

Estimating a Structural Model of Herd Behavior in Financial Markets

Online Appendix

Marco Cipriani and Antonio Guarino

Abstract

In this addendum we provide further results for the paper “Estimating a Structural Model of Herd Behavior in Financial Markets.” In Section *A*, we prove 1) Proposition 1, and 2) that the second condition in Definition 1 is redundant (or, equivalently, that contrarianism cannot occur in our economy). In Section *B*, we provide robustness checks for Table 6. In Section *C*, we provide parameter and herding estimates for several stocks. In Section *D* we show that the martingale assumption for Ashland Inc. is supported by the data. In Section *E* we provide robustness checks for our estimates, by allowing for time dependence in the behavior of noise traders. In Section *F* we show, through a regression analysis, that the probability of a buy or sell order is affected by the trade imbalance in event days, but not in no-event days. In Section *G* we illustrate the behavior of the thresholds and of other variables (e.g., the price) for each of the 252 days of trading in our sample; we also characterize the asymptotic behavior of the thresholds. In Section *H* we compare our herding results with the traditional measure of herd behavior by Lakonishok, Shleifer and Vishny (1992).

A Proofs

A.1 Proof of Proposition 1

Consider the belief $E(V_d|h_2^d = \{buy_1^d\})$ of a trader who observes a buy order at time 1. Let $E(V_d|h_2^d = \{buy_1^d\}) - v_{d-1} = \kappa$, where $\kappa > 0$. Consider now the belief $E(V_d|h_2^d = \{buy_1^d\}, s_2^d = 0.5 - \vartheta)$ 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 $E(V_d|h_2^d = \{buy_1^d\}, s_2^d = 0.5 - \vartheta) - v_{d-1} > 0$. 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 being in an event day). 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 misdirected herd-buy occurs with positive probability. The proof for herd-sell is analogous.

A.2 Proof that the Second Condition in Definition 1 is Redundant (No Contrarianism)

In this subsection, we show that if at time t an informed trader buys upon receiving a bad signal, i.e.,

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

then, the price of the asset must be higher than at time 1, that is,

$$E(V_d|h_t^d) \equiv p_t^d > p_1^d \equiv v_{d-1}.$$

Similarly, if an informed trader sells upon receiving a good signal, then the price of the asset must be lower than at time 1. That is, the second condition in Definition 1 is redundant.

As we explained in the paper, the proof is tantamount to showing that contrarianism never arises in equilibrium.¹ Therefore, we first define contrarianism formally and then we prove that it cannot occur in equilibrium.

Definition 3 An informed trader engages in contrarian-buying at time t of day d if he buys upon receiving a bad signal, that is,

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

and the price has decreased since time 1, that is,

$$p_t^d < p_1^d = v_{d-1}.$$

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

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

and the price has increased since time 1, that is,

$$p_t^d > p_1^d = v_{d-1}.$$

Before providing the formal proof, let us discuss intuitively why, in equilibrium, herding can occur but contrarianism cannot. Let us consider a day in which buy orders arrive in the market and, as a result, the market maker updates the prices upward. As we explained in the paper, herd behavior occurs because the market maker updates the price “too slowly.” Therefore, after a series of buy orders, some traders may find it profitable to buy despite a bad signal because the price has not increased enough; that is, herding occurs. The same logic explains why contrarianism never arises: when buy orders arrive in the market, the market maker updates the price too slowly, and a trader with a good signal always finds it profitable to buy. This occurs both because he has good information on the asset, and because he knows that

¹In theory, a trader might buy with a bad signal for a price *equal* to the asset expected value. Therefore, strictly speaking, proving that contrarianism cannot occur does not make the second condition in the herding definition redundant. As we make clear in the following footnote, however, the same logic of the proof can be used to show that this never occurs. That is, in order for a trader to buy (sell) against his signal, the price of the asset must have *strictly* increased (decreased).

the price increase until time t does not fully reflect the information content of the order flow.

The logic of the proof presented below is simple. First, observe that, conditional on knowing that an event has occurred, the only difference between the expectations of the market maker and of an informed trader is the signal. This means that the expectation of an informed trader with a good signal is always higher than the market maker's expectation conditional on an information event having occurred. Second, consider the fact that for contrarian selling to occur the price must have gone up with respect to the beginning of the day, that is, the order flow must convey good information to the market. The price (i.e., the market maker's expectation), however, always increases by a smaller amount than the market maker's expectation *conditional* on an event having occurred, since the market maker attaches a positive probability to the fact that no event occurred. By the first observation, an informed trader with a good signal has an expectation higher than that of the market maker conditional on an event having occurred. *A fortiori*, his expectation is higher than the price the market maker sets conditional only on the positive information coming from the order flow, without knowing whether the event has occurred.

We now turn to the formal analysis.

Proposition 2 In equilibrium an informed trader never engages in contrarian buying or in contrarian selling.

We prove the proposition for the case of contrarian selling. The proof for contrarian buying is analogous. We start by proving a simple claim: the expected value of the asset for the market maker at time t of a day is higher than at time 1 (i.e., at the beginning of the day) when, conditional on knowing that an event has occurred, the probability that he attaches to the good event is higher than at time 1.

Claim 1: The condition $p_t^d = E(V_d|h_t^d) > p_1^d = v_{d-1}$ is satisfied if and only if $\Pr(V_d = v_d^H|h_t^d, I^E = 1) > \delta$, where I^E is an indicator function taking value 1 if there has been an informational event and 0 if not.

Proof of Claim 1: Note that

$$\begin{aligned} E(V_d|h_t^d) &= v_{d-1} \Pr(V_d = v_{d-1}|h_t^d) + v_d^H \Pr(V_d = v_d^H|h_t^d) + v_d^L \Pr(V_d = v_d^L|h_t^d) = \\ &= v_{d-1} + \lambda^H \Pr(V_d = v_d^H|h_t^d) + \lambda^L \Pr(V_d = v_d^L|h_t^d) = \\ &= v_{d-1} + (\lambda^H \Pr(V_d = v_d^H|h_t^d, I^E = 1) + \lambda^L \Pr(V_d = v_d^L|h_t^d, I^E = 1)) \Pr(I^E = 1|h_t^d). \end{aligned}$$

Hence,

$$E(V_d|h_t^d) > v_{d-1},$$

if and only if

$$\Pr(V_d = v_d^H|h_t^d, I^E = 1)\lambda^H + \Pr(V_d = v_d^L|h_t^d, I^E = 1)\lambda^L > 0.$$

Recall that $\delta\lambda^H + (1-\delta)\lambda^L = 0$. Therefore, the inequality is satisfied if and only if $\Pr(V_d = v_d^H|h_t^d, I^E = 1) > \delta$. That is, for the price to have increased with respect to the start of the day, it must be the case that the market maker has updated upward the probability of a good event, conditional on the fact that an event has occurred. ■

To complete the proof of Proposition 2, we now show that if $\Pr(V_d = v_d^H|h_t^d, I^E = 1) > \delta$, then $E(V_d|h_t^d, s_t^d) > E(V_d|h_t^d)$, for any $s_t^d > 0.5$. This means that if the price has increased with respect to the beginning of the day, a trader with a good signal always has an evaluation higher than the market maker's; that is, he does not engage in contrarian selling. The inequality $E(V_d|h_t^d, s_t^d) > E(V_d|h_t^d)$ can be rewritten as

$$\begin{aligned} & v_{d-1} + \lambda^H \Pr(V_d = v_d^H|h_t^d, s_t^d) + \lambda^L \Pr(V_d = v_d^L|h_t^d, s_t^d) > \\ & v_{d-1} + \lambda^H \Pr(V_d = v_d^H|h_t^d) + \lambda^L \Pr(V_d = v_d^L|h_t^d), \end{aligned}$$

or as

$$\begin{aligned} & \lambda^H \Pr(V_d = v_d^H|h_t^d, s_t^d) + \lambda^L \Pr(V_d = v_d^L|h_t^d, s_t^d) > \\ & (\lambda^H \Pr(V_d = v_d^H|h_t^d, I^E = 1) + \lambda^L \Pr(V_d = v_d^L|h_t^d, I^E = 1)) \Pr(I^E|h_t^d). \end{aligned}$$

This is equivalent to

$$\begin{aligned} & \Pr(V_d = v_d^H|h_t^d, s_t^d) - \frac{\delta}{(1-\delta)} \Pr(V_d = v_d^L|h_t^d, s_t^d) > \\ & \left(\Pr(V_d = v_d^H|h_t^d, I^E = 1) - \frac{\delta}{(1-\delta)} \Pr(V_d = v_d^L|h_t^d, I^E = 1) \right) \Pr(I^E|h_t^d). \end{aligned}$$

Now note that, since $s_t^d > 0.5$, $\Pr(V_d = v_d^H|h_t^d, s_t^d) > \Pr(V_d = v_d^H|h_t^d, I^E = 1)$. Moreover, as observed in Claim 1, if $E(V_d|h_t^d) > v_{d-1}$, then $\Pr(V_d = v_d^H|h_t^d, I^E = 1) > \delta$. Since both $\Pr(V_d = v_d^H|h_t^d, s_t^d)$ and $\Pr(V_d = v_d^H|h_t^d, I^E = 1)$ are greater than δ , $\Pr(V_d = v_d^L|h_t^d, s_t^d)$ and $\Pr(V_d = v_d^L|h_t^d, I^E = 1)$ are smaller than $(1-\delta)$. Therefore, both the left and the right hand sides of

the inequality are positive. A sufficient condition for the inequality to hold is that

$$\Pr(V_d = v_d^H | h_t^d, s_t^d) - \frac{\delta}{(1-\delta)} \Pr(V_d = v_d^L | h_t^d, s_t^d) > \\ \Pr(V_d = v_d^H | h_t^d, I^E = 1) - \frac{\delta}{(1-\delta)} \Pr(V_d = v_d^L | h_t^d, I^E = 1),$$

which is obviously true for $s_t^d > 0.5$, since $\Pr(V_d = v_d^H | h_t^d, s_t^d) > \Pr(V_d = v_d^L | h_t^d, I^E = 1)$ for $s_t^d > 0.5$.²

As a final remark, note that the last part of the proof does not contradict the possibility that herding may arise. Indeed, even if $s_t^d < 0.5$, it is still possible that $E(V_d | h_t^d, s_t^d) > E(V_d | h_t^d)$ because the right hand side of the second to last inequality in the proof is multiplied by $\Pr(I^E | h_t^d)$. That is, when the probability that the market maker attaches to the event is low enough, the expectation of a trader with a bad signal may be higher than that of a market maker, and the trader may therefore engage in herd-buying behavior.

B Robustness Checks for Table 6 when the Criterion to Classify Days is Set to 0.75, 0.8 and 0.85

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

²Using the same logic, one can show that if $p_t^d = v_{d-1}$, then $\Pr(V_d = v_d^H | h_t^d, I^E = 1) = \delta$ (note that $\Pr(I^E | h_t^d)$ is never equal to zero if $\alpha > 0$). From the inequality in the text, it follows that $E(V_d | h_t^d, s_t^d) > E(V_d | h_t^d)$ for $s_t^d > 0.5$, that is, a trader with a good signal will never sell. Therefore, for a trader to sell with a good signal (herd selling), the asset price *must have decreased* with respect to the asset's unconditional expected value.

	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%

Table B.1: Percentage of Herders.

The tables show the mean, the standard deviation and the maximum of the percentage of herd-buyers and herd-sellers when we classify a day as a good (bad) event day if, at the end of the day, the posterior probability of the day being a good (bad) event day was higher than 0.75, 0.8 and 0.85, respectively.

C Estimates for Other Stocks

In the paper, we have studied herd behavior in Ashland Inc. during 1995. In this section, we briefly summarize similar results for other stocks. Specifically, we consider AMR Corporation (NYSE symbol: AMR), Franklin Resources Inc. (BEN), Duke Energy Corporation (DUK), W.W. Grainger Inc. (GWW), McGraw Hill (MHP), NiSource Inc. (NI), Sherwin-Williams Company (SHW). These companies operate in various industries: transport, investment management and banking, energy, manufacturing, publishing, media and services. All the stocks are traded in the New York Stock Exchange, the American Stock Exchanges, and the consolidated regional exchanges.

Similarly to the analysis in the paper, the data we use refer to the entire year 1995. All the data are obtained from the TAQ (Trades and Quotes) dataset. We used the same algorithms discussed in the paper to classify trades. We obtained the no-trade interval by computing, for each stock, the ratio between the total trading time in a day and the average number of buy and sell orders over the 252 days of trading.

The stocks were traded with different frequencies, going from an average of 77 trades (buy or sell orders) per day for the case of NiSource Inc. to an average of 293 trades for the case of AMR Corporation.

Stock	α	δ	μ	τ	ε
AMR	0.28(0.01)	0.73(0.02)	0.40(0.004)	0.61(0.01)	0.55(0.0004)
BEN	0.24(0.03)	0.73(0.05)	0.39(0.01)	0.64(0.03)	0.56(0.0012)
DUK	0.29(0.02)	0.83(0.04)	0.43(0.01)	0.26(0.02)	0.57(0.0012)
GWW	0.26(0.03)	0.57(0.06)	0.40(0.01)	0.47(0.02)	0.57(0.0015)
MHP	0.25(0.02)	0.73(0.05)	0.44(0.01)	0.49(0.03)	0.56(0.0016)
NI	0.40(0.02)	0.42(0.05)	0.47(0.01)	0.53(0.02)	0.57(0.0021)
SHW	0.43(0.04)	0.91(0.04)	0.36(0.01)	0.52(0.02)	0.55(0.0019)

Table C.1: Estimation Results.

The table shows the estimates for the five parameters of the model for all stocks. The standard errors are in parenthesis.

C.1 Results

Table C.1 reports the estimated parameters for all stocks. There is, of course, no reason why these structural parameters should be similar across stocks. Nevertheless, it is interesting to observe some common characteristics. For all stocks, an information event occurred quite frequently, between 24% and 43% of the trading days. Good events were more likely than bad events for all these stocks, with the exception of NI. The proportion of informed traders is approximately 40% for all stocks. The parameter τ , which measures the informativeness of the signal, is always lower than 1. In other words, for all these stocks, beliefs were “bounded.” More importantly, since τ is clearly less than infinity, the sequence of trades (and not just the number of buy and sell orders during the day) clearly matters. The parameter ε is relatively stable across stocks.

Finally, we report the frequency of herding for the different stocks. Table C.2 summarizes how frequently the buy threshold β_t^d was below 0.5 (herd buy) or the sell threshold σ_t^d was above 0.5 (herd sell).

The results are not very dissimilar across stocks, with both herd buying and herd selling varying from approximately 15% to approximately 27%.

Stock	Herd Buy	Herd Sell
AMR	0.23	0.18
BEN	0.25	0.26
DUK	0.19	0.17
GWW	0.25	0.27
MHP	0.26	0.25
NI	0.18	0.20
SHW	0.26	0.15

Table C.2: Estimation Results.

The table shows the mean percentage of herd-buyers and herd-sellers for all stocks.

D The Martingale Process of Ashland Inc.’s Price

In our theoretical model, we assume that $(1-\delta)\lambda^L = \delta\lambda^H$, which is equivalent to assuming that the closing daily price is a martingale. The assumption is standard in market microstructure models, since stock prices usually behave as martingales, at least at the horizons relevant to the market microstructure literature. In this section, we test the martingale hypothesis for Ashland Inc.’s closing daily prices during 1995. We test the hypothesis with respect to the history of past prices (i.e., only with respect to the information available in our dataset, and not with respect to other macroeconomic or financial variables). As standard in the literature, we estimate the autocorrelations of both changes in the levels and in the logarithms of Ashland Inc.’s closing daily prices and test for zero autocorrelations at all lags with the Ljung-Box Q statistic. As Tables *D.1* (showing the autocorrelogram for changes in levels) and *D.2* (showing the autocorrelogram for changes in logarithms) illustrate, the null of zero autocorrelation cannot be rejected at any conventional significance levels at all lags. Therefore, there is support for the martingale assumption in the data.

Note that, in the theoretical model, the martingale assumption refers to the stochastic process of the asset fundamental value (V_d). In modeling this process, we abstract from public information, whereas the market price is affected by it. However, for the market price to be a martingale when V_d is not, it would have to be the case that the arrival of public information is time-dependent in such a way so as to perfectly offset the time dependence

in the arrival of private information, a far-fetched hypothesis.

Lag	AC	Q-Stat	Prob
1	0.062	0.9793	0.322
2	0.015	1.0372	0.595
3	0.046	1.5711	0.666
4	-0.049	2.1924	0.700
5	0.009	2.2153	0.819
6	0.078	3.8096	0.702
7	0.037	4.1587	0.761
8	-0.064	5.2324	0.732
9	-0.104	8.0948	0.525
10	-0.103	10.894	0.366
11	0.026	11.073	0.437
12	0.051	11.776	0.464
13	-0.035	12.097	0.520
14	0.002	12.099	0.598
15	0.018	12.183	0.665
16	0.054	12.982	0.674
17	0.035	13.314	0.715
18	0.085	15.304	0.641
19	0.045	15.863	0.666
20	-0.046	16.442	0.689
21	0.041	16.913	0.716
22	0.023	17.053	0.760
23	-0.067	18.311	0.740
24	-0.115	21.986	0.580

Table D.1: Autocorrelations of changes of Ashland Inc.’s closing prices. The table shows the autocorrelations of changes in Ashland Inc.’s closing daily prices for the first 24 lags. Column 2 reports the autocorrelations, column 3 the Q statistic, and column 4 the p-values.

Lag	AC	Q-Stat	Prob
1	0.058	0.8630	0.353
2	0.013	0.9033	0.637
3	0.062	1.8856	0.596
4	-0.049	2.4960	0.645
5	0.009	2.5177	0.774
6	0.070	3.7905	0.705
7	0.026	3.9647	0.784
8	-0.054	4.7347	0.786
9	-0.107	7.7167	0.563
10	-0.106	10.671	0.384
11	0.020	10.773	0.462
12	0.045	11.309	0.503
13	-0.024	11.461	0.572
14	-0.002	11.463	0.649
15	0.014	11.512	0.716
16	0.050	12.184	0.731
17	0.038	12.576	0.764
18	0.088	14.695	0.683
19	0.054	15.500	0.690
20	-0.040	15.929	0.721
21	0.040	16.373	0.748
22	0.021	16.499	0.790
23	-0.078	18.213	0.746
24	-0.108	21.451	0.612

Table D.2: Autocorrelations of log-changes of Ashland Inc.’s closing prices. The table shows the autocorrelations of changes in Ashland Inc.’s closing daily prices (in logarithms) for the first 24 lags. Column 2 reports the autocorrelations, column 3 the Q statistic, and column 4 the p-values.

E Time Dependence in Noise Trading Activity and Herd Behavior

In the paper we have shown that rational traders sometimes find it optimal to follow the established pattern of trade. In the analysis, we have assumed that the behavior of noise traders is i.i.d. One may wonder to what extent our results on herd behavior are robust to different specifications on the decisions

of noise traders. In this section we provide such a robustness check.

Clearly, any assumption on the behavior of noise traders is to some extent arbitrary. Nevertheless, one can think of two ways in which liquidity trading could be time dependent: one is that there is heterogeneity in liquidity traders' activity *across days* (e.g., because of an aggregate liquidity shock every day); the other is that there is time-dependence *within a day* (e.g., because a common liquidity shock occurs at some points during the day). Therefore, we proceed by considering two alternative specifications of our model that can be estimated relatively easily. In the first specification, we assume that every day there is an aggregate liquidity shock which is *known* to market participants. In the second, we assume that the decision of a noise trader at time t depends on the decision at time $t - 1$, independently of the type of trader trading in $t - 1$. Both models use shortcuts to model time dependence in noise trading: respectively the fact that the aggregate shock is known to market participant, and that the behavior of a noise trader at time t depends on the trade at time $t - 1$ (as opposed to the behavior of the previous noise trader's trading in the market). These shortcuts allow us to write the likelihood function and estimate the model (indeed, the second specification is already used in the literature, see, e.g., Easley, Kiefer and O'Hara, 1997a and Easley et al., 2008). We believe that both specifications serve well our purpose of checking the robustness of our results. Since, as we shall see, in neither case herd behavior decreases (compared to our main model), this analysis provides additional support for our findings.

E.1 Model Specification 1

In our first specification, we allow for the liquidity or hedging motive to trade to be different day by day. In some days there is a "common aggregate liquidity shock" that leads most noise traders to sell; in other days there is a "common aggregate liquidity shock" that leads most noise traders to buy. In other words, in the first type of day, a noise trader buys with a higher probability than he sells; in the second, the opposite is true. If market participants did not know in which of the two types of days they are, they would have to learn it through the observation of the order flow during the day. This would lead to a complicated inference problem since the market participants (in particular the market maker) would have to learn not only about the occurrence of a (good or bad) event, but also about the type of liquidity shock. This model would be very different from our current one,

and would generate a very complex and hard to estimate likelihood function. As mentioned above, we assume, instead, that market participants do know the nature of the liquidity shock each day. As a result, this new model can be estimated with relative ease.

Let us describe the model formally. The only change to the model presented in the paper concerns the behavior of noise traders. We assume that there are two types of days, occurring with probability θ and $(1 - \theta)$. In the first type of day, if a noise trader trades, he buys with probability $\eta \geq 0.5$; in the second type of day, if a noise trader trades, he sells with probability η . That is, in the first type of day, the probability of a noise trader buying, selling or not trading is equal to $\varepsilon\eta$, $\varepsilon(1 - \eta)$ and $1 - \varepsilon$, respectively. Similarly, in the second type of day, the probabilities are $\varepsilon(1 - \eta)$, $\varepsilon\eta$ and $1 - \varepsilon$.³ We will refer to the first type of day as a “high noise-buying” day and to the second as a “high noise-selling” day. All agents in the market know whether they are in a day of the first or second type. Obviously, whether a day is of the first or of the second type is independent of the information event.

Parameter	Estimate	S.D.
α	0.34	0.0203
δ	0.78	0.0401
μ	0.38	0.0146
τ	0.47	0.0261
ε	0.56	0.0016
η	0.53	0.0049
θ	0.001	0.0756

Table E.1: Estimation Results.

The table shows the estimates for the seven parameters of model specification 1 and their standard deviations.

Table *E.1* reports the parameter estimates. The first thing to notice is that essentially all days are classified as high noise-selling days (the parameter θ is close to 0). Second, the difference between the probabilities of buying and selling is small ($\eta = 0.53$), indicating that noise trading activity is balanced every day. It is also interesting to observe that the other parameter estimates are not far from those of our main model. In particular, the parameter τ (crucial for estimating herding and estimated as 0.45 in the main model) is

³The assumption $\eta \geq 0.5$ is obviously necessary for the model to be identified.

equal to 0.47, indicating, as in the main model, bounded beliefs.

Given that agents in the market are assumed to know in which day (high noise-selling or high noise-buying) they are whereas the econometrician does not, to compute the measures of herding we should first classify each day of trading as a high noise-buying or a high noise-selling day. In practice, since in our case all days are high noise-selling, the analysis is straightforward. Table E.2 contrasts the frequency of herding in this model specification with that in the main model.

	Main Model	Model Specification 1
Herd Buy	0.30	0.30
Herd Sell	0.37	0.34

Table E.2: Herd Behavior.

The table shows the proportion of herd buy and herd sell in the main model and in model specification 1.

The results are virtually the same, indicating that our findings on the occurrence of herding by informed traders do not change after accounting for the possibility of daily aggregate liquidity shocks.

E.2 Model Specification 2

In this second specification, we use a different strategy to allow for time dependence in noise trading. This strategy has already been used in the market microstructure literature (see, e.g., Easley, Kiefer and O'Hara, 1997a; and Easley et al., 2008). The probability of a decision by a noise trader at time t depends on whether the previous transaction is a buy or a sell order, independently of whether that transaction comes from a noise or from an informed trader. That is, a noise trader at time t buys with a different probability after a purchase, a sale or a no trade at time $t - 1$. This modeling strategy, although used in the literature, is a shortcut, since there is no clear reason why a liquidity trader should change his behavior on the basis of the previous trade (the predecessor could in fact be an informed trader). The shortcut, however, keeps the likelihood function simple, as the time dependence is a function of an observed variable (the order in previous period). The likelihood function can therefore be constructed in a similar way to our main model.

Formally, the only change to the model presented in the paper is the following. We assume that if at time $t - 1$ there was a buy order, then a noise trader acting at time t buys with probability $\varepsilon\gamma$, sells with probability $\varepsilon(1 - \gamma)$, and does not trade with probability $1 - \varepsilon$. Similarly, if at time $t - 1$ there was a sell order, then a noise trader acting at time t buys with probability $\varepsilon(1 - \gamma)$, sells with probability $\varepsilon\gamma$, and does not trade with probability $1 - \varepsilon$. Finally, if at time $t - 1$ there was a no trade, then a noise trader acting at time t buys with probability $\frac{\varepsilon}{2}$, sells with probability $\frac{\varepsilon}{2}$, and does not trade with probability $1 - \varepsilon$.⁴

Parameter	Estimate	S.D.
α	0.21	0.0321
δ	0.72	0.0852
μ	0.32	0.0252
τ	0.39	0.0636
ε	0.58	0.0015
γ	0.69	0.0034

Table E.3: Estimation Results.

The table shows the estimates for the six parameters of model specification 2 and their standard deviations.

Table *E.3* reports the parameter estimates. The estimate for γ indicates that there is some positive time dependence in the behavior of noise traders: a buy (sell) order is more likely if there was a buy (sell) order at time $t - 1$ (there would be time independence for $\gamma = 0.5$). Given these parameters, we computed the buy and sell thresholds defining informed traders' strategies. We then estimated their correlation with the thresholds of our benchmark model: for both the buy and sell threshold the average daily correlation is above 0.9, thus suggesting that our characterization of the behavior of informed traders is robust to our assumption on the pattern of noise traders' activity.

Table *E.4* reports the frequency of herding in this model specification. Herd buy and herd sell do not decrease; they slightly increase compared to our original estimates. At a first glance this result may look surprising. One could expect the positive time dependence in the noise trading activity to capture some of the time dependence in the data that we attribute to herd

⁴These are also the probabilities of noise buying, selling and no trading at time 1.

	Main Model	Model Specification 2
Herd Buy	0.30	0.36
Herd Sell	0.37	0.44

Table E.4: Herd Behavior.

The table shows the proportion of herd buy and herd sell in the main model and in model specification 2.

behavior by informed traders. The rationale behind the result is, however, simple: if after a buy order a noise trader buys with higher probability (which is the case in this model), then after the buy order the market maker increases the price even less than if the decisions of noise traders were i.i.d. (because he attaches higher probability to the event that the new buy order comes from another noise trader). As a result, the incentive for informed traders to engage in herd buying is even higher. In other words, the market maker is slower in learning that an information event has occurred and in adjusting the price accordingly so as to prevent herding.

Overall, the results of these two model specifications show that our findings on herding by rational traders are not fragile to our assumption on the behavior of noise traders. Even allowing noise trading to be time dependent does not eliminate (indeed barely affects) herding.

F Trade Imbalance and the Probability of Buy and Sell Orders

In this section we study what pattern in the data produces the statistical rejection of a model with signals that are always correct, such as Easley, Kiefer and O'Hara (1997), in favor of a model in which the signal is noisy, such as ours.

Let us consider a good-event day. According to Easley, Kiefer and O'Hara (1997), at the starting of the trading activity, there will be a higher probability of buy orders and a lower probability of sell orders compared to a no-event day. The proportion of buy and sell orders will, however, remain constant throughout the day, independently of the order flow (and, therefore, of the trade imbalance). Similarly, in a bad-event day, the probability of a sell (buy) order is higher (lower) than in a no-event day and remains constant throughout the day.

To test this prediction of the model, we proceeded in the following way. We classified each day using the Easley, Kiefer and O'Hara (1997) model; in particular, consistently with the classification in the paper, we classified a day as an informed day or a no-event day if $\Pr(V_d \in \{v_d^H, v_d^L\} | h_{T_d}) > 0.9$ (informed day) or $\Pr(V_d = v_{d-1} | h_{T_d}) > 0.9$ (no-event day). That is, we classified a day as an event or a no-event event day if at the end of the day the posterior probability of the day being an event or a no event day was higher than 0.9.

For both types of days separately, we then ran a logit regression of the sign of the trade at each trading time⁵ on a second order polynomial of the contemporaneous level of the trade imbalance.⁶ The results are reported in Tables *F1* and *F2*. According to the Easley, Kiefer and O'Hara (1997) model, the trade imbalance should be statistically non significant in event and in no-event days alike. We find that the trade imbalance is, indeed, non significant in no event days (which is consistent with both our model and Easley, Kiefer and O'Hara (1997) since in both models all trading activity stems from noise traders). However, in event days, the trade imbalance significantly affects the

⁵We only considered buy and sell orders (+1 and -1), disregarding trading times when there was a no trade.

⁶In the regression for event days, we have two dummy variables, one for good and the other for bad event days, to take into account the fact that the unconditional probabilities of a buy and of a sell order are different in the two types of days.

probability of a buy or sell order, which contradicts the prediction of Easley, Kiefer and O'Hara (1997). This empirical finding is, instead, consistent with the idea that since information is not always correct (as in our model), the order flow matters, and affects the probability of buy and sell orders (because traders may engage in herd behavior). Note also that the coefficient on the level of the trade imbalance is positive, which is consistent with the idea that in event days (both bad and good), buy (sell) orders are more likely to be followed by other buy (sell) orders.

Parameter	Estimate	Std. Err.	P-value
Good-day dummy	0.3998	0.0454	0.00
Bad-day dummy	-0.5928	0.0686	0.00
Trade Imbalance	0.0040	0.002	0.05
Trade Imbalance Square	0.0002	0.00005	0.00

Table F.1: Estimation Results for Event Days.

The table shows the estimates of a logit regression for the direction of trade in event days as a function of the trade imbalance and the square of the trade imbalance

Parameter	Estimate	Std. Err.	P-value
Constant	0.0993	0.0239	0.00
Trade Imbalance	-0.0012	0.0030	0.70
Trade Imbalance Squared	-0.00005	0.00025	0.85

Table F.2: Estimation Results for No-Event Days.

The table shows the estimates of a logit regression for the direction of trade in no-event days as a function of the trade imbalance and the square of the trade imbalance

G The Thresholds

In this section we report the evolution of the asset price (p_t^d) and of the market maker's belief on the occurrence of an event (Figures *G4 – G13*), and the evolution of the thresholds characterizing the informed traders' equilibrium strategy (Figures *G14–G23*) for each of the 252 days of trading in our sample. In each figure we also report the trade imbalance (dashed line, measured on the left axis).

In the following paragraph we discuss the thresholds' asymptotic behavior. We do so for no information event days and information event days separately.

- No informational event days

In these days the market maker gradually learns that there was no information event (i.e., the probability of an event conditional on the history of trades goes to zero). As a result, the price converges to the asset's unconditional expected value v_{d-1} and the bid-ask spread converges to zero.⁷ The thresholds in this case will converge to each other. The reason is simple. While the price has converged to the unconditional asset value and the spread to zero, an informed trader would know that the true asset value has in fact changed. Therefore, he would never choose not to trade whatever his signal, which means that the buy and sell thresholds have to be the same. These thresholds will fluctuate together between 0 and 1 as a result of new trades arriving in the market. Since the price is essentially fixed at the unconditional expected value, a series of, e.g., buy orders pushes (almost) all informed traders' expectations above the (ask) price, which means that informed traders would herd buy. When sell orders arrive in the market, they break the herd buy and create a herd sell.⁸ This pattern is in fact what we observe in the standard herding model with “fixed price” (Bikhchandani, Hirshleifer and Welch, 1992): herding arises very easily as the number of buy orders exceeds the number of sell orders, or vice versa, and is fragile. As an illustration, Figure *G1* reports the simulation of the thresholds for a day with a balanced trading activity,⁹ which is what we would expect in a no-event day. Towards the end of the day (after period 160) a series of 20 buy orders and then 20 sell orders arrive in the market. The thresholds (the second panel in the figure) converge to each other, in fact they cannot be distinguished. They, however, swing with the arrival of buy and sell orders at the end of the day. After

⁷Almost-sure convergence of the price to the asset's fundamental value and convergence (in probability) of the bid-ask spread to zero are standard results in all Glosten and Milgrom types of model. For this reason we do not provide a proof.

⁸During a herd buy period, a series of sell orders is very informative for informed traders (because it is likely to come from informed traders with a very negative signal).

⁹There are 2 buy orders, 2 sell orders and 6 no trades in each block of 10 trading times. The simulation is carried out using as parameter values the estimates of our model ($\alpha = 0.28$, $\delta = 0.62$, $\mu = 0.42$, $\tau = 0.45$, $\varepsilon = 0.57$).

time 160 they converge to 0 (almost all informed traders herd buy), and, eventually, after sell orders arrive in the market, they converge to 1 (almost all informed traders herd sell).

- Informational event days

In these days, the market maker's belief on an event having occurred goes to one. The price converges to the true new asset value. The bid-ask spread converges to zero. In contrast with the case of a no-event day, the thresholds do not converge to each other (for illustration, see Figures *G2* and *G3*, which refer to a good and to a bad event day respectively). Instead, the buy threshold eventually settles above 0.5, and the sell threshold settles below 0.5. In other words, eventually there will be no herding. Why is this pattern different from that of a no-event day? Why are the thresholds not converging to each other while the bid-ask spread converges to zero? And why is there no herding? Let us provide some intuition. In an event day, the market maker and the informed traders both learn, eventually, the true asset value. The asymmetric information which is at the root of herding in our model vanishes. Since the market maker attaches a probability very close to zero to being in a no-event day, we are basically back to a standard Glosten and Milgrom model (or Avery and Zemsky's *IS1* information setup) in which there is no herding.¹⁰ Furthermore, although the spread is converging to zero, also the information content of the private signal is becoming small, which explains why for some signal realizations informed traders choose not to trade, that is, the buy and sell thresholds do not coincide.

Finally, observe that when the price converges to the high value (in a good information event day), the information contained in a sell order affects the market maker's belief more than that contained in a buy order (because of the mechanism of Bayesian updating, private information that goes against the information accumulated in the market is more valuable). For this reason, the market maker will want to protect himself more against a sell order than against a buy order. That is, the sell threshold will be further below 0.5 than the buy threshold is above 0.5 (i.e., they will settle asymmetrically around 0.5, as showed in Figures *G2* and *G3*).

¹⁰The intuition is as follows. By contradiction, suppose there exists a time t' after which the buy threshold is strictly lower than 0.5, say it is $\beta_t^d = 0.5 - \rho$ with $\rho > 0$. We know that the price of the asset converges (almost surely) to the true asset value. So, in an event day, the probability that an event has not occurred tends to 0. Hence there exists a time t'' such that for any $t > t'' > t'$, $\frac{\Pr(V_d = v_d^H | h_t^d, s_t^d = 0.5)}{\Pr(V_d = v_d^L | h_t^d, s_t^d = 0.5)} - \frac{\Pr(V_d = v_d^H | h_t^d)}{\Pr(V_d = v_d^L | h_t^d)} < a$, for any $a > 0$. Therefore, for a sufficiently small, $\frac{\Pr(V_d = v_d^H | h_t^d, s_t^d = 0.5 - \rho)}{\Pr(V_d = v_d^L | h_t^d, s_t^d = 0.5 - \rho)} = \frac{g^H(0.5 - \rho | v_d^H)}{g^L(0.5 - \rho | v_d^L)} \frac{\Pr(V_d = v_d^H | h_t^d, s_t^d = 0.5)}{\Pr(V_d = v_d^L | h_t^d, s_t^d = 0.5)} < \frac{\Pr(V_d = v_d^H | h_t^d)}{\Pr(V_d = v_d^L | h_t^d)}$, which contradicts that for $s_t^d = 0.5 - \rho$, $E(V_d = v_d^L | h_t^d, s_t^d = 0.5 - \rho) = a_t^d \geq p_t^d$.

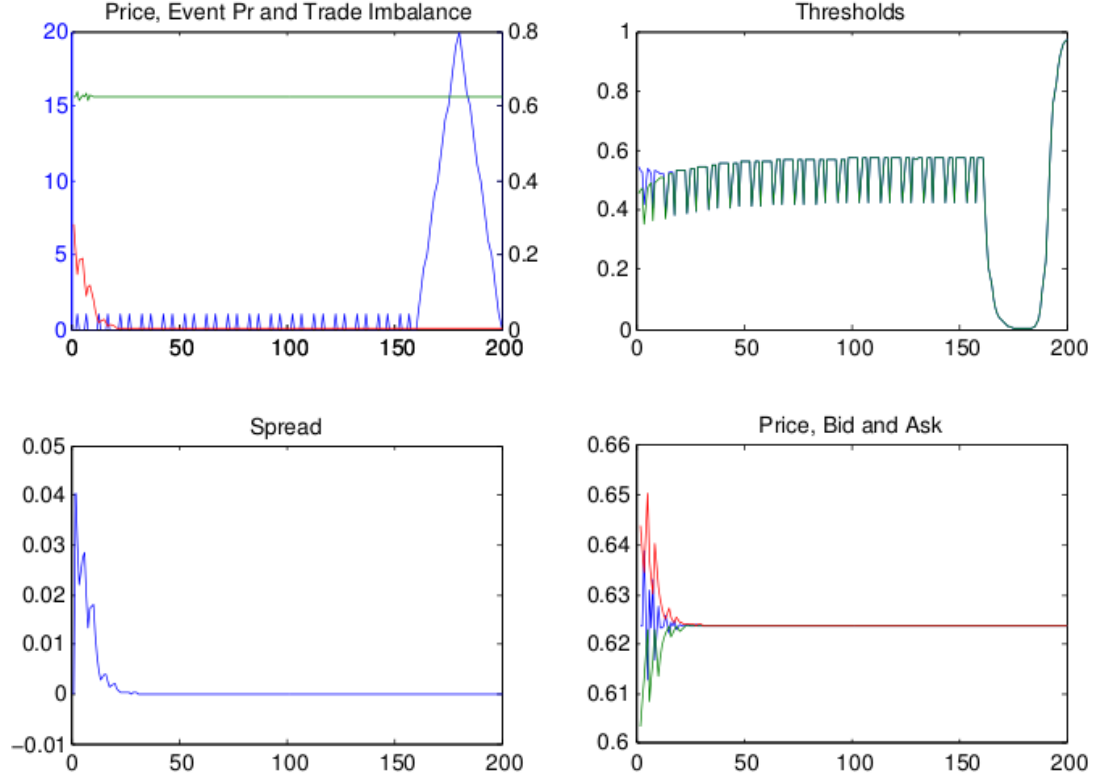


Figure G1 : The figure reports results of a simulated day of trading in which, for the first 160 periods, there are 2 buy orders, 2 sell orders and 6 no trades in each block of 10 trading times; after period 160, there are 20 buy orders followed by 20 sell orders. The simulation is carried out using as parameter values the estimates of our model ($\alpha = 0.28$, $\delta = 0.62$, $\mu = 0.42$, $\tau = 0.45$, $\varepsilon = 0.57$). The first panel shows the trade imbalance (blue line), the price (green line) and the market maker's belief on the occurrence of an event (red line). The second panel reports the buy threshold (blue line) and the sell threshold (green line). The third panel reports the bid-ask spread. The fourth line reports the price, the bid and the ask.

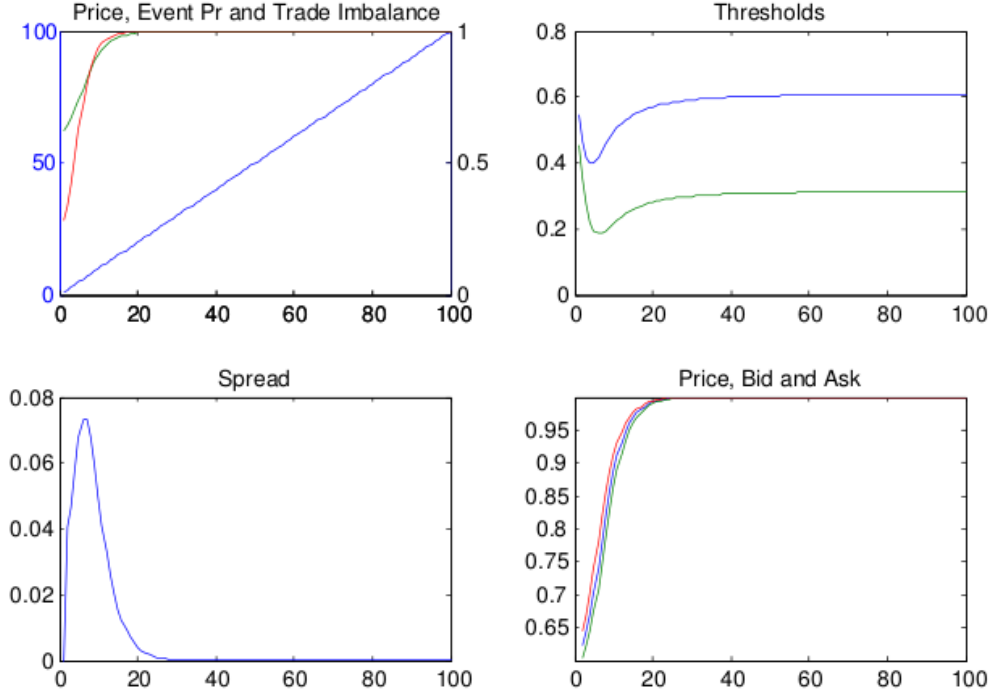


Figure G2 : The figure reports results of a simulated day of trading in which all trades are buy orders. The simulation is carried out using as parameter values the estimates of our model ($\alpha = 0.28$, $\delta = 0.62$, $\mu = 0.42$, $\tau = 0.45$, $\varepsilon = 0.57$). The first panel shows the trade imbalance (blue line), the price (green line) and the market maker's belief on the occurrence of an event (red line). The second panel reports the buy threshold (blue line) and the sell threshold (green line). The third panel reports the bid-ask spread. The fourth line reports the price, the bid and the ask.

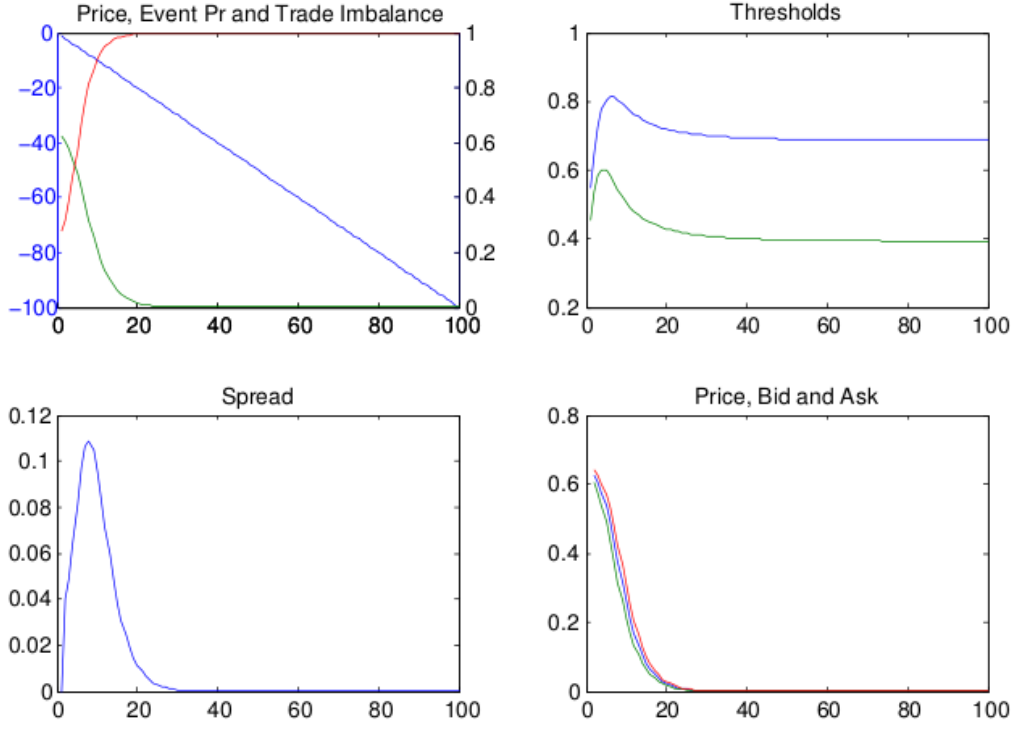


Figure G3 : The figure reports results of a simulated day of trading in which all trades are sell orders. The simulation is carried out using as parameter values the estimates of our model ($\alpha = 0.28$, $\delta = 0.62$, $\mu = 0.42$, $\tau = 0.45$, $\varepsilon = 0.57$). The first panel shows the trade imbalance (blue line), the price (green line) and the market maker's belief on the occurrence of an event (red line). The second panel reports the buy threshold (blue line) and the sell threshold (green line). The third panel reports the bid-ask spread. The fourth line reports the price, the bid and the ask.

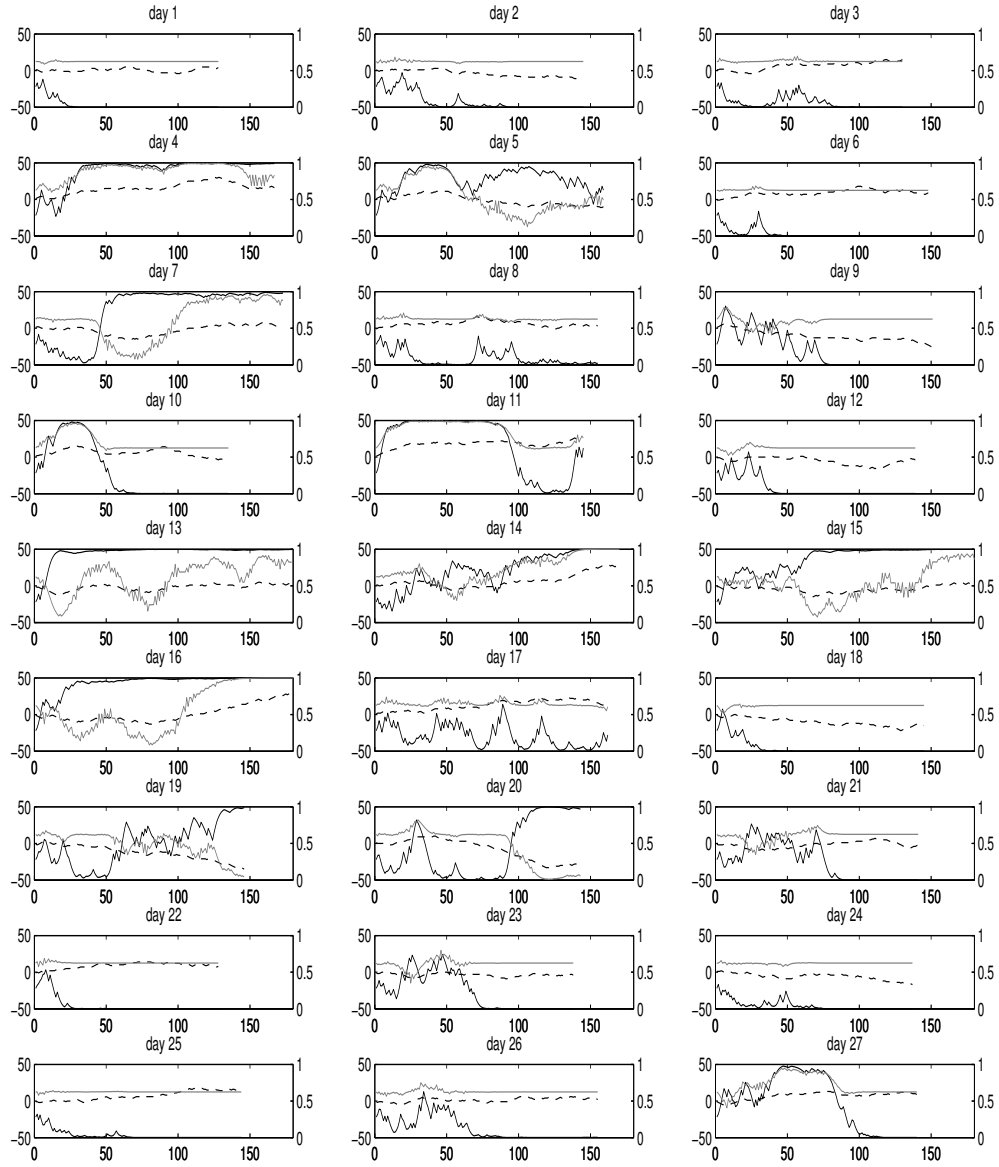


Figure G4 :The figure reports in each panel the daily evolution of: trade imbalance (dashed line), price (grey line), and market maker's belief on the probability of an event (black line).

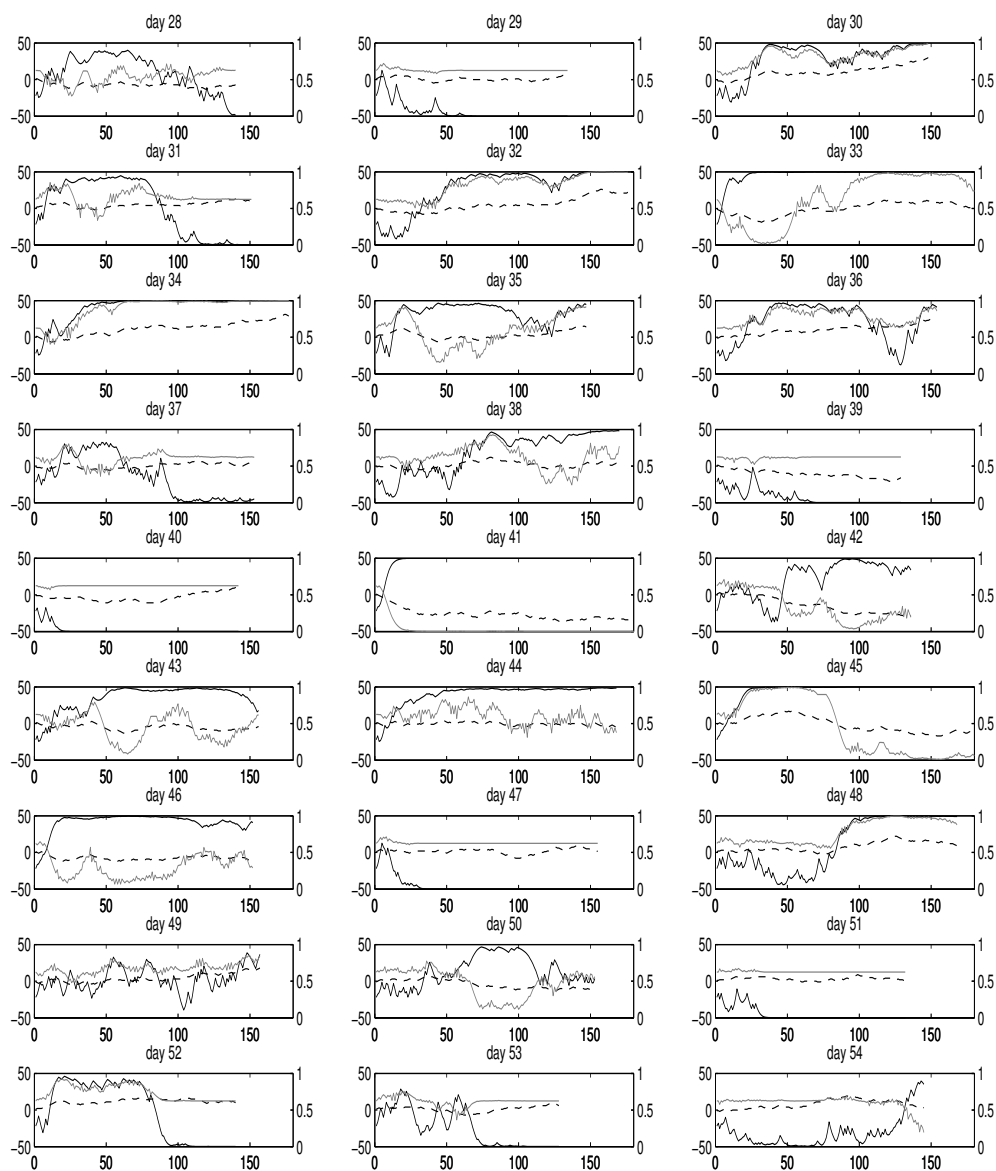


Figure G5

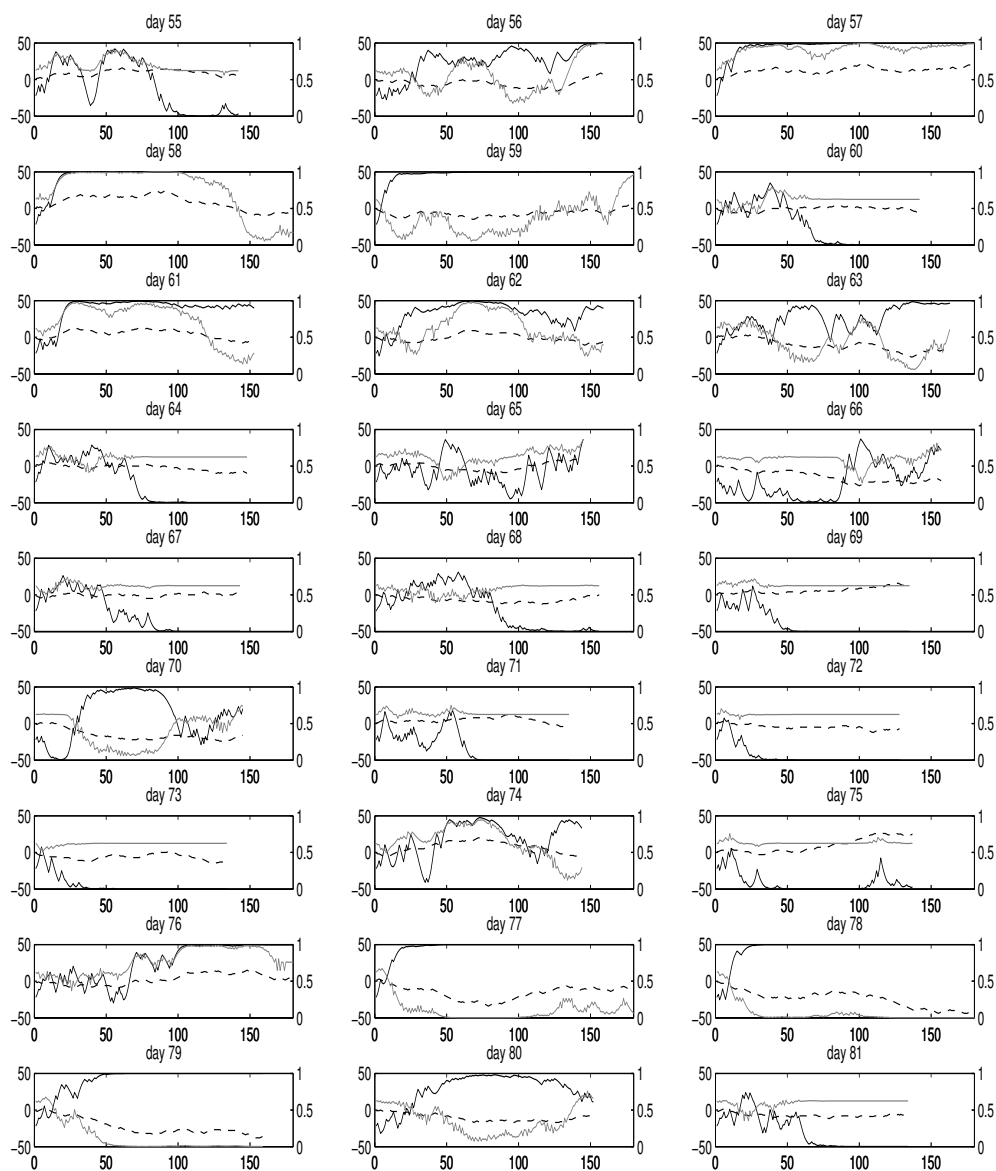


Figure G6

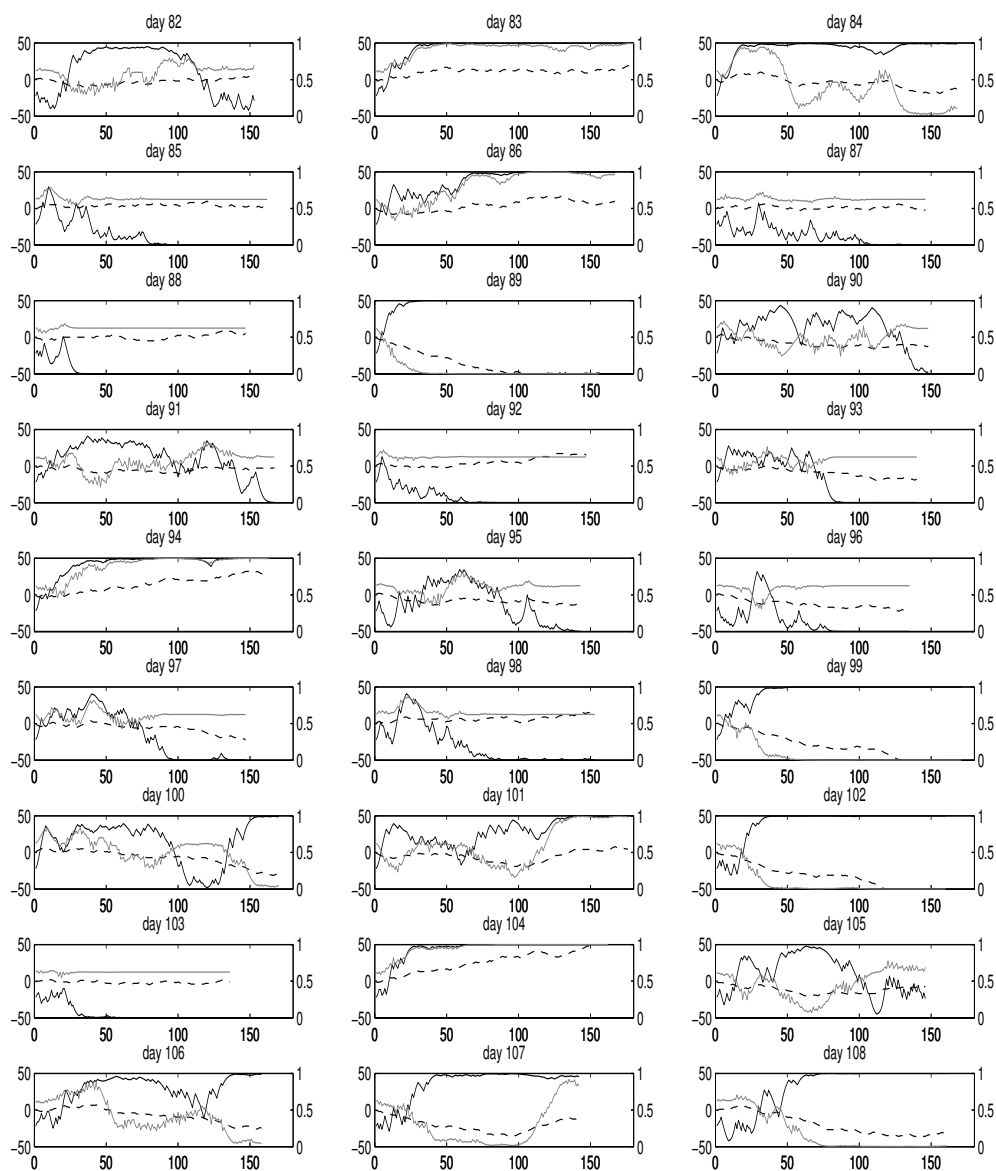


Figure G7

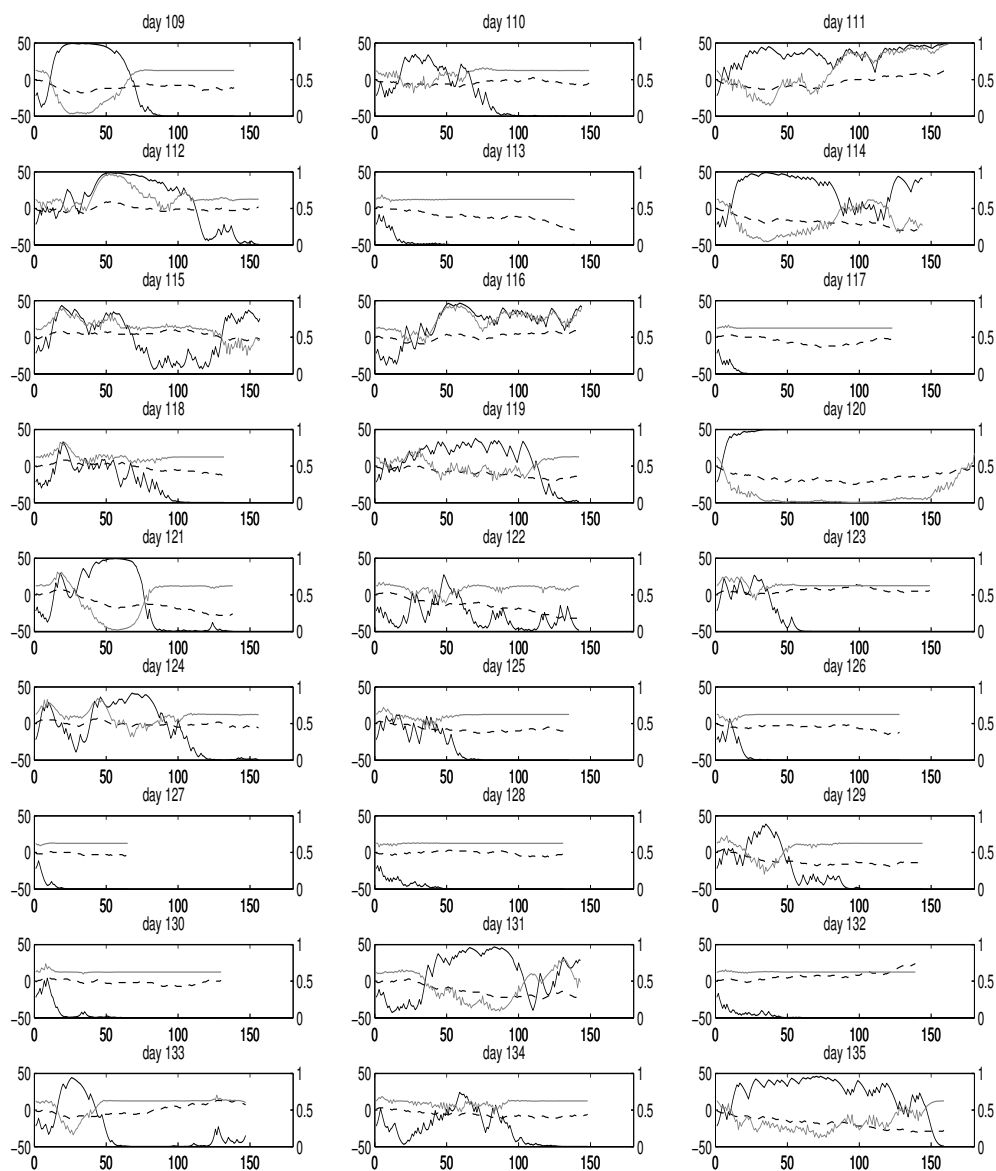


Figure G8

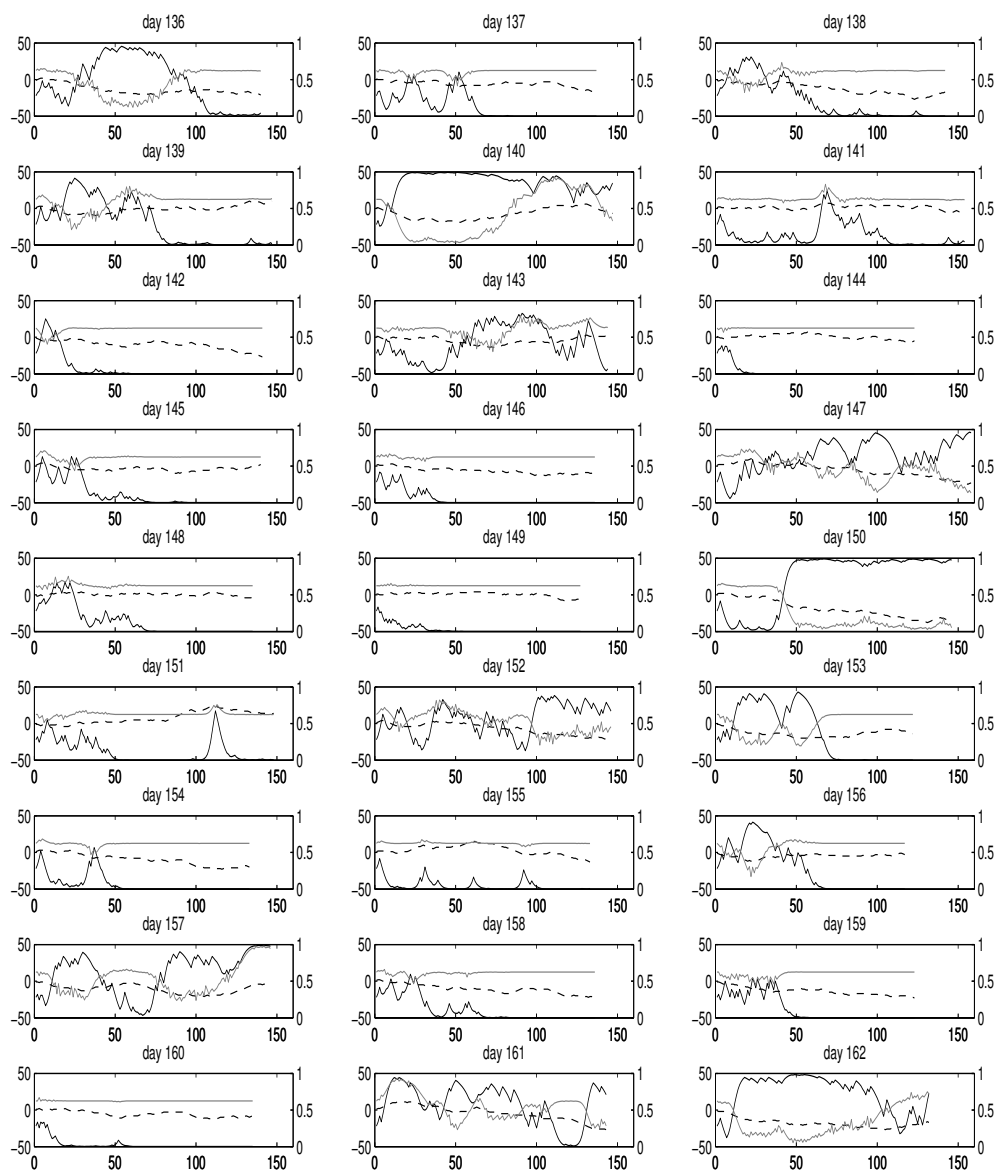


Figure G9

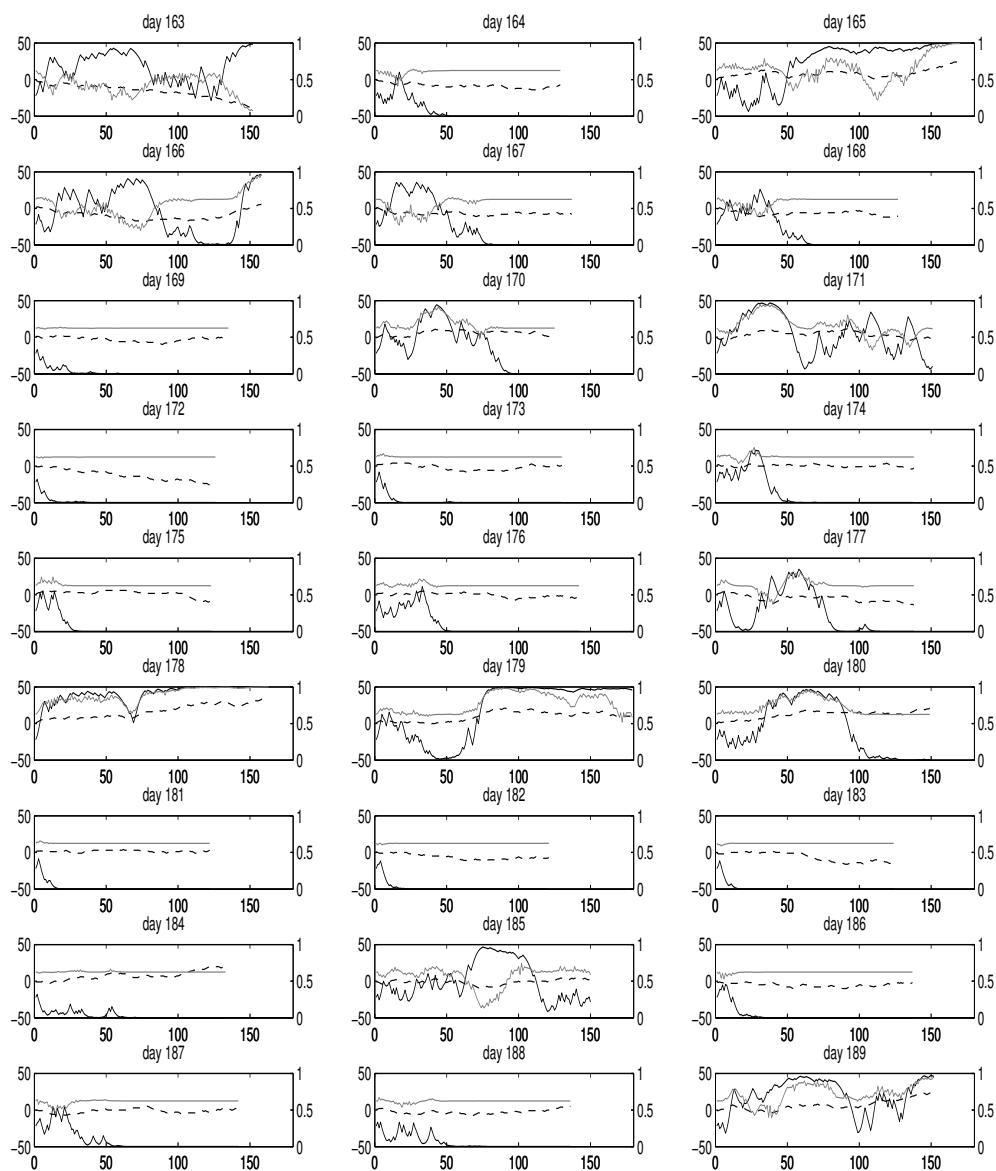


Figure G10

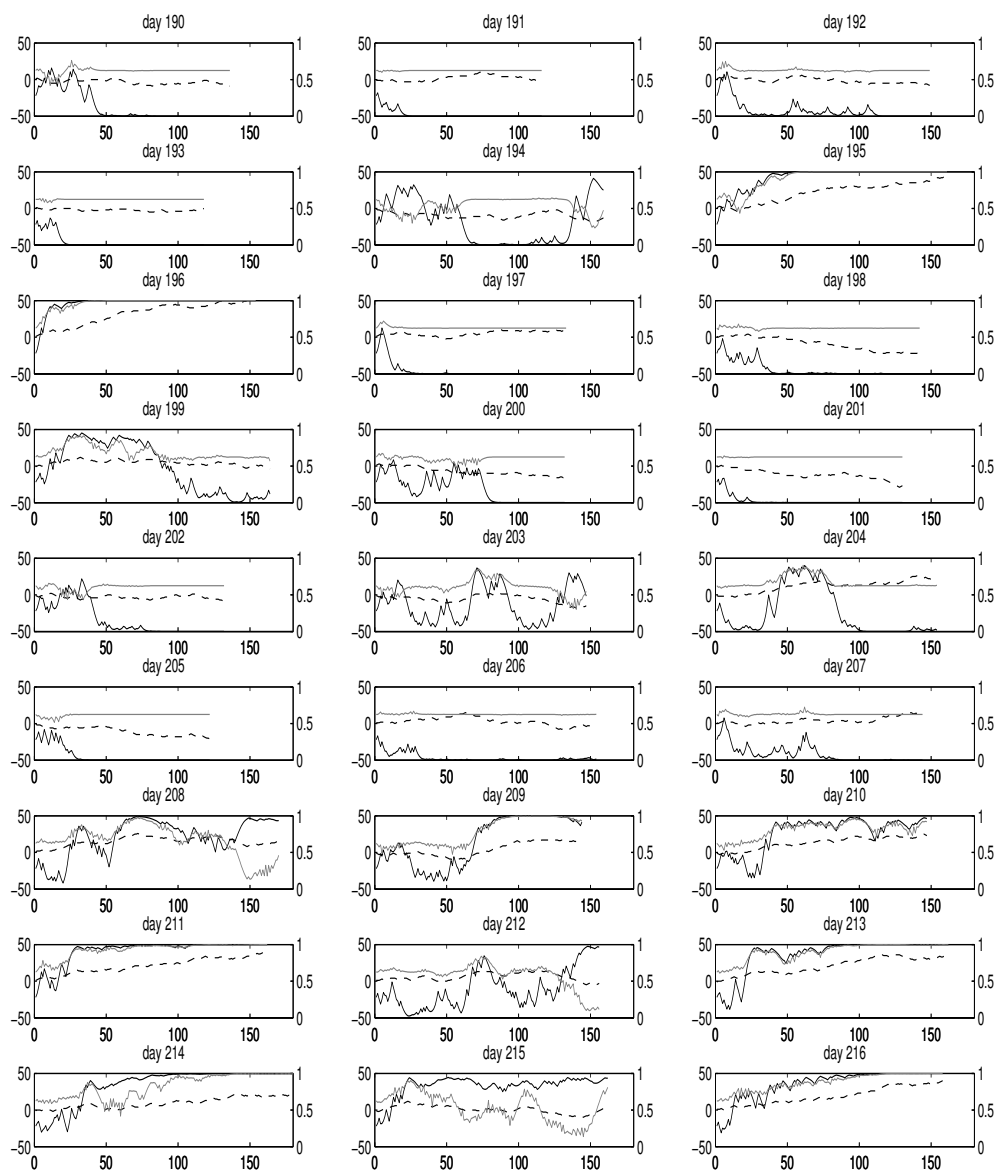


Figure G11

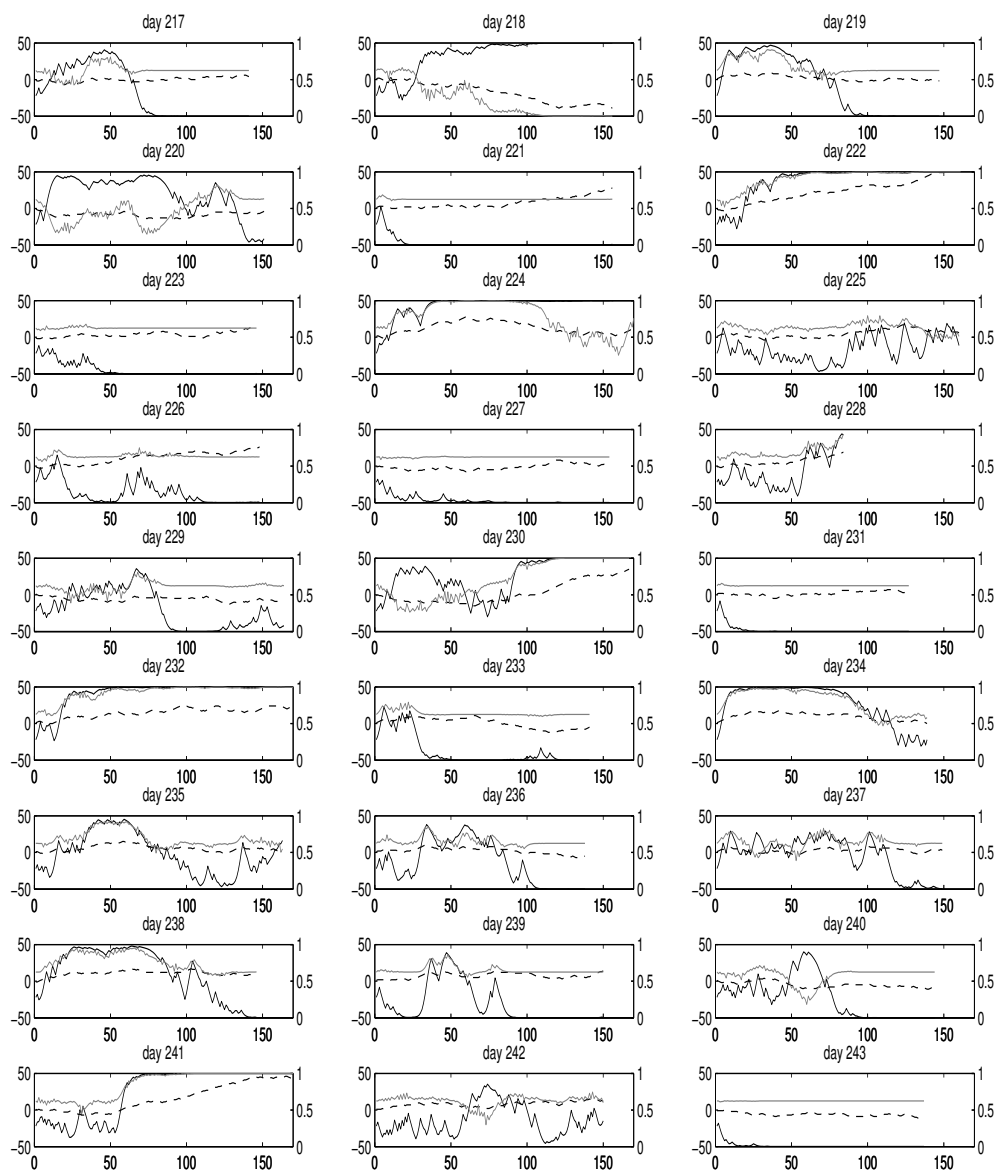


Figure G12

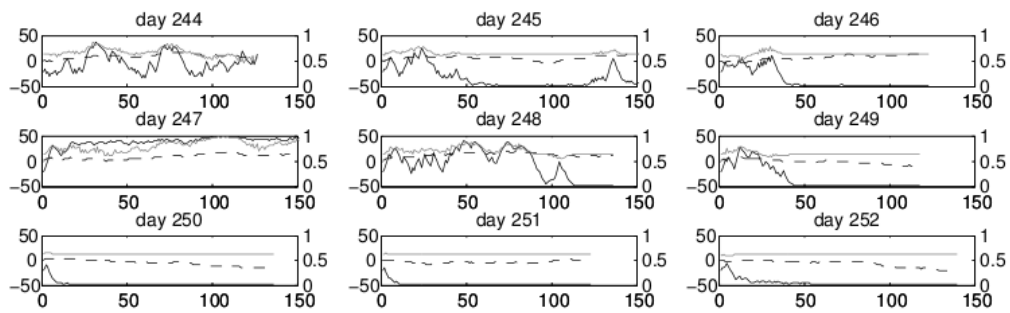


Figure G13

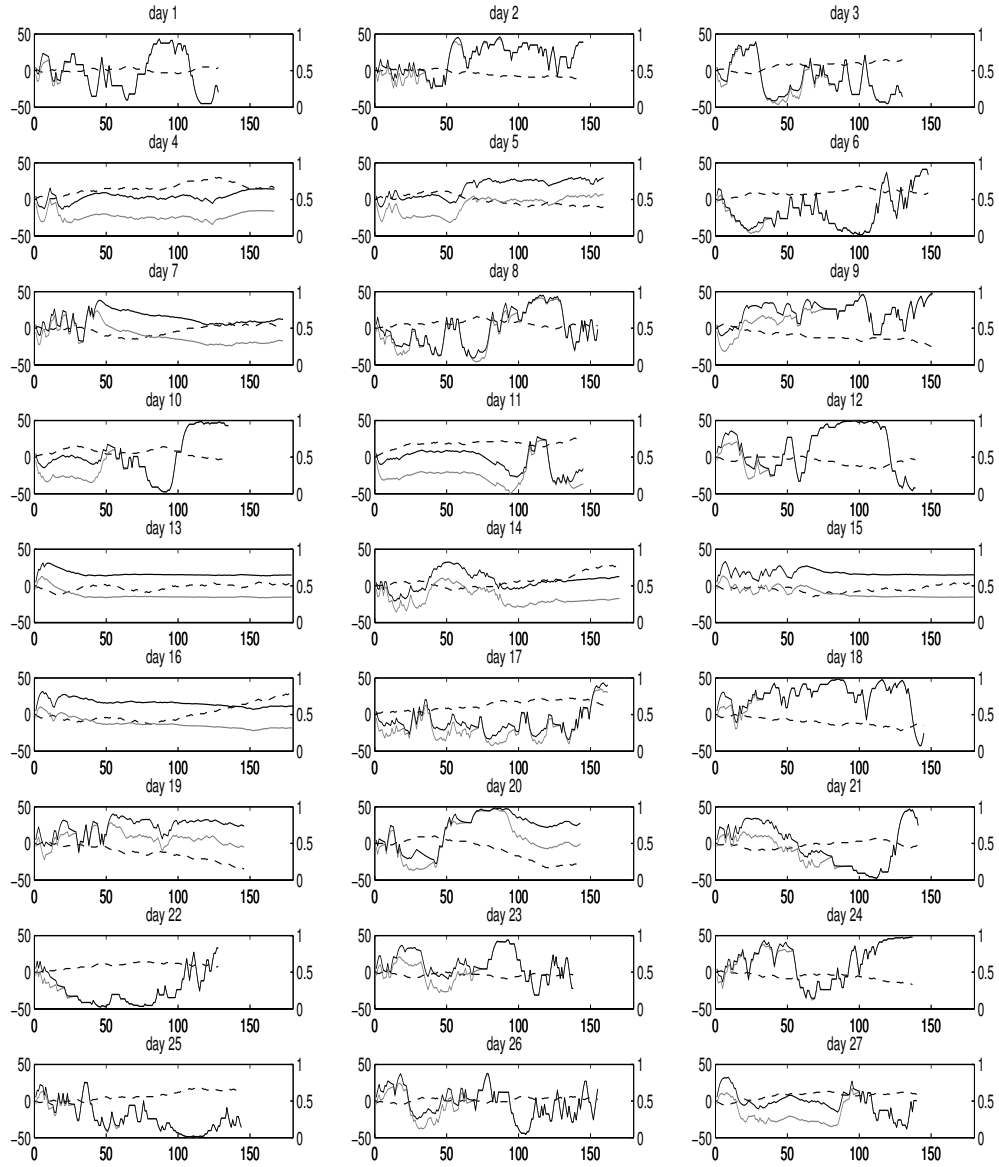


Figure G14 : The figure reports in each panel the daily evolution of: the trade imbalance (dashed line), the buy threshold (black solid line) and the sell threshold (grey solid line).

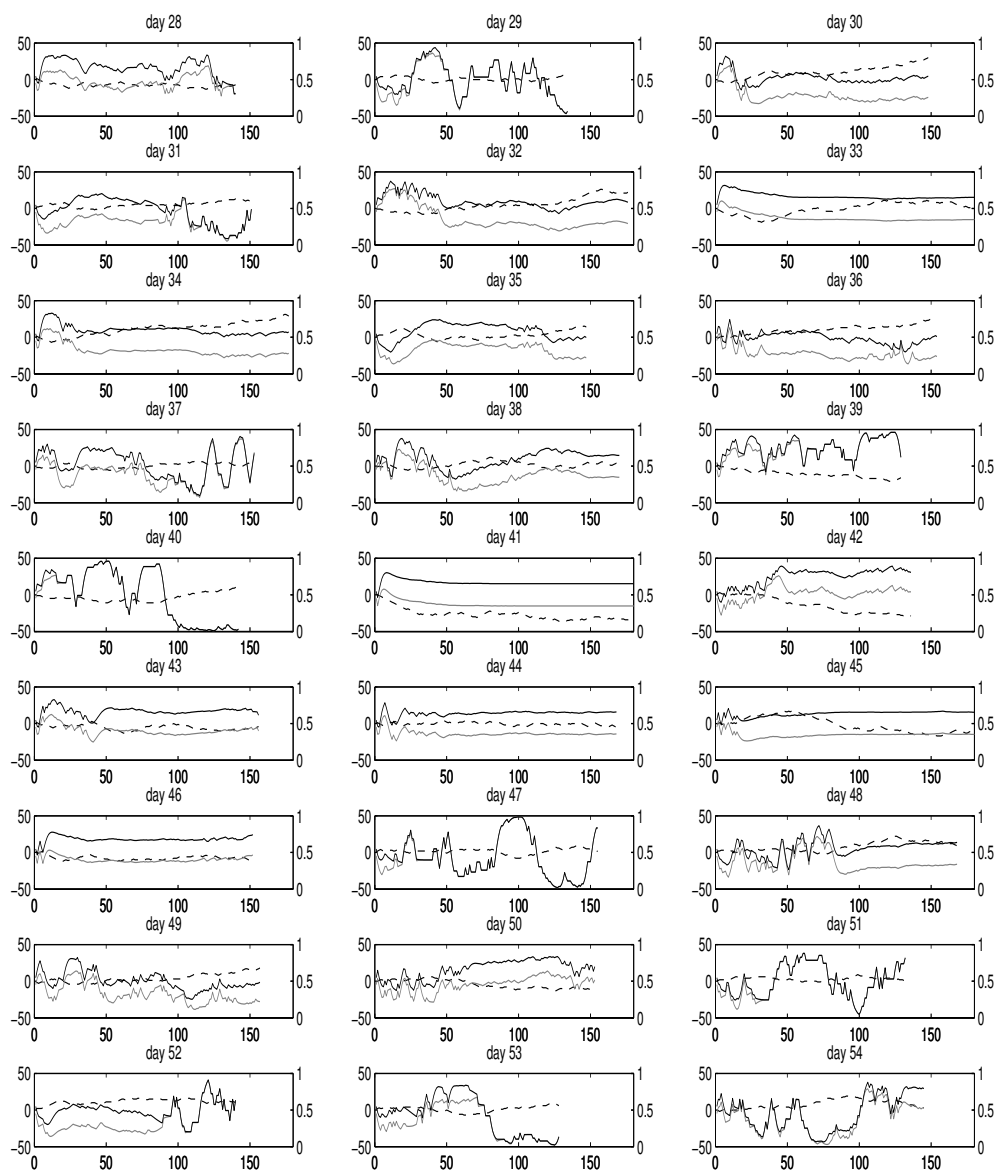


Figure G15

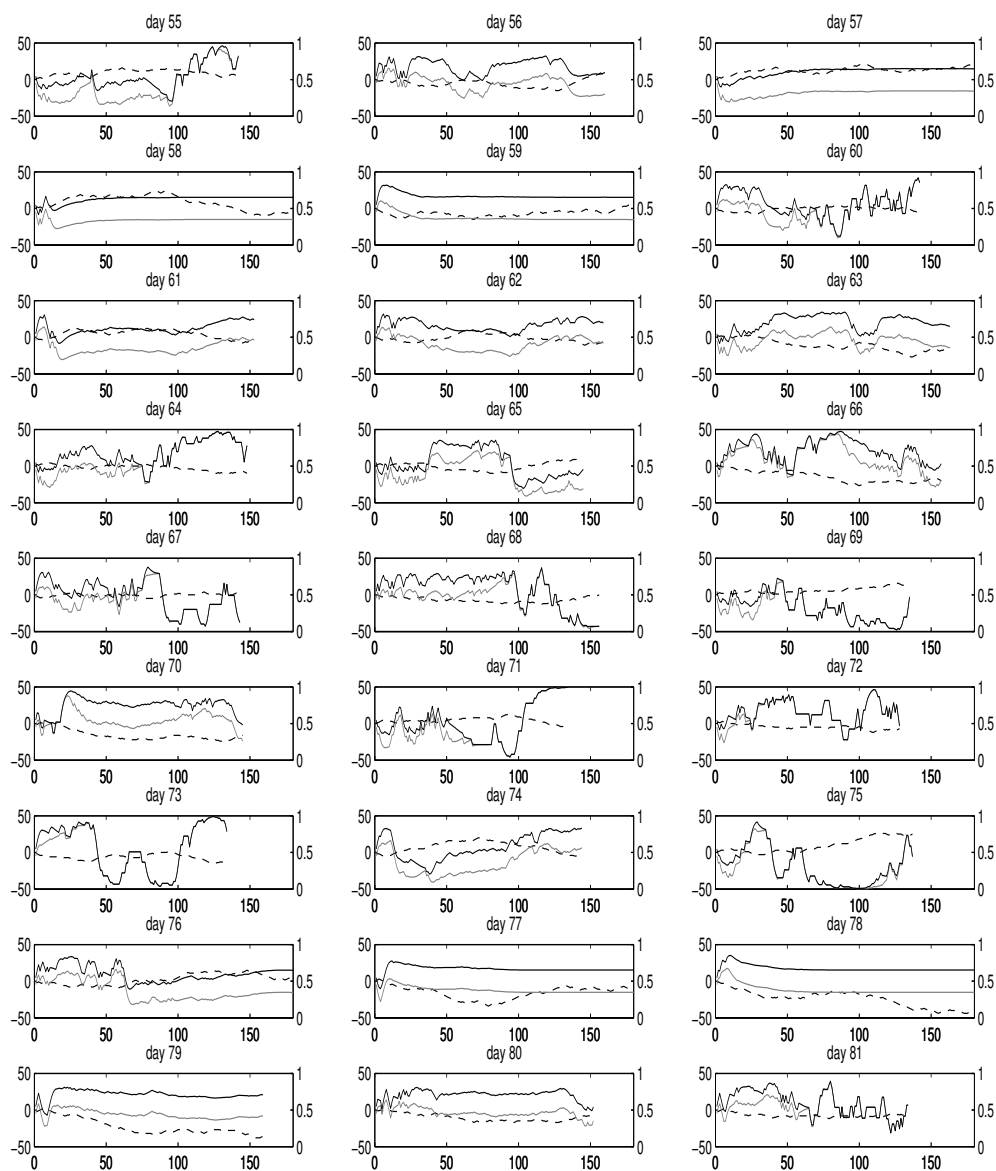


Figure G16

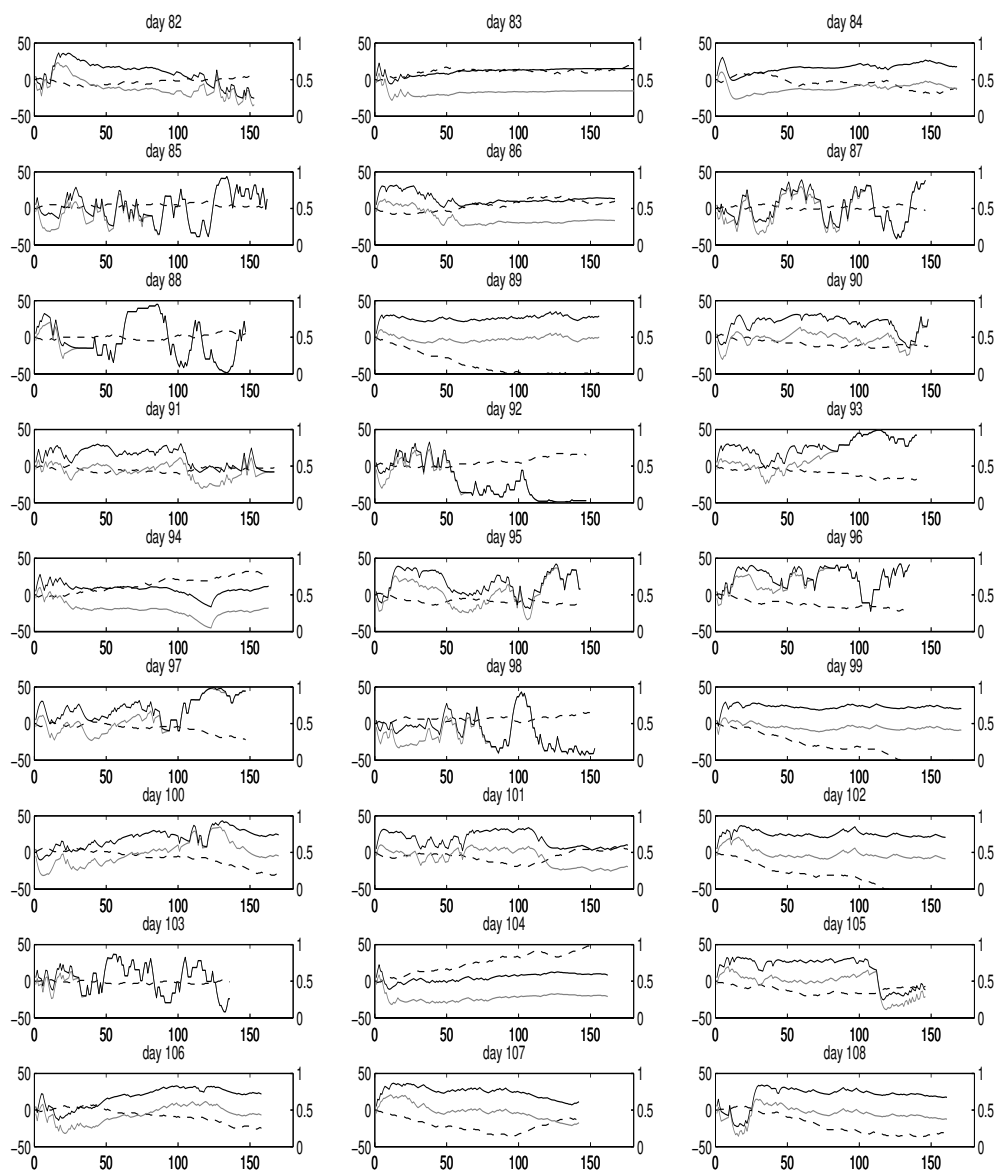


Figure G17

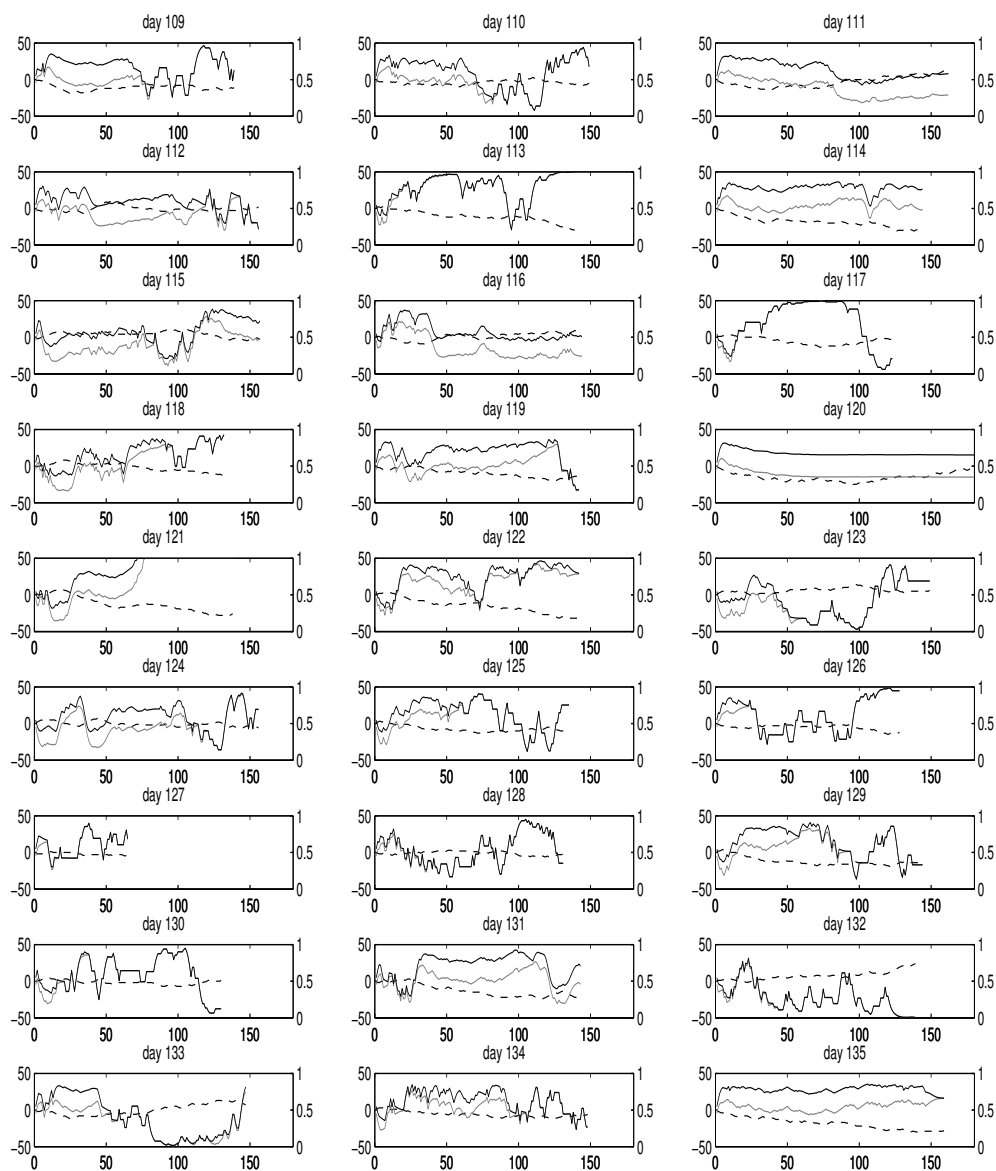


Figure G18

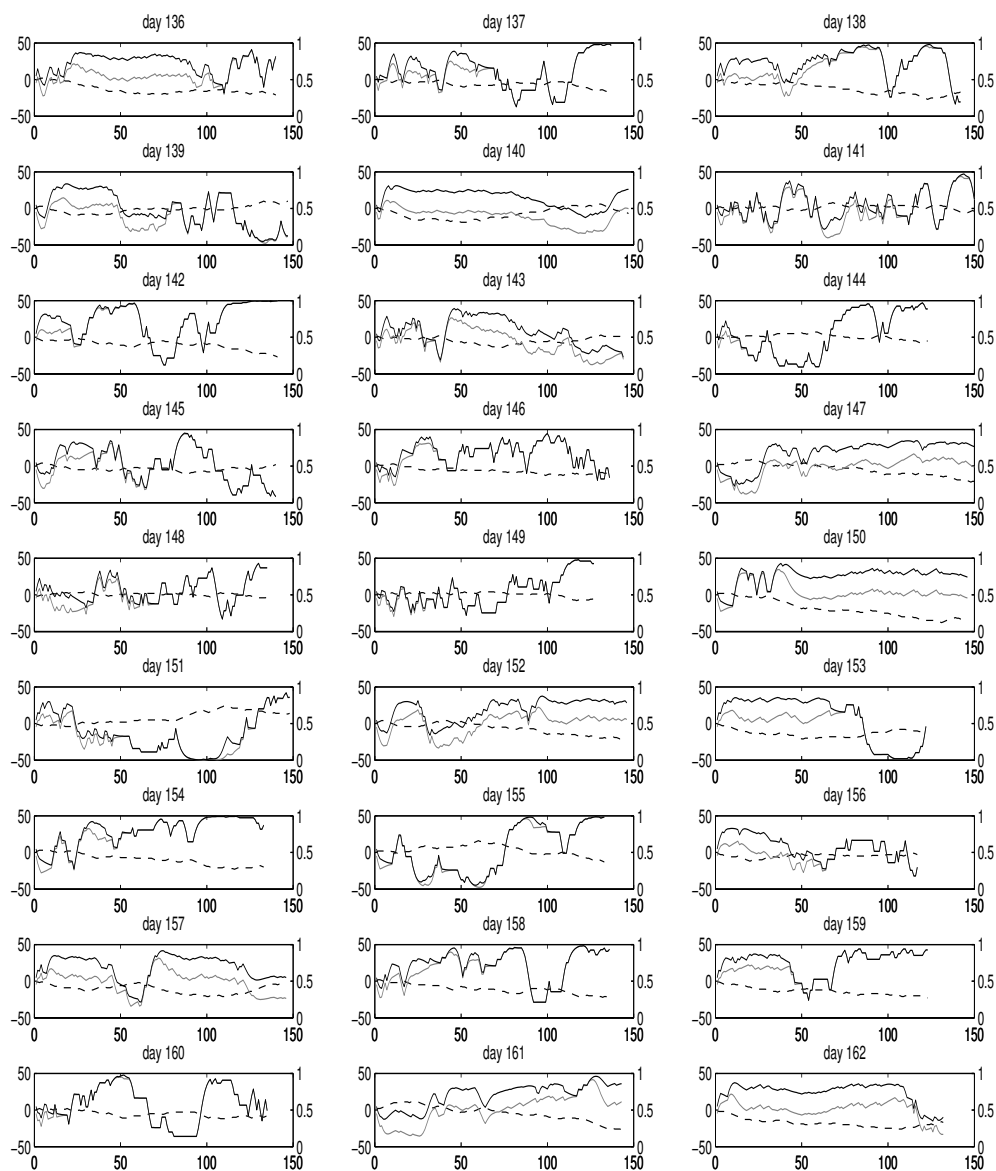


Figure G19

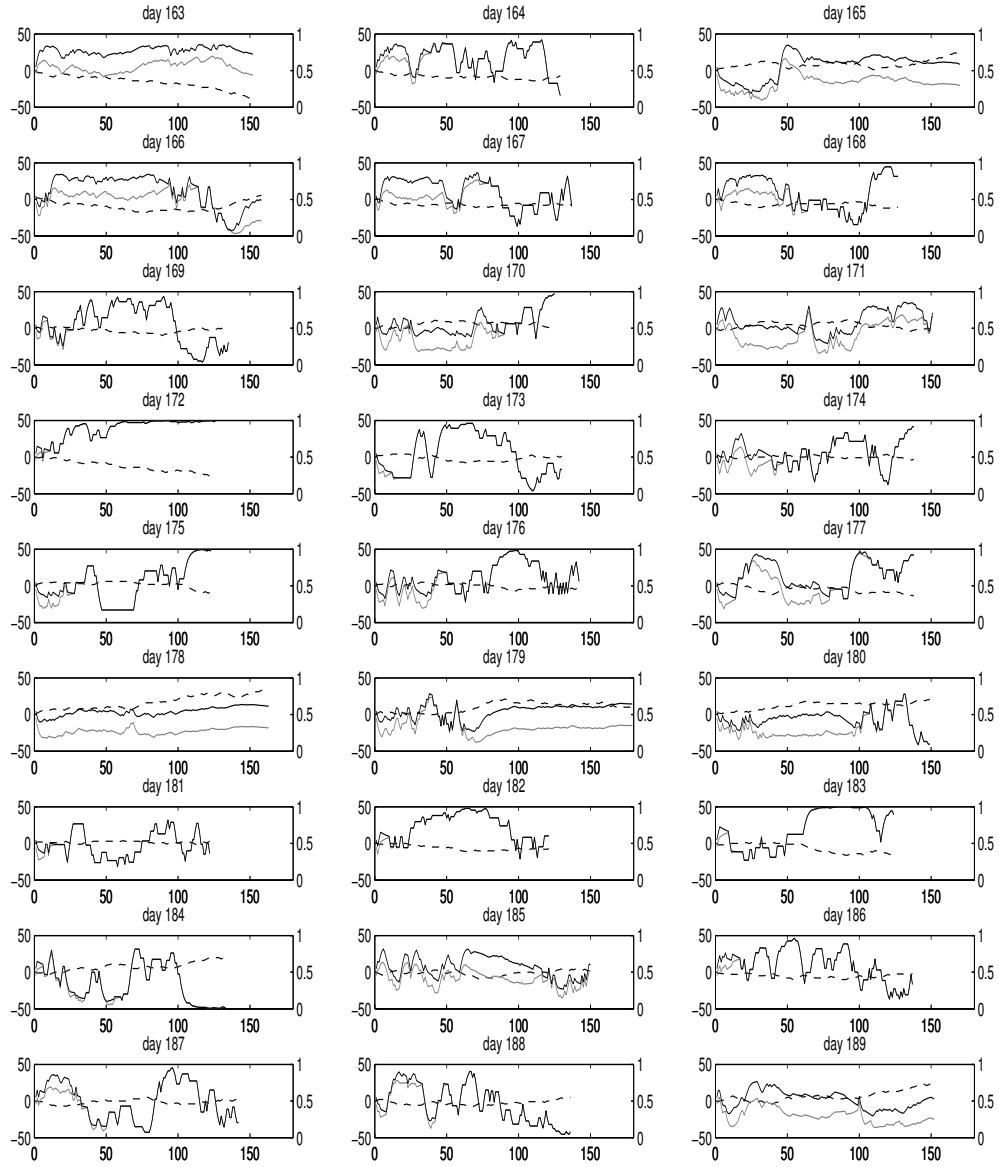


Figure G20

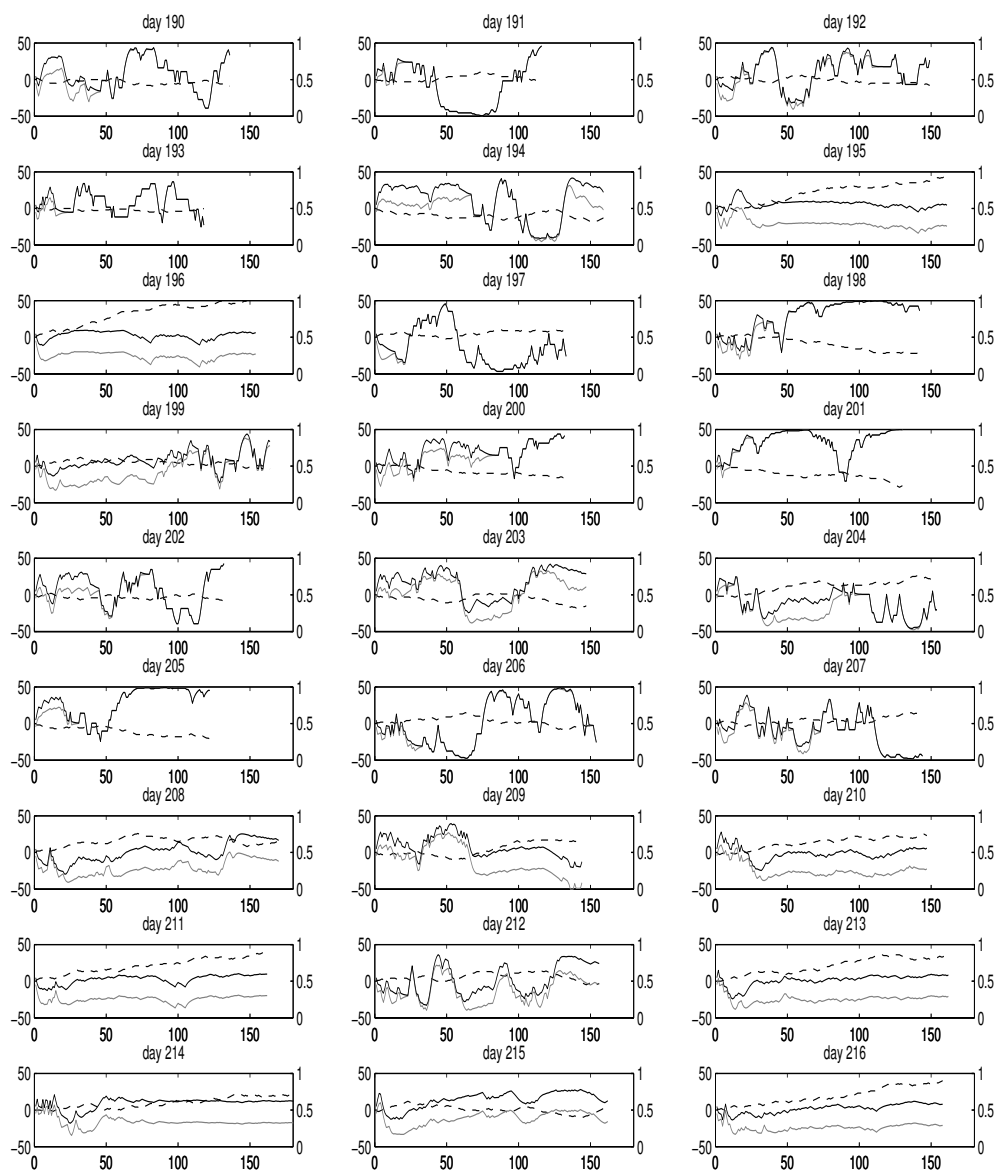


Figure G21

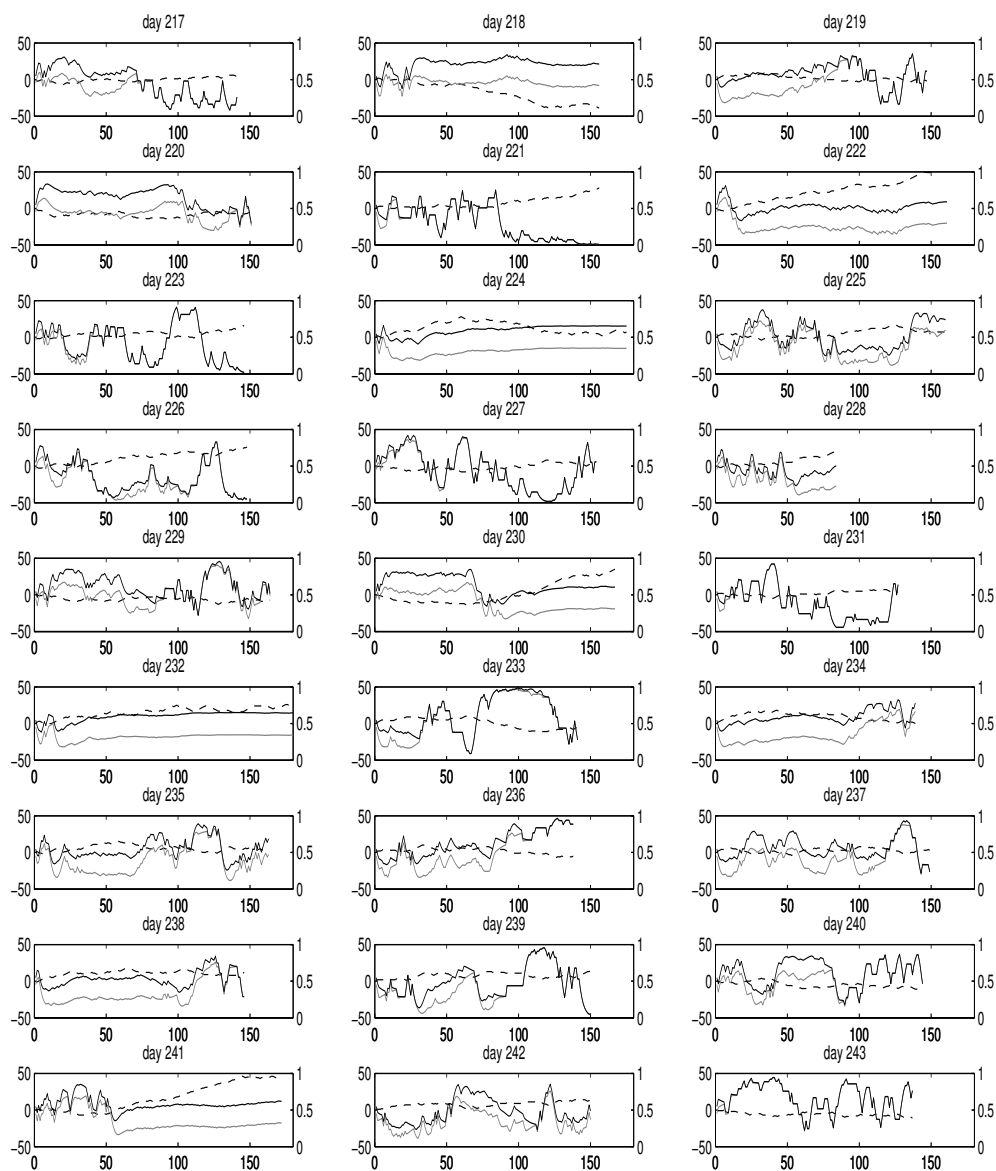


Figure G22

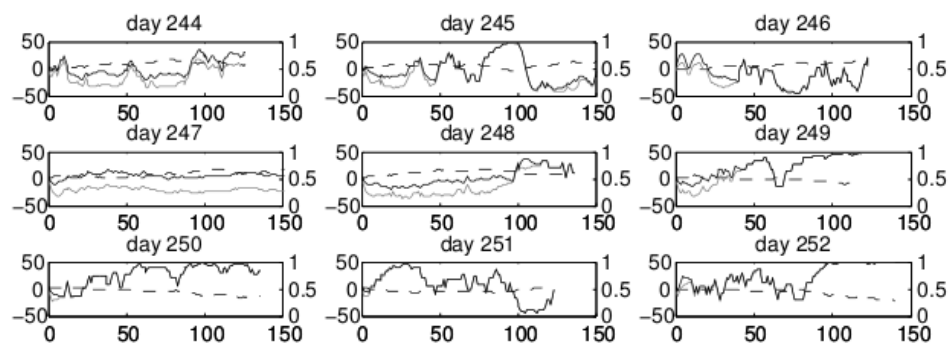


Figure G23

H A Comparison of our Results with the Lakonishok, Shleifer and Vishny (1992) Measure of Herding

H.1 Introduction

In this section we illustrate the measure of “herd behavior” introduced in the literature by Lakonishok, Shleifer and Vishny (1992) (LSV, from now on), and discuss its relation with our analysis. We proceed in the following way:

- we review the LSV methodology;
- we explain how it has been applied to the TAQ dataset;
- we present the LSV measure for Ashland Inc. transaction data (i.e., the data we use in our structural estimation);
- we compare the LSV measure with a “Maximum likelihood-based” (“ML-based”) LSV statistic that we obtain from our structural estimation. We show that the LSV measure is much higher than our ML-based LSV statistics;
- to shed more light on the previous finding, we prove analytically that in a market microstructure model in which by construction there is no herding (because private information is always correct, such as in the Easley, Kiefer and O’Hara, 1997), the LSV measure is nevertheless positive.

In summary, one should be cautious in using the traditional LSV measure with transaction-level data as a measure of informational herding; the reason is that the LSV measure does not disentangle true informational herding from trading based on the same piece of private information (sometimes referred to in the literature as “spurious herding,” see, e.g., Bikhchandani and Sharma, 2001).

H.2 The LSV measure

In most of the empirical herding literature, the LSV measure is used to estimate the extent to which a group of portfolio managers cluster their

portfolio decisions. This clustering (in excess of what would be predicted by a model of independent decision making) is referred to as herding. Let us consider an example. Let us say that we have data on a group of fund managers' holdings of stocks at the quarterly frequency. Suppose that, in a quarter, across all stocks and fund managers, 70 percent of portfolio decisions consisted of an increase in holdings of a given stock with respect to the previous quarter, and 30 percent of a decrease. Now, let us focus on a single stock, let us say GE. Suppose that 70 percent of fund managers increased their holdings of GE, whereas 30 percent decreased it. Then, we would say that there is no "herding" on GE, since, as a group, managers' decisions on GE are in line with their market-wide behavior. Suppose instead, that 90 percent of managers increased their holding of GE, and only 10 percent decreased them. Then, we would say that fund managers are "herding" on GE since, as a group, they buy GE in excess of their market-wide behavior.

Notice that this informal definition of herding is very different from that presented in our paper. Not only is it traditionally applied to a very different type of data (portfolio holdings as opposed to transaction data), it is also an inherently cross-sectional measure of herding (since to compute herding on one stock, fund managers' decisions on that stock are compared to their "average" market-wide behavior). More importantly, for a given stock, fund managers could be buying or selling simply because they received the same stock-specific private information, and not because they were imitating each other. That is, the measure does not distinguish informational herding from the reaction to the same piece of private information

We will come back to this point. For the time being, let us introduce the LSV formal measure of herding. Let us denote B_{iq} the number of fund managers who increase their holdings of stock i in quarter q . Similarly, S_{iq} is the number of fund managers who decrease their holdings of stock i in quarter q . The LSV measure of herding for stock i in quarter q is

$$H_{iq} = \left| \frac{B_{iq}}{B_{iq} + S_{iq}} - p_q \right| - E \left(\left| \frac{B_{iq}}{B_{iq} + S_{iq}} - p_q \right| \right),$$

where p_q is the proportion of managers increasing their holdings across all stocks in quarter q . The second addendum is an adjustment factor, to correct for the fact that because of the absolute value, $E \left| \frac{B_{iq}}{B_{iq} + S_{iq}} - p_q \right|$ is greater than zero under the assumption of no herding. "Absence of herding" is defined by LSV in purely statistical terms: there is no herding on stock i if fund

managers make their decisions of increasing or decreasing their holdings of the stock independently with probabilities p_q and $1 - p_q$. In this case, the decision of each fund manager on stock i can be modelled as an independent Bernoulli distribution with probability p_q . Therefore, the probability of observing exactly B_{iq} holding increases out of n portfolio choices, is given by $\binom{n}{B_{iq}} p_q^{B_{iq}} (1 - p_q)^{n - B_{iq}}$. Note that for a large number of fund managers trading stock i , $\frac{B_{iq}}{B_{iq} + S_{iq}}$ converges almost surely to p_q ; therefore, as the number of fund managers goes to infinity, the adjustment factor $E\left(\left|\frac{B_{iq}}{B_{iq} + S_{iq}} - p_q\right|\right)$ converges to 0.

How do we interpret H_{iq} ? Suppose we find $H_{iq} = 4$ percent. This means that, in quarter q , there are 4 percent more fund managers increasing (or decreasing) stock i 's holding than we would expect if there were no herding on the stock. In other words, managers are more on one side of the market for stock i than we could expect looking at their changes in positions across all stocks.¹¹ Lakonishok, Shleifer and Vishny (1992) report average quarterly herding measures across stock holdings.

H.3 The LSV measure applied to the TAQ dataset

Recently, the LSV measure has been applied to the TAQ dataset by Christoffersen and Tang (2010). At the daily level, herding for stock i in day d is measured by

$$H_{id} = \left| \frac{B_{id}}{B_{id} + S_{id}} - p_d \right| - E\left(\left| \frac{B_{id}}{B_{id} + S_{id}} - p_d \right|\right),$$

where B_{id} and S_{id} are the numbers of buy and sell orders in day d for stock i . Note that p_d is the average proportion of buy orders for the portfolio of all stocks that traders trade. Christoffersen and Tang (2010) approximate this measure by considering the average proportion of buy orders for all the stocks traded on the NYSE that are included in the TAQ dataset. They report the average herding estimate across stocks and across time. Additionally, following Wermers (1999), Christoffersen and Tang (2010) also report average

¹¹Note that the 4 percent is not necessarily in excess of 50 percent since in a given quarter managers may be in aggregate increasing or decreasing their stock holdings (i.e., p_q may be different from 0.5).

herding measures over those days in which $\frac{B_{id}}{B_{id} + S_{id}} - p_d > 0$, which they interpret as a measure of “herd buy” behavior; analogously, for herd sell.¹²

Instead of considering the entire universe of traders, one can consider more homogenous groups of traders.¹³ Christoffersen and Tang (2010) *also* computed the measure of herding for institutional and retail investors. Although the TAQ dataset does not contain information on the traders’ identity, there are algorithms in the literature to identify the trades coming from institutions or retailers. In particular, Campbell, Ramondarai and Schartz (2009) and Lee and Radhakrishna (2000) have developed algorithms to classify trades as retail or institutional on the basis of the trade size, an information available in the TAQ dataset. Finally, the analysis needs not be done at the daily frequency; in addition to daily herding statistics, Christoffersen and Tang (2010) produce monthly and intra-day herding measures.

H.4 LSV measure of herding for Ashland Inc.

Following Christoffersen and Tang (2010), we computed the LSV measure of herding for our stock, Ashland Inc. in 1995. The data are all taken from the TAQ dataset.¹⁴ To compute p_d , the overall proportion of buy orders in the market, we used two portfolios: the stocks included in the S&P100 index and those included in the Dow Jones industrial average index.¹⁵ Given that we are interested in computing the LSV measure for Ashland Inc. in 1995 to compare it to our results, we focused on three horizons: intradaily, daily and

¹²Note that in the case of, e.g., herd buy, the adjustment factor is computed as $E\left(\frac{B_{id}}{B_{id}+S_{id}} - p_d \mid \frac{B_{id}}{B_{id}+S_{id}} - p_d > 0\right)$.

¹³The importance of considering a homogeneous group of traders is emphasized by Lakonishok, Shleifer and Vishny (1992) themselves (see, e.g., their considerations at pages 2 and 31).

¹⁴To classify transactions as buy or sell orders we used the same classification rules discussed in the paper for Ashland Inc.

¹⁵While the two indices obviously do not comprise all equities potentially traded by traders of Ashland Inc., they represent a large fraction of the capitalization of the US equity market. Of course, we used the stocks that were part of the two indices in 1995. For the S&P100 index, we excluded stocks that in 1995 were not traded in the NYSE (e.g., because traded over the counter or in the Nasdaq). We also excluded three stocks for which the TAQ dataset contained transactions data only for some days and not for the entire year — e.g., because the stock became a constituent of the index during the year). The complete list of stocks that we used in the analysis is available from the authors.

weekly.¹⁶ Table *H.1* reports the results.

Time Interval	Benchmark: DJ30			Benchmark: S&P100		
	Herd	Herd buy	Herd sell	Herd	Herd buy	Herd sell
Intraday	4.8%	2.0%	6.3%	4.6%	2.0%	6.0%
Day	4.8%	1.8%	5.9%	4.2%	1.9%	5.2%
Week	4.0%	0.0%	5.2%	3.7%	0.5%	5.0%

Table H.1: LSV measures of herding for Ashland Inc. for 1995.

The table shows the LSV measure of herding for three time intervals: intradaily, daily, weekly. The LSV measure is computed using two benchmarks: Dow30 constituents and SP100 constituents.

The first thing to notice is that the results are very similar for the two benchmarks. Overall, at the daily level, the LSV measure of herding is above 4 percent. This means that, on average, during 1995, there were more than 4 percent of buy (or sell) orders in excess of the buy (sell) orders in the rest of the market. The results for intradaily and weekly intervals are not dissimilar from those at the daily level, although one can notice that the shorter the interval, the higher the measure of herding, a finding also present in Christoffersen and Tang (2010).

In Table *H.2*, we report the measures of herding restricting the analysis to institutional investors. In the first two columns of Table *H.2*, a trade is classified as institutional if its size is higher than \$50,000 (as proposed by Lee and Radhakrishna, 2000), whereas in the last two columns a trade is classified as institutional if its size is either lower than \$2,000 or higher than \$30,000 (as proposed by Campbell, Ramondarai and Schartz, 2009).

The results are very similar, that is, the two methods of classification of institutional trades do not produce significantly different outcomes. Institutional traders seem to herd slightly less than other traders. The measure of herding is again decreasing with the time interval.

Finally, in Table *H.3*, we report the herding measures for retail trades (following Lee and Radhakrishna, 2000, a trade is classified as retail when its size is lower than \$5,000).

¹⁶We cannot reliably compute monthly or quarterly herding measures since we only have one year of data.

Time Interval	DJ30	S&P100	DJ30	S&P100
	Herd	Herd	Herd	Herd
Intraday	3.7%	3.7%	3.6%	3.7%
Day	3.3%	3.3%	3.3%	3.4%
Week	2.3%	2.5%	2.0%	2.4%

Table H.2: LSV herding measures for institutional investors (Ashland Inc., 1995).

The table shows the LSV measure of herding for institutional investors for three time intervals: intraday, day, week. The LSV measure is computed using two benchmarks: Dow30 constituents and SP100 constituents. In the first two columns, trades are classified as institutional when their value is higher than 50,000 dollars; in the last two columns, trades are classified as institutional when their value is lower than 2,000 or higher than 30,000 dollars.

Time Interval	Benchmark: DJ30	Benchmark: S&P100
	Herd	Herd
Intraday	4.3%	3.5%
Day	4.1%	2.7%
Week	5.7%	3.4%

Table H.3: LSV herding measures for retail investors (Ashland Inc., 1995).

The table shows the LSV measure of herding for retail investors for three time intervals: intraday, day, week. The LSV measure is computed using two benchmarks: Dow30 constituents and SP100 constituents. Trades are classified as retail when their value is lower than 5,000 dollars.

H.5 LSV and informational herding

As mentioned above, the statistical measure of herding proposed by Lakonishok, Shleifer and Vishny (1992) and the one we present in our paper are very different. We focus on just one stock, and estimate both the frequency of herding periods and the “proportion of herders,” that is, the proportion of informed traders who, when called to trade in a herding period, would disregard their private information in order to follow their predecessors. The LSV measure, instead, estimates the proportion of traders buying a given stock in excess of the proportion of traders who buy the market portfolio. Comparing the LSV measure directly with the estimates of herding that we report in our paper would, therefore, not be correct since the two measures

capture different phenomena.

In order to compare the LSV results to ours, in this section we study the implications of our measures of herding in terms of “excessive buying” or “excessive selling” (which is what LSV measures). We use our maximum likelihood estimates to construct what we call an “ML-based” LSV measure, which can be directly compared with the traditional LSV measure. We show that this “ML-based” LSV measure is much lower than the traditional LSV measure. Our result suggests that the traditional LSV measure overestimates the amount of informational herding in the market. To shed more light on why this is the case, we conclude the section by proving analytically that in a (multi-asset) market microstructure model the LSV measure of herding applied to the TAQ dataset is positive even in the absence of informational herding.

H.5.1 Comparing the LSV Measure with Our Measure of Informational Herding

We construct the ML-based LSV measure as the difference between the actual proportion of buy orders and the proportion of buy orders that we would have observed had traders not engaged in informational herding (given our estimated maximum likelihood parameters). Specifically, for each good-event day d for Ashland Inc. we compute

$$H_d^{Good} = \left| \frac{B_d^{Good}}{B_d^{Good} + S_d^{Good}} - p_{Good}^{ML} \right| - E \left(\left| \frac{B_d^{Good}}{B_d^{Good} + S_d^{Good}} - p_{Good}^{ML} \right| \right),$$

and similarly for each bad-event day, we compute

$$H_d^{Bad} = \left| \frac{B_d^{Bad}}{B_d^{Bad} + S_d^{Bad}} - p_{Bad}^{ML} \right| - E \left(\left| \frac{B_d^{Bad}}{B_d^{Bad} + S_d^{Bad}} - p_{Bad}^{ML} \right| \right),$$

where B_d^{Good} and S_d^{Good} are the numbers of buy and sell orders on Ashland Inc. in good-event days, B_d^{Bad} and S_d^{Bad} the numbers of buy and sell orders in bad-event days, and p_{Good}^{ML} and p_{Bad}^{ML} are the proportions of buy orders that we would observe in good and bad event days if traders did not herd, given our maximum likelihood estimates. The expectations (in the adjustment factors)

are computed in the same way as in the traditional LSV measure. We can then compute the ML-based LSV measure as¹⁷

$$H^{ML} = \alpha \delta \frac{\sum_{d \in \mathcal{G}} H_d^{Good}}{N^{Good}} + \alpha(1 - \delta) \frac{\sum_{d \in \mathcal{B}} H_d^{Bad}}{N^{Bad}}.$$

where $\mathcal{G}$ is the set of good-event days (with cardinality N^{Good}) and $\mathcal{B}$ is the set of bad-event days (with cardinality N^{Bad}).

As one can immediately see, the difference between our ML-based LSV measure and the standard one is that to compute the ML-based LSV measure we use the proportions p_{Good}^{ML} and p_{Bad}^{ML} instead of p_d (the average proportion of buy orders in the market). Given our structural estimation, we can compute p_{Good}^{ML} and p_{Bad}^{ML} , the proportions of buy and sell orders in the absence of informational herding, in a straightforward way. Let us neglect for a moment the presence of the bid-ask spread. In the absence of informational herding traders with a good signal would buy and traders with a bad signal would sell. Therefore, in a good-event day, the proportion of buy orders would be

$$p_{Good}^{ML} = \frac{\mu(0.5 + 0.25\tau) + (1 - \mu)\frac{\varepsilon}{2}}{\mu + (1 - \mu)\varepsilon},$$

and, similarly, in a bad-event day it would be

$$p_{Bad}^{ML} = \frac{\mu(0.5 - 0.25\tau) + (1 - \mu)\frac{\varepsilon}{2}}{\mu + (1 - \mu)\varepsilon}$$

By using our estimated parameters, we obtain $p_{Good}^{ML} = 0.56$ and $p_{Bad}^{ML} = 0.44$. The analytical expressions derived above are an approximation, since we are neglecting the bid and ask spread. To take the spread into account, we also simulated the model under the assumption that the market maker knows whether there has been an informational event, and sets the price accordingly. Obviously, as discussed in the paper (see the second benchmark for the efficiency analysis), the market maker's knowledge of the event implies that the price is set in such a way that informed traders never find it optimal to herd. The results of the simulation analysis are almost identical to the theoretical approximations: $p_{Good}^{ML} = 0.57$ and $p_{Bad}^{ML} = 0.44$.¹⁸

¹⁷Note that in no-event days by definition there is no informational herding, since there are no informed traders.

¹⁸We simulated the trading process given our parameter estimates for 1,000,000 days. That the approximation is almost identical to the simulated result is not surprising. In the analytical formulas, we are approximating both the proportion of buy orders and the proportion of sell orders. The two approximations compensate each other.

We find that $H^{ML} = 0.01$, that is, the ML-based LSV measure is only 1 percent. In contrast, recall that the traditional LSV measure (reported in Table H.1) is higher than 4 percent. Why the difference? The reason is actually very simple. In any event day, informed traders may buy (or sell) more Ashland Inc. than they buy (or sell) the other stocks not because they are herding, but mainly because they receive mostly signals of the same type (e.g., good news in good event days). In this situation, the traditional LSV measure will be high not because traders engage in informational herding, but because there is some stock-specific information that makes most informed traders buy (sell) the stock. Analogously, in a no-event day, traders may buy the asset more (or less) than they buy other stocks (which results in a high LSV measure) because the asset has not been affected by an informational event that has affected many other stocks (e.g., a systemic event). In other words, the LSV measure, which is a measure of clustering of actions, cannot disentangle spurious herding from true herd behavior. We can do it with our structural estimation, hence the difference between the LSV measure of 4 percent and our ML-based one of 1 percent.

H.5.2 LSV in the absence of informational herding

To conclude, in this section we build a very simple multi-asset market microstructure model, in which herding cannot occur in equilibrium because informed traders know the state of the world (as in Easley, Kiefer and O'Hara, 1997). We show that, nevertheless, the LSV measure is positive.

Let us consider a multi-asset economy in which an asset (a stock) is modelled as in Easley, Kiefer and O'Hara (1997). Let us also assume that asset values are independent, and the parameters governing the trading process $(\alpha, \delta, \mu, \varepsilon)$ are the same across assets. For simplicity, let us assume that traders are restricted to trade on a given stock only (e.g., uninformed traders have liquidity reasons to buy or sell a specific stock in the economy; similarly, informed traders receive a stock-specific piece of private information, and they trade to buy or sell the stock on which they are informed).

As we explained in the paper, in such a market there is no herding, since informed traders all receive the correct information. What would the LSV measure be in this economy? For simplicity, assume that each day there are sufficiently many trades and sufficiently many stocks that we can use the strong law of large numbers both across days and across stocks. In that case, in a good event day, the proportion of buy orders (over the total number of

orders) for stock i converges almost surely (a.s.) to

$$\frac{B_{id}}{B_{id} + S_{id}} = \frac{(1 - \mu)\frac{\varepsilon}{2} + \mu}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}}.$$

In a bad event days, it converges a.s. to

$$\frac{B_{id}}{B_{id} + S_{id}} = \frac{(1 - \mu)\frac{\varepsilon}{2}}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}},$$

and in a no-event day it converges a.s. to

$$\frac{B_{id}}{B_{id} + S_{id}} = \frac{(1 - \mu)\frac{\varepsilon}{2}}{(1 - \mu)\frac{\varepsilon}{2} + (1 - \mu)\frac{\varepsilon}{2}} = \frac{1}{2}.$$

Given that events are i.i.d., the proportion of buy orders in a day is

$$\begin{aligned} p_d &= \frac{\sum_{i=1}^n B_{id}}{\sum_{i=1}^n (B_{id} + S_{id})} = \\ &\alpha \delta \left(\frac{(1 - \mu)\frac{\varepsilon}{2} + \mu}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}} \right) + \\ &\alpha(1 - \delta) \frac{(1 - \mu)\frac{\varepsilon}{2}}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}} + \\ &(1 - \alpha) \frac{1}{2}. \end{aligned}$$

Notice that since we have assumed that there are many trades for each stock, the adjustment factor in the LSV formula, $E \left(\left| \frac{B_{id}}{B_{id} + S_{id}} - p_d \right| \right)$, converges to 0. Therefore, for a good day, the LSV measure of herding converges to

$$H_{id}^{Good} = \left| \left(\frac{(1 - \mu)\frac{\varepsilon}{2} + \mu}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}} \right) - p_d \right|,$$

in a bad day, it converges to

$$H_{id}^{Bad} = \left| \left(\frac{(1 - \mu)\frac{\varepsilon}{2}}{((1 - \mu)\frac{\varepsilon}{2} + \mu) + (1 - \mu)\frac{\varepsilon}{2}} \right) - p_d \right|,$$

whereas in a no-event day, it converges to

$$H_{id}^{NoEvent} = \left| \frac{1}{2} - p_d \right|.$$

Given the parameter values we estimated for Ashland Inc. for the Easley, Kiefer and O'Hara (1997) model ($\alpha = 0.33$, $\delta = 0.60$, $\mu = 0.17$, $\varepsilon = 0.58$), we obtain that $p_d = 0.51$, $H_{id}^{Good} = 0.122$, $H_{id}^{Bad} = 0.139$ and $H_{id}^{NoEvent} = 0.009$. That is, on event days, the LSV measure is very high although informed traders do not herd. The intuition for this result is immediate. Whenever there is, e.g., a good event for a stock, all informed traders buy the stock. As a result, we will observe a proportion $\frac{(1-\mu)\frac{\varepsilon}{2}+\mu}{((1-\mu)\frac{\varepsilon}{2}+\mu)+((1-\mu)\frac{\varepsilon}{2})} = 0.63$ of buy orders for that stock, whereas in the market as a whole the proportion of buy orders is only $p_d = 0.51$. Agents trading stock i do not imitate each other; they cluster their decisions only because they receive the same information. The LSV measure of herding is positive, although there is no informational herding.

If we consider all (i.e., event and no-event) days, as we have done in Table H.1 for Ashland Inc., the average LSV daily measure would be $H_i = 4.8$ percent.¹⁹ The fact that the LSV measure of herding may capture the reaction of traders to the same information is well known in the literature, even beyond the specific application to the TAQ dataset (see, e.g., Choe, Bong-Chan and Stulz, 1999, and the reviews of the literature by Bikhchandani and Sharma, 2001, and Hirshleifer and Teoh, 2009). With a structural model as the one presented in this Addendum, however, we can actually quantify the amount of spurious herding captured by LSV.

References

- [1] Campbell, John, Tarun Ramondarai, and Allie Schartz, 2009, "Caught on Tape: Institutional Trading, Stock Returns, and Earning Announcements," *Journal of Financial Economics*, vol. 92, pp. 66—91.

¹⁹Of course, the result that $H_i = 4.8$ percent depends on our assumption on the structure of the economy; in particular, for simplicity's sake we assumed the same structural parameters for all stocks, and i.i.d. information events across stocks. *Caeteris paribus*, negative (positive) correlation among events would increase (decrease) the LSV measure of herding. Nevertheless, even if information events were correlated across stocks, the LSV measure of herding in an economy with no informational herding would still be positive.

- [2] Choe Hyuk, Kho Bong-Chan and Rene Stulz, 1999, “Do Foreign Investors Destabilize Stock Markets? The Korean Experience in 1997,” *Journal of Financial Economics*, vol. 54, pp. 227—264.
- [3] Lee, Charles and Balkrishna Radhakrishna, 2000, “Inferring Investor Behavior: Evidence from TORQ data,” *Journal of Financial Markets*, vol. 3, 83—111.