Estimating a Structural Model of Herd Behavior in Financial Markets

Marco Cipriani, and Antonio Guarino*

June 10th, 2006

Abstract

We develop and estimate a structural model of informational herding in financial markets. In the model, a sequence of traders exchanges an asset with a market maker. Herd behavior, i.e, the choice to follow the actions of one's predecessors, can arise as the outcome of a rational choice because there are multiple sources of asymmetric information in the economy. We estimate the model using transaction data on a NYSE stock in the first quarter of 1995. We are able to detect the periods of the trading day in which traders herd, and find that they account for 15% of trading periods. Moreover, we find that in more than 10% of days, herding account*s* for more than 30% of all trading activity. Finally, by simulating the model, we estimate the informational inefficiency generated by herding. On average, because of herding, the actual price is 0.4% distant from the full informtion price. Moreover, in 2.8% of trading periods, the distance between actual and full information prices is larger than 10%. This suggests that the informational inefficiency caused by herding, although not extremely large on average, is very significant in certain days.

*Cipriani: Department of Economics, George Washington University (e-mail: mciprian@gwu.edu); Guarino: Department of Economics and ELSE, University College London (e-mail: aguarino@ucl.ac.uk). We thank Orazio Attanasio, Andrew Chesher, Douglas Gale, Robert Engle, Christopher Flinn, Lars Nesheim, Henri Pages, Nicola Pavoni and Jean Marc Robin for very helpful comments and suggestions. We gratefully acknowledge the financial support of the Fondation Banque de France. Guarino gratefully acknowledges the financial support of the ESRC and of the Leverhulme Trust. We are responsible for any error.

1 Introduction

In recent years there has been much interest in herd behavior in financial markets. Especially after the financial crises of the 1990s, many scholars have suggested that herd behavior may be a reason for excess price volatility and financial systems fragility. This interest has led researchers to look for both theoretical explanations and empirical evidence.[1]

The first theoretical work on herd behavior dates from the beginning of the 90's with the seminal papers of Banerjee (1992), Bikhchandani et al. (1992), and Welch (1992). These papers do not discuss herd behavior in financial markets, but in an abstract environment, in which agents with private information make their decisions in sequence. They show that, after a finite number of agents have chosen the same action, all following agents will disregard their own private information and imitate that action. More recently, a number of papers (see, e.g., Avery and Zemsky, 1998, Lee, 1998, Cipriani and Guarino, 2001, Cipriani and Guarino, 2006) have focused on herd behavior in financial markets. In particular, all these studies analyze a market where agents sequentially trade a security of unknown value. The price of the security is efficiently set by a market maker according to the order flow. The presence of a price mechanism makes herding more difficult to arise. Still there are cases in which it occurs. In Avery and Zemsky (1998) people can herd when there is uncertainty not only on the value of the asset but also on other parameters of the model. In Cipriani and Guarino (2001) agents herd because they have other reasons to trade, in addition to informational motives. In this model, not only agents herd, but, in contrast with Avery and Zemsky, a complete informational cascade arises.[2] Whereas the theoretical research has tried to identify the mechanisms through which herd behavior can arise, the empirical literature has followed a different track. The existing work (see, e.g., Lakonishok et al., 1992, Grinblatt et al. , 1995, Wermers, 1999, and the other papers cited in the survey of Hirshleifer and Teoh, 2003) does not test these models directly, but analyzes the presence of

[1] In this paper we only study informational herding. Therefore, we do not discuss herd behavior due to reputational concerns or payoff externalities. For critical surveys of the literature on herd behavior see Gale (1996), Hirshleifer and Teoh (2003), and the book by Chamley (2004).

[2] We will define the concept of herd behavior formally later in the paper. Here we note that herd behavior refers to conformity in actions (e.g., all traders buy); informational cascade refers to the actions being completely uninformative to the other agents.

herding in financial markets through statistical measures of clustering. These papers find that in some markets fund managers tend to cluster their investment decisions more than if they acted independently. The existing empirical research on herding is important, as it sheds light on the behavior of financial market participants and in particular on whether they act in a coordinated fashion. As the authors themselves emphasize, however, decision clustering may or may not be due to herding (for instance, it may be the result of a common reaction to public announcements). These papers cannot distinguish spurious herding from true herd behavior, i.e., the decision to disregard one's private information to follow the behavior of others (see Bikhchandani and Sharma, 2000, and Welch, 2000). Testing informational models of herd behavior is a difficult task. In these models herding in the financial market consists in trading independently of private information. The problem that empiricists face in the task of detecting herding is that there are no data on the private information available to the traders and, therefore, it is difficult to understand whether traders make similar decisions because they disregard their own information and imitate or because of other reasons.

The purpose of this paper is to overcome this problem and offer an empirical analysis of herd behavior which is not purely statistical. We present a theoretical model of herding and estimate it using financial market transaction data. We are able to identify the periods in the trading day in which traders act as herders according to the model. This is the first empirical paper on informational herding that, instead of using a statistical, a-theoretical approach, estimates a theoretical model.[3]

Our theoretical analysis is inspired by the work of Avery and Zemsky (1998), who use a sequential trading model *à la* Glosten and Milgrom (1985) to show the conditions under which herding can arise in financial markets. In their model, traders trade an asset of unknown value with a market maker. Traders receive private information on it. The market maker is uninformed and sets the price of the asset on the basis of the buy and sell orders that he receives. Avery and Zemsky show that, if the private information only concerns the asset fundamental value, traders will always find it optimal to trade on the difference between their own information (the history of trades and the private signal) and the commonly available information (the history

[3] While there are no direct empirical tests of herding models, there is experimental work that tests these models in the laboratory: see Cipriani and Guarino (2005) and Drehman et al (2005).

only). Therefore, it will never be the case that agents neglect their information and imitate previous traders' decisions. They also show, however, that when are multiple sources of asymmetric information between the traders and the market maker (i.e., asymmetric information not only on the asset value but also on other model parameters) herd behavior arises.

In our model, herding arises for a mechanism similar to that exposed by Avery and Zemsky. Whereas Avery and Zemsky were interested in providing some theoretical examples in which herd behavior can arise, our aim is to estimate the importance of herding in real financial market. For this purpose, we build a financial market model of herding that can be estimated using market data. An asset is traded over many days. At the beginning of each day, an informational event can occur or not. In the former case, the fundamental asset value changes with respect to the value of the previous day. It can be higher (in the case of a good informational event) or lower (in the case of a bad informational event) than the value in the previous day. In the latter case, instead, it remains unchanged. If an event has occurred, some agents will receive private information on the new asset value.[4] These agents will trade the asset to exploit their informational advantage on the market maker. On the contrary, if no event has occurred, all traders in the market will be uninformed: they will trade for non-information reasons only, e.g., liquidity or hedging motives. While the informed traders know that they are in a market in which there is private information (since they themselves are informed) the market maker does not know whether he is in an informed or uninformed market for that day. This asymmetry in information determines a different way of updating the beliefs on the asset value by the traders and the market maker. The market maker will move the price "slowly" since he has to take into account the possibility that the asset value has not changed, the market is uninformed and all orders are coming from uninformed (noise) traders. His interpretation of the history of trades will be different from the traders'. There can be times in which, irrespective of his private signal, a trader will value the asset more than the market maker and, therefore, will find it optimal to buy; or he will value the asset less than the market maker and will find it optimal to sell. These are the periods when herd buy or herd sell arises.

To estimate our model, we will use a strategy proposed by Easley, Kiefer

[4] The event is called "informational" precisely because some traders in the market will receive private information on it.

and O'Hara (1997). They show how to use transaction data to estimate the parameters of the Glosten and Milgrom model with maximum likelihood. We will construct the likelihood function for the trading of an asset over many days. Our function takes into account that for some histories of trade agents herd. This means that in these histories, the probability of a buy or a sell order will be different from that when each agent follows his own private information. Our task will be more complicated than Easley, Kiefer and O'Hara's. In their set up, informed traders are perfectly informed on the value of the asset. Given that their signal is perfectly informative, these traders' decisions will never be affected by the previous decisions and they will never herd. Therefore, the only thing that matters is the *number* of buys, sells and no trades during the day. The *sequence* in which orders arrive is irrelevant. In contrast, in our framework, history matters. Sequences with the same number of buys and sells may occur with different probabilities depending on the order in which buy and sell decisions arrive in the market. Therefore, we cannot limit our analysis to the number of orders but we have to consider the entire sequence of trades. Differently from Easley, Kiefer and O'Hara we will estimate the model through a Bayesian analysis, i.e., we will start from some prior for the parameters and we will compute their posteriors conditional on the data.

We applied our methodology to the trading activity of Ashland Oil stock, a stock traded in the NYSE, in the first quarter of 1995. We find that 7.3% of the trading periods are characterized by herd buy and 8% by herd sell. In 15 (out 63) days of trade, herding periods were higher than 30% of the total periods of trade. Therefore, although, on average herding is not so common in the market, there are particular days in which it heavily characterizes traders' behavior. In those days, herding behavior generated a significant deviation of the price path from the full information level. Through simulation results, we find that, because of herding, 2.8% of the times the price is more than 10% farther away from the price that would have prevailed if traders never herded.

The paper is organized as follows. Section 2 describes the theoretical model. Section 3 characterizes herd behavior and shows how it can arise. Section 4 describes the estimation strategy. Section 5 describes the data. Sections 6 presents the results, and Section 7 concludes.

2 The model

Our model is based on Glosten and Milgrom (1985) and on Easley and O'Hara (1992), who generalize Glosten and Milgrom to an economy where trading happens over many days.

In our economy there is one asset traded by a sequence of traders who interact with a market maker. Trade occurs over many trading days, indexed by $d = 1, 2, 3, ...$Time within each day is discrete and represented by a countably infinite set of trading dates indexed by $t = 1, 2, 3, ...$

The asset

The fundamental value of the asset in day d, V_d, is a random variable. Such a value does not change *during* the day. In day 1 the value V_1 is distributed on the support $\{v_1^L, v, v_1^H\}$ with the following probabilities: $\Pr(V_1 = v) = (1 - \alpha)$, $\Pr(V_1 = v_1^L) = \alpha(1 - \delta)$, and $\Pr(V_1 = v_1^H) = \alpha\delta$ (with $0 < \alpha < 1$ and $0 < \delta < 1$). In any future day $d \geq 2$, the asset value can remain the same as in the previous day, or change. In particular, V_d is equal to v_{d-1} with probability $(1 - \alpha)$ and changes with probability α. In the latter case, it decreases to the value $v_{d-1} - \Delta L$ with probability $\alpha(1-\delta)$, and increases to $v_{d-1} + \Delta^H$ with probability $\alpha\delta$, where $\Delta L > 0$ and $\Delta^H > 0$. Note that we are assuming that informational events are independent over days. To alleviate notation, we define $v_d^H := v_{d-1} + \Delta^H$ and $v_d^L := V_{d-1} - \Delta L$. Finally, we assume that $(1 - \delta)\Delta L = \delta\Delta^H$ As will become clear in the next pages, we required such a condition for the price to be a martingale.

The market

The asset is exchanged in a specialist market. Its price is set by a competitive market maker (the specialist) who interacts with a sequence of traders. At any time $t = 1, 2, 3...$during the day a trader is randomly chosen to act and can buy, sell or decide not to trade. Each trade consists of the exchange of one unit of the asset for cash. The trader's action space is, therefore, $\mathcal{A} = \{buy, sell, no\ trade\}$. We denote the action of the trader at time t in day d by X_t^d. Moreover, we denote the history of trades and prices until time $t - 1$ by H_t^d.

The market maker

At any time t of day d, the market maker sets the prices at which a trader can buy or sell the asset. When posting these prices, he must take into account the possibility of trading with agents who have some private information on the asset value. He will set different prices for buying and for selling, i.e., there will be a bid-ask spread. We denote the ask price (i.e., the

price at which a trader can buy) at time t by a_t^d and the bid price (i.e., the price at which a trader can sell) by b_t^d.

Due to unmodeled potential competition, the market maker makes zero expected profits by setting the ask and bid prices equal to the expected value of the asset conditional on the information available at time t and on the chosen action, i.e.,

$$\begin{aligned} a_t^d &= E(V_d|h_t^d, X_t^d = buy, a_t^d, b_t^d), \\ b_t^d &= E(V_d|h_t^d, X_t^d = sell, a_t^d, b_t^d). \end{aligned}$$

We also define the "price" of the asset at time t as the market maker's expected value of the asset *before* the time-t trader has traded, i.e., $p_t^d = E(V_d|h_t^d)$.[5]

The traders

There are a countably infinite number of traders. Traders act in an exogenously determined sequential order. Each trader is chosen to take an action only once. Traders can be of two types, informed and noise. The trader's type is not publicly know, i.e., it is his private information.

Noise traders trade for unmodeled (e.g., liquidity) reasons: they buy with probability $\frac{\varepsilon}{2}$, sell with probability $\frac{\varepsilon}{2}$ and do not trade with probability $(1-\varepsilon)$ (with $0 < \varepsilon < 1$).

Informed traders have private information on the asset value. They are present in the market in a day d only if an event occurred at the beginning of the day that made the asset value go up or down with respect to the previous day. Informed traders receive a private binary signal on the new asset value and maximize their expected profit based on that signal (i.e., they are risk neutral). The signal is a random variable S_t^d distributed on $\{s^L, s^H\}$. We denote the conditional probability function of S_t^d given a realization v_d of V_d by $\sigma(s_t^d|v_d)$. We assume that, conditional on the asset value v_d, the random variables S_t^d are independent and identically distributed across time. In particular, we assume that

$$\sigma(s^L|v_d^L) = \sigma(s^H|v_d^H) = q \in (0.5, 1).$$

[5] We use capital letters for random variables and small letters for their realizations. For instance, h_t^d is a particular history realized at time t, while H_t^d denotes any possible history until that time. Furthermore, from now on, to simplify the notation, we write $E(.|y)$ to mean $E(.|Y = y)$, i.e., the expected value conditional on the realization y of the random variable Y.

An informed agents knows that an event has occurred and his signal is informative on whether the event is good or bad. Nevertheless, he is not completely sure of the effect of the event on the asset value. For instance, he knows that there has been a change in the investment strategy of a company, but he cannot be completely sure that this change will affect the asset value in a positive or negative way. The precision of the signal (q) can also be interpreted as measuring the ability of traders to process the private information that they receive.

When no event occurs, there is nothing agents can learn in that day and, therefore, all agents in the market are noise. In contrast, when there is an event, some agents receive a private signal on the new asset value and go the market to exploit it. In this case, the proportion of informed traders is $\mu \in (0,1)$. At each time t in an informed day, an informed trader is chosen with probability μ and a noise trader with probability $1-\mu$.

In addition to his signal, a trader at time t observes the history of trades and prices and the current price. Therefore, his expected value of the asset is $E(V_d|h_t^d, s_t^d)$.

His payoff function $U : \{v_d^L, v_{d-1}, v_d^H\} \times \mathcal{A} \times [v_d^L, v_d^H]^2 \longrightarrow \mathbf{R}^+$ is defined as

$$U(v_d, X_t^d, a_t^d, b_t^d) = \begin{cases} v_d - a_t^d & \text{if } X_t^d = buy, \\ 0 & \text{if } X_t^d = no\ trade, \\ b_t^d - v_d & \text{if } X_t^d = sell. \end{cases}$$

The trader chooses X_t^d to maximize $E(U(V_d, X_t^d, a_t^d, b_t^d)|h_t^d, s_t^d)$. Therefore, he finds it optimal to buy whenever $E(V_d|h_t^d, s_t^d) > a_t^d$, and sell whenever $E(V_d|h_t^d, s_t^d) < b_t^d$. He chooses not to trade when $b_t^d < E(V_d|h_t^d, s_t^d) < a_t^d$. Finally, he is indifferent between buying and no trading when $E(V_d|h_t^d, s_t^d) = a_t^d$ and between selling and no trading when $E(V_d|h_t^d, s_t^d) = b_t^d$.

2.1 Herd Behavior

Let us discuss now the predictions of our model. We start by considering the equilibrium prices.

Proposition 1 *At any time t, there exists a unique bid and ask price for the asset, which satisfies $b_t^d \leq p_t^d \leq a_t^d$.*

Proof. See the Appendix.

The market maker takes into account that buying or selling orders contain private information and sets a spread between the price at which he is willing

to sell and buy (Glosten and Milgrom, 1985). Equilibrium prices always exist because noise traders are willing to accept any loss and, therefore, the market will never shut down.

In order to discuss how herding arises in the financial market, let us introduce the formal definition of herd behavior.

Definition 1 *There is herd buying in equilibrium at time t of day d when, if an informed trader trades at that time, he buys with probability* 1*, i.e.,* $E(V_d|h_t^d, s_t^d) > a_t^d$*, for any* $s_t^d \in \{s^L, s^H\}$*. There is herd selling in equilibrium at time t of day d when, if an informed trader trades at that time, he sells with probability* 1*, i.e.,* $E(V_d|h_t^d, s_t^d) < b_t^d$*, for any* $s_t^d \in \{s^L, s^H\}$*.*

Definition 2 *There is herd behavior in equilibrium between times t' and t'' ($t'' > t'$) of day d if, for $t = t', ..., t''$, there is either herd buying or herd selling.*

Herding arises when informed traders act alike, i.e., they choose the same action, independently of their private signal. Our definition of herding is standard in the literature (see, e.g., Gale, 1996, and Smith and Sørensen, 2000). Agents herd when they act independently of their own private information. Because of this, there is conformity of action in the market.

Traders who herd can make the wrong decision. For instance, it is possible that in a trading day characterized by a positive informational event, traders in a period of herding neglect their positive information on the true asset value and decide to sell it. In the cases in which agents neglect the correct signal and take the opposite action, we say that herding is *misdirected*.

We can prove the following proposition:

Proposition 2 *During an informed trading day, herd behavior arises with positive probability. Furthermore, herd behavior can be misdirected.*

Proof. See the Appendix.

Intuitively, the reason for the occurrence of herding is that the price, although efficiently set by the market maker, moves "too slowly." When informed traders and market maker look at the past history of trades, they interpret it in different ways. Suppose that a sequence of buy orders arrives in the market. Informed traders, knowing that there has been an informational

event, attach a certain probability to the fact that these orders come from previous traders with positive signals. The market maker attaches a lower probability to this event, as he has to take into account the possibility that there was no event at the beginning of the day and that all traders in the market trade for non-speculative reasons. Therefore, after a sequence of buys, he will update the price up, but less than if he knew that an event had occurred for sure. As a result, even a trader with negative information can have an expected value of the asset higher than the ask price and, therefore, can neglect his signal to herd buy. The same logic explains herd selling: after a sequence of sells, the bid price may be high enough that also traders with positive information find it optimal to sell. Avery and Zemsky (1998) have analyzed cases in which herd behavior can arise that are similar in the spirit to that presented here.[6] Our analysis differs from theirs in that they discuss herding in the context of theoretical examples that are not suitable for an empirical analysis. But the logic of their argument to explain why traders can neglect private information is similar to that presented here.

The presence of herding in the market is, of course, important for the informational efficiency of prices. During periods of herd behavior, private information is not aggregated by the price. This happens because traders do not make use of the private information they have and, as a result, the price cannot aggregate such information.

While the price does not aggregate private information efficiently, even during a period of herding, the market maker does learn something on the true asset value. Indeed, even in a period of herding, he updates his belief on the fact that there has been an informational event or not, i.e., that the asset value has changed or not. Therefore, the bid and ask prices are updated even in a period of herding. To continue on the same example above, if the market maker sees more traders to buy the asset, he will give more and more weight to the fact that these traders are informed (liquidity traders would indeed buy or sell with the same probability). Hence, he will post higher prices. Because of this movement in prices, herd behavior will eventually disappear. Agents will no longer find it optimal to neglect their signal and private information will be aggregated by the market price.

Essentially, while in our model there is herd behavior, there is no *informational cascade*. An informational cascade requires that the action be

[6] See their IS2 and IS3 information setups.

independent of the asset value.[7] In a situation of informational cascade, the market maker is unable to infer the traders' private information from their actions and, hence, is unable to update his beliefs on the asset value. In other words, in an informational cascade trades do not convey any information on the asset value. This is not the case in our model, since, while informed traders do not use their private signals, still the traders' actions are informative on whether an event occurred or not.

Our model is a theory of temporary intraday herd behavior. In our model, herd behavior arises because in some periods the prices, although efficiently set by the market maker, are such that traders have an incentive to buy (or to sell) independently of their private signals. Figure 1 shows the case in which, following a series of buy orders, buy herding arise. In the figure, we drew the bid and ask prices set by the market maker, the beliefs of a trader with a positive signal and those of a trader with a negative signal. After a long enough series of buys, the prices become lower than the expectations of the trader with a negative signal; when this happens herding arise. The bid and ask prices are lower than the expectations of a trader with a negative signal since traders and market makers interpret the past history of trades differently. In particular, the market maker must take into account the possibility of being in an uninformed day (whereas traders know that there was an information event): in such a day there is no information coming from trades and, therefore, the price should not be revised at all after a buy or a sell. As a result, past buys convey more positive news on the asset values to traders than to the market maker.

During a period of herding, the informed traders do not update their beliefs at all. They already know that they are in an informed day, and they also know that the actions taken are independent of the private signal. The market maker, instead, keeps updating his belief during, since the trades change his posterior probability of whether an information event has occurred. Since the market maker gradually learns that there was an information event, he will gradually start interpreting the history of past trades more and more similarly to the traders: as a result, herding will eventually stop. This is illustrated in Figure 2, where, after some more trades, the bid and ask prices cross the expectation of a trader with a negative signal. When this happens, traders with a negative signal start selling and the herd is broken. This result

[7] Formally, nn informational cascade arises at time t when $\Pr[X_t^d = x | h_t^d, a_t^d, b_t^d, s_t^d] = \Pr[X_t^d = x | h_t^d, a_t^d, b_t^d]$ for all $x \in \mathcal{A}$ and for all $s_t^d \in \{s^L, s^H\}$.

is formalized in the following proposition:

Proposition 3 *Suppose herd behavior starts at time t. Herd behavior cannot last forever, i.e., it stops with probability* 1 *at a time $t' \in (t, \infty)$.*

Proof. See the Appendix.

Of course, during an informed day, herd behavior can start and break more than once, in different times of the day. Nertheless, each period of herding has a limited life.

Given that information always flows to the market (even during time of herding behavior) and given that herding does not last forever, the price is able to aggregate the information that traders receive. Since this information is on average correct, the price will converge to the true asset value.[8] We prove this result formally in the next proposition:

Proposition 4 *In any day d the asset price converges almost surely to the realized value of the asset.*

Proof. See the Appendix.

Although information aggregation is slowed down during periods of herding, eventually the market maker learns whether in a day the market is characterized by a good or bad event or, on the contrary, is uninformed.

3 The Likelihood Function

One of the characteristic of the model is that it is possible to write the likelihood function for the sequence of buys, sells and no trades over many trading days. This likelihood function depends on the five parameters of the model. Let us define the complete history of trades in a single day as $h^d := \left\{h_j^d\right\}_{j=1}^{T_d}$, where T_d is the number of trading periods in day d. We denote the likelihood function by

$$\mathcal{L}(\Psi) = \Pr\left(\{h^d\}_{d=1}^{D} | \Psi\right)$$

[8] Recall that we have assumed that $(1-\delta)\Delta_d^L = \delta\Delta_d^H$. This implies that $E[V_{d+1}|V_d] = V_d$. Since the price converges to the fundamental value almost surely, this guarantees that also the conditional expected price for day $d+1$ be equal to the closing price in day d, i.e., it guarantees that the martingasle property of prices is satisfied.

where $\Psi = \{\alpha, \delta, q, \mu, \varepsilon\}$ is the vector of parameters that characterize our economy. In particular, let us recall that

α is the probability that there is an information event in any given day;

δ is the probability that the information event drives the fundamental value of the asset up;

q is the precision of the signal;

μ is the probability that a trader is informed in an information day;

ε it the probability that an uninformed trader trades.

Note that we write the likelihood function as the conditional probability of the history of trades given the parameters, i.e., we omit the history of (bid and ask) prices. In the model described above there is no public information (all information is private); for this reason there is a one-to-one mapping from trades to prices (and, therefore, the likelihood of trades and the joint likelihood of trades and prices are the same). If we add public information to the model, the one to one mapping from trades to prices breaks down (since prices also reflect the accumulation of public information, whereas trades do not). Nevertheless, since we are only interested in estimating the parameters governing the arrival of private information, we can write the likelihood function only in terms of transactions.

Informational events at the beginning of each day are independent of each other. Trades in a day only depend on the value of the asset that day. For this reason, a particular history of trades over multiple days can be written as the product of the probability of the history of each single day. Then,

$$\mathcal{L}(\Psi) = \Pr\left(\{h^d\}_{d=1}^{D} | \Psi\right) = \Pi_{d=1}^{D} \Pr(h^d | \Psi).$$

In order to understand the whole likelihood function, we must therefore understand the likelihood of each single day. Note that

$$\begin{aligned} \Pr(h^d|\Psi) &= (1-\alpha)\Pr(h^d|V_d = v_{d-1}, \Psi) \\ &+ \alpha(1-\delta)\Pr(h^d|V_d = v_d^L, \Psi) + \alpha\delta \Pr(h^d|V_d = v_d^H, \Psi). \end{aligned}$$

This means that we need to describe the probability of a history of trades in any given day, conditional on that day being a day without an information event or with positive or negative information event. Let us start from no information event days. During a non informed day, the probability of each buy or sell is $\frac{\varepsilon}{2}$, while the probability of each no trade is $(1-\varepsilon)$. Let us denote by B_d, S_d and N_d the number of buys, sells and no trades in day d. The probability of a history in day d can be written as:

$$\Pr(h^d|\Psi, V_d = v_{d-1}) = K\left(\frac{\varepsilon}{2}\right)^{B_d+S_d}(1-\varepsilon)^{N_d},$$

where K is the number of permutations of B_d buys, S_d sells and N_d no trades. From the above formula, it is clear that in a no-event day, we may expect liquidity traders to buy or sell in a balanced way. In contrast, given that informed traders follow an informative signal, when there is an informational event there will be either a prevalence of buys or a prevalence of sells. This is the feature of the trading process that allows us to distinguish between information event day and no-information days.

Let us now compute the likelihood for a day in which there was a positive information event. Since, in a day of information event, the probability of a trade at any time depends on the evolution of beliefs until then, the probability of a history of trades must be computed as

$$Pr(h_t^d|V_d = v_d^H) = \Pi_{s=1}^t \Pr(x_s^d|h_s^d, V_d = v_d^H),$$

i.e., the probability of each trade depends on the previous history of trades.[9]

Since, in an information day, the probability of a trade at time t depends on the sequence of trades until time t, our likelihood function cannot be written as a simple function of the number of buys, sells and no trades in each day, as done in Easley, Kiefer and O'Hara (1997). According to our model, the *sequence* of trades, not just the *number* of transactions, conveys information. Having many buy orders at the beginning of the day is not necessarily equivalent to having the same number of buy orders spread during the day. In fact, if there is a concentration of, for instance, buys at the beginning of the day, this may create herd behavior. The market maker in periods of herding will have to update his quotes (beliefs) in a different way than in the absence of herding. Furthermore, the probability of a particular sequence of trades in such a period of herding is different from the probability of the same sequence in the absence of herding. Therefore, to estimate our three parameters, we will not only use the number of buys, sells, and no trades. Rather, we will use the entire sequence of trades during each day.

In order to compute $\Pr(x_s^d|h_s^d, V_d = v_d^H)$, we need to distinguish those periods when agents follow their own signal from those periods in which they herd. Let us start from the case in which there is no herding (i.e., informed

[9] Recall that x_s^d deoted the action at time s of day d.

traders follow their own signal). In this case, the probability of observing a buy, a sell or a no trade at time t in a positively informed day are

$$\Pr(buy_t^d|h_t^d, V_d = v_d^H) = \mu q + (1-\mu)\frac{\varepsilon}{2},$$

$$\Pr(sell_t^d|h_t^d, V_d = v_d^H) = \mu(1-q) + (1-\mu)\frac{\varepsilon}{2},$$

$$\Pr(nt_t^d|h_t^d, V_d = v_d^H) = (1-\mu)(1-\varepsilon).$$

As we illustrated in the previous paragraph, after a prevalence of buys, the expectation of a trader with a negative signal can be higher than the ask price. In such a case, an informed trader will buy independently of his signal. In this case, the probability of observing a buy, a sell or a no trade are

$$\begin{aligned}
\Pr(buy_t^d|h_t^d, V_d &= v_d^H) = \mu + (1-\mu)\frac{\varepsilon}{2}, \\
\Pr(sell_t^d|h_t^d, V_d &= v_d^H) = (1-\mu)\frac{\varepsilon}{2}, \\
\Pr(nt_t^d|h_t^d, V_d &= v_d^H) = (1-\mu)(1-\varepsilon).
\end{aligned}$$

Similarly, after a prevalence of sells, the expectation of a trader with a positive signal can be lower than the bid price. In this case there will be herd selling and the probabilities of each action will be

$$\begin{aligned}
\Pr(buy_t^d|h_t^d, V_d &= v_d^H) = (1-\mu)\frac{\varepsilon}{2}, \\
\Pr(sell_t^d|h_t^d, V_d &= v_d^H) = \mu + (1-\mu)\frac{\varepsilon}{2}, \\
\Pr(nt_t^d|h_t^d, V_d &= v_d^H) = (1-\mu)(1-\varepsilon).
\end{aligned}$$

If we define three indicator functions I, I_{hb}, I_{hs} for each case above (i.e., I_{hb} takes value 1 if there is herd buy and 0 otherwise), we can write the probability of, for instance, a buy order in the case of a good event day as:[10]

$$\Pr(buy_t^d|h_t^d, V_d = v_d^H) =$$
$$\left[\mu + (1-\mu)\frac{\varepsilon}{2}\right] I_{hb} + \left[(1-\mu)\frac{\varepsilon}{2}\right](I_{hs} + I_{mb}) + [\mu q + (1-\mu)](I + I_{ms}).$$

[10] Besides the normal case and those of herd buying and herd selling, there are two intermediate cases . The first occurs when, after a positive trade imbalance, $b_t^d < E(V|h_t^d, s^L) < p_t^d$. In this case, in equilibrium, the market maker computes the bid assuming that only a noise trader would sell while an agent receiving a negative signal would not trade, and this

We can write the probability of any other action conditional on a good event in a similar way. The analysis for the case of a bad event is identical.

To conclude our description, we need to understand how, for any values of the parameters, we can determine whether we are in a situation of no herd, herd buy or herd sell. Herd buying arises when an informed trader values the asset more than the ask price posted by the market maker, independently of his private information. Therefore, in order to check whether there is herd selling at time t, we simply need to compare the expectation of the trader with that of the market maker.

One could believe that in order to compare the traders' and the market maker's beliefs, and decide in which of the cases illustrated above we are at any time t, we would need data on the magnitude of the shock of the information event that buffets the asset's fundamental (i.e., that we would need to estimate Δ^H and ΔL). We can easily show that this is not the case. The expected value of the asset for a trader at time t is

$$E(V_d|h_t^d, s_t^d) = v_d^L \Pr(V_d = v_d^L|h_t^d, s_t^d) + v_d^H \Pr(V_d = v_d^H|h_t^d, s_t^d) =$$
$$v_{d-1} - \Delta L \Pr(V_d = v_d^L|h_t^d, s_t^d) + \Delta^H \Pr(V_d = v_d^H|h_t^d, s_t^d).$$

---

is indeed what happens. In this instance, the probabilities of a trade in a good information day will be

$$\Pr(buy_t^d|h_t^d, V_d = v_d^H) = \mu q + (1-\mu)\frac{\varepsilon}{2},$$
$$\Pr(sell_t^d|h_t^d, V_d = v_d^H) = (1-\mu)\frac{\varepsilon}{2},$$
$$\Pr(nt_t^d|h_t^d, V_d = v_d^H) = (1-\mu)(1-\varepsilon) + \mu(1-q).$$

The other case occurs when, after a sequence of sells, $p_t < E(V|h_t^d, s^H) < a_t^d$, and in equilibrium the market maker sets the ask assuming that only a noise trader would buy, as it is indeed the case. Therefore, in a good information day

$$\Pr(buy_t^d|h_t^d, V_d = v_d^H) = (1-\mu)\frac{\varepsilon}{2},$$
$$\Pr(sell_t^d|h_t^d, V_d = v_d^H) = \mu(1-q) + (1-\mu)\frac{\varepsilon}{2},$$
$$\Pr(nt_t^d|h_t^d, V_d = v_d^H) = (1-\mu)(1-\varepsilon) + \mu q.$$

The probabilities of each action on the case of a bad information day are computed similarly. *casi indifferenza*

On the other hand, the expected value of the market maker is

$$E(V|h_t^d) = v_d^L \Pr(V_d = v_d^L|h_t^d) + v_{d-1} \Pr(V_d = v_d^L|h_t^d) + v_d^H \Pr(V_d = v_d^H|h_t^d) = \\ -\Delta L \Pr(V_d = v_d^L|h_t^d) + v_{d-1} + \Delta^H \Pr(V_d = v_d^H|h_t^d),$$

Therefore, the difference between the two expectations is

$$\begin{aligned} &\left[v_{d-1} - \Delta L \Pr(V_d = v_d^L|h_t^d) + \Delta^H \Pr(V_d = v_d^H|h_t^d)\right] - \\ &\left[v_{d-1} - \Delta L \Pr(V_d = v_d^L|h_t^d, s_t^d) + \Delta^H \Pr(V_d = v_d^H|h_t^d, s_t^d)\right] \\ &= \left[-\frac{1-\delta}{\delta}\Delta^H \Pr(V_d = v_d^L|h_t^d) + \Delta^H \Pr(V_d = v_d^H|h_t^d)\right] - \\ &\left[-\frac{1-\delta}{\delta}\Delta^H \Pr(V_d = v_d^L|h_t^d, s_t^d) + \Delta^H \Pr(V_d = v_d^H|h_t^d, s_t^d)\right] = \\ &\Delta^H \left[\frac{1-\delta}{\delta}\left(\Pr(V_d = v_d^L|h_t^d, s_t^d) - \Pr(V_d = v_d^L|h_t^d)\right) - \right. \\ &\left. \Pr(V_d = v_d^H|H_t^d, s_t^d) + \Pr(V_d = v_d^H|H_t^d)\right] \end{aligned}$$

whose sign is independent of how big the positive or negative news is.[11] This means that, in order to understand whether we are in a period of herding or not, we do not need to estimate the magnitude of the information shock that hits the asset.

4 Bayesian Estimation of the Model

The main objective of our work is to estimate the structural model of herd behavior that we have just illustrated. After estimating the model, we will be able to detect the herding periods during the trading days.

We carry out the estimation in a Bayesian framework, simulating the posterior distribution of the parameters conditional on the data. All the five parameters in the model are assumed to be random variables with a given prior distribution. For all the parameters, except q, the prior distribution is assumed to be uniform on the interval $[0, 1]$; the prior distribution of q is assumed to be uniform in the interval $[1/2, 1]$.

[11] For simplicity, we have studied the difference between the trader 's expectation and the asset price (i.e., the market maker's expectation before trader t trades). It is straightforward to repeat the argument for the bid and the ask prices.

In order to draw from the posterior distributions of the parameters, we generate random samples using a Markov-Chain Monte Carlo (MCMC) procedure. In particular, using the "Metropolis-Hasting " method, we sequentially sample from the posterior distribution of one parameter conditional to the other four.[12]

Under regularity conditions (see Geweke, 1996 and 1997) here satisfied, the Markov chain so produced converges, and yields a sample from the joint posterior distribution of the parameters conditional on the data. In the Section "Results" we present the average and the standard deviation of 1,000,000 draws from the posterior distributions of each parameter.[13] In computing the average, we discard the first 100,000 draws from a "burn in" phase.

Finally, let us remark that to estimate all our parameters we clearly need data on more than one trading day. The parameters α and δ define the probability that there is an event at the beginning of a trading day and that the event is good or bad. Therefore, if we used data on one day only, this would be equivalent to observing just one draw from a joint distribution with parameters α and δ. Clearly we would be unable to estimate these two parameters. For this reason, we will use data for many days.

5 Data

We use data for Ashland Oil, a stock traded in the New York Stock Exchange.[14] We took data for this stock from the TAQ dataset. This dataset reports a complete list of the posted prices (quotes), the price at which the transactions occurred (trade), the size of the transactions and, of course, the time when the quotes were posted and the transaction occurred. We use transactions data on this stock for the first quarter of 1995. In this period there were 63 trading days.

In order to extract the history of trades from this dataset, we had to make several transformations. First, these data do not say who initiated the trade, i.e., whether the transaction was a sell or a buy. In order to classify a

[12] As starting values we use the mode of the posterior distribution (which, since the priors are flat, is the maximum of the likelihood function). We estimate the mode using a genetic algorithm optimization method.

[13] The results we obtain are not significantly different if we use the posterior mode instead of the posterior mean.

[14] This is the same stock studied by Easley, Kiefer and O 'Hara 1997.

trade as a sell or buy, we had to compare the trade with the quotes that were posted at the time of the transaction. For this purpose, we used the standard algorithm proposed by Lee and Ready (1991). We compared the transaction price with the quotes that were posted just before the transaction occurred. Any trade above the midpoint was classified as a buy and every trade below the midpoint was classified as a sell; trades at the midpoint were classified as buy or sell according to whether this price had increased or decreased with respect to the previous transaction price.[15] [16] Given that transaction prices are reported with a delay, we followed Lee and Ready (1991) suggestion of moving each quote ahead in time for five seconds.

Second, clearly these data do not contain any direct information on "no trades." But, of course, there is a significant difference between an hour of trading in which many transactions occur and an hour in which there is none. We used the convention of classifying any period of five minutes in which no transaction occurred as a no trade. For instance, we considered a period of 20 minutes between two transactions as four no trades. The choice of five minutes is, of course, arbitrary, as it would be for any other integer. As a robustness check, we also used other alternative intervals.

As we said, we considered a sample period with 63 trading days. On average, we observed 100 transactions (either buys or sells) a day; our five minutes rule implied on average 33 no trades a day. There were more buys than sells in our sample: 38.7% of periods were buys, 35.7% were sells and 25.6% were no trades.

6 Results

Table 1 shows the average and standard deviation of the posterior distribution of the model parameters:

Table 1: Estimation Results

Parameter	Estimate	Standard deviation
α	0.79	0.057
δ	0.68	0.097
μ	0.24	0.018
ε	0.66	0.027
q	0.69	0.006

[15] There were only 16 trades at the midpoint in our dataset.

[16] If there was no change, then we looked at the previous price movement and so on.

If the specification we adopted is correct, informational events are quite frequent: 79% of trading days are in fact classified as days in which some trading activity was motivated by private information. There is a certain imbalance between good and bad news. Good informational events account for almost 70% of the informed days. During informed days, the proportion of traders with private information is, on average, 24%. The remaining trading activity is explained by noise traders. Noise traders traded 66% of the time, and did not the remaining 34%. The precision of private information is just below 70%. Such a precision of the signal, strictly lower than 1, opens the door to herd behavior. On the basis of these parameters, we tracked down the beliefs of the traders (with a positive or negative signal) and the beliefs of the market maker (i.e., the bid and ask prices) during each trading day. By comparing such beliefs we can detect periods in which, according to our model, there was herd behavior in the market. These are periods in which an informed trader would have made the same decision independently of the signal he received. For instance, when the belief of the trader is higher than the equilibrium ask even if he received a negative signal, we classify this period as herd buying. Similarly, when the belief of the trader is lower than the equilibrium bid even if he received a positive signal, we classify this period as herd selling. We found that 7% of trading periods are of herd buying and 8% are of herd selling. It is important to remark that herding periods are relevant for the informational (in)efficiency of the market. Indeed, during herding periods, the market is unable to learn whether the traders received a positive or a negative signal. Although the market maker still learns something from the trading activity (namely, whether he is in an informed or uninformed day), information is aggregated less efficiently and the price converges more slowly to the fundamental asset value. For the period under analysis, we found that herding was pronounced in 23 days (out of 63). In such days, at least 15% of the trading periods were characterized by herding behavior. In 7 days, in particular, herding was very pronounced, since it characterized more than half of the trading periods.

Table 2: Number of days in which herding periods were frequent.

$> 50\%$	$> 30\%$	$> 15\%$
7	15	23

To have a better understanding of the inefficiency produced by herding behavior, we simulated $10,000$ days of trading in a financial market with

our estimated parameters. We then simulated the same days of trading with the same parameters but assuming that traders, instead of behaving rationally as in our model, always followed private information. We took this as a benchmark case, since in this case all private information would be revealed by the trading decisions. We compared the price paths under the two scenarios. We considered the absolute difference at each time of every trading day between the simulated price and the full information price. We found that the average distance between the two prices is 0.4%. In other words, the presence of herding determined a deviation of the price from the full information level of 0.4% on average during each day. Moreover, in 2.8% of trading periods, the distance between the two prices was larger than 10%. This suggests that there are times when intraday herding affects the informational properties of the price in a very significant manner.

7 Conclusion

We estimated a model of herding behavior in financial markets. We used transaction data for a stock traded in the NYSE. We estimated the parameters of the structural model and detected periods in each trading day in which, according to our model, informed traders chose the same action independently of whether they had positive or negative private information on the value of the stock, i.e., they herded. We found that herding is present in the market. In some days of trade it is fairly pervasive. In our future research we will apply our methodology to verify whether herding is more pronounced in particular markets, and in particular times (like during financial crises).

References

[1] Avery, Christopher and Zemsky, Peter. "Multidimensional Uncertainty and Herd Behavior in Financial Markets." *American Economic Review*, September 1998, 88 (4), pp.724-748.

[2] Banerjee, Abhijit, "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, August 1992, 107 (3), pp. 787-818.

[3] Bikhchandani, Sushil; Hirshleifer, David and Welch Ivo. "A Theory of Fads, Fashion, Custom and Cultural Change As Informational Cascades." *Journal of Political Economy,* October 1992, 100 (5), pp. 992-1027.

[4] Bikhchandani, Sushil and Sharma Sunil. "Herd Behavior in Financial Markets: A Review." 2000, IMF Working paper.

[5] Chamley, Christophe. *Rational Herds. Economic Models of Social Learning.* Cambridge: Cambridge University Press, 2004.

[6] Cipriani, Marco and Guarino Antonio. "Herd Behavior and Contagion in Financial Markets." mimeo, NYU, 2001.

[7] Cipriani, Marco and Guarino Antonio. "Herd Behavior in a Laboratory Financial Market." *American Economic Review*, December 2005, pp. 1427-1443.

[8] Cipriani, Marco and Guarino Antonio. "Transaction Costs and Informational Cascades in Financial Markets: Theory and Experimental Evidence." ELSE Discussion Paper, 2006.

[9] Drehmann, Mathias; Oechssler, Jörg and Rider, Andreas. "Herding and Contrarian Behavior in Financial Markets - An Internet Experiment." *American Economic Review*, December 2005.

[10] Easley D. and M. O'Hara, 1987, "Price, Trade Size, and Information in Securities Markets", *Journal of Financial Economics*, 19, 69-90.

[11] Easley David and Maureen O'Hara. "Time and the Process of Security Price Adjustment." *Journal of Finance*, 1992

[12] Easley David; Kiefer, Nicholas and Maureen O'Hara, 1997, "One Day in the Life of a Very Common Stock," *Revie of Financial Studies*, 10 (3), 805-835.

[13] Easley David, Maureen O'Hara, Robert Engle and Liuren Wo, 2001, "Time-Varying Arrival Rates of Informed and Uninformed Trades." mimeo, NYU and Cornell.

[14] Gale, D., 1996, "What Have We Learned from Social Learning", *European Economic Review*, 40, 617-628.

[15] Geweke, John, 1996 "Monte Carlo Simulation and Numerical Integration," in Hans M. Amman, David A. Kendrick, and John Rust, eds., Handbook of computational economics, volume 1. Amsterdam: North-Holland, 1996, pp. 731–800.

[16] Geweke, John, 1997 "Posterior Simulators in Econometrics," in David M. Kreps and Kenneth F. Wallis, eds., Advances in economics and econometrics: Theory and applications. Cambridge: Cambridge University Press,1997, pp. 128– 65.

[17] Glosten, Lawrence and Milgrom Paul. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, March 1985, 14 (1), pp. 71-100.

[18] Grinblatt Mark, Sheridan Titman and Russ Wermers, "Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior." *American Economic Review*, December 1995

[19] Hirshleifer D. and S.H. Teoh, 2003, "Herd Behavior and Cascading in Capital Markets: a Review and Synthesis," *European Financial Management*, 9, 25-66.

[20] Lakonishok, Joseph, Shleifer Andrei and Vishny Robert W. "The Impact of Institutional Trading on Stock Prices," *Journal of Financial Economics*, 1992, 32, pp. 23-43.

[21] Lee, C., M. Ready, 1991. "Inferring Trade Direction from Intraday Data." *Journal of Finance 46*, 733-746.

[22] Milgrom, Paul and Stokey, Nancy. "Information, Trade and Common Knowledge." *Journal of Economic Theory*, February 1982, 26 (1), pp. 17-27.

[23] Smith, L. and P. Sørensen. "Pathological Outcomes of Observational Learning." *Econometrica, 2000, 68,* pp. *371-398.*

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

[25] Welch, Ivo. "Herding Among Security Analysts." *Journal of Financial Economics* 58-3, December 2000, 369-396.

8 Appendix

8.1 Proof of Proposition 1

First, we prove the existence of the ask price. Because of unmodeled potential Bertrand competition, the ask price at time t, a_t^d, must satisfy the condition

$$a_t^d = E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d].$$

Let us define I_t^d as a random variable that takes value 0 if the agent at time t in day d is noise and 1 if he is informed. The expected value of the asset at time t in day d, given a buy order at the ask price a_t^d, is

$$\begin{aligned} E[V|h_t^d, X_t^d &= 1, a_t^d, b_t^d] = \\ E[V|h_t^d, X_t^d &= 1, a_t^d, b_t^d, I_t^d = 1] \Pr[I_t^d = 1|h_t^d, X_t^d = 1, a_t^d, b_t^d] + \\ E[V|h_t^d, a_t^d, b_t^d, I_t^d &= 0] \Pr[I_t^d = 0|h_t^d, X_t^d = 1, a_t^d, b_t^d]. \end{aligned}$$

Let us consider the correspondence $\psi : [v_d^L, v_d^H] \rightrightarrows [v_d^L, v_d^H]$ defined as $\psi(a_t^d) := E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$, and let us make the following observations:

1) If $a_t^d > E[V|h_t^d, S_t^d = s^H]$, $\Pr[I_t^d = 1|h_t^d, X_t^d = 1, a_t^d, b_t^d] = 0$ and $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d] = E[V|h_t^d, X_t^d = 1, I_t^d = 0]$.

2) If $E[V|h_t^d, S_t^d = s^L] < a_t^d < E[V|h_t^d, S_t^d = s^H]$, then $E[V|h_t^d, X_t = 1, a_t^d, b_t^d] = E(V|h_t^d, I_t^d = 1, S_t^d = s^H) \Pr(I_t^d = 1, S_t^d = s^H|h_t^d)+$

$E[V|h_t^d, I_t^d = 0]\left(1 - \Pr(I_t^d = 1, S_t^d = s^H|h_t^d)\right)$.

3) If $a_t^d < E[V|h_t^d, S_t^d = s^L]$, then $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d] = E(V|h_t^d, I_t^d = 1)\Pr(I_t^d = 1|h_t^d) + E(V|h_t^d, I_t^d = 0)(1 - \Pr(I_t^d = 1|h_t^d))$,

where, of course, since the ask would be lower than the expectation of a trader with a negative signal, the conditional expected values and probabilities are computed assuming that an informed trader buys whatever signal he receives.

Finally note that

4) If $a_t^d = E[V|h_t^d, S_t^d = s^H]$, then the informed trader receiving a positive signal can randomize between buying and not trading. If he buys with probability 0, then

$E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ is equal to the expression indicated in Observation 1. If he buys with probability 1, then $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ is equal to the expression indicated in Observation 2. If he buys with any probability belonging to the interval $(0, 1)$, then $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ takes any value between these two expressions.

5) Similarly, if $a_t^d = E[V|h_t^d, S_t^d = s^L]$, then the informed trader receiving a negative signal can randomize between buying and not trading. If he buys with probability 0, then

$E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ is equal to the expression indicated in Obseration 2. If he buys with probability 1, then $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ is equal to the expression indicated in Observation 3. If he buys with any probability belonging to the interval $(0, 1)$, then $E[V|h_t^d, X_t^d = 1, a_t^d, b_t^d]$ takes any value between these two expressions.

Observations 1, 2, 3, 4, and 5 imply that the correspondence ψ is piecewise constant. Furthermore, for $a_t^d = E[V|h_t^d, S_t^d = s^L]$ and $a_t^d = E[V|h_t^d, S_t^d = s^H]$, $\psi(a_t^d)$ takes all the values belonging to the intervals indicated above in observations 4 and 5. Therefore, it is immediate to see that the correspondence $\psi(a_t^d)$ is non empty, convex-valued and has a closed graph. By Kakutani's fixed point theorem, the correspondence has a fixed point. If there is more than one fixed point, the ask price will be equal to the minimum of them (it is straightforward to show that the other fixed points do not represent an equilibrium, due to the potential Bertrand competition that the market maker faces).

The proof of the existence and uniqueness of the bid price is analogous.

The proof that $b_t^d \leq p_t^d \leq a_t^d$ follows immediately from the proof in Glosten and Milgrom (1985, p. 81).

8.2 Proof of Proposition 2

Suppose $V_d = V^L$. After a series of buy orders, the traders attach a higher probability to $V_d = V^H$ (for whatever signal they receive). Such a probability goes to 1 when the number of buys goes to infinity. Therefore, we can always find a sufficiently high t' such that, after a histoy of t' buy orders, $E[V_d|h_{t'+1}^d, s^L] > V_{d-1} + \epsilon$ where $\epsilon > 0$. Let us assume that in these t' periods herding has not arisen (otherwise the proposition is already proven). Now note that a no trade does not affect the traders' beliefs, since it only comes from noise traders. That is, $E[V_d|h_{t'+1}^d, nt_{t'+1}, s] = E[V_d|h_{t'+1}^d, nt_{t'+1}, s]$. In contrast, a no trade always increases the probability that the market maker attaches to $V_d = V_{d-1}$. Indeed, after a no trade, $\Pr[V_d = V_{d-1}|h_t^d, X_t^d = 0, a_t^d, b_t^d] =$ $\frac{(1-\varepsilon)\Pr[V_d=V_{d-1}|h_t^d]}{(1-\mu)(1-\varepsilon)\left[\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d]\right]+(1-\varepsilon)\Pr[V_d=V_{d-1}|h_t^d]}$ which is obviously greater than $\Pr[V_d = V_{d-1}|h_t^d]$.

Therefore, for any $\epsilon > 0$ and for any set of values for the parameters $\{\alpha, \delta, \mu, q, \varepsilon\}$, we can always find a number of t'' no trades sufficiently high that, in equilibrium, $E[V_d|h_{t'+t''+1}^d, X_{t'+t''+1}^d = 1, a_{t'+t''+1}^d, b_{t'+t''+1}^d] < V_{d-1} + \epsilon$. At this point, herd behavior arises and is misdirected. Such a history $h_{t'+t''}^d$ occurs with positive probability, because of noise traders. This shows that (misdirected) herd buying occurs with positive probability. The proof for (misdirected) herd selling is analogous.

8.3 Proof of Proposition 3

Let us consider herd buying. During a situation of herd buying, informed traders buy independently of their signal. Suppose the period of herd buying is never broken, i.e., it lasts for ever after it has started at some time t. In such a case, the probability of each action after t is always the same. The conditional probabilities of a buy are given by $\Pr[X_t^d = 1|h_t^d, a_t^d, b_t^d, V_d = V^H] = \Pr[X_t^d = 1|h_t^d, a_t^d, b_t^d, V_d = V^L] = [\mu + (1-\mu)\frac{\varepsilon}{2}]$,

$\Pr[X_t^d = 1|h_t^d, a_t^d, b_t^d, V_d = V_{d-1}] = \frac{\varepsilon}{2}$. The probabilities of selling are $\Pr[X_t^d = -1|h_t^d, a_t^d, b_t^d, V_d = V^H] = \Pr[X_t^d = -1|h_t^d, a_t^d, b_t^d, V_d = V^L] = (1-\mu)\frac{\varepsilon}{2}$,

$\Pr[X_t^d = -1|h_t^d, a_t^d, b_t^d, V_d = V_{d-1}] = \frac{\varepsilon}{2}$. Finally, the probabilities of not trading are $\Pr[X_t^d = 0|h_t^d, a_t^d, b_t^d, V_d = V^H] = \Pr[X_t^d = 0|h_t^d, a_t^d, b_t^d, V_d = V^L] = (1-\mu)(1-\varepsilon)$,

$\Pr[X_t^d = 0|h_t^d, a_t^d, b_t^d, V_d = V_{d-1}] = (1-\varepsilon)$. Let us denote by β,σ,ν, the number of buys, sells and no trades observed during a certain period after herding has started in t. Then,

$$\Pr[V_d = V^H|h_t^d, \beta, \sigma, \nu] = \frac{[\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d=V^H|h_t^d]}{K},$$

$$\Pr[V_d = V^L|h_t^d, \beta, \sigma, \nu] = \frac{[\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d=V^L|h_t^d]}{K},$$

$$\Pr[V_d = V_{d-1}|h_t^d, \beta, \sigma, \nu] = \frac{\left(\frac{\varepsilon}{2}\right)^{\beta+\sigma}(1-\varepsilon)^\nu \Pr[V_d=V_{d-1}|h_t^d]}{K},$$

where $K = [\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \left(\Pr[V_d = V^H|h_t^d] + \Pr[V_d = V^L|h_t^d]\right) + \left(\frac{\varepsilon}{2}\right)^{\beta+\sigma}(1-\varepsilon)^\nu \Pr[V_d = V_{d-1}|h_t^d]$.

Therefore, the probability that an event occurred at the beginning of day d is

$$\Pr[V_d = V^H|h_t^d, \beta, \sigma, \nu] + \Pr[V_d = V^L|h_t^d, \beta, \sigma, \nu] = \frac{[\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{K}.$$

The likelihood ratio between an event occurring or not is

$$\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]} = \frac{[\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{[\frac{\varepsilon}{2}]^{\beta+\sigma}[(1-\varepsilon)]^\nu \Pr[V_d=V_{d-1}|h_t^d]}.$$

The loglikelihood ratio can be expressed in the following way:

$$\log \frac{[\mu+(1-\mu)\frac{\varepsilon}{2}]^\beta[(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{\left(\frac{\varepsilon}{2}\right)^{\beta+\sigma}(1-\varepsilon)^\nu \Pr[V_d=V_{d-1}|h_t^d]} = \log \frac{(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{\Pr[V_d=V_{d-1}|h_t^d]} + \beta \log \frac{\mu+(1-\mu)\frac{\varepsilon}{2}}{\frac{\varepsilon}{2}} + \sigma \log(1-\mu) + \nu \log(1-\mu).$$

By dividing both sides by the total number of buys, sells, and no trades, we obtain

$$\frac{1}{\beta+\sigma+\nu} \log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]} =$$

$\frac{1}{\beta+\sigma+\nu}\log\frac{(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{\Pr[V_d=V_{d-1}|h_t^d]}+$
$\frac{\beta}{\beta+\sigma+\nu}\log\frac{\mu+(1-\mu)\frac{\varepsilon}{2}}{\frac{\varepsilon}{2}}+\frac{\sigma}{\beta+\sigma+\nu}\log(1-\mu)+\frac{\nu}{\beta+\sigma+\nu}\log(1-\mu)$.

Now, suppose there has been an event (i.e., $V_d = V^L$ or $V_d = V^H$) and let $(\beta+\sigma+\nu)\to\infty$. Then,
$\frac{1}{\beta+\sigma+\nu}\log\frac{(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{\Pr[V_d=V_{d-1}|h_t^d]}\to 0$,
$\frac{\beta}{\beta+\sigma+\nu}\overset{a.s.}{\to}[\mu+(1-\mu)\frac{\varepsilon}{2}]$,
$\frac{\sigma}{\beta+\sigma+\nu}\overset{a.s.}{\to}[(1-\mu)\frac{\varepsilon}{2}]$,
$\frac{\nu}{\beta+\sigma+\nu}\overset{a.s.}{\to}[(1-\mu)(1-\varepsilon)]$,
where the convergence almost surely just results from the law of large numbers. Hence, as time goes to infinity (and herd never stops):
$\frac{1}{\beta+\sigma+\nu}\log\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]}\to$
$[\mu+(1-\mu)\frac{\varepsilon}{2}]\log\frac{\mu+(1-\mu)\frac{\varepsilon}{2}}{\frac{\varepsilon}{2}}+(1-\mu)\frac{\varepsilon}{2}\log(1-\mu)+(1-\mu)(1-\varepsilon)\log(1-\mu)=$
$[\mu+(1-\mu)\frac{\varepsilon}{2}]\log\frac{\mu+(1-\mu)\frac{\varepsilon}{2}}{\frac{\varepsilon}{2}}+(1-\mu)(1-\frac{\varepsilon}{2})\log(1-\mu)$.

It is easy to show that the RHS is positive.

Therefore,
$\frac{1}{\beta+\sigma+\nu}\log\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]}$ converges to a positive constant.
This implies that
$\log\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]}\to+\infty$,
that is,
$\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]+\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]}\to+\infty$.

If herd buys keeps forever, the belief that there has been no event converges to 0. During the period of herding, the informed traders do not update their beliefs, i.e., $\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu,s]=\Pr[V_d=V^H|h_t^d,s]$ for $s=\{s^L,s^H\}$. Hence, when $\Pr[V_d=V_{d-1}|h_t^d,\beta,\sigma,\nu]\to 0$, $E[V_d|h_t^d,\beta,\sigma,\nu,s^L]<E[V_d|h_t^d,\beta,\sigma,\nu]<E[V_d|h_t^d,\beta,\sigma,\nu,s^H]$, which contadicts that herd buying keeps forever.

The proof for the case of herd selling is analogous.

8.4 Proof of Proposition 4

Let us consider the case in which $V_d = V^H$. We know from the proof of Proposition x that during periods of herding the market maker only learns about the probability of an informational event, while the loglikelihood ratio between the event being good or bad does not change. Indeed, from the

proof of Proposition x we immediately obtain that, during periods herding, $\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]} = \frac{\Pr[V_d=V^H|h_t^d]}{\Pr[V_d=V^L|h_t^d]}$. Now we show that during the (infinitely many) periods of non herding, $\frac{\Pr[V_d=V^H|h_t^d]}{\Pr[V_d=V^L|h_t^d]} \to \infty$.

Suppose that during periods of non herding there are β buys, σ sells and ν no trades. In such periods

$\Pr[V_d = V^H|h_t^d, \beta, \sigma, \nu] =$
$\frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d=V^H|h_t^d]}{K}$,

$\Pr[V_d = V^L|h_t^d, \beta, \sigma, \nu] =$
$\frac{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d=V^L|h_t^d]}{K}$,

$\Pr[V_d = V_{d-1}|h_t^d, \beta, \sigma, \nu] =$
$\frac{\left(\frac{\varepsilon}{2}\right)^{\beta+\sigma}(1-\varepsilon)^\nu \Pr[V_d=V_{d-1}|h_t^d]}{K}$,

where $K = [\mu q + (1-\mu)\frac{\varepsilon}{2}]^\beta[\mu(1-q) + (1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d = V^H|h_t^d]+$
$[\mu(1-q) + (1-\mu)\frac{\varepsilon}{2}]^\beta[\mu q + (1-\mu)\frac{\varepsilon}{2}]^\sigma[(1-\mu)(1-\varepsilon)]^\nu \Pr[V_d = V^L|h_t^d]+$
$\left(\frac{\varepsilon}{2}\right)^{\beta+\sigma}(1-\varepsilon)^\nu \Pr[V_d = V_{d-1}|h_t^d]$.

The likelihood ratio between an event being good or bad is
$\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]} =$
$\frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\sigma \Pr[V_d=V^H|h_t^d]}{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\sigma \Pr[V_d=V^L|h_t^d]}$.
The loglikelihood ratio can be expressed in the following way:
$\log \frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\sigma \Pr[V_d=V^H|h_t^d]}{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\sigma \Pr[V_d=V^L|h_t^d]} =$
$\log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]} + \beta \log \frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta}{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta} + \sigma \log \frac{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]}{[\mu q+(1-\mu)\frac{\varepsilon}{2}]}$.
By dividing both sides by the total number of observed buys, sells and no trades, we obtain
$\frac{1}{\beta+\sigma+\nu} \log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]} =$
$\frac{1}{\beta+\sigma+\nu} \log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}+$
$\frac{\beta}{\beta+\sigma+\nu} \log \frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta}{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta} + \frac{\sigma}{\beta+\sigma+\nu} \log \frac{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]}{[\mu q+(1-\mu)\frac{\varepsilon}{2}]}$.
Let $(\beta+\sigma+\nu) \to \infty$. Then,
$\frac{1}{\beta+\sigma+\nu} \log \frac{(\Pr[V_d=V^H|h_t^d]+\Pr[V_d=V^L|h_t^d])}{\Pr[V_d=V_{d-1}|h_t^d]} \to 0$,
$\frac{\beta}{\beta+\sigma+\nu} \overset{a.s.}{\to} [\mu q + (1-\mu)\frac{\varepsilon}{2}]$,
$\frac{\sigma}{\beta+\sigma+\nu} \overset{a.s.}{\to} [\mu(1-q) + (1-\mu)\frac{\varepsilon}{2}]$,

$\frac{\nu}{\beta+\sigma+\nu} \overset{a.s.}{\rightarrow} [(1-\mu)(1-\varepsilon)]$,

where the convergence almost surely just results from the law of large numbers. Hence, during the infinitely many periods of non herding:

$\frac{1}{\beta+\sigma+\nu} \log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]} \rightarrow$

$[\mu q + (1-\mu)\frac{\varepsilon}{2}] \log \frac{[\mu q+(1-\mu)\frac{\varepsilon}{2}]^\beta}{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]^\beta} + [\mu(1-q) + (1-\mu)\frac{\varepsilon}{2}] \log \frac{[\mu(1-q)+(1-\mu)\frac{\varepsilon}{2}]}{[\mu q+(1-\mu)\frac{\varepsilon}{2}]}$.

It is easy to show that the RHS is positive.

Therefore,

$\frac{1}{\beta+\sigma+\nu} \log \frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}$ converges to a positive constant, i.e., $\Pr[V_d = V^L|h_t^d,\beta,\sigma,\nu] \rightarrow 0$.

Since during periods of herding $\frac{\Pr[V_d=V^H|h_t^d,\beta,\sigma,\nu]}{\Pr[V_d=V^L|h_t^d,\beta,\sigma,\nu]}$ remains constant, this result immediately shows that the market maker learns that the asset value cannot be $V_d = V^L$. An analogous proof shows that, during periods of non-herding, $\Pr[V_d = V_{d-1}|h_t^d,\beta,\sigma,\nu] \rightarrow 0$. From the proof of Proposition x we know that, if there were infinitely many times of herding, $\Pr[V_d = V_{d-1}|h_t^d,\beta,\sigma,\nu] \rightarrow 0$. Therefore, this immediately implies that the market maker learns that the asset value cannot be $V_d = V_{d-1}$.

The proof of convergence for the case in which $V_d = V^L$ is analogous. The converges for the case in which $V_d = V_{d-1}$ is proven by invoking the law of large numbers, as done above (since the probability of each action is the same at each time t).