}_t^{j^*}$, the probability of a buy order differs from the probability in other type-signal paths, as traders herd and buy regardless of their signal. In particular,

$$\begin{aligned} \Pr\left(a_t = B \mid V=1, \mathcal{G}_t^i\right) &= \left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\,p\right] \text{ for } i \neq j^*, \\ \Pr\left(a_t = B \mid V=1, \mathcal{G}_t^i\right) &= \left[\frac{\lambda\mu}{2} + (1-\lambda\mu)(p+1-p)\right] \text{ for } i = j^*. \end{aligned}$$

We can the rewrite (17) as

$$\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]\underbrace{\sum_{i=1}^{2^t}\Pr\left(V=1\mid\mathcal{G}_t^i\right)\Pr\left(\mathcal{G}_t^i\mid\mathcal{F}_t\right)}_{V_{t-1}^m}+(1-\lambda\mu)\,(1-p)\underbrace{\Pr\left(V=1\mid\mathcal{G}_t^{j^*}\right)}_{V_{t-1}^B}\Pr\left(\mathcal{G}_t^{j^*}\mid\mathcal{F}_t\right).$$

It follows that

$$\begin{aligned}&\Pr\left(V=1\mid a_t=B,\mathcal{F}_t\right)\\ =\;&\frac{\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m+(1-\lambda\mu)\,(1-p)\,V_{t-1}^B\eta_t^B}{\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m+(1-\lambda\mu)\,(1-p)\,V_{t-1}^B\eta_t^B+\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m+(1-\lambda\mu)\,p\left(1-V_{t-1}^B\right)\eta_t^B}\end{aligned}$$

where $\eta_t^B=\Pr\left(\mathcal{G}_t^{j^*}\mid\mathcal{F}_t\right)$ is the probability of the path leading to herding/contrarianism. Q.E.D.

A.2 Proof of Proposition 2

$V_t^i\left(\mathcal{G}_t^j,\sigma^i=L\right)>A_t^{Naive}$ is equivalent to

$$\frac{(1-p)\,V_{t-1}^i}{(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)}>\frac{\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m}{\frac{\lambda\mu}{2}+(1-\lambda\mu)\left[pV_{t-1}^m+(1-p)\left(1-V_{t-1}^m\right)\right]},$$

rearranging,

$$(1-p)\,V_{t-1}^i\left\{\frac{\lambda\mu}{2}+(1-\lambda\mu)\left[pV_{t-1}^m+(1-p)\left(1-V_{t-1}^m\right)\right]\right\}>\left[(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)\right]\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m.$$

Adding $\eta_t\,(1-\lambda\mu)\,(1-p)\,V_{t-1}^i\left\{(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)\right\}$ to both sides,

$$\begin{aligned}&(1-p)\,V_{t-1}^i\left\{\frac{\lambda\mu}{2}+(1-\lambda\mu)\left[pV_{t-1}^m+(1-p)\left(1-V_{t-1}^m\right)\right]+\eta_t\,(1-\lambda\mu)\left[(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)\right]\right\}\\ >\;&\left[(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)\right]\left\{\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m+\eta_t\,(1-\lambda\mu)\,(1-p)\,V_{t-1}^i\right\}.\end{aligned}$$

Rearranging,

$$\begin{aligned}&\frac{(1-p)\,V_{t-1}^i}{(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)}\\ >\;&\frac{\left[\frac{\lambda\mu}{2}+(1-\lambda\mu)\,p\right]V_{t-1}^m+\eta_t\,(1-\lambda\mu)\,(1-p)\,V_{t-1}^i}{\frac{\lambda\mu}{2}+(1-\lambda\mu)\left[pV_{t-1}^m+(1-p)\left(1-V_{t-1}^m\right)\right]+\eta_t\,(1-\lambda\mu)\left\{(1-p)\,V_{t-1}^i+p\left(1-V_{t-1}^i\right)\right\}},\end{aligned}$$

which is equivalent to $V_t^i\left(\mathcal{G}_t^j,\sigma^i=L\right)>E\left[V\mid\mathcal{F}_t,a_t=\text{buy}\right]$. Hence, $A_t=E\left[V\mid\mathcal{F}_t,a_t=\text{buy}\right]$ is a rational expectations price. Q.E.D.

A.3 Proof of Proposition 3

$A_t > A_t^{naive}$ is equivalent to

$$\begin{aligned} & \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right] V_{t-1}^m + \eta_t (1-\lambda\mu)(1-p)\, V_{t-1}^i}{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[pV_{t-1}^m + (1-p)\left(1-V_{t-1}^m\right)\right] + \eta_t (1-\lambda\mu)\left\{(1-p)\, V_{t-1}^i + p\left(1-V_{t-1}^i\right)\right\}} \\ > & \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right] V_{t-1}^m}{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[pV_{t-1}^m + (1-p)\left(1-V_{t-1}^m\right)\right]}. \end{aligned}$$

Rearranging,

$$\begin{aligned} & \left\{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[pV_{t-1}^m + (1-p)\left(1-V_{t-1}^m\right)\right]\right\} \eta_t (1-\lambda\mu)(1-p)\, V_{t-1}^i \\ > & \ \eta_t (1-\lambda\mu)\left\{(1-p)\, V_{t-1}^i + p\left(1-V_{t-1}^i\right)\right\}\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right] V_{t-1}^m, \end{aligned}$$

which is equivalent to

$$\frac{(1-p)\, V_{t-1}^i}{(1-p)\, V_{t-1}^i + p\left(1-V_{t-1}^i\right)} > \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right] V_{t-1}^m}{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[pV_{t-1}^m + (1-p)\left(1-V_{t-1}^m\right)\right]},$$

which concludes the proof. Q.E.D.

A.4 Proof of Theorem 1

Herd buying occurs when $E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i = L\right] > A_t$. By proposition (2), this is equivalent to $E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i = L\right] > A_t^{naive}$ which, in the case of first time herding, can be written as

$$\begin{aligned} & \frac{p^{(h_{t-1}-l_{t-1}-1)}\left[\frac{\mu}{2} + (1-\mu)\, p\right]^{b_{t-1}^i - s_{t-1}^i}}{p^{(h_{t-1}-l_{t-1}-1)}\left[\frac{\mu}{2} + (1-\mu)\, p\right]^{b_{t-1}^i - s_{t-1}^i} + (1-p)^{(h_{t-1}-l_{t-1}-1)}\left[\frac{\mu}{2} + (1-\mu)(1-p)\right]^{b_{t-1}^i - s_{t-1}^i}} \\ > & \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right]^{b_{t-1}^m - s_{t-1}^m + 1}}{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right]^{b_{t-1}^m - s_{t-1}^m + 1} + \left[\frac{\lambda\mu}{2} + (1-\lambda\mu)(1-p)\right]^{b_{t-1}^m - s_{t-1}^m + 1}}. \end{aligned} \tag{18}$$

As long as no path-dependent behavior has occurred yet, given the number of and type of buys and sells, their sequence does not change the market maker's and traders' valuations at t. Setting $\delta_t = h_{t-1} - l_{t-1}$, $\delta_t^i = b_{t-1}^i - s_{t-1}^i$, $\delta_t^m = b_{t-1}^m - s_{t-1}^m$ and $\gamma_t = \frac{\delta_t - 1}{\delta_t + 1}$ and

$$K(p) = \left[\frac{\left[\frac{\mu}{2} + (1-\mu)\, p\right]\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)(1-p)\right]}{\left[\frac{\mu}{2} + (1-\mu)(1-p)\right]\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right]}\right]^{\frac{\delta_t^i}{\delta_t + 1}},$$

the previous condition is equivalent to

$$\Delta\left(p, \lambda, \mu, \delta_t, \delta_t^i\right) = p^{\gamma_t}\frac{\lambda\mu}{2}K(p) + p^{\gamma_t}(1-p)(1-\lambda\mu)\, K(p) - (1-p)^{\gamma_t}\frac{\lambda\mu}{2} - (1-p)^{\gamma_t}(1-\lambda\mu)\, p > 0.$$

Notice that $0 < K(p) \leq 1$ for every $p \in \left[\frac{1}{2}, 1\right]$. As $K\left(\frac{1}{2}\right) = 1$, we have that $\Delta\left(\frac{1}{2}\right) = 0$ and $\Delta(1) > 0$ for every λ, μ, δ_t and δ_t^i. The derivative of Δ with respect to p is

$$\begin{aligned}\Delta'\left(p,\mu,\delta_t,\delta_t^i\right) &= \gamma_t p^{\gamma_t-1}\frac{\lambda\mu}{2}K(p)+p^{\gamma_t}\frac{\lambda\mu}{2}K'(p)+\gamma_t p^{\gamma_t-1}(1-p)(1-\lambda\mu)K(p)-p^{\gamma_t}(1-\lambda\mu)K(p)\\ &\quad +p^{\gamma_t}(1-p)(1-\lambda\mu)K'(p)+\gamma_t(1-p)^{\gamma_t-1}\frac{\lambda\mu}{2}-(1-p)^{\gamma_t}(1-\lambda\mu)+\gamma_t(1-p)^{\gamma_t-1}(1-\lambda\mu)p.\end{aligned}$$

At $p=\frac{1}{2}$ this is equal to

$$\Delta'\left(\frac{1}{2},\lambda,\mu,\delta_t,\delta_t^i\right)=\left(\frac{1}{2}\right)^{\gamma_t-1}\left\{\gamma_t-(1-\lambda\mu)-\frac{\delta_t^i}{\delta_t+1}\mu(1-\lambda)\right\},$$

whose sign behaves as follows:

$$\begin{aligned}\Delta'\left(\frac{1}{2}\right) &\geq 0 \quad \text{when } \delta_t\geq\frac{2}{\lambda\mu}-1+\delta_t^i\left(\frac{1}{\lambda}-1\right),\\ \Delta'\left(\frac{1}{2}\right) &< 0 \quad \text{when } \delta_t<\frac{2}{\lambda\mu}-1+\delta_t^i\left(\frac{1}{\lambda}-1\right).\end{aligned}$$

The function $\Delta\left(p,\lambda,\mu,\delta_t,\delta_t^i\right)=0$ implicitly defines $\delta_t=\varphi_t\left(p,\lambda,\mu,\delta_t^i\right)$, everywhere but at $\left(\frac{1}{2},\lambda,\mu,\delta_t^i\right)$. We can find $\varphi_t\left(p,\lambda,\mu,\delta_t^i\right)$ explicitly:

$$\varphi_t\left(p,\lambda,\mu,\delta_t^i\right)=\frac{L+L^\lambda}{L-L^\lambda}+\delta_t^i\frac{L^\lambda-L^\mu}{L-L^\lambda}. \tag{19}$$

Define $l=\partial L/\partial p$. Differentiating (19) we respect to p we find:

$$\frac{\partial\varphi_t}{\partial p}=\frac{\left(l^\lambda+l\right)\left(L-L^\lambda\right)-\left(L^\lambda+L\right)\left(l-l^\lambda\right)+\delta_t^i\left[\left(l^\lambda-l^\mu\right)\left(L-L^\lambda\right)-\left(l-l^\lambda\right)\left(L^\lambda-L^\mu\right)\right]}{\left(L-L^\lambda\right)^2}$$

The sing of $\frac{\partial\varphi_t}{\partial p}$ depends on the value of δ_t^i. In particular, $\frac{\partial\varphi}{\partial p}$ is positive if and only if

$$\delta_t^i<2\frac{lL^\lambda-l^\lambda L}{l^\lambda L-lL^\lambda-l^\mu\left(L-L^\lambda\right)+L^\mu\left(l-l^\lambda\right)}=D(p,\lambda,\mu),$$

where $D(p,\lambda,\mu)<0$ for every p, λ and μ and $\frac{\partial D(p,\lambda,\mu)}{\partial p}>0$.

Case 1: $\delta_t<\frac{2}{\lambda\mu}-1+\delta_t^i\left(\frac{1}{\lambda}-1\right)$. When $p=\frac{1}{2}$ we found that $\Delta'\left(\frac{1}{2}\right)<0$. Since $\Delta(1)>0$ for every λ, μ, δ_t and δ_t^i, we can conclude that $\Delta\left(p,\lambda,\mu,\delta_t,\delta_t^i\right)$ cuts the x-axis at least once. Suppose that $\Delta\left(p^*,\lambda,\mu,\delta_t,\delta_t^i\right)=0$ and that $\delta_t^i<D(p^*,\lambda,\mu)$. This means that if we increase p from p^* to $p'\in(p^*,1]$, the level of δ_t needed to keep Δ at zero increases. As $\Delta\left(p,\lambda,\mu,\delta_t,\delta_t^i\right)$ is increasing in δ_t, this means that, for given δ_t and δ_t^i, if Δ is cutting the x-axis at p^* it must be doing it from above. However, as $\Delta(1)>0$, it must be the case that it is cutting it again from below. This last fact is not possible, as $\frac{\partial D(p,\lambda,\mu)}{\partial p}>0$, which implies that for any other $p'\in(p^*,1]$, we continue to have $\delta_t^i<D(p',\lambda,\mu)$: any crossing of the x-axis as p increases must occur from above. We conclude that we must have $\delta_t^i>D(p^*,\lambda,\mu)$, and any crossing of the x-axis for $p\in\left(\frac{1}{2},1\right]$ must be occurring once and from below.

Case 2: $\delta_t\geq\frac{2}{\lambda\mu}-1+\delta_t^i\left(\frac{1}{\lambda}-1\right)$. As $\Delta'\left(\frac{1}{2}\right)\geq 0$ and $\Delta(1)>0$, and given that we have just established that $\Delta\left(p,\mu,\lambda,\delta_t,\delta_t^i\right)$ can never cross the x-axis from above, it follows that $\Delta\left(p,\mu,\lambda,\delta_t,\delta_t^i\right)>0$ for every $p\in\left(\frac{1}{2},1\right]$.

The analysis of the previous two cases shows that path-dependent buying occurs either for high values or for any value of the signal precision. If the conditions on the type-signal path are satisfied and this type-signal path leads to $b_t^m>s_t^m$ ($b_t^m<s_t^m$) we incur in herd (contrarian) buying.

To study the effect of λ on p^*, notice that $\Delta>0$ if and only if $\Gamma=(\delta_t-1)L+\delta_t^iL^\mu-(\delta_t^m+1)L^\lambda>0$. As L^λ is decreasing in λ, Γ increases in λ whenever $b_t^m<s_t^m$, and it decreases in λ when $b_t^m>s_t^m$, for every $p\in\left(\frac{1}{2},1\right]$. It follows that $\frac{\partial p^*}{\partial\lambda}>0$ when $b_t^m<s_t^m$, and that $\frac{\partial p^*}{\partial\lambda}<0$ when $b_t^m>s_t^m$. Moreover, when condition (11) is satisfied, Γ is increasing in μ for every $p\in\left(\frac{1}{2},1\right]$. It follows that $\frac{\partial p^*}{\partial\mu}<0$.

A.5 Proof of Theorem 2

Consider $p = \frac{1}{2}$. Then both the market maker and the traders' valuations will be equal to $\frac{1}{2}$ for every $\mathcal{F}_t$, every σ and every $\lambda, \mu \in \left[\frac{1}{2}, 1\right]$. Since we have defined herd buying only in the case where buying is strictly preferred to any other actions, then herding is not possible.

Consider $p = 1$. Then for any $\mathcal{F}_t$ and every $\lambda, \mu \in \left[\frac{1}{2}, 1\right]$, $E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i\right] = 1$ if $\sigma^i = H$ (and $\mathcal{F}_t$, $E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i\right] = L$ if $\sigma^i = L$) and no herding is possible.

For $p \in \left(\frac{1}{2}, 1\right)$, we have from (18) that herd buying occurs at t whenever

$$h_{t-1} - l_{t-1} > \frac{L + L^\lambda}{L - L^\lambda} + \left(b^i_{t-1} - s^i_{t-1}\right) \frac{L^\lambda - L^\mu}{L - L^\lambda}$$

whose left hand side is maximized for $l_{t-1} = 0$ and whose right hand side it minimized for $b^i_{t-1} = 0$. The path just described, where all buys are type I and the sells are type II, is a unique type-signal path compatible with $\mathcal{F}_t$, and the one with the highest possible traders' valuation of the asset associated to it. It follows that herd buying has positive probability of happening at t whenever

$$b^m_{t-1} > \frac{L + L^\lambda}{L - L^\lambda} - s^m_{t-1} \frac{L^\lambda - L^\mu}{L - L^\lambda} \tag{20}$$

We cannot appeal to noise trading in order to say that a trading history such that the previous inequality is satisfied has positive probability and this is enough to prove existence of herding because if, in fact, the trading history were generated by at least one noise buy, then herding could not occur anymore. However, if the market maker assigns positive probability to herd buying given the trading history, then consistency implies that, from an ex-ante perspective, this probability must be positive.

A.6 Proof of Proposition 4

The statement is certainly satisfied when the possibility of herding arises for the first time, since all informed traders agree on $E\left[V \mid \mathcal{G}^j_t\right]$ for some $\mathcal{G}^j_t$. Suppose that herding, in fact, occurs at t: then traders will not update and $V^i_t = E\left[V \mid \mathcal{G}^j_{t+1}\right] = E\left[V \mid \mathcal{G}^j_t\right]$ for every i. Hence, if there is still a possibility for herding at $t+1$, traders will all agree on the value V^i_t.

Suppose, instead that we were not on $\mathcal{G}^j_t$, that there was no herding at t and that we are in fact on $\mathcal{G}^k_t$. It is obvious that $E^i\left[V \mid \mathcal{G}^k_t\right] < E^i\left[V \mid \mathcal{G}^j_t\right]$. If that were not the case and $E^i\left[V \mid \mathcal{G}^k_t\right] \geq E^i\left[V \mid \mathcal{G}^j_t\right]$ herding would have occurred on $\mathcal{G}^k_t$ to start with. If a type I buy occurs at t leading to the possibility of herding at $t+1$ on the k-path, we have $E^i\left[V \mid \mathcal{G}^k_t, \sigma^i = H\right] = E^i\left[V \mid \mathcal{G}^j_t\right] = E^i\left[V \mid \mathcal{G}^j_{t+1}\right]$, so herding at $t+1$ happens for the same traders' prior on the $\mathcal{G}^k_t$ and on the $\mathcal{G}^j_t$ paths.

Suppose that herding is again possible at $t+2$ on both $\mathcal{G}^k_{t+2}$ and $\mathcal{G}^j_{t+2}$ and suppose we are in fact on $\mathcal{G}^l_{t+2}$. It follows that $E^i\left[V \mid \mathcal{G}^l_t\right] < E^i\left[V \mid \mathcal{G}^k_t\right] < E^i\left[V \mid \mathcal{G}^j_t\right]$ must hold. If then herding is possible on $\mathcal{G}^l_{t+2}$, it must be the case that $E^i\left[V \mid \mathcal{G}^l_{t+2}\right] = E^i\left[V \mid \mathcal{G}^k_{t+2}\right] = E^i\left[V \mid \mathcal{G}^k_{t+1}\right] = E^i\left[V \mid \mathcal{G}^j_t\right]$. And so on. Q.E.D.

A.7 Proof of Theorem 3

Herd buying occurs when $E\left[V \mid \mathcal{F}_t, \mathcal{T}_t, \sigma^i = L\right] > A_t$. This is equivalent to

$$\begin{aligned} & \frac{(1-p)\, V^i_{t-1}}{(1-p)\, V^i_{t-1} + p\left(1 - V^i_{t-1}\right)} \\ > \; & \frac{\left[\frac{\lambda\mu}{2} + (1-\lambda\mu)\, p\right] V^m_{t-1} + \eta_t\, (1-\lambda\mu)\, (1-p)\, V^i_{t-1}}{\frac{\lambda\mu}{2} + (1-\lambda\mu)\left[p V^m_{t-1} + (1-p)\left(1 - V^m_{t-1}\right)\right] + \eta_t\, (1-\lambda\mu)\left[(1-p)\, V^i_{t-1} + p\left(1 - V^i_{t-1}\right)\right]}. \end{aligned}$$

We can re-write the above inequality as:

$$\Delta_t(p) = p^2\left(V_{t-1}^i - V_{t-1}^m\right)(1-\lambda\mu) + \left(V_{t-1}^i - V_{t-1}^i V_{t-1}^m\right)\left(1 - \frac{\lambda\mu}{2}\right)$$
$$+p\frac{\lambda\mu}{2}\left(V_{t-1}^i - V_{t-1}^m\right) - 2p\left(V_{t-1}^i - V_{t-1}^m V_{t-1}^i\right)\left(1 - \frac{\lambda\mu}{2}\right) > 0.$$

It is easy to check that $\Delta_t(1) < 0$ and $\Delta_t\left(\frac{1}{2}\right) > 0$, and that $\Delta_t'(p) < 0$ for every $p \in \left[\frac{1}{2}, 1\right]$. We can conclude that there exists a cut-off $\overline{p}\left(V_{t-1}^i, V_{t-1}^m, \lambda, \mu\right)$ such that herd buying occurs whenever $p < \overline{p}\left(V_{t-1}^i, V_{t-1}^m, \lambda, \mu\right)$ .Q.E.D.

References

[1] AVERY, C. and ZEMSKY, P. (1998) "Multidimensional uncertainty and herd behavior in financial markets", *American Economic Review*, **88**, 724-48.

[2] BIKHCHANDANI, S., HIRSHLEIFER, D. and WELCH, I. (1992) "A theory of fads, fashion, custom and cultural change in informational cascades", *Journal of Political Economy*, **100**, 992-1026.

[3] CHAMLEY, C. P. (2004) "*Rational Herds: Economic Models of Social Learning*", Cambridge University Press.

[4] GLOSTEN, L. R. and MILGROM, P. R. (1985) "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders", *Journal of Financial Economics*, **14**, 71-100.

[5] Andreas PARK, A. and SABOURIAN, H. (2011) "Herding and Contrarian Behavior in Financial Markets", *Econometrica*, **79**, 973-1026.

[6] VIVES, X. (2008) *Information and Learning in Markets*, Princeton University Press.