WP/10/288

IMF Working Paper

Estimating a Structural Model of Herd Behavior in Financial Markets

Marco Cipriani and Antonio Guarino

INTERNATIONAL MONETARY FUND

 WP/10/288

IMF Working Paper

INS

Estimating a Structural Model of Herd Behavior in Financial Markets

Prepared by Marco Cipriani and Antonio Guarino*

Authorized for distribution by Ling Hui Tan

December 2010

Abstract

We develop a new methodology to estimate the importance of herd behavior in financial markets: we build a structural model of informational herding that can be estimated with financial transaction data. In the model, rational herding arises because of information-event uncertainty. We estimate the model using data on a NYSE stock (Ashland Inc.) during 1995. Herding often arises and is particularly pervasive on some days. The proportion of herd buyers (sellers) is 2 percent (4 percent) and is greater than 10 percent in 7 percent (11 percent) of information-event days. Herding causes important informational inefficiencies, amounting, on average, to 4 percent of the expected asset value.

JEL Classification Numbers: G14, D82, C13

Keywords: Herd Behavior, Market Microstructure, Structural Estimation

Author's E-Mail Address: mcipriani@imf.org

* IMF Institute and University College London respectively. We thank Orazio Attanasio, Pedro Carneiro, Andrew Chesher, Burkhard Drees, Robert Engle, Christopher Flinn, Ana Fostel, Douglas Gale, Toru Kitagawa, Mico Loretan, Lars Nesheim, Sam Ouliaris, Henri Pages, Nicola Pavoni, Jean Marc Robin, Ling Hui Tan, and seminar participants in various institutions for comments and suggestions. We thank Silvia Camussi, Minyu Chen, Thibaut Lamadon, and Yuki Sato for excellent research assistance. We gratefully acknowledge the financial support of the Fondation Banque de France, the ERC and the ESRC. This work was partially conducted while Guarino was a visiting scholar at the IMF Institute; Guarino is grateful to the IMF Institute for its hospitality. We are responsible for any error.

Contents Page

I. INTRODUCTION

In recent years there has been much interest in herd behavior in financial markets. This interest has led researchers to look for theoretical explanations and empirical evidence of herding. There has been, however, a substantial disconnect between the empirical and theoretical literatures: the theoretical work has identified motives for herding in abstract models that cannot easily be brought to the data; the empirical literature has mainly looked for atheoretical, statistical evidence of trade clustering, which is interpreted as herding.

This paper takes a novel approach: we develop a theoretical model of herding in financial markets that can be estimated with financial markets transaction data. This methodology allows us to measure the quantitative importance of herding, to identify when it happens, and to assess the informational inefficiency that it generates.

The theoretical work on herd behavior started with the seminal papers of Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992), and Welch (1992).[1] These papers model herd behavior 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 disregard their own private information and imitate their predecessors. More recently, a number of papers (see, among others, Avery and Zemsky, 1998; Lee, 1998; Cipriani and Guarino, 2008) have focused on herd behavior in financial markets. In particular, these studies analyze a market where informed and uninformed traders sequentially trade a security of unknown value. The price of the security is set by a market maker according to the order flow. The presence of a price mechanism makes it more difficult for herding to arise. Nevertheless, there are cases in which it occurs. In Avery and Zemsky (1998), for instance, herd behavior can occur when there is uncertainty not only about the value of the asset but also about the occurrence of an information event or about the model parameters.

As mentioned above, 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, Shleifer, and Vishny, 1992; Grinblatt, Titman, and Wermers, 1995; and Wermers, 1999) does not test the theoretical herding models directly, but analyzes the presence of herding in financial markets through statistical measures of clustering.[2] These papers find that, in some markets, fund managers tend to cluster their investment decisions more than would be expected if they acted independently. This 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

[1] We only study informational herding. Therefore, we do not discuss herd behavior due to reputational concerns or payoff externalities. For an early critical assessment of the literature on herd behavior see Gale (1996). For recent surveys of herding in financial markets see Bikhchandani and Sharma (2001), Vives (2008) and Hirshleifer and Teoh (2009).

[2] See also the recent paper by Dasgupta, Prat, and Verardo (forthcoming), who study the effect of institutional herding on long term returns, and the literature cited therein.

spurious herding from true herd behavior, that is, the decision to disregard one's private information to follow the behavior of others (see Bikhchandani and Sharma, 2001; and Hirshleifer and Teoh, 2009).

Testing models of informational herd behavior is difficult. In such models, a trader herds if he trades against his own private information. The problem that empiricists face is that there are no data on the private information available to traders and, therefore, it is difficult to know when traders decide not to follow it. Our purpose in this paper is to present a methodology to overcome this problem. We develop 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 and to measure the informational inefficiency that this generates. This is the first paper on informational herding that, instead of using a statistical, atheoretical approach, brings a theoretical social learning model to the field data.[3]

Our theoretical analysis builds on 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. Avery and Zemsky (1998) show that, if the only source of uncertainty is the asset's 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 only). Therefore, it will never be the case that agents neglect their information to imitate previous traders' decisions (i.e., they herd). In contrast, when there are multiple sources of asymmetric information between the traders and the market maker (e.g., asymmetric information on the asset's volatility) herd behavior may arise.[4]

In our model, herding arises for a mechanism similar to that exposed by Avery and Zemsky (1998). However, whereas they were interested in providing theoretical examples of herding, our aim is to provide an empirical methodology to gauge the importance of herding in actual financial markets. For this purpose, we build a model of herding that can be estimated with financial market transaction data. In the model, an asset is traded over many days; at the beginning of each day, an informational event may occur, which causes the fundamental asset value to change with respect to the previous day. If an informational event has occurred, some traders receive private information on the new asset value.[5] These traders trade the asset to exploit their informational advantage over the market maker. If no event has occurred, all traders in the market are noise traders, that is, they trade for non-information reasons only (liquidity or hedging motives). Whereas the informed traders know that they are in a market with private information (since they themselves are informed), the market maker does not. This asymmetry of information between traders and the market maker implies that the market maker moves the price too "slowly" in order to take into account the possibility that the asset value may have not changed (in which case all trading activity is due to non-informational

[3]Whereas there are no direct empirical tests of herding models, there is experimental work that tests these models in the laboratory (see, e.g., Cipriani and Guarino, 2005 and 2009; and Drehmann, Oechssler, and Rider, 2005).

[4]A similar mechanism is also present in Gervais (1997).

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

motives). As a result, after, for instance, a history of buys, a trader, even with a bad signal, may value the asset more than the market maker does. He will, therefore, trade against his own private information and herd-buy.

We estimate the model with stock market transaction data via maximum likelihood, using a strategy first proposed by Easley, Kiefer, and O'Hara (1997) to estimate the parameters of the Glosten and Milgrom (1985) model. There is an important difference, however, between Easley, Kiefer, and O'Hara's (1997) methodology and ours. In their set up, informed traders are perfectly informed about the value of the asset; as a result, their decisions are never affected by the decisions of previous traders, and they never herd. Therefore, only the total *number* of buys, sells and no trades in each day matters; *the sequence* in which these trades arrives is irrelevant. In contrast, in our framework, the precision of private information is one of the parameters that we estimate. This opens the possibility that informed traders may receive noisy signals, and that they may find it optimal to ignore them and engage in herd behavior. In this circumstance, the sequence by which trades arrive in the market does matter: in contrast to Easley, Kiefer, and O'Hara (1997), we cannot estimate our model using only the number of buy or sell orders in a given day, but we must consider the whole history of trading activity in each day of trading.

As an illustration of the methodology, we estimate the model using transaction data for a NYSE stock (Ashland Inc.) during 1995. The restriction that private signals are perfectly precise is rejected by the data, which implies both that herd behavior arises in equilibrium and that there is information content in the sequence of trades. In particular, we find that informed traders receive incorrect information 40 percent of the time.

This has important consequences for estimates of trading informativeness. A large literature has studied the information content of trading activity using a measure (usually called the PIN, an acronym for **P**rivate **IN**formation-based trading) based on the Easley, Kiefer, and O'Hara (1997) methodology. Using that methodology, the measure of information-based activity in our sample would be 9 percent. Using our methodology, instead, we obtain 19 percent. The difference is due to the fact that in the previous literature incorrect trades (e.g., selling in a good-event day) can only be due to exogenous, non-informative (e.g., liquidity) reasons, whereas in our setup we do not exclude the possibility that they may be due to informed traders who either receive incorrect information or herd.

Given our estimated parameters, we study how traders' beliefs evolve during each day of trading. By comparing these beliefs to the prices, we are able to identify periods of the trading day in which traders herd. In most of the trading periods, a positive (albeit small) measure of informed traders herd. In an information-event day, on average, between 2 percent (4 percent) of informed traders herd-buy (sell).

Herd behavior generates serial dependence in trading patterns, a phenomenon documented in the empirical literature. Herding also causes informational inefficiencies in the market. On average, the misalignment between the price we observe and the price we would observe in the absence of herding is equal to 4 percent of the asset's unconditional fundamental value.

The rest of the paper is organized as follows. Section II describes the theoretical model. Section III presents the likelihood function. Section IV describes the data. Section V presents the results. Section VI concludes. An Appendix contains the proofs.

II. THE MODEL

Following Easley and O'Hara (1987), we generalize the original Glosten and Milgrom (1985) model to an economy where trading happens over many days.

An asset is traded by a sequence of traders who interact with a market maker. Trading days are indexed by $d = 1, 2, 3, \ldots$. Time within each day is discrete and indexed by $t = 1, 2, 3, \ldots$.

The asset

We denote the fundamental value of the asset in day d by V_d. The asset value does not change during the day, but can change from one day to the next. At the beginning of the day, with probability $1-\alpha$ the asset value remains the same as in the previous day $(V_d = v_{d-1})$, and with probability α it changes.[6] In the latter case, since as we will see, there are informed traders in the market, we say that an information event has occurred. If an information event occurs, with probability $1-\delta$ the asset value decreases to $v_{d-1} - \lambda^L$ ("bad informational event"), and with probability δ it increases to $v_{d-1} + \lambda^H$ ("good informational event"), where $\lambda^L > 0$ and $\lambda^H > 0$. Informational events are independently distributed over the days of trading. To simplify the notation, we define $v_d^H := v_{d-1} + \lambda^H$ and $v_d^L := v_{d-1} - \lambda^L$. Finally, we assume that $(1-\delta)\lambda^L = \delta\lambda^H$, which, as will become clear later, guarantees that the closing price is a martingale.

The market

The asset is exchanged in a specialist market. Its price is set by a market maker who interacts with a sequence of traders. At any time $t = 1, 2, 3\ldots$ 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, $\mathscr{A} = \{buy, sell, no\ trade\}$. We denote the action of the trader at time t in day d by X_t^d, and the history of trades and prices until time $t-1$ of day d 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 traders who (as we shall see) have some private information on the asset value. He will set different prices for buying and for selling, that is, there will be a bid-ask spread (Glosten and

[6]Note that v_{d-1} is the realization of the random variable V_{d-1}. Throughout the text, we will denote random variables with capital letters and their realizations with lower case letters.

Milgrom, 1985). We denote the ask price (the price at which a trader can buy) at time t by a_t^d and the bid price (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, that is,

$$a_t^d = E(V_d | h_t^d, X_t^d = buy, a_t^d, b_t^d),$$
$$b_t^d = E(V_d | h_t^d, X_t^d = sell, a_t^d, b_t^d).$$

The traders

There are a countable number of traders. Traders act in an exogenous sequential order. Each trader is chosen to take an action only once, at time t of day d. Traders are of two types, informed and noise. The trader's own type is private information.

In no-event days, all traders in the market are noise. In information-event days, at any time t an informed trader is chosen to trade with probability μ and a noise trader with probability $1-\mu$, with $\mu \in (0,1)$.

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 receive a private signal on the new asset value and observe the previous history of trades and prices, and the current prices. The private signal S_t^d has the following value-contingent densities:

$$g^H(s_t^d | v_d^H) = 1 + \tau(2s_t^d - 1),$$

$$g^L(s_t^d | v_d^L) = 1 - \tau(2s_t^d - 1),$$

with $\tau \in (0,\infty)$. (See Figure 1.)

For $\tau \in (0,1]$, the support of the densities is $[0,1]$. In contrast, for $\tau > 1$, the support shrinks to $[\frac{\tau-1}{2\tau}, \frac{\tau-1+2\sqrt{\tau}}{2\tau}]$ for g^H and to $[\frac{\tau+1-2\sqrt{\tau}}{2\tau}, \frac{\tau+1}{2\tau}]$ for g^L (in order for the density functions to integrate to one). Note that, given the value of the asset, the signals S_t^d are i.i.d. The signals satisfy the monotone likelihood ratio property. At each time t, the likelihood ratio after receiving the signal, $\frac{\Pr(V_d=v_d^H|h_t^d,s_t^d)}{\Pr(V_d=v_d^L|h_t^d,s_t^d)} = \frac{g^H(s_t^d|v_d^H)}{g^L(s_t^d|v_d^L)} \frac{\Pr(V_d=v_d^H|h_t^d)}{\Pr(V_d=v_d^L|h_t^d)}$, is higher than that before receiving the signal if $s_t^d > 0.5$, and lower if $s_t^d < 0.5$. For this reason we refer to a signal larger than 0.5 as a "good signal" and to a signal smaller than 0.5 as a "bad signal."

The parameter τ measures the informativeness of the signals. When $\tau \longrightarrow 0$, the densities are uniform and the signals are completely uninformative. As τ increases, the signals become

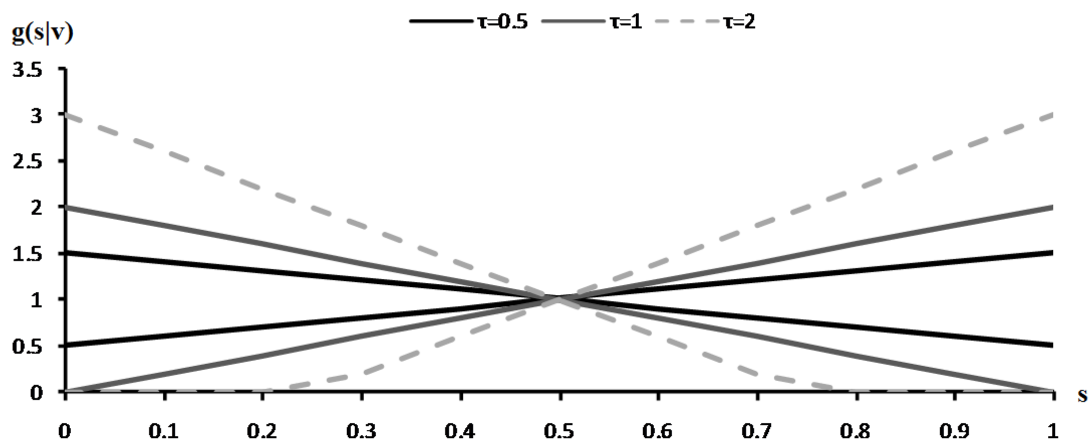

Figure 1. The signal. **Signal state-contingent density functions for different values of τ.**

more and more informative. For $\tau \in [0,1)$, the support of the distribution of the likelihood ratio is bounded away from 0 and infinity, while for $\tau \geq 1$ it is not. Following Smith and Sørensen (2000), in the first case we say that beliefs are bounded, and in the second case, that they are unbounded. With bounded beliefs, no signal realizations (even the most extreme ones) reveal the asset value with probability one. With unbounded beliefs, in contrast, some high (low) signal realizations are only possible when the asset value is high (low), and therefore, the signal can be perfectly informative.[7] As τ tends to infinity, the measure of perfectly informative signals tends to one.

An informed trader knows that an information event has occurred, and that, as a result the asset value has changed with respect to the previous day. Moreover, his signal is informative on whether the event is good or bad. Nevertheless, according to the signal realization that he receives and to the precision τ, he may not be completely sure of the effect of the event on the asset value. For instance, he may know that there has been a change in the investment strategy of a company, but not be sure of whether this change will affect the asset value in a positive or negative way. The parameter τ can be interpreted as measuring the precision of the information that the trader receives, or the ability of the trader to process such private information. Finally, note that, given our signal structure, informed traders are heterogenous, since they receive signal realizations with different degrees of informativeness about the asset's fundamental value.

An informed trader's payoff function, $U : \{v_d^L, v_d^H\} \times \mathscr{A} \times [v_d^L, v_d^H]^2 \longrightarrow \mathbf{R}^+$, is defined as

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

[7] In particular, any signal greater than or equal to $\frac{\tau+1}{2\tau}$ reveals that the asset value is v_d^H, whereas a signal lower than or equal to $\frac{\tau-1}{2\tau}$ reveals that the asset value is v_d^L.

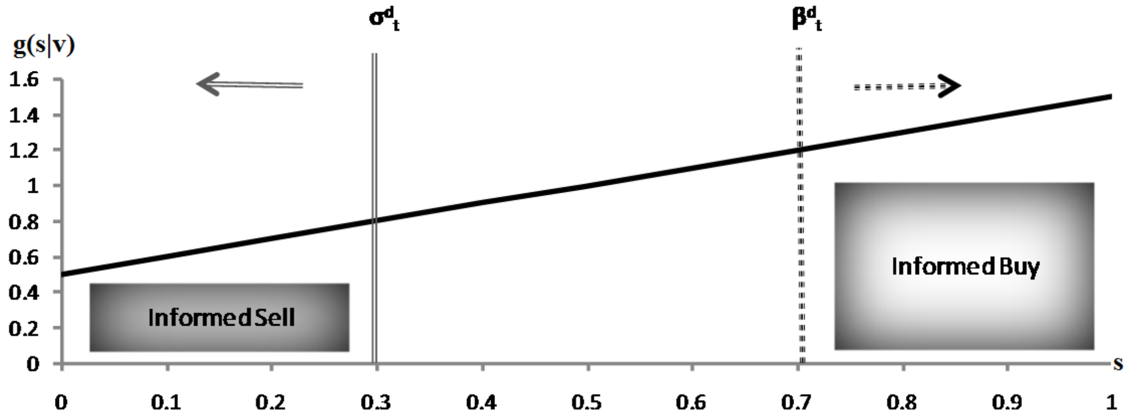

Figure 2. Informed trader's decision. **The figure illustrates the signal realizations for which an informed trader decides to buy or sell when $V_d = v_d^H$ (the signal density function is conditional on v_d^H).**

An informed trader chooses X_t^d to maximize $E(U(V_d, X_t^d, a_t^d, b_t^d)|h_t^d, s_t^d)$ (i.e., he is risk neutral). Therefore, he finds it optimal to buy whenever $E(V_d|h_t^d, s_t^d) \geq a_t^d$, and sell whenever $E(V_d|h_t^d, s_t^d) \leq b_t^d$. He chooses not to trade when $b_t^d < E(V_d|h_t^d, s_t^d) < a_t^d$.

Note that at each time t, the trading decision of an informed trader can be simply characterized by two thresholds, σ_t^d and β_t^d, satisfying the equalities

$$E\left[V_t^d|h_t^d, \sigma_t^d\right] = b_t^d$$

and

$$E\left[V_t^d|h_t^d, \beta_t^d\right] = a_t^d.$$

An informed trader will sell for any signal realization smaller than σ_t^d and buy for any signal realization greater than β_t^d. Obviously, the thresholds at each time t depend on the history of trades until that time and on the parameter values.[8]

Figure 2 (drawn for the case of a good informational event) illustrates the decision of informed traders. An informed trader buys the asset with a signal higher than the threshold value β_t^d, sells it with a signal lower than σ_t^d, and does not trade otherwise. The measure of informed traders buying or selling is equal to the areas (labelled as "Informed Buy" and "Informed Sell") below the line representing the signal density function.

[8] Since noise traders buy and sell with probabilities bounded away from zero, standard arguments prove that both the bid and ask prices, and the informed traders' signal thresholds exist and are unique.

Herd Behavior

To discuss herd behavior, let us start by introducing some formal definitions.

Definition An informed trader engages in herd-buying at time t of day d when he buys upon receiving a bad signal, that is,

$$E(V_d|h_t^d, s_t^d) > a_t^d \text{ for } s_t^d < 0.5.$$

Similarly, an informed trader engages in herd-selling at time t of day d when he sells upon receiving a good signal, that is,

$$E(V_d|h_t^d, s_t^d) < b_t^d \text{ for } s_t^d > 0.5.$$

In other words, a trader herds when he trades against his own private information. Since traders in our model receive different signals, it may well be (and typically will be the case) that, at a given point in time, traders with less informative signals (i.e., close to 0.5) will herd, whereas traders with more informative signals (close to the extremes of the support) will not. We are interested in periods of the trading day in which traders engage in herd behavior for at least some signal realizations. At any given time t, we can detect whether an informed trader herds for a positive measure of signals by comparing the two thresholds σ_t^d and β_t^d to 0.5. Since a trader engages in herd-buying behavior if he buys despite a bad signal $\left(s_t^d < 0.5\right)$, there is a positive measure of herd-buyers whenever $\beta_t^d < 0.5$.[9] A similar condition holds for herd-sellers.

Definition There is herd behavior at time t of day d when there is a positive measure of signal realizations for which an informed trader either herd-buys or herd-sells, that is, when

$$\beta_t^d < 0.5 \text{ or } \sigma_t^d > 0.5.$$

Figures 3 and 4 show an example of herd-buy and herd-sell, respectively, in a day with a good information event. The areas below the signal density function and between the thresholds and 0.5 represent the measures of informed traders who herd-buy and herd-sell.

The reason why herd behavior arises is that prices move "too slowly" as buy and sell orders arrive in the market. Suppose that, at the beginning of an information-event day, there is a sequence of buy orders. Informed traders, knowing that there has been an information event, attach a certain probability to the fact that these orders come from informed traders with good signals. The market maker, however, attaches a lower probability to this event, as he takes into account the possibility that there was no information event, and that all the buys came

[9] We identify an informed trader with the signal he receives: thus, "a positive measure of herd-buyers" means "a positive measure of signal realizations for which an informed trader herd-buys."

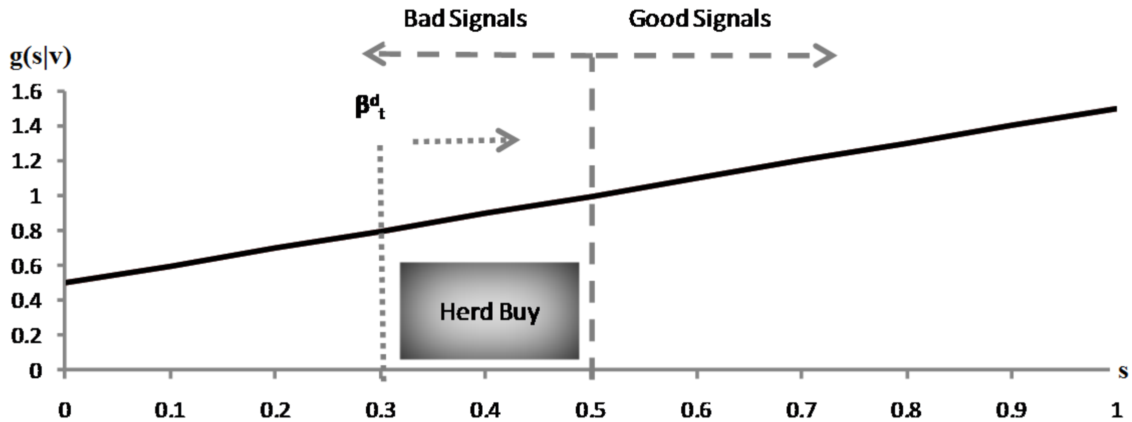

Figure 3. Herd-buy. In the figure, an informed trader buys even upon receiving a bad signal (higher than 0.3).

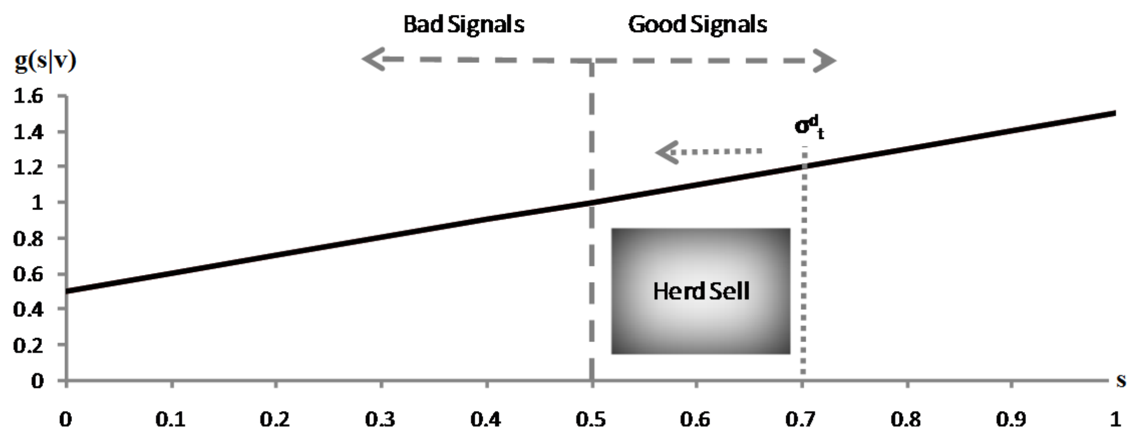

Figure 4. Herd sell. In the figure, an informed trader sells even upon receiving a good signal (lower than 0.7).

from noise traders. Therefore, after a sequence of buys, he will update the prices upwards, but by less than the movement in traders' expectations. Because traders and the market maker interpret the history of trades differently, the expectation of a trader with a bad signal may be higher than the ask price, in which case he herd-buys. Obviously, traders who receive signals close to 0.5 will be more likely to herd, since the history of trades has more weight in forming their beliefs.

We state this result in the next proposition:

Proposition For any finite τ, herd behavior arises with positive probability. Furthermore, herd behavior can be misdirected, that is, an informed trader can engage in herd-buy (sell) in a day of bad (good) information event.

Proof See the Appendix.

Avery and Zemsky (1998) have shown how herding can arise because of uncertainty on whether an information event has occurred (see their IS2 information setup). In our model herding arises for the same reason. Our contribution is to embed this theoretical reason to herd in a model that is suitable to empirical analysis.

When $\tau > 1$, extreme signals reveal the true value of the asset, and traders receiving them never herd. In the limit case of τ tending to infinity, all signal realizations become perfectly informative, with the result that no informed trader herds. Therefore, while our model allows for herd behavior, it also allows for the possibility that some traders (when $\tau > 1$) or all traders (when $\tau \longrightarrow \infty$) only rely on their private information and never herd.

The probability of herding depends on the parameter values. To take an extreme example, when α (the probability of an information event) is arbitrarily close to zero, the market maker has a very strong prior that there is no information event. He barely updates the prices as trades arrive in the market, and herding arises as soon as there is an imbalance in the order flow, as happens in the seminal model of Bikhchandani, Hirshleifer, and Welch (1992). In contrast, if α is close to 1, the market maker and the informed traders update their beliefs in very similar manners, and herding rarely occurs.

Herding is important also for the informational efficiency of the market. During periods of herd behavior, private information is aggregated less efficiently by the price as informed traders with good and bad signals may take the same action. The most extreme case is when traders herd for all signal realizations (e.g., traders herd-buy even for $s_t^d = 0$). In such a case, the market maker is unable to make any inference on the signal realization from the trades. The market maker, however, updates his belief on the asset value, since the action remains informative on whether an information event has occurred.[10] Since the market maker never stops learning, he gradually starts interpreting the history of past trades more and more similarly to the traders and, as a result, the measure of herders shrinks.

During an information-event day, the measure of herders changes with the sequence of trades, and can become positive more than once at different times of the day. Given that information always flows to the market, however, the bid and ask prices converge to the asset value almost surely.[11] Eventually the market maker learns whether a good event, a bad event, or no event occurred.[12]

[10]The market maker learns since in periods of herding, the proportion of buys and sells is different from that in an uninformed day. Essentially, whereas in our model there is herd behavior, there is never an informational cascade.

[11]The proof of convergence is standard and we omit it.

[12]Recall that we have assumed that $(1-\delta)\lambda^L = \delta\lambda^H$. This implies that $E(V_{d+1}|V_d = v_d) = v_d$. Since the price converges to the fundamental value almost surely, this guarantees that the martingale property of prices is satisfied.

III. The Likelihood Function

To estimate the herding model presented above, we have to specify its likelihood function. Let us denote the history of trades at the end of a trading day by $h^d := h^d_{T_d}$, where T_d is the number of trading dates in day d. We denote the likelihood function by

$$\mathscr{L}(\Phi; \{h^d\}_{d=1}^D) = \Pr\left(\{h^d\}_{d=1}^D | \Phi\right),$$

where $\Phi := \{\alpha, \delta, \mu, \tau, \varepsilon\}$ is the vector of parameters.

Note that we write the likelihood function for the history of trades only, disregarding bid and ask prices. In our model there is no public information: for this reason, there is a one-to-one mapping from trades to prices, and adding prices would be redundant.

The one-to-one mapping from trades to prices breaks down in the presence of public information, since price changes may be the result of public information arrival (as opposed to being only determined by the order flow). Nevertheless, our likelihood function for the history of trades would still be correctly specified. The reason is that the probability of any given trade only depends on whether the trader is informed, and, in such a case, on whether his belief is higher or lower than the market maker's; neither event is affected by the arrival of public information (since this would affect traders' and the market maker's beliefs in the same way, shifting all beliefs by the same amount).

Remember that information events are assumed to be independent. Moreover, the probability of a given sequence of trades in a day only depends on the value of the asset that day. For this reason, the likelihood of a history of trades over multiple days can be written as the product of the likelihoods of the histories of trades for each day:

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

Let us focus on the probability of a history of trades in a single day. As noted earlier, the *sequence* of trades, and not just the *number* of trades, conveys information. Having many buy orders at the beginning of the day is not equivalent to having the same number of buy orders spread out during the day. In fact, a particular sequence of buy or sell orders may create herd behavior: in periods of herding, the probability of a trade depends on the measure of informed traders who herd and is different from the probability in the absence of herding. Therefore, we have to compute the probability of a history of trades recursively, that is,

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

where the probability of an action at time t of day d, $\Pr(x_t^d | h_t^d, \Phi)$, depends on the measure of informed traders who buy, sell or do not trade after a given history of trades h_t^d.

Using the law of total probability, at each time t, we compute $\Pr(x_t^d|h_t^d,\Phi)$ in the following way:

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

Let us consider, first, the probability of an action conditional on a good-event day. As illustrated above, at each time t, in equilibrium there is a signal threshold β_t^d such that an informed trader buys for any signal realization greater than β_t^d, that is,

$$E(V_d|h_t^d,\beta_t^d) = a_t^d = E(V_d|h_t^d,X_t^d = buy,a_t^d,b_t^d),$$

which can be written as

$$v_{d-1}+\lambda^H\Pr(v_d^H|h_t^d,\beta_t^d)-\lambda^L\Pr(v_d^L|h_t^d,\beta_t^d) =$$
$$v_{d-1}+\lambda^H\Pr(v_d^H|h_t^d,buy_t^d)-\lambda^L\Pr(v_d^L|h_t^d,buy_t^d),$$

or, after some manipulations, as[13]

$$\Pr(v_d^H|h_t^d,\beta_t^d)-\Pr(v_d^H|h_t^d,buy_t^d) = \frac{\delta}{1-\delta}(\Pr(v_d^L|h_t^d,\beta_t^d)-\Pr(v_d^L|h_t^d,buy_t^d)).$$

The probabilities in this equation can easily be expressed as a function of the traders' and market maker's beliefs at time $t-1$ and of the parameters. (See the Appendix for further details.)

Note that at time $t = 1$, the prior beliefs of the traders and the market maker are a function of the parameters only. Therefore, we can easily compute β_1^d as the solution to the equation above. After observing x_1^d, we can update the market maker's and traders' beliefs, repeat the same procedure for time 2 and compute β_2^d.[14] We do so recursively for each time t, always conditioning on the previous history of trades.

From β_t^d, we can compute the probability of a buy order in a good-event day. For the sake of exposition, let us focus on the case in which $\tau \in [0,1)$, that is, let us concentrate on the case of bounded beliefs. In this case:

$$\Pr(buy_t^d|h_t^d,V_d = v_d^H,\Phi) =$$
$$\mu\int_{\beta_t^d}^1(1+\tau(2s_t^d-1))ds_t^d+(1-\mu)\left(\frac{\varepsilon}{2}\right) =$$
$$\left(\left(\tau(1-\beta_t^{d^2})+(1-\tau)(1-\beta_t^d)\right)\mu+(1-\mu)\left(\frac{\varepsilon}{2}\right)\right).$$

[13] To simplify the notation, we omitted a_t^d and b_t^d in the conditioning. More importantly, note that the magnitudes of the shocks that buffet the asset's value (λ^L and λ^H) do not appear in this equation (they cancel out since $(1-\delta)\lambda^L = \delta\lambda^H$). This is important, since it implies that we do not need to estimate them.

[14] To update the beliefs of traders and the market maker after x_1^d we also need to know how to compute the probability of a sell and of a no-trade, which we discuss below.

We use a similar procedure to compute the probability of a sell, that is,

$$\Pr(sell_t^d|h_t^d, V_d = v_d^H, \Phi) =$$
$$\mu \int_0^{\sigma_t^d} (1 + \tau(2s_t^d - 1))ds_t^d + (1-\mu)\left(\frac{\varepsilon}{2}\right) =$$
$$\left(\left((1-\tau)\sigma_t^d + \tau\sigma_t^{d^2}\right)\mu + (1-\mu)\left(\frac{\varepsilon}{2}\right)\right).$$

Finally, the probability of a no trade is simply the complementary to the probabilities of a buy and a sell.

The analysis for the case of a bad information event ($V_d = v_d^L$) follows the same steps. The case of unbounded beliefs, where $\tau \geq 1$, can be dealt with in a similar manner. The only changes are the extremes of integration when computing the probability of a trade.

The case of a no-event day ($V_d = v_{d-1}$) is easy, since the probabilities of a buy or sell is $\frac{\varepsilon}{2}$ and the probability of a no-trade is $1-\varepsilon$.

Finally, to compute $\Pr(x_t^d|h_t^d, \Phi)$, we need the conditional probabilities of V_d given the history until time t, that is, $\Pr(V_d = v|h_t^d, \Phi)$ for $v = v_d^L, v_{d-1}, v_d^H$. These can also be computed recursively by using Bayes's rule. This completes the description of the likelihood function.

To conclude this section, let us give an intuition of how the model is identified. For simplicity's sake let us consider only the number of buys, sells and no trades in each day.[15] Similarly to analogous structural models of market microstructure, our model classifies days into high-volume days with a prevalence of buys ("good-event" days), high-volume days with a prevalence of sells ("bad-event" days) and low-volume days ("no-event" days). The parameter α is identified by the proportion of high-volume days. The direction of the imbalance in the event days identifies δ. No-event days allow us to identify ε, since in no-event days only noise traders trade. Finally, in good-event days, the ratio between buys and sells is determined by the proportion of traders who trade in the right direction (i.e., buy when the there is a good event), which depends on μ and τ. An analogous argument holds for bad-event days. Any given estimate of μ and τ corresponds to only one predicted ratio between buys and sells in the two types of days.[16]

IV. DATA

The aim of our study is mainly methodological. We are not trying to gauge the importance of herding across different stocks or different markets. Our goal is to develop a methodology to carry out a structural estimation of herding based on a market microstructure model.

[15]In our empirical estimation we use much more information than that, since we take into account the entire sequence of trades when constructing the likelihood function.

[16]For a further argument for identification, see footnote 21.

For this reason, we perform our empirical analysis on a single stock that has been studied before. Our choice is Ashland Inc., a stock traded in the New York Stock Exchange and used in Easley, Kiefer, and O'Hara (1997).[17] We obtained the data from the TAQ (Trades and Quotes) dataset.[18] The dataset contains the posted bid and ask prices (the "quotes"), the prices at which the transactions occurred (the "trades"), and the time when the quotes were posted and when the transactions occurred. We used transactions data on Ashland Inc. in 1995, for a total of 252 trading days. The data refer to trading in the New York Stock Exchange, the American Stock Exchanges, and the consolidated regional exchanges.

The TAQ dataset does not sign the trades, that is, it does not report whether a transaction was a sale or a purchase. To classify a trade as a sell or a buy order, we used the standard algorithm proposed by Lee and Ready (1991). We compared the transaction price with the quotes that were posted just before a trade occurred.[19] Every trade above the midpoint was classified as a buy order, and every trade below the midpoint was classified as a sell order; trades at the midpoint were classified as buy or sell orders according to whether the transaction price had increased (uptick) or decreased (downtick) with respect to the previous one. If there was no change in the transaction price, we looked at the previous price movement, and so on.[20]

TAQ data do not contain any direct information on no trades. We used the established convention of inserting no-trades between two transactions if the elapsed time between them exceeded a particular time interval (see, e.g., Easley, Kiefer, and O'Hara, 1997). We obtained this interval by computing the ratio between the total trading time in a day and the average number of buy and sell trades over the 252 days (see, e.g., Chung, Li, and McInish, 2005). In our 252 trading day window, the average number of trades per day was 90.2. We divided the total daily trading time (390 minutes) by 90.2, and obtained a unit-time interval of 259 seconds (i.e., on average, a trade occurred every 259 seconds). If there was no trading activity for 259 seconds or more, we inserted one or more no-trades to the sequence of buy and sell orders. The number of no-trades that we inserted between two consecutive transactions was equal to the number of 259-second time intervals between them. To check the robustness of our results, we also replicated the analysis for other no-trade time intervals (2, 3, 4, 5, 6 and 7 minutes).

Our sample of 252 trading days contained on average 149 decisions (buy, sell, or no-trade) per day. The sample was balanced, with 30 percent of buys, 31 percent of sells and 40 percent of no trades.

[17] The name of the stock is slightly different, since the company changed name in 1995, and Easley *et al.* (1997) use 1990 data.

[18] Hasbrouck (2004) provides a detailed description of this dataset.

[19] Given that transaction prices are reported with a delay, we followed Lee and Ready (1991)'s suggestion of moving each quote ahead in time of five seconds. Moreover, following Hasbrouck (1991, p. 581), we ignore quotes posted by the regional exchanges.

[20] We classified all trades, with the exception of the opening trades, since these trades result from a trade mechanism (an auction) substantially different from the mechanism of trade during the day (which is the focus of our analysis).

V. RESULTS

We first present the estimates of the model parameters, and then illustrate the importance of herd behavior in the trading activity of Aschland Inc. during 1995.

A. Estimates

We estimated the parameters through maximum likelihood, using both a direct search method (Nelder-Mead simplex) and the Genetic Algorithm.[21] The two methods converged to the same parameter values. Table 1 presents the estimates and the standard deviations for the five parameters of the model.[22]

Parameter	Estimate	S.D.
α	0.28	0.03
δ	0.62	0.06
μ	0.42	0.01
τ	0.45	0.02
ε	0.57	0.00

Table 1. Estimation Results.

The table shows the estimates for the five parameters of the model and their standard deviations.

Information events are relatively frequent: from the estimate of α, we infer that the probability of an information event is 28 percent, that is, in almost a third of trading days trading activity is motivated by private information. There is a small imbalance between good and bad-event days: the probability of a good information event is 62 percent (although the parameter δ has a relatively high standard deviation).[23] During event days, the proportion of traders with private information is 42 percent. The remaining trading activity comes from noise traders, who trade 57 percent of the time. Moreover, private information is noisy (that is, it is not perfectly informative). The estimate for τ is 0.45, which means that the probability of receiving an "incorrect signal" — i.e., a signal below 0.5 when we are in a good-

[21]We also simulated the theoretical model and verified that we could recover the model's parameters. Both methods converged to the true parameter values, which provides further evidence in favor of identification.

[22]The standard deviation for ε is 0.00172, rounded to 0.00 in the table. The standard deviations for ε are also rounded to 0.00 in Tables 2 and 3.

[23]Note that δ is greater than 0.5, although in the sample the number of buys and sells is essentially balanced. This happens because among the days with high trading volume (classified as event days), a higher number of days have a positive trade imbalance than a negative one. To see this, consider the posterior beliefs of δ and α at the end of each day. In 22 percent of days the posterior belief of both α and δ is above 0.5 (i.e., we are in an good-event day), whereas in only 12 percent of days the posterior belief of α is above 0.5 and that of δ is below 0.5 (i.e., we are in a bad-event day).

information event day or a signal above 0.5 when we are in a bad-information event day — is 39 percent.[24]

As explained above, we constructed our dataset adding a no-trade after each 259 seconds of trading inactivity. As a robustness check, we repeated the estimation on several other datasets, where we added a no trade for different intervals of trading inactivity. We report these estimates in Table 2.

	NT=120	S.D.	NT=180	S.D	NT=240	S.D.
α	0.20	0.02	0.25	0.02	0.27	0.02
δ	0.72	0.02	0.67	0.04	0.61	0.01
μ	0.26	0.01	0.36	0.01	0.41	0.01
τ	0.44	0.01	0.40	0.03	0.44	0.02
ε	0.33	0.00	0.45	0.00	0.54	0.00
β	0.64		0.63		0.62	
Γ	0.32		0.33		0.34	

	NT=300	S.D.	NT=360	S.D.	NT=420	S.D.
α	0.30	0.03	0.27	0.03	0.27	0.04
δ	0.60	0.03	0.61	0.03	0.55	0.07
μ	0.40	0.01	0.39	0.01	0.37	0.01
τ	0.52	0.03	0.59	0.03	0.67	0.01
ε	0.62	0.00	0.69	0.00	0.74	0.00
β	0.62		0.62		0.62	
Γ	0.32		0.31		0.30	

Table 2. Robustness Checks for Different No-trade Intervals.

The table shows the estimates for various no-trade intervals, from 2 to 7 minutes. The last two rows report two more statistics derived from the estimated parameters and explained in the text.

The estimates of the probability of an information event (α) and of a good event (δ) are fairly similar over the different numbers of seconds defining a no-trade interval. The estimate of ε increases with the the no-trade interval: this is expected since the number of no-trades in the sample (and also in the no-event days) becomes smaller and smaller. To have a better description of the trading activity in no-event days, we computed the probability of observing at least one trade during a 5-minute interval in a no-event day: $\beta = 1-(1-\varepsilon)^{\frac{300}{Seconds}}$ (where "Seconds" is the no-trade interval). Table 2 shows this probability to be independent of the choice of the no-trade interval.

The parameter μ is quite stable across samples, whereas τ increases. To understand this, it is useful to observe that if both τ and μ were constant, as ε increases the estimated proportion of trading activity due to traders not having correct information (either because they are noise

[24] Given the signal density functions, the probability of an incorrect signal is given by $0.5-0.25\tau$.

or because their signal is incorrect) would increase. In contrast, this proportion should obviously be independent of our choice of no-trade interval. This is indeed the case. To show this we computed the parameter

$$\Gamma = \frac{\mu(0.5+0.25\tau)}{(1-\mu)\varepsilon+\mu},$$

which represents the proportion of correctly informed traders (i.e., informed traders with a signal greater than 0.5 in a good-event day) over the sum of all informed traders and the noise traders who trade. In other words, Γ is approximately equal to the fraction of trades coming from informed traders with the correct signal.[25] It is remarkable that Γ, which equals 0.34 when the no-trade interval is 259 seconds, is constant across all the different datasets that we used to estimate the model's parameters. This shows the robustness of our results to the choice of the no-trade interval.

Let us now discuss how our results compare to different specifications of the model. A natural comparison is with a model in which the signal precision is not estimated, but is restricted to be perfectly informative (i.e., $\tau \longrightarrow \infty$). This is the case studied by Easley, Kiefer, and O'Hara (1997). In this case, all informed traders follow their own private information, the sequence of trades has no informational content beyond the aggregate numbers of buys, sells and no trades, and herding never arises. As a result, the likelihood function does not need to be computed recursively (see Easley, Kiefer, and O'Hara, 1997, for a detailed description). Table 3 presents the estimated parameters.

Parameter	Estimate	S.D.
α	0.33	0.04
δ	0.60	0.06
μ	0.17	0.01
ε	0.58	0.00

Table 3. Parameter Estimates for the Easley, Kiefer, and O'Hara (1997) Model.

The table shows the estimates for the four parameter model of Easley, Kiefer, and O'Hara (1997), in which informed traders know the true asset value. The no-trade interval is 259 seconds.

The estimates for α and δ are very close to those we obtained for our model. This shows that the classification of days is not affected by the specification of the signal structure. Similarly, the estimates for ε in the two models are almost the same. This is not surprising since ε captures the trading activity of noise traders, and is not affected by assumptions on the structure of private information. The parameter μ is smaller in the restricted model, which is intuitive since this model imposes that informed traders all receive the correct signal (i.e., they know whether a good or a bad information event occurred).

The restriction in Easley, Kiefer, and O'Hara (1997) is not supported by the data. The likelihood ratio test overwhelmingly rejects the hypothesis of perfectly informative signals, with a

[25]The approximation is due to the fact that we are ignoring that, because of the bid-ask spread, a small measure of informed traders may not trade. Easley, Kiefer, and O'Hara (1997) report a similar composite parameter when analyzing their results for different no-trade intervals.

LR statistic of 272.15 (and a p-value of zero).[26] This is important for our aims, since the fact that signals are not perfectly informative implies that the sequence in the order flow matters. In other words, the number of buys, sells and no-trades at the end of the day is not a sufficient statistic for the pattern of trading activity. Depending on the sequence, herd behavior by informed traders may occur in equilibrium.

In the market microstructure literature, a great deal of attention has been given to the PIN, a measure of the probability that a trade comes from an informed trader (see, among others, Easley and others, 1996, and the literature cited in Chung, Li, and McInish, 2005). This measure is given by PIN$= \frac{\alpha\mu}{\alpha\mu+\varepsilon(1-\alpha\mu)}$, where the numerator is the beginning-of-the-day probability that a trade is information based and the denominator is the probability that a trade occurs. With the estimated parameters of our model, the PIN equals 19 percent, whereas for the Easley, Kiefer, and O'Hara (1997) model it is only 9 percent.[27] If we adjust for the fact that in our model the information may not be correct (i.e., we multiply the numerator by the probability of a correct signal $0.5+0.25\tau$), the measures of information-based trading given by the two models become almost the same. Since the null that the signal is perfectly precise is rejected by the data, our results suggest that the PIN, as usually computed, measures the proportion of informed-based trading coming from traders receiving the correct information and not the overall proportion of information-based trading. Essentially, the difference is due to the fact that in the previous literature incorrect trades (e.g., sell in a good-event day) can only be due to exogenous, non-informative (e.g., liquidity) reasons to trade, whereas in our setup we do not exclude that they may come from informed traders who either receive the incorrect information or herd.

Finally, note that a 99 percent confidence interval for τ does not include 1.[28] This means that there is evidence in our sample that there are no realizations of the signal that reveal the true asset value with probability one. In the jargon of the social learning literature, signals are bounded.

B. Herd Behavior

The estimates of the parameters α and τ imply that herd behavior can occur in our sample. Since the estimate of α is clearly lower than 1,[29] there is information uncertainty in the market, which is a necessary condition for the mechanism of herd behavior highlighted in Section II to work. Moreover, the estimate $\tau = 0.44$ means that traders receive a signal that is noisy

---

[26] A note of warning on the result of the test is needed here, since the null hypothesis is on the boundary of the parameter space (see Andrews, 2001).

[27] We compute the PIN for our model using the same formula as Easley and others (1996). They interpret the PIN as the probability of a trade coming from an informed trader at the beginning of the day. In our model, since the signal is continous, the interpretation is correct only if we ignore the bid-ask spread (otherwise, some informed traders may decide not to trade because their expectations fall inside the bid-ask spread.) We use this approximation for simplicity's sake and to keep comparability with the existing work on the PIN.

[28] Since the parameter's standard deviation is 0.02, this is the case for any reasonable confidence interval.

[29] The parameter's standard deviation is 0.03. See the argument in the previous footnote.

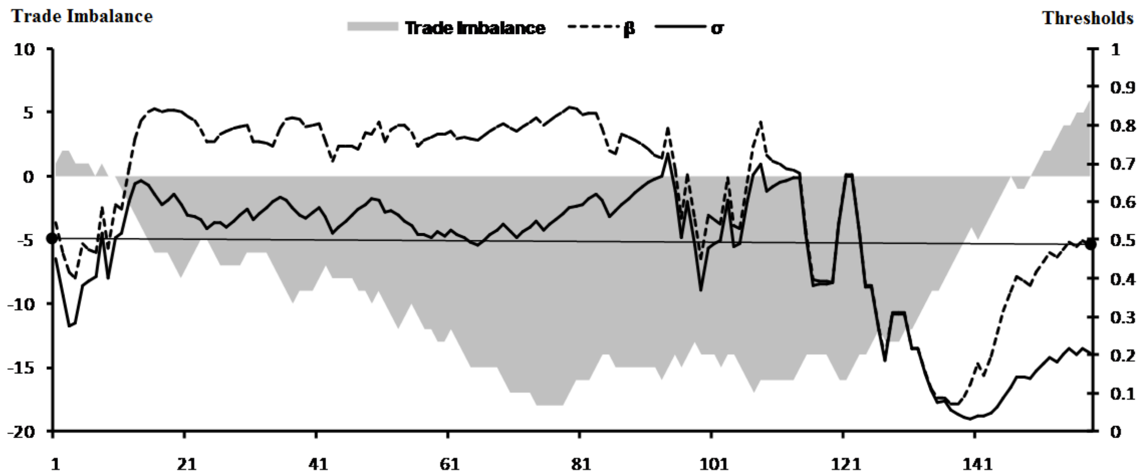

Figure 5. A day of trading. The figure reports the evolution of the trade imbalance (shaded line), buy threshold (dashed line) and sell threshold (solid line) in one day of trading. The thresholds are measured on the right vertical axis, and the trade imbalance on the left vertical axis. Herd-buy occurs when the solid line is above 0.5 (indicated by a horizontal line) and herd-sell when the dashed line falls below 0.5.

(i.e., not perfectly informative) and may decide to act against it (i.e., buy upon receiving a bad signal or sell upon receiving a good one).

The Frequency of Herding

Recall that there is herd behavior at time t of day d when there is a positive measure of signal realizations for which an informed trader either herd-buys or herd-sells, that is, when, in equilibrium, either $\beta_t^d < 0.5$ (herd-buy) or $\sigma_t^d > 0.5$ (herd-sell). To gauge the frequency of herd behavior in our sample, for each trading day we computed the buy thresholds (β_t^d) and the sell thresholds (σ_t^d) given our parameter estimates. As an illustration, Figure 5 shows the thresholds (on the right vertical axis) for one day out of the 252 days in the sample. Whenever the buy threshold (dotted line) drops below 0.5 or the sell threshold (solid line) goes above 0.5, there is herd behavior. The shaded area (measured on the left vertical axis) represents the trade imbalance, that is, at eacht time t, the number of buys minus the number of sells arrived in the market from the beginning of the day until time $t-1$. As one can see, herd-buying occurs at the beginning of trading activity, as the trade imbalance is positive, that is, as more buy orders arrive in the market. This is followed by a long stretch of herd-sells, as sell orders arrive and the trade imbalance becomes negative. At the very end of the day, herd behavior effectively disappears.

In this section we look at the frequency of herding, asking how often the buy threshold is below 0.5, or the sell threshold is above 0.5. It should be clear that this analysis is relevant both for event and no-event days. In both cases, the existence of signal realizations for which informed traders herd (if an information event has occurred, which neither the market maker nor an external observer knows at the moment of the trade) modifies the way the market maker updates the price, which, as we shall see, affects the market's informational efficiency.

In our sample, herding happens quite frequently: over the 252 days of trading, β_t^d is below 0.5 in 30 percent of trading periods, and σ_t^d is above 0.5 in 37 percent of trading periods. Moreover, there are some days where herding is very pronounced. Table 4 reports the proportion of days in which the buy or sell thresholds (i.e., β_t^d or σ_t^d) cross 0.5 at least 10, 30, and 50 percent of the time. Herd-buying was observed in more than 50 percent of the trading times in 23 percent of the 252 days in our sample. Similarly, herd-selling was observed in more than 50 percent of the trading times in 35 percent of the 252 days in our sample.

	$> 10\%$	$> 30\%$	$> 50\%$
Herd-Buy	0.79	0.47	0.23
Herd-Sell	0.83	0.58	0.35

Table 4. Days with High Frequency of Herd Behavior.

The table shows the proportion of days in which the percentage of trading periods with herd behavior was higher than 10, 30 or 50. For instance, in 23 percent of days, herd-buy periods were at least 50 percent of the total.

	Herd-Buy	Herd-Sell
Mean Time	60	61
Mean Length	9	10
Length S.D.	13	15
Max Length	99	120

Table 5. Time and Length of a Herd.

The first row of the table shows the mean trading period in which herd behavior occurred. The other rows show the mean, standard deviation and maximum for the number of consecutive trading periods in which there was herd behavior.

Table 5 reports the mean period of the day when we observed herding, that is, approximately the 60^{th} trading period for both herd-buy and herd-sell (which, in clock time, is roughly after 2 hours and 36 minutes of trading).[30] It is also interesting to ask how long periods of herding last, that is, for how many trading periods after a herd starts we observe a positive measure of herders. Herd-buys last on average 9 trading periods (corresponding to about 24 minutes) and herd-sells last on average for 10 periods. There is, however, pronounced variability in the length of herds, with a standard deviation of 13 and 15 trading periods. The longest herd-buy

[30] On average there are 149 decisions in a day (390 minutes). Therefore, the average length of a trading period is 2.6 minutes.

lasted for 99 periods (about 257 minutes) and the longest herd-sell lasted 120 periods (312 minutes).

Proportion of Herders

The previous analysis helps us gauge how often herding occurs in our sample. The fact that at a given time the buy (sell) threshold is lower (higher) than 0.5, however, does not tell us how likely it is for an informed trader to herd at that time. This is captured by the measure of signal realizations for which an informed trader herds. As we will discuss in detail in the next subsection, this measure is also very important for the informational efficiency of the market. The higher the measure of signal realizations for which traders herd, the lower the informational efficiency.

To compute the measure of herd-signal realizations, however, we need to know the distribution of signals on any given day. This, in turn, depends on whether the day of trading was a good-event or a bad-event day. To this purpose, we classified a day as a good-event day (bad-event day) if two conditions hold: a) $\Pr(V_d \neq v_{d-1}|h_{T_d}) > 0.9$ and b) either $\Pr(V_d = v^H|h_{T_d}, V_d \neq v_{d-1}) > 0.9$ (good-event day) or $\Pr(V_d = v^L|h_{T_d}, V_d \neq v_{d-1}) > 0.9$ (bad-event day). That is, we classified a day as a good (bad) event day if at the end of the day, the posterior probabilities of an information event occurring, and of the event being good (bad) were both higher than 0.9.[31]

We concentrated our analysis on the days classified as good-event or bad-event days. For each trading period, we computed the proportion of bad signals for which informed traders would herd-buy, and the proportion of good signals for which informed traders would herd-sell. Consider, for instance, a good-event day. On such a day, the signal is distributed according to $g^H(s_t^d|v_d^H) = 1 + 0.44(2s_t^d - 1)$. For each trading period in which $\beta_t^d < 0.5$, we computed the measure of signals between β_t^d and 0.5 (i.e., the measure of signals for which an informed trader herd-buys). We then divided this measure by the measure of signals between 0 and 0.5 (the measure of all bad signals, i.e., all signals for which the informed trader could potentially herd-buy). We refer to this ratio as the "proportion of herd-buyers," since it is the proportion of informed traders with a bad signal who would nevertheless herd-buy were they to trade in that period. The proportion of herd-sellers was calculated in a similar way. We report the average results in Tables 6 and 7.

	Average	S.D.	Max
Herd-Buyers	2%	3%	11%
Herd-Sellers	4%	6%	29%

Table 6. Percentage of Herders.

The table shows the mean, the standard deviation and the maximum of the percentage of herd-buyers and herd-sellers.

[31] Of course, the 0.9 threshold is arbitrary. As a robustness check we repeated the calculations for 0.75, 0.8 and 0.85 and obtained very similar results (see the Appendix).

	$> 1\%$	$> 5\%$	$> 10\%$
Herd-Buyers	0.53	0.16	0.07
Herd-Sellers	0.55	0.22	0.11

Table 7. Event Days with a High Percentage of Herders.

The table shows the proportion of days in which the percentage of herd-buyers or herd-sellers was higher than 1, 5 or 10. For instance, in 11 percent of days, the percentage of herd-sellers was at least 10.

On average, across all event days, the proportion of herd-buyers was 2 percent and that of herd-sellers 4 percent. These proportions are, however, quite variable across days, reaching a maximum of 11 percent for herd-buy and 29 percent for herd-sell. Misdirected herding (i.e., herd-buying in a bad-event day and herd-selling in a good-event day) does occur: on average, in a bad-event day the proportion of herd-buyers was 1 percent; in a good-event day, the proportion of herd-sellers was 2 percent.

As Table 7 shows, there are a substantial number of days where the percentage of herd-buyers or sellers is significant: for instance, in 7 percent of event days, the proportion of informed traders who herd-buy was higher than 10 percent; similarly, in 11 percent of event days, the proportion of informed traders who herd-sell was higher than 10 percent. This confirms the result of the previous section that herding behavior seems to be particularly concentrated in some days of trading.

An important question is whether herding usually happens after buy or sell orders have accumulated in the market. Table 8 shows that this is actually the case. The first row of Table 8 reports the average level of the trade imbalance in periods of herd-buy (i.e., when $\beta_t^d < 0.5$). The trade imbalance is on average positive both in good and in bad-event days. This means that herd-buy usually happens when there has been a preponderance of buys. Similarly, the second row of Table 8 shows that herd-selling usually occurs when there has been a preponderance of sells.

	Good-Event Days	Bad-Event Days
$\beta_t^d < 0.5$	9	3.5
$\sigma_t^d > 0.5$	-8.2	-12.9

Table 8. Trade Imbalance in Periods of Herding.

The table shows the average level of the trade imbalance in periods of herd-buy and herd-sell.

	Good-Event Days	Bad-Event Days
$\Pr(buy\|\beta_t^d < 0.5)$	0.48	0.31
$\Pr(buy)$	0.43	0.26
$\Pr(sell\|\sigma_t^d > 0.5)$	0.39	0.48
$\Pr(sell)$	0.27	0.45

Table 9. Proportion of Buys and Sells in Periods of Herding.

The table shows the proportion of buys and sells in periods of herd-buy, in periods of herd-sell and in the whole sample.

Obviously, by definition, herd-buying increases the proportion of buys and herd-selling increases the proportion of sells. Table 9 illustrates this point, by showing the frequencies of buy orders and of sell orders that we observe in periods of herd-buying and herd-selling and contrasting them with the overall frequencies. In good-event days, for instance, the overall frequency of buy orders is 43 percent. This frequency goes up to 48 percent when there is herd-buying. It is important to note that the results of Tables 8 and 9 taken together imply that higher positive (negative) levels of the trade imbalance increase the probability of a buy (sell) order. That is, herd behavior generates serial dependence in the trading pattern during the day.[32]

Informational Inefficiency

In periods of herd behavior, a proportion of informed traders do not trade according to their private information; as a result, information is aggregated less efficiently by the asset price. It is easy to show, in fact, that trades convey the maximum amount of information when informed traders buy upon receiving a good signal and sell upon receiving a bad one.[33] In periods of herding, in contrast, traders may buy even with a bad signal or sell even with a good one.

To quantify the informational inefficiency caused by herding, we proceeded in the following way. We simulated the history of trades and prices over many days for our theoretical model, using our estimates of the parameter values. We then compared the simulation results with two benchmarks that capture the price behavior in an informationally efficient market. In the first benchmark, we simulated the model forcing informed traders to buy (sell) upon a good (bad) signal. In other words, informed traders (irrationally) never herded and always followed their private information. As a second benchmark, we considered the case in which there is

[32]In a related strand of literature, Hasbrouck (1991, 1991a) and Chung, Li, and McInish (2005), among others, provide evidence of autocorrelation in trades. Easley, Kiefer, and O'Hara (1997a) and Easley and others (2008) recognize the importance of the sequence of trades in conveying information. To capture it, they allow for path dependence in noise trading which is due to unmodeled reasons. In contrast, the sequence becomes important in our setup for reasons dictated by economic theory.

[33]By this, we mean that the strategy $X_t^d = -1$ for $s_t^d < 0.5$ and $X_t^d = 1$ for $s_t^d > 0.5$ minimizes $E[(E(V_d|h_t^d, x_t^d) - V_d)^2]$. Note that, for simplicity's sake, we are assuming that traders can only buy or sell, since in equilibrium they abstain from trading only when their expectation is inside the bid-ask spread.

no information uncertainty, that is, the market maker knows whether there has been an informational event. As a result, he updates his beliefs (and prices) exactly as informed traders do, and, because of this, informed traders never herd. Essentially, in the first benchmark, the market is efficient because traders (irrationally) follow their signals; in the second benchmark, the market is efficient because the informational asymmetry between traders and market makers due to event uncertainty is eliminated. The difference between the two scenarios is caused by the bid-ask spread. When, in the first benchmark, we force informed traders to buy (sell) upon a good (bad) signal, we disregard the incentive not only to herd, but also to abstain from trading (because a trader's expectation may fall within the bid-ask spread).[34]

We simulated the price paths for $100,000$ days of trading (with 149 trading periods per day) for our theoretical model and for the two benchmarks. Then, at each time t of any day d, we computed the distance (i.e., the absolute value of the difference) between the public belief $E(V_d|h_t^d)$ in our model and that in the benchmark.[35]

Table 10 presents the average distance taken over all trading periods as a percentage of the expected value of the asset. For the first benchmark, we present the average distance both over all days and over informed days only. For the second benchmark, by construction, only the average distance over event days is meaningful.

No-Herd Benchmark		No Event Uncertainty Benchmark
All Days	Event Days	Event Days
4%	10%	7%

Table 10. Informational Inefficiency

The table shows the average distance between the public belief in the model and that in the benchmarks as a percentage of the expected value of the asset.

Av. Distance	No-Herd Benchmark		No Event Uncertainty Benchmark
	All Days	Event Days	Event Days
$\geq 10\%$	0.13	0.37	0.21
$\geq 15\%$	0.07	0.20	0.11
$\geq 20\%$	0.04	0.11	0.06

Table 11. Days of High Informational Inefficiency

The table shows the proportion of days in which the average distance between the public belief in the model and that in the benchmarks is higher than 10, 15, or 20 percent of the expected asset value.

[34] See previous footnote.

[35] The public belief $E(V_d|h_t^d)$ is always between the bid and ask prices. It is common in the literature to interpret it as the price (abstracting from the bid-ask spread).

The average distance between the prices in our model and in the first benchmark over all days amounts to 4 percent of the asset expected value. If we focus our attention on event days, the distance is higher (10 percent), since the imbalance between buys and sells causes herding to arise more often. The average distance between the public beliefs using the second benchmark is similar, equal to 7 percent. In Table 11 we repeated these computations for days in which herding was more pronounced. In 7 percent of days (20 percent of event days) the distance between the price and the non-herding price is greater 15 percent. This suggests that there are days when intraday herding affects the informational properties of the price in a very significant manner.

VI. Conclusion

We developed a theoretical model of herd behavior in financial markets amenable to structural estimation with transaction data. We estimated the model using data for a NYSE stock (Ashland Inc.) in 1995, and identified the periods in each trading day when informed traders herd. We found that herding was present in the market and fairly pervasive on some trading days. Moreover, herding generated important informational inefficiencies.

The main contribution of this paper is methodological: it provides an empirical strategy to analyze herding within a structural estimation framework. This contrasts with existing empirical studies of herding, which are based on atheoretical, statistical measures of trade clustering.

In future research, we plan to use our methodology to investigate the importance of herding for a large number of stocks, by analyzing how herding changes with stock characteristics (e.g., large stocks versus small ones) and with the macroeconomic environment (e.g., crises versus tranquil periods). We also plan to contribute to the existing literature on information and asset pricing, both by studying whether our measure of market informativeness (which takes into account that information may not be perfectly precise) improves the performance of the information factor, and by seeing whether herding itself is a risk factor priced in the market.

Finally, whereas our interest was in learning in financial markets, our methodology could be fruitfully used in fields other than financial economics. The voluminous and growing theoretical literature on social learning has been intensively tested in laboratory experiments. It is, however, generally acknowledged that these models cannot be easily studied with field data, because we lack data on private information. Our methodology shows how this problem can be overcome.

References

Avery, Christopher and Peter Zemsky, 1998, "Multidimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, vol. 88, pp. 724-748.

Andrews, Donald, 2001, "Testing When a Parameter is on the Boundary of the Maintained Hypothesis," *Econometrica*, vol. 69, pp. 683—734.

Banerjee, Abhijit, 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, vol. 107, pp. 787—818.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, 1992, "A Theory of Fads, Fashion, Custom and Cultural Change As Informational Cascades," *Journal of Political Economy,* vol. 100, pp. 992—1027.

Bikhchandani, Sushil and Sunil Sharma, 2001, "Herd Behavior in Financial Markets: A Review," IMF Working paper, 47, 3.

Chung, Kee, Mingsheng Li, and Thomas H. McInish, 2005, "Information-based Trading, Price Impact of Trades, and Trade Autocorrelation," *Journal of Banking and Finance*, vol. 29, pp. 1645—1669.

Cipriani, Marco and Antonio Guarino, 2005, "Herd Behavior in a Laboratory Financial Market," *American Economic Review*, vol. 95, pp. 1427—1443.

Cipriani, Marco and Antonio Guarino, 2008, "Herd Behavior and Contagion in Financial Markets," *B.E. Journal of Theoretical Economics*, 8, Article 24.

Cipriani, Marco and Antonio Guarino, 2009, "Herd Behavior in Financial Markets: An Experiment with Financial Market Professionals," *Journal of the European Economic Association*, vol. 7, pp. 206—233.

Dasgupta, Amil, Andrea Prat, and Michela Verardo, forthcoming, "Institutional Trade Persistence and Long-Term Equity Returns," *Journal of Finance*.

Drehmann, Mathias, Jörg Oechssler, and Andreas Rider, 2005, "Herding and Contrarian Behavior in Financial Markets - An Internet Experiment," *American Economic Review*, vol. 95, pp. 1403—1426.

Easley, David and Maureen O'Hara, 1987, "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, vol. 19, pp. 69—90.

Easley, David and others, 1996, "Liquidity, Information and Infrequently Traded Stocks," *Journal of Finance*, vol. 10, pp. 805—835.

Easley, David, Nicholas Kiefer, and Maureen O'Hara, 1997, "One Day in the Life of a Very Common Stock," *Review of Financial Studies*, vol. 10, pp. 805—835.

Easley, David, Nicholas Kiefer, and Maureen O'Hara, 1997a, "The Information Content of the Trading Process," *Journal of Empirical Finance*, vol. 4, pp. 159—186.

Easley, David and others, 2008, "Time-Varying Arrival Rates of Informed and Uninformed Trades," *Journal of Financial Econometrics*, vol. 6, pp. 171—207.

Gale, Douglas, 1996, "What Have We Learned from Social Learning," *European Economic Review*, vol. 40, pp. 617—628.

Gervais, Simon, 1997, "Market Microstructure with Uncertain Information Precision: A Multi-Period Analysis," mimeo, Wharton School.

Glosten, Lawrence and Paul Milgrom, 1985, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, vol. 14, pp. 71—100.

Grinblatt Mark, Sheridan Titman, and Russ Wermers, 1995, "Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior," *American Economic Review*, vol. 85, pp. 1088—1105.

Hasbrouck, Joel, 1991, "The Summary Informativeness of Stock Trades: An Econometric Analysis," *Review of Financial Studies*, vol. 4, pp. 571—595.

Hasbrouck, Joel, 1991a, "Measuring the Information Content of Stock Trades," *Journal of Finance*, vol. 46, pp. 179—207.

Hasbrouck, Joel, 2004, "Working with the WRDS (TAQ and CRSP) Data," mimeo, NYU Stern School of Business.

Hirshleifer, David and Siew Teoh, 2009, "Thought and Behavior Contagion in Capital Markets," in Reiner Schenk-Hoppé K., and Hens T., ed. *Handbook of Financial Markets: Dynamics and Evolution* Elsevier, 1—56.

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

Lee, Charles and Marc Ready, 1991, "Inferring Trade Direction from Intraday Data," *Journal of Finance*, vol. 46, pp. 733—746.

Milgrom, Paul and Nancy Stokey, 1982, "Information, Trade and Common Knowledge," *Journal of Economic Theory*, vol. 26, pp. 17—27.

Smith, Lones and Peter Sørensen, 2000, "Pathological Outcomes of Observational Learning," *Econometrica*, vol. 68, pp. 371—398.

Vives, Xavier, 2008, *Information and Learning in Markets: The Impact of Market Microstructure,* Princeton University Press, Princeton.

Welch, Ivo, 1992, "Sequential Sales, Learning, and Cascades," *Journal of Finance*, vol. 47, pp. 695—732.

APPENDIX

1. Proof of Proposition 1

Consider the belief $E(V_d|h_2^d = \{buy_1\})$ of a trader who observes a buy order at time 1. Let $E(V_d|h_2^d = \{buy_1\}) - v_{d-1} = \kappa$, where $\kappa > 0$. Consider now the belief $E(V_d|h_2^d = \{buy_1\}, s_2^d = 0.5 - \vartheta)$ of a trader with a bad signal $s_2^d = 0.5 - \vartheta$ (with $\vartheta > 0$) after he observes a buy order at time 1. For any κ there is a ϑ such that $E(V_d|h_2^d = \{buy_1\}, s_2^d = 0.5 - \vartheta) - v_{d-1} > 0$. Let us denote this difference by ϕ.

Consider now a history of trades in which in each odd period there is a buy order and in each even period there is a sell order. Given the presence of noise traders, such a history has positive probability. For any ϕ, there exists an (odd) t such that at time $t+1$, in equilibrium, $a_{t+1}^d - v_{d-1} < \phi$ (since the market maker after such a history attaches a lower and lower probability to the day being informed). Finally, note that at $t+1$ the belief of an informed trader is exactly as at time 2, since he knows that an event has occurred and the equal number $\frac{t-1}{2}$ of buy and sell orders before time t offset each other in the updating of the belief about the event being good or bad. Therefore, a trader with signal $s_{t+1}^d \in [0.5 - \vartheta, 0.5)$ has an expectation $E(V_d|h_{t+1}^d, s_{t+1}^d) > v_{d-1} + \phi > a_{t+1}^d$ and herds. Since this history occurs with positive probability when $V_d = v_d^L$, this shows that incorrect herd-buy occurs with positive probability. The proof for herd-sell is analogous.

2. Further computations for the likelihood function

The probabilities in the equation in the text can easily be expressed as a function of β_t^d and of the parameters. For instance, the probability that an informed trader with signal β_t^d attaches to the good informational event can be computed as

$$\Pr(v_d^H|h_t^d, \beta_t^d) = \frac{2\beta_t^d \Pr(v_d^H|h_t^d, V_d \neq v_{d-1})}{2\beta_t^d \Pr(v_d^H|h_t^d, V_d \neq v_{d-1}) + 2\left(1-2\beta_t^d\right)(1-\Pr(v_d^H|h_t^d, V_d \neq v_{d-1}))}.$$

(Note that $\Pr(v_d^H|h_t^d, V_d \neq v_{d-1})$ is conditioned on $V_d \neq v_{d-1}$ since the informed trader knows that an event has occurred.) The probability that the market maker attaches to the good event can, instead, be computed as

$$\Pr(v_d^H|h_t^d, buy_t^d) = \frac{\left(\left(1-\beta_t^{d^2}\right)\mu + \left(\frac{\varepsilon}{2}\right)(1-\mu)\right) \Pr(v_d^H|h_t^d, V_d \neq v_{d-1}) \Pr(V_d \neq v_{d-1}|h_t^d)}{D},$$

where

$$D = \left(\left(1-\beta_t^{d^2}\right)\mu + \left(\frac{\varepsilon}{2}\right)(1-\mu)\right)\Pr(v_d^H|h_t^d, V_d \neq v_{d-1})\Pr(V_d \neq v_{d-1}|h_t^d)+$$
$$\left(\left(\beta_t^{d^2} - 2\beta_t^d + 1\right)\mu + \left(\frac{\varepsilon}{2}\right)(1-\mu)\right)(1-\Pr(v_d^H|h_t^d, V_d \neq v_{d-1}))\Pr(V_d \neq v_{d-1}|h_t^d)+$$
$$\left(\frac{\varepsilon}{2}\right)\Pr(V_d = v_{d-1}|h_t^d).$$

3. Robustness checks for Table 6 for 0.75, 0.8 and 0.85

	Average	S.D.	Max
Herd-buyers	3%	4%	20%
Herd-sellers	3%	5%	29%

	Average	S.D.	Max
Herd-buyers	3%	4%	20%
Herd-sellers	3%	5%	28%

	Average	S.D.	Max
Herd-buyers	3%	4%	20%
Herd-sellers	3%	5%	29%