What drives the herding behavior of individual investors?*

Maxime Merli[†] Tristan Roger[‡]

October 2011

Abstract

This article intends to provide answers concerning what drives individual investor herding behavior. Our empirical study uses transaction records of 87,373 French individual investors for the period 1999-2006. In a first part, we show - using both the traditional Lakonishok *et al.* (1992) and the more recent Frey *et al.* (2007) measures - that herding is prevalent and strong among French individual investors. We then show that herding is persistent: stocks on which investors concentrate their trades at time t are more likely to be the stocks on which investors herd at time t+1. In a second part, we focus on the motivations of individual herding behavior. We introduce an investor specific measure of herding which allows us to track the persistence in herding of individual investors. Our results highlight that this behavior is influenced by investor-specific characteristics. We also reveal the fact that individual herding behavior is strongly and negatively linked with investors' own past performance.

*We are grateful to the brokerage firm that provided us with the data for this study. We thank Patrice Fontaine, Sonia Jimenez-Garces, Russ Wermers and Raphaëlle Bellando, as well as seminar participants at the 2011 Academy of Behavioral Finance and Economics, Grenoble University and Strasbourg University, for comments and suggestions.

†Professor of Finance, LaRGE Research Center, EM Strasbourg Business School, University of Strasbourg. E-mail: merli@unistra.fr

‡**Corresponding author.** PhD student, EUROFIDAI-CERAG (CNRS). University of Grenoble. E-mail: tristan.roger@eurofidai.org

1 Introduction

Herding behavior, defined in a broad way as the fact for an investor to imitate others' actions, has been widely documented for professionals but less studied for individual investors. In this paper, we analyze herding behavior by studying a large trading record of 87,373 French individual investors from a major European broker house over the period 1999-2006. To answer the question posed by the title of the paper, we investigate in an original way the herding behavior and its persistence over time, both at an asset and at an individual level. This new approach leads us to analyze the influence of individuals' attributes, such as sophistication, on the herding behavior.

In a strict sense, herding is defined as the fact of irrationally imitating other agents. This type of herding is extremely difficult to capture empirically as it is driven by fashion and fads. In the context of a quantitative study, we focus on the second type of herding, namely rational herding. Devenow and Welch (1996), in a survey of the herding literature, emphasize three reasons for rational herding. The first one is payoff externalities (the outcome of an action is increasing in the number of agents undertaking this same action). Such payoff externalities are at the source of trading patterns caused by liquidity issues. It has been documented that investors tend to trade at the same time in order to benefit from a deeper liquidity (Admati and Pfleiderer, 1988; Dow, 2005). Reputational effects are the second reason for rational herding. They are particularly important in the context of principal-agent models. It can be said of a manager that he is "hiding in the herd". The idea behind this metaphor is that the performances of institutional traders are very often considered relative to a benchmark (the average performance of other managers or the performance of a market/industry index). By following closely the benchmark, the manager sacrifices a potential to perform better than average but hedges himself against bad relative performances. Models of herding caused by reputational concerns can be found in Scharfstein and Stein (1990), Rajan (1994) or Graham (1994). Finally, the third explanation for rational herding is informational externalities. In Bikhchandani *et al.* (1992) and Welch (1992), investors acquire (noisy) information by observing other agents' actions. Information externalities can be so strong that an investor can decide to ignore completely his own signal. In an extreme case of information externalities, individuals' actions do not carry information anymore as those actions result only from the imitation of others'. In that case, an informational cascade occurs.

Early studies such as Lakonishok *et al.* (1992) investigated a way to empirically measure correlated trading across groups of investors. The idea underlying their measure (LSV hereafter) is to analyze the buying pressure on a given asset for a homogeneous subgroup (pension funds, mutual funds, individual investors). For the market as a whole, each purchase is balanced by a sale. Thus, the number of buyers equals the number of sellers. However, for a given subgroup of investors and a given asset, there can be an excess of buyers or sellers, indicating that the investors composing the subgroup exhibit herding behavior.

Following the seminal work of Lakonishok *et al.* (1992), herding among investors has been the subject of several empirical studies. In particular, the mimetic behavior of U.S. mutual funds and institutional investors has been studied with scrutiny (Lakonishok *et al.*, 1992; Grinblatt *et al.*, 1995; Wermers, 1999). Similar studies have been performed outside the U.S in particular in Germany (Oehler, 1998; Frey *et al.*, 2007; Kremer et Nautz, 2011), France (Arouri *et al.*, 2010), United Kingdom (Wylie, 2005), Portugal (Loboa and Serra, 2002) and Poland (Voronkova and Bohl, 2005). The number of studies targeting individual investors is slightly lower. Such studies have been performed in Germany (Dorn *et al.,* 2008), Israel (Venezia *et al.*, 2010), China (Feng and Seasholes, 2004) and the US (Barber *et al.*, 2009).

Despite the fact that the presence of herding for individual investors has been demonstrated in some countries, no such research has been yet carried out in France. Our paper fits this loophole by investigating the daily transaction records of a sample of 87,373 French individual investors from a major European broker house over the period 1999-2006. Our contribution is then the first one on the French market and one of the most comprehensive in the European context.

Even if the LSV measure has been widely used to study the influence of stock characteristics on the investors herding behavior, it has been shown that this measure is not exempt of criticisms and suffers from drawbacks. In this paper, we choose to focus on the two main criticisms addressed to this measure. First, the recent papers of Frey *et al.* (2007) and Bellando (2010) demonstrate that the LSV measure is biased downward. They also prove that the bias declines with the number of active traders. These properties could have an impact on the empirical results, especially for empirical studies dealing with individual investors. Second, the LSV measure does not permit to evaluate the herding level of an investor and thus fails to evaluate the herding persistence over time for a given investor.

Concerning the first limit, we analyze the level of herding behavior using both LSV and Frey *et al.* (2007) measures (FHW hereafter). This double estimation leads us to build intervals containing the “true” value of herding. We then check whether asset-specific characteristics such as industry classification, market capitalization and volume of trading explain part of the herding captured by these measures. In addition, following the methodology of Barber *et al.* (2009), our results highlight a strong persistence of herding over time, whatever the measure used. Briefly speaking, the correlations of herding are 32.80% for a horizon of one month and about 10% at a 6 month horizon.

In order to overcome the second criticism of the LSV measure and to pay attention to individual characteristics, we build an original investor specific-measure of herding (called IHM, Individual Herding Measure). We propose a measure, inspired by the one of Grinblatt *et al.* (1995), that we define as the weighted sum of the signed LSV measures for the assets on which changes in holdings for the quarter under consideration are observed. This new approach leads us to investigate the differences in the herding behavior between individuals. In particular, we demonstrate that sophisticated investors tend to herd less and that bad past performances increase the propensity of investor to herd in the next quarter.

This paper is structured as follows. Section 2 is dedicated to the data. In section 3, we introduce the different measures and estimate herding at the stock level. In section 4, we describe the individual herding measure and examine factors that affect individual herding. The last section concludes.

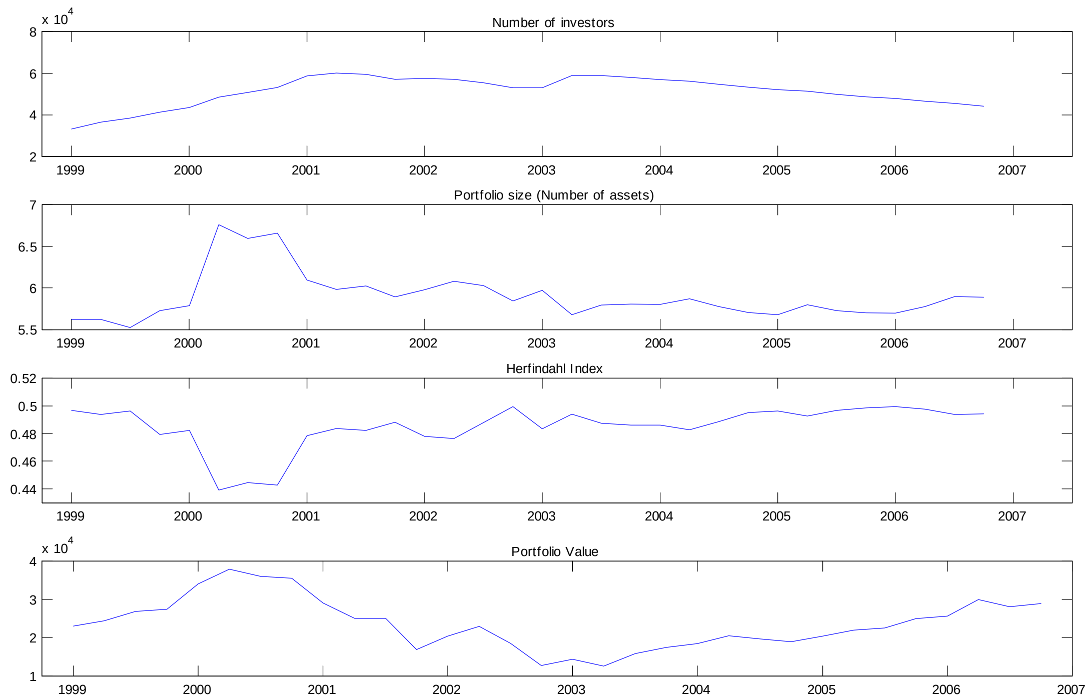
2 Data and descriptive statistics

The primary data set used in this study is a record of the daily transactions of 87,373 French investors at a major European broker house. From this record, we computed the daily stock portfolio of each investor for the period January 1999-December 2006. We are therefore able to calculate the daily realized returns. In order to do that, we extracted the closing prices (adjusted for splits and dividends) of the traded securities from Bloomberg (1180 stocks) and Eurofidai[1] (1311 stocks). A little over a thousand securities were ignored because of missing data. However, they accounted only for 1.51% of the total number of transactions. On the 2491 stocks under consideration, there are 1,190 French stocks. The remaining are from the U.S. (1,020), Great Britain (62), Canada (35), Netherlands (34), Germany (31), Italy (15) and others (104). It should be noted that the trading volume across the different countries is not homogeneous: French stocks represent more than 90% of the trading volume while U.S. stocks accounts for less than 1%.

[1] European Institute for Financial Data

In order to compute the LSV herding measure, we consider portfolios at the beginning of each quarter (January, April, July and October) for years 1999 to 2006. For a given quarter, we exclude the investors that have no investment in stocks. On average, there are 51,243 investors with at least one position. The average number of stocks held by investors is 5.9, the median 4 and a maximum of 484. The average Herfindahl index of diversification is 0.4851. The average portfolio value is 28,969 €, the median 7620 € and a maximum of 11,976,000 €. It appears that the sample of investors contains a few very wealthy individuals. Figure 1 below shows the evolution, from January 1999 to December 2006, of the number of investors, the average number of assets, the average Herfindahl Index and the average portfolio value. In order to get a deeper look on the structure of the data, we present in Table 1 the distribution of portfolio values conditioned on the number of assets held, at three points in time.

Figure 1: Summary Statistics



3 Measuring the herding behavior at the asset level

3.1 The LSV and FHW measures

One of the first herding measures was introduced by Lakonishok *et al.* (1992). This measure aimed at evaluating the herding behavior among pension funds. The underlying idea is that herding can be measured as the tendency for traders to accumulate on the same side of the market for a given stock and a given period. In order to determine on which side (buy or sell) of the market the investor is, we observe the difference between the number of shares held at time t and $t-1$[2]. We note $n_{i,j,t}$ the number of shares of asset j held by investor i at time t. If the difference $n_{i,j,t} - n_{i,j,t-1}$ is positive (respectively negative), investor i increased (decreased) her holdings and thus is on the buy (sell) side. For a given asset j, the purchase intensity $p_{j,t}$ is defined as the number of investors that increased their holdings divided by the number of investors that traded the asset. We write:

$$p_{j,t} = \frac{\sum_{i=1}^{I_{j,t}} b_{i,j,t}}{\sum_{i=1}^{I_{j,t}} (b_{i,j,t} + s_{i,j,t})} = \frac{1}{I_{j,t}} \sum_{i=1}^{I_{j,t}} b_{i,j,t},$$

where $I_{j,t}$ is the number of active traders over the period $[t-1; t]$ and $b_{i,j,t}$ ($s_{i,j,t}$) is a binary variable that takes the value 1 if the investor i increased (decreased) her holdings of asset j between $t-1$ and t, and 0 otherwise.

It follows that it is possible to compute the purchase intensity, and thus the LSV measure, only for a subgroup of investors as, for the whole universe of investors, the number of purchases equals the number of sales. Formally, the LSV herding measure of asset j at time t is written as

$$LSV_{j,t} = |p_{j,t} - p_t| - AF_{j,t},$$

[2] We stress the fact that the variations in holdings between $t-1$ and t correspond to the variations in number of shares and not in weights , as price variations would incur artificial increases and decreases. It is also important to point out that corporate actions such as splits, new issues, etc... must be taken into account.

where p_t is the purchase intensity across all stocks and $AF_{j,t}$ is an adjustment factor due to the absolute value in the definition of $LSV_{j,t}$ and the fact that the number of traders $I_{j,t}$ is varies across stocks and over time.

The term p_t is subtracted in order to account for liquidity shocks. To illustrate this point, let assume that for the majority of assets, individual investors aggregate on the buy side. This does not necessarily mean that they herd. It can be the result of a new fiscal disposition favoring investments in the stock market rather than traditional saving accounts. It results in a high buying pressure among individual investors as they withdraw their money from the saving accounts and invest it in the stock market. By subtracting p_t, we take into consideration the aggregate shifts in an out of the stock market and separate them from the herding behavior.

Under the null hypothesis of no herding, we have:

$$H_0 : p_{j,t} = p_t, \forall j.$$

Each $b_{i,j,t}$ is a Bernoulli variable with parameter p_t. It follows that the number of purchases $\sum_{i=1}^{I_{jt}} b_{i,j,t}$ is binomially distributed with probability p_t and $I_{j,t}$ independent draws. We have $p_{j,t} = p_t + \varepsilon_{j,t}$ where $\varepsilon_{j,t}$ is an iid error term with zero mean and a variance equal to $\frac{p_t(1-p_t)}{I_{j,t}}$ Under the null hypothesis H_0, the LSV measure is:

$$\begin{aligned} LSV_{j,t} &= |p_t + \varepsilon_{j,t} - p_t| - AF_{j,t} \\ &= |\varepsilon_{j,t}| - AF_{j,t}, \end{aligned}$$

In order to obtain an unbiased measure in the case of no herding, the adjustment coefficient needs to satisfy $AF_{j,t} = E\left[|\varepsilon_{j,t}|\right]$. We thus write:

$$\begin{aligned} AF_{j,t} &= E\left[\left|\frac{\sum_{i=1}^{I_{j,t}} b_{i,j,t}}{I_{j,t}} - p_t\right|\right] = \frac{1}{I_{j,t}} E\left[\left|\sum_{i=1}^{I_{j,t}} b_{i,j,t} - p_t I_{j,t}\right|\right] \\ &= \sum_{k=0}^{I_{j,t}} \binom{I_{j,t}}{k} (p_t)^k (1-p_t)^{I_{j,t}-k} \left|\frac{k}{I_{j,t}} - p_t\right|. \end{aligned}$$

As mentioned before, the LSV measure suffers from a few drawbacks and has, as such, been exposed to a certain number of criticisms. Bikhchandani and Sharma (2001) first point out that the LSV measure captures both intentional and unintentional (or spurious) herding. According to their definition, an investor is said to herd intentionally if, by observing other investors' actions, he prevents himself from making an investment he would have made otherwise (and conversely he undertakes an investment he would not have done otherwise). In other words, intentional herding corresponds to the fact of deliberately imitating others' actions. Alternatively, spurious herding occurs when investors with similar preference sets are provided with the same information. Separating these two types of herding is important as the latter is an efficient outcome whereas the former can destabilize markets and increase volatility.

A second issue discussed by Bikhchandani and Sharma (2001) is that the LSV measure considers only the number of traders and ignores the amount that is bought or sold. Oehler (1998) and Wermers (1999) propose derived measures that aim to remedy this problem. This issue has important consequences when studying the impact of herding on the market. However, as we adopt a more behavioral approach and focus on the drivers of the herding behavior, this issue does not have important consequences for our results.

Also, the LSV measure does not allow us to observe the intertemporal herding behavior of investors. We are able to follow how investors herd over time on a given asset but we cannot observe the persistence in herding of a given investor. The second part of this paper will deal with this issue by introducing an investor-specific herding measure.

Finally, two recent papers (Frey *et al.*, 2007; Bellando, 2010) show that under the alternative hypothesis of herding, the measure is biased downward. Jensen inequality implicates that the difference between $E[|x|]$ and $E[x]$ decreases with $|E[x]|$. Therefore, as the adjustment factor does not depend on the herding level, the LSV measure is biased downward and this bias increases with the herding level. They also prove that the bias declines with the number of active traders $I_{j,t}$. We will see in the empirical results that the level of herding rises when we impose a minimum number of active traders. This observation has crucial consequences for the interpretation of empirical results. For example, Dorn *et al.* (2008) establish a link between differences in opinion (proxied by trading activity) and herding behavior as they observe a very important positive correlation between trading activity and herding. It seems actually that the properties of the adjustment factor might explain part of the observed correlation. Indeed, the higher the trading activity, the lower the bias and the higher the herding measure is. Even if trading activity and herding behavior were independent, a positive correlation would appear.

In order to remedy this problem, Frey *et al.* (2007) propose to use square values instead of absolute ones in the expression of the LSV measure. Formally, their new measure is defined as:

$$FHW_{j,t}^2 = \left((p_{j,t} - p_t)^2 - E\left[(p_{j,t} - p_t)^2\right]\right) \frac{I_{j,t}}{I_{j,t} - 1},$$

where the notations are the same as in previous equations.

For a given time period t and a universe of J stocks, the average FHW measure is computed as:

$$\overline{FHW} = \sqrt{\frac{1}{J}\sum_{j=1}^{J} FHW_{j,t}^2}.$$

Monte-Carlo simulations show that this new measure does not suffer from the bias that exists for the LSV measure. Frey *et al.* (2007) show that for varying values of the number of active traders and/or of the level of herding, their measure is unbiased and possesses good statistical properties.

However, Bellando (2011) shows that the measure is unbiased only in the particular setting considered by Frey *et al.* (2007). The idea is that the purchase intensity of a given stock $p_{j,t}$ can take three realizations. With a probability $\pi_{0,t}$, we have $p_{j,t} = p_t$ and there is no herding. The stock is subject to buy-herding (respectively sell-herding) when we have $p_{j,t} > p_t$ ($p_{j,t} < p_t$) with a probability $\pi_{b,t}$ ($\pi_{s,t}$). Frey *et al.* (2007) considers the particular setting where $\pi_{b,t} = \pi_{s,t} = 0.5$. Bellando (2011) shows that as soon as the probability of no herding is not null or when some asymmetry is introduced, the measure is biased upward[3]. It follows that both measures are biased and potentially misstate the true value of herding. However, it is possible to show that the true value of herding lies between the LSV and FHW values (we refer the reader to Bellando, 2011 for a complete explanation).

3.2 General results

Table 2 provides the values of the semiannually, quarterly and monthly LSV and FHW measures for all stocks (line 1) for the whole period. These measures are then computed when we assign to each stock a level of capitalization (Large, Medium, Small), a level

[3]This bias comes from the aggregation process when computing the average measure $\overline{FHW}$.

of volume of trading (High, Medium, Low) and finally an industry classification (based on ICB industry classification). Table 3 shows the results for the two measures for each quarter between 1999 and 2006.

At a general level, the monthly average value of the LSV measure for all stocks is 12.63%. Briefly speaking, this means that for a given stock during a given month about 13% more investors are in the same side than what could be predicted if decisions were randomly taken. This result supports previous findings that individual investors herd more than institutional investors. For instance, on US market, Lakonishok *et al.* (1992) give an average value for institutional investors of 0.02 and Wermers (1999) reports a value of 0.036. More recently, Venezia *et al.* (2011) calculate an average herding measure of 0.058 for the Israeli market and Arouri *et al.* (2010) report a value of 0.065 for the French market.

Our findings indicate that French individual investors exhibit a high level of herding. Our results are in line with the findings for US individual investors (Barber *et al.*, 2009) but are slightly higher than those of Dorn *et al.* (2008) for Germany. More precisely, the monthly average value of the LSV measure for all stocks is 0.1279 in US against 0.064 in Germany. As in Dorn *et al.* (2008), the results highlight correlated trading across all horizons, all industries and the correlation is higher for longer observation intervals. Concerning the impact of the capitalization, our results confirm the findings of Dorn *et al.* (2008) and are in contrast to Barber *et al.* (2009) and to previous study of institutional investors who demonstrate that investors herd more on small firm stocks (Wermers, 1999, for example). In fact, we find that correlated trading is higher for larger capitalizations. Note that this result is not obtained for all quarters (21/31 quarters, see Table 3, Columns "Market capitalization").

Finally, the LSV measure takes a higher value for stocks ranked in the "high volume of trading" category. Even if further investigations are needed, this result could be due to a concentration of purchases in attention-grabbing stocks (Barber and Odean, 2008; Barber *et al.*, 2009) or to informational signals. Note that this result is effective for all quarters (see Table 3, Columns "Volume of trading"). Considering these findings, it is natural to wonder how the downward bias of the LSV measure (see previous section) could impact our results. Comparing the level of the two measures (Tables 2 and 3), it is apparent that the value of the FHW is sharply higher whatever the category (or the quarters) under study. At a general level, the monthly average value of the FHW measure for all stocks is 21.75%. The herding behavior is also 1.75 times stronger when this last measure is implemented. Note that this difference is stable when the observation intervals are modified (6 months or

3 months). Finally, for monthly observation intervals, the true value of herding for French individual investors is high and takes a value between 12.63% and 21.75%.

To go one step further, we conduct in the next section some tests in the spirit of Barber *et al.*, (2009) to analyze the persistence of the herding behavior over time.

3.3 Persistence in herding

In this section, we adopt another approach (following the methodology used by Barber *et al.*, 2009) in order to test whether investors' trading decisions are correlated. We also analyze the persistence, at the asset level, of the herding behavior. It is said to be persistent if the autocorrelation of the purchase intensity $p_{j,t}$ is high: A high (respectively low) level of purchase intensity at time t is followed by a high (low) level in the consecutive periods.

For each month, we part the population of investors into two equally sized random groups. We then calculate the assets monthly purchase intensity $p_{j,t}^{G_1}$ (respectively $p_{j,t}^{G_2}$) that results from the transactions of group 1 (group 2). If the investors' trading decisions are independent, we should observe no correlation between the purchases intensities $p_{j,t}^{G_1}$ and $p_{j,t}^{G_2}$. The transaction records span over 8 years resulting in a time-series of 96 contemporaneous correlations between purchase intensities. We then compute the average correlation and employ a t-test in order to check whether the average correlation is significantly different from 0. As explained by Barber *et al.* (2009), the null hypothesis of no correlation is similar to the null hypothesis of no herding in the LSV and FHW herding measures. As in the previous analysis, it is not possible to distinguish between spurious and intentional herding. The rejection of the null hypothesis only indicates that trading decisions are correlated but does not allow us to verify whether investors intentionally herd.

Once we showed that investors engage into correlated trading, we aim to see if they tend to herd on the same assets over time. A high persistence in the herding behavior would tend to indicate that it is influenced by characteristics that do not change much over time such as industry classification, index membership and market capitalization. On the contrary, a low persistence might indicate that herding is dynamic and is a direct reaction to new information, new market conditions or new fashion (for example, positive feedback trading strategy).

In order to measure the persistence, we first compute for each month the correlation between stock purchase intensities at time τ and time $t+\tau$ with $\tau = 0, ..., 36$. For $\tau = 1$, it consists in measuring each month the correlation between the purchase intensities between month t and the consecutive month. We thus obtain a time series of 95 correlations that we average to get the general persistence for a horizon equal to 1. It follows that we have a time-series of 94 correlations for $\tau = 2$, ..., and a time-series of 60 correlations for $\tau = 36$. We first compute these correlations for the whole set of investors. In a second time, we compute this persistence for two random groups of investors (in the fashion of the analysis for contemporaneous correlations which is actually the particular case where $\tau = 0$). That is, we compute the correlation between the purchase intensities obtained from the transactions of group 1 at time t, and the purchases intensities obtained from the transactions of group 2 at time $t + \tau$.

Table 4 presents contemporaneous and time-series correlations of purchase intensities. The first row ($\tau = 0$) indicates the contemporaneous correlation of purchase intensities between groups 1 and 2. We observe that the average correlation is very strong (a little over 85%), indicating that investors' trading decisions are highly correlated. Our correlation is 10 points higher than the one found by Barber *et al.* (2009). This is coherent with the fact that we also obtain slightly higher values for the LSV measure. It follows that by knowing the purchase intensities associated with one group, we are able to explain more than 2/3 of the variations in purchase intensities of the second group. The rest of the table presents the correlations between purchase intensities at time t and time $t + \tau$ where $\tau = 1, ..., 36$. The persistence between two consecutive months is expressed by an average correlation of 32.80%. The average correlations are all significantly different from zero up to a horizon of $\tau = 19$. In comparison to Barber *et al.* (2009), the correlations are slightly lower (32.80% instead of 46.7% for a horizon of one month) and the persistence fades away at a fastest rate (the correlation at a 6 months horizon is 9.46% in our study compared to 16.4% in Barber *et al.*, 2009).

4 Measuring herding at the investor level

4.1 The Investor Herding Measure (IHM)

One of the drawbacks of the LSV measure is that it is not possible to compute an investor specific measure. Thus, we cannot determine whether only a part of investors herd, whether some investor-specific characteristics influence the herding behavior or to observe the persistence in herding of investors. In order to analyze the tendency of individual investors to herd, we first need to discriminate between buy herding ($p_{j,t} > p_t$) and sell herding ($p_{j,t} < p_t$). Following Grinblatt *et al.* (1995) and Wermers (1999), we consider the signed herding measure defined by:

$$\begin{aligned} SLSV_{j,t} &= \begin{cases} LSV_{j,t} \mid p_{j,t} > p_t \\ -LSV_{j,t} \mid p_{j,t} < p_t \end{cases} \\ &= \begin{cases} p_{j,t} - p_t - AF_{j,t} \\ p_{j,t} - p_t + AF_{j,t} \end{cases} \end{aligned}$$

Grinblatt *et al.* (1995) introduced the Fund Herding Measure (FHM) defined as:

$$FHM_{i,t} = \sum_{j=1}^{J} (\omega_{i,j,t} - \omega_{i,j,t-1})\, SLSV_{j,t}$$

where $\omega_{i,j,t}$ is the weight of asset j in the portfolio of the i^{-th} fund at time t.

This measure is quite appealing but poses the problem of whether an investor can be seen as herding on an asset on which he does not trade. Indeed, a transaction on one asset only causes the weights of all the other assets in the portfolio to change.

We propose to introduce a new measure, the Investor Herding Measure (IHM) that considers herding only on the assets that are actually traded by the investor. For a given transaction, there are six possible scenarios represented below:

	Purchase	**Sale**
SLSV > 0	Herding	Anti-Herding
SLSV < 0	Anti-Herding	Herding
SLSV $= 0$	No Herding	No Herding

If an investor trades only one asset, her herding value will be equal to the signed LSV measure of the asset if the transaction is a purchase (the investor increases her holdings in this asset) and to minus the signed LSV measure otherwise. When trading several assets, computing the individual herding value is less obvious. A first approach would be to sum the signed herding measures of every asset purchased and to subtract the ones of the assets that were sold. This solution has two drawbacks. First, it does not consider the size of the transactions. Second, the measure so-built is not bounded (in the sense that it is not independent with the number of assets traded by the investor). The first drawback has as a consequence that zero-herding assets are not taken into consideration (a situation where an investor increased substantially her holdings on zero-herding assets and only slightly on a high buy-herding asset will result in a high individual herding measure). To illustrate the second drawback, let us consider an investor which makes only one purchase on asset 0 with a signed herding measure equal to $SLSV_0 > 0$. As stated above, her individual herding measure will be equal to $SLSV_0$ Now, let consider another investor who purchases assets $1, ..., n$ with equal herding measures $SLSV_1, ..., SLSV_n = SLSV_0$. The second investor will achieve an individual herding measure of $n \times SLSV_0$, that is, n-times the herding measure of the first investor. We adopt a solution that resolves both of these problems: the herding value of an asset is weighted by the size of its transaction and the sum of the weighted herding measure is then divided by the total sum of the transactions of the investor over the period. Formally, we write:

$$IHM_{i,t} = \frac{\sum_{j=1}^{J} \left(n_{i,j,t} - n_{i,j,t-1}\right) \overline{P}_{j,t} SLSV_{jt}}{\sum_{j=1}^{J} \left|n_{i,j,t} - n_{i,j,t-1}\right| \overline{P}_{j,t}},$$

where $n_{i,j,t}$ is the number (adjusted for corporate actions) of shares of asset j held by investor i at time t, $\overline{P}_{j,t}$ is the average price of asset j over the period $[t-1;t]$. It follows that $\sum_{j=1}^{J} \left(n_{i,j,t} - n_{i,j,t-1}\right) \overline{P}_{j,t}$ is the average value of asset j transaction and the denominator in the formula is the total value of all the transactions[4] made by investor i in the considered period.

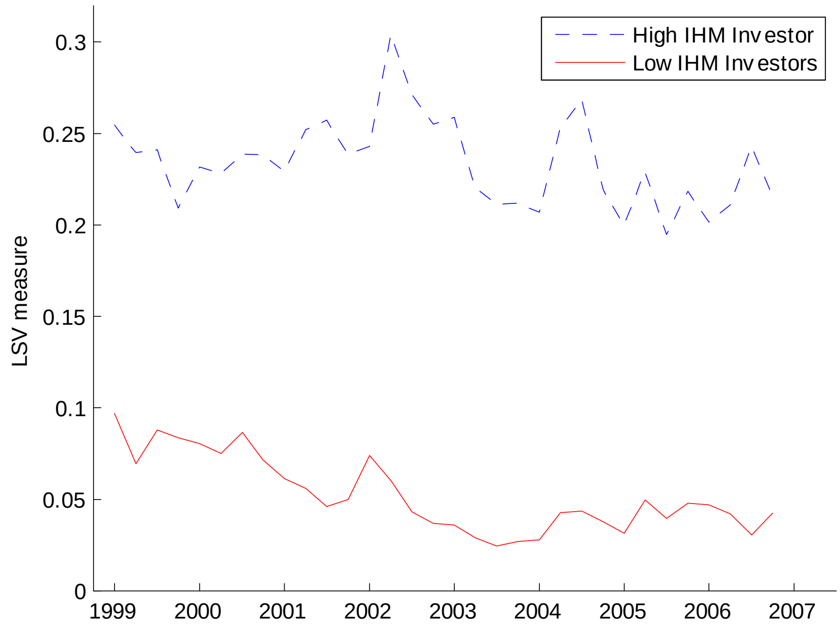
In this way, we account for the herding coefficient of assets only on the ones that are traded during the quarter and we weight them by the size (euros-volume) of the transactions. The IHM measure indicates that investor i is herding if it takes a positive value and that he goes against the herd if the value is negative.

A first confirmation of the validity of such a measure is to separate the population of investors into two equally sized subgroups: low and high IHM investors. We then compute the standard LSV measure for both subgroups. We observe in Figure 2 that the difference between the two subgroups is quite important and highly significant[5] which seems to support the validity of our measure to evaluate herding at the individual level.

[4] We only observe the number of shares at time t and $t-1$ but not the sequence of transactions during the period under study. Hence, we chose to use the average price to evaluate the value by which the investor increased or decreased her holdings.

[5] Significance tests were done using Monte-Carlo simulations. Results are not reported here.

Figure 2: LSV measure for high IHM investors and low IHM investors

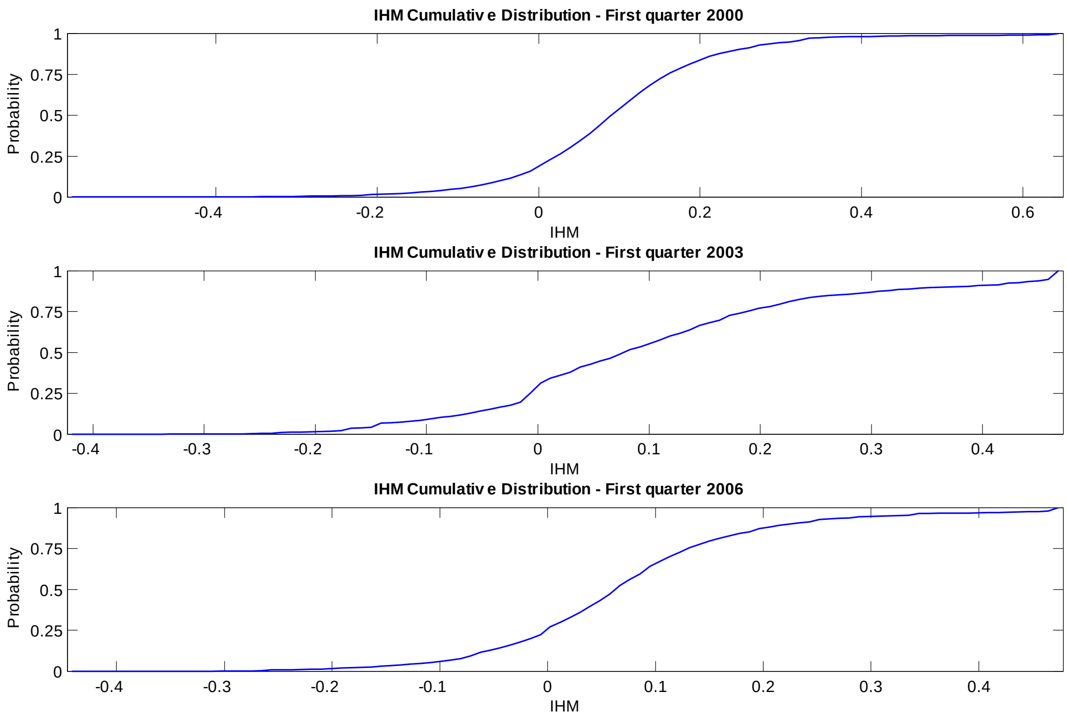


4.2 General results and persistence

We first give a brief overview of the computed IHM values. Figure 3 gives the distribution of IHM at three time points (first quarter of 2000, 2003 and 2006). Not surprisingly, we observe that most individuals have a positive IHM value. The average IHM value is equal to 0.1010 for the first quarter of 2000, 0.1055 for the first quarter of 2003 and 0.0768 for the first quarter of 2006. The medians are respectively 0.0960, 0.0838 and 0.0668. In the first part of the article, we showed that LSV and FHW values were much higher in the beginning of the sample period. The computed IHM values are coherent with these first results.

Using the same methodology than the one employed to measure the persistence at the asset level, we check if there is a significant autocorrelation in the investor herding behavior. That is, we verify if a high herding (anti-herding) behavior at a quarter t is followed by high herding (anti-herding) in the following quarters. The presence of a strong autocorrelation would tend to indicate that some investors are more prone to herd, regardless of the time-period considered. The results in Table 5 give an average correlation of 12.51% between the IHM values of two consecutive quarters. The correlations appear to

Figure 3: Empirical cumulative distribution of IHM



be significant for a horizon up to four years with a minimum of 4.60%. It follows that the herding behavior shows some signs of persistence. However, this persistence is relatively weak and these results call for a deeper investigation of the components of the individual herding behavior.

4.3 Performance, investor attributes and Investor Herding Measure

In this section, we focus on whether the investor's profile determinates part of the observed herding behavior. The baseline assumption is that some investors might be more prone to herd than others (regardless of market conditions or other time-varying variables). We test different characteristics such as the gender, the sophistication and the wealthiness of individuals. Gender differences in investment behavior are now well-documented. For instance, Barber and Odean (2001) investigate overconfidence by using a "gender approach" and show that men are more overconfident than women, leading them to trade 45% more than women. This behavior has the consequence to hurt their portfolio performance and to reduce their net returns. It follows that it is a natural choice to test whether the herding intensity differs from women to men. Our second hypothesis is that more sophisticated investors herd, in average, less. A number of researchers have documented the role played by sophistication on trading behavior. For instance, the individual differences in the disposition effect - which describes the tendency of investors to more readily sell winners stocks than losers - are significantly attributed to financial sophistication (Feng and Seasholes, 2005; Dhar and Zhu, 2006). As sophisticated investors have a better ability to obtain and to treat information (or at least they have the impression they do), the need to rely on others' information is less pronounced. The sophistication is proxied by three variables. First, an investor is seen as being more sophisticated if he trades warrants in addition to common stocks. Second, the degree of sophistication increases with the total number of transactions the investor made over the sample period. Third, the wealthier the investor, the more sophisticated he is. The wealthiness of the investor is proxied by his average portfolio value. Of course, this proxy is imperfect and reflects only partially the real wealth (or the individual income) of the investor. Indeed, positions such as bonds, cash and derivatives are ignored. The three last measures are valid under the assumption that investors' attributes are stable over time. From a methodological point of view, the fact that we use data from $t + \tau$ in order to discriminate investors at time t can seem startling or even wrong. However, under the assumption that characteristics do not change radically over the sample period, these measures give us proxies that do not vary in time and carry little noise. A change of behavior (for exogenous reasons) in one unique quarter will have a low impact on the measures we use to capture the investors profile.

The results are presented in Table 6. For the gender attribute, we report the average IHM values of male and female investors. The warrant characteristic discriminates investors between the ones that trade warrants and the ones that do not. For the average portfolio value, the first subgroup contains investors with an average portfolio value below 5000 €. The second subgroup is formed with the ones whose value is above 100,000 €. For the last characteristic, we distinguished between those who accomplished less than 100 trades and the ones that did more than 200 transactions. The reported p-values are associated with the test of no difference between the average IHM of the two subgroups for one given characteristic. As the theoretical distribution is unknown, we turned to Monte-Carlo simulations in order to estimate the empirical distribution. For a given characteristic and a given quarter, we have the average IHM values of the two subgroups $\overline{IHM_1}$ and $\overline{IHM_2}$. $\overline{IHM_1}$ (respectively $\overline{IHM_2}$) is the average of the n_1 (n_2) IHM values of the investors that belong to the first (second) subgroup. We part randomly the population of investors into two subgroups of size n_1 and n_2. We then take the average IHM of each subgroup and we compute the absolute value of the difference. This last step is then repeated 1000 times in order to obtain the empirical distribution of the difference.

The quarterly results are given in Table 6. It appears that, in average, men herd more than women. The average IHM value for men is 0.0789 compared to a value of 0.0783 for women. However, the reported p-values indicate that, for most quarters, the difference is not significant. The results on sophistication reveal that investors who trade warrants have, in average, a lower herding intensity than investors who do not. Individuals with a low number of transactions exhibit a much higher herding behavior than investors who trade a lot. For both sophistication attributes, the differences are highly significant. In particular, when considering the number of transactions, we observe a very high magnitude (up to 4 points) of the difference between the two subgroups' average IHM values. The average IHM value for the subgroup associated to a low number of transactions is 0.0854 whereas the value for the subgroup associated with a high number of transactions is only 0.0654. Finally, we observe differences between the two subgroups when discriminating by the portfolio average value. Although, these differences are significant for most quarters, their sign varies over the different quarters and prevents us from drawing any clear conclusion.

To go further in our analysis, we want to evaluate the influence of investors' past performance on herding and conversely the influence of herding on investors' subsequent returns.

In our first analysis, we compute, for each quarter, the Spearman rank correlation between investor's IHM and the first three moments of investors' portfolio past returns. The results in Table 7 indicate that there exists a strong rank correlation between past average returns and investors' herding (all but five coefficients are significant at a 5% level). However, the sign of these coefficients varies over time without any clear pattern. The coefficients for the Spearman correlation between IHM and portfolio's standard deviation are all significant and negative. This means that the less risky investors are the ones that herd the most. The results for skewness[6] are less clear as only 75% of the coefficients are significant at a 5% level and the sign changes over time. The results for the correlation between IHM and subsequent returns are equivocal though we still observe a clear negative relation between herding and standard deviation.

So far, we are not able to determine precisely how the investors' own past performance influence their herding behavior. However, it appears clearly that there exists a relationship. We now wish to exploit both the cross-section and time dimensions of our database.

[6] Mitton and Vorkink (2007) show that individual investors have heterogeneous preference for skewness. This heterogeneity helps explaining why individual investors are underdiversified.

For each quarter, we compute the investors' IHM value, past performance, level of diversification, and portfolio value. We then have an unbalanced panel data[7]. We aim at testing the influence of past performances which vary across individuals and over time. We thus run a panel data regression. The results of the Hausman test reject the null hypothesis of random effects. We therefore choose to include both investor and time fixed effects. We estimate the past performances by using the normed past return, that is, the return of the portfolio divided by its standard deviation. The formulation of the regression is the following:

$$\begin{aligned} IHM_{i,t} &= \gamma_0 IHM_{i,t-1} + \gamma_1 IHM_{i,t-2} + \sum_{\tau=0}^{3} \beta_\tau \frac{\mu_{i,t-\tau}}{\sigma_{i,t-\tau}} \\ &\quad + \theta DIV_{i,t} + \lambda PV_{i,t} + \alpha_1 IFE_i + \alpha_2 TFE_t + \varepsilon_{it}, \end{aligned}$$

where $IHM_{i,t}$ is the herding value of investor i in quarter t, $\frac{\mu_{i,t-\tau}}{\sigma_{i,t-\tau}}$ is the performance of investor i over the quarter preceding quarter $t-\tau$. $DIV_{i,t}$ and $PV_{i,t}$ are respectively the Herfindahl Index and the Portfolio value of investor i at the beginning of quarter t. IFE_i are the individual fixed effects and TFE_t are the time fixed effects.

We do not add any lag for the diversification DIV and the portfolio value PV because individual investors in our sample do not trade much. By incorporating several lags, we would include multicollinearity in the regression. Also, we include only two lags for IHM as more would reduce too dramatically the size of our sample. The results (IFE and TFE not reported) are presented in Table 8. The lags of the herding measure appear to be significant and negatively correlated with the herding measure. The estimates of the coefficients are respectively -0.068270 and -0.032374. The coefficients for the performance over the preceding quarter and the quarter before that take the negative values -0.000356 and -0.000873 and are significant. Further lags are not significant. It confirms our hypothesis that bad past performance gives incentives to herd. Also, we note that, while the portfolio value is not significant, the diversification - measured by the Herfindahl Index (which takes a value of 1 if the investor holds only one asset and tends toward 0 when the investor is perfectly diversified) - is significant and appears to be positive.

[7] The panel is unbalanced because investors are excluded from the quarters where they do not trade.

5 Conclusion

Most studies focus on stock characteristics to explain the individual or institutional investors herding behavior. Despite important drawbacks, these results are generally based on the implementation of the well-known LSV measure proposed by Lakonishok *et al.* (1992). In this paper, dedicated to the herding behavior of 87,373 French individual investors, we extend the existing literature in two original ways.

First, at an asset level, the herding behavior is analyzed using both the traditional LSV measure and the more recent Frey *et al.* (2007) measure (FHW measure). Our results show that French individual investors are prone to herding behavior and that the level of herding is sharply stronger when the FHW measure is implemented. Moreover, this behavior exhibits a strong persistence over time.

Second, we introduce an original individual herding measure. This new approach allows us to track the herding persistence of a given agent and to highlight the role played by individuals' attributes. More precisely, based on this new methodology, we demonstrate that the level of individual herding depends on the investor sophistication degree (trading derivative assets, for instance). Furthermore, based on a dynamic panel data analysis, we establish a link between investors' portfolio performance and herding behavior.

6 References

Admati A. and P. Pfleiderer, (1988), "A Theory of Intraday Patterns: Volume and Price Variability", *Review of Financial Studies 1(1)*, 3-40.

Arouri M., Bellando R. Ringuedé S. and A.-G. Vaubourg, (2010), "Herding by Institutional Investors: Empirical Evidence from French Mutual Funds", *HAL Working paper*, 26 pages.

Barber B. and T. Odean, (2001), "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment", *Quarterly Journal of Economics 116(1)*, 261-292.

Barber B. and T. Odean, (2008), "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors", *Review of Financial Studies 21(2)*, 785-818.

Barber B., Odean T. and N. Zhu, (2009), "Systematic Noise", *Journal of Financial Markets 12(4)*, 547-569.

Bellando R., (2010), "Measuring Herding Intensity: A Hard Task", *HAL Working Paper*, 26 pages.

Bikhchandani S, Hirshleifer D. and I. Welch, (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy 100(5)*, 992-1026.

Bikhchandani S. and S. Sharma, (2001), "Herd Behaviour in Financial Markets", *IMF Staff Papers 47(3)*, 279-310.

Dhar R. and N. Zhu., (2006), "Up Close and Personal: Investor Sophistication and the Disposition Effect", *Management Science 52(5),* 726-740.

Devenow A. and I. Welch, (1996), "Rational Herding in Financial Economics", *European Economic Review 40(3-5)*, 603-615.

Dorn D., Huberman G. and P. Sengmueller, (2008), "Correlated Trading and Returns", *Journal of Finance 63(2)*, 885-920.

Dow J., (2004), "Is Liquidity Self-Fulfilling?", *Journal of Business 77(4)*, 895-908.

Feng L. and M. Seasholes, (2004), "Correlated Trading and Location", *Journal of Finance 59(5)*, 2117-2144.

Feng, L. and M. Seasholes, (2005), “Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets?”, *Review of Finance 9(3)*, 305-351.

Frey S., Herbst P. and A. Walter, (2007), “Measuring Mutual Fund Herding - A Structural Approach”, *SSRN Working Paper*, 46 pages.

Graham J., (1999), “Herding among Investment Newsletters: Theory and Evidence”, *Journal of Finance 54(1)*, 237-268.

Grinblatt M., Titman S. and R. Wermers, (1995), “Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior”, *American Economic Review 85(5)*, 1088-1105.

Kremer S. and D. Nautz, (2011), “Short-Term Herding of Institutional Traders: New Evidence from the German Stock Market”, *SFB 649 Discussion Paper 2011-015*, 29 pages.

Lakonishok J., Shleifer A. and R. Vishny, (1992), “The Impact of Institutional Trading on Stock Prices”, *Journal of Financial Economics 32(1)*, 23-43.

Loboa J. and A. Serra, (2007), “Herding Behavior Evidence from Portuguese Mutual Funds” in *Diversification and Portfolio Management of Mutual Funds, Greg N. Gregoriou, ed.,pp. 167-197, Palgrave MacMillan, New York 2007.*

Mitton T. and K. Vorkink, (2007), “Equilibrium Underdiversification and the Preference for Skewness”, *Review of Financial Studies 20(4)*, 1255-1288.

Ohler A., (1998), “Do Mutual Funds Specializing in German Stocks Herd?”, *Finanzmarkt and Portfolio Management 12*, 452-465.

Rajan R., (1994), “Why Credit Policies Fluctuate: A Theory and some Evidence”, *Quaterly Journal of Economics 436*, 399-442.

Scharfstein and Stein, (1990), “Herd Behavior and Investment”, *American Economic Review 80(3)*, 465-479.

Venezia I., Nashikkar A. and Z. Shapira, (2010), “Firm Specific and Macro Herding by Professional and Amateur Investors and their Effects on Market Volatility”, *Journal of Banking & Finance 35(7)*, 1599-1609.

Voronkova S. and M. Bohl, (2005), “Institutional Traders’ Behavior in an Emerging Stock Market: Empirical Evidence on Polish Pension Fund Investors”, *Journal of Business Finance & Accounting 32(7-8)*, 1537-1560.

Welch I., (1992), "Sequential Sales, Learning, and Cascades", *Journal of Finance 47(2)*, 695-732.

Wermers R., (1999), "Mutual Fund Herding and the Impact on Stock Prices", *Journal of Finance 54(2)*, 581-622.

Wylie S., (2005), "Fund Manager Herding: A Test of the Accuracy of Empirical Results using U.K. Data", *Journal of Business 78(1)*, 381-403.

Table 1: Descriptive Statistics

This table presents descriptive statistics of the individual investor portfolios in the dataset. The dataset consists in the transaction records of 87,373 investors at a major European broker. Investor portfolios are sorted with respect to the number of stocks held in the portfolio.

		Portfolio Value (€)			
Portfolio Size	***Nb. of Observations***	***Mean***	25^{th} ***percentile***	***Median***	75^{th} ***percentile***
		Panel A: Portfolios as of January 2000			
1	9020	6736	720	1607	3844
2	6798	9576	1993	3674	7467
3	5324	14965	3556	6034	11807
4	4069	19374	5175	8696	16267
5	3151	24115	7101	11860	20641
6-9	7663	40330	10971	18285	33944
10+	7654	101678	26851	47241	89308
All	43679	33143	3124	9105	25682
		Panel B: Portfolios as of January 2003			
1	11353	2291	235	533	1414
2	7929	3975	738	1506	3306
3	6057	6895	1388	2655	5621
4	4814	7962	2160	3911	7955
5	3745	10724	3099	5458	10404
6-9	9302	17651	5173	9067	17034
10+	9977	46717	13957	25598	48371
All	53177	15196	1229	4248	13252
		Panel C: Portfolios as of January 2006			
1	11349	4279	389	1007	2515
2	7432	7982	1258	2802	6361
3	5475	11000	2488	4877	10385
4	4141	16210	3805	7221	14646
5	3346	21329	5307	9675	19493
6-9	8079	31705	8983	16395	31654
10+	8067	84915	24226	45011	88592
All	47889	26055	1926	6879	21960

Table 2: LSV and FHW measures

LSV measure for stock j in period t is computed as $LSV_{j,t} = |p_{j,t} - p_t| - AF_{j,t}$, where $p_{j,t}$ is the purchase intensity for the stock j, p_t is the purchase intensity across all stocks and $AF_{j,t}$ is an adjustment factor. With the same notations, FHW measure is computed as, $FHW_{j,t}^2 = \left((p_{j,t} - p_t)^2 - E\left[(p_{j,t} - p_t)^2 \right] \right) \frac{I_{j,t}}{I_{j,t} - 1}$, where $I_{j,t}$ is the number of active traders. We consider a minimum number of 10 active traders per stock. Stocks with less than 10 active traders in period t are excluded from the analysis for this period. Average semiannually, quarterly and monthly LSV and FHW measures are calulated for all stocks over the period 1999-2006. The LSV and FHW measures are calculated for 3 levels of stock capitalization ("Market capitalization"). Large (small) capitalizations correspond to the 30 % top (bottom) capitalizations. The medium category contains the remaining observations. The LSV and FHW measures are computed for 3 levels of trading volume in euros ("Volume of trading"). High trading volume (low trading volume) corresponds to the 30 % top (bottom) volume. The medium category contains the remaining observations. The herding measures of the different industries ("Industry") are the average herding measures of stocks that belong to the industry (using the Industry Classification Benchmark, ICB).

	Semiannualy		Quaterly		Monthly	
	LSV	FHW	LSV	FHW	LSV	FHW
All stocks	13.92	22.97	13.09	22.00	12.63	21.75
Market capitalization						
Large capitalization	14.96	22.13	14.09	21.23	13.67	21.07
Medium capitalization	12.30	20.98	11.63	20.32	11.37	20.42
Small capitalization	12.51	22.56	12.24	22.15	12.18	22.32
Volume of trading						
High volume of trading	14.81	21.54	13.87	20.57	13.45	20.49
Medium volume of trading	11.89	20.31	11.37	19.89	11.04	19.86
Low volume of trading	13.24	23.89	12.81	23.28	12.84	23.50
Industry						
Oil & Gas	12.88	20.23	12.56	19.13	12.43	19.46
Basic Materials	14.29	23.39	13.48	22.33	13.66	22.73
Industrials	13.81	23.03	12.83	21.89	12.30	21.43
Consumer Goods	13.80	22.81	13.14	22.25	12.90	22.01
Health Care	13.02	21.81	11.89	20.55	11.81	20.79
Consumer Services	13.95	22.71	13.32	21.87	12.89	21.63
Telecommunications	18.44	27.68	16.73	25.24	13.94	21.62
Utilities	15.65	22.79	14.28	20.45	12.89	18.86
Financials	15.25	24.55	14.12	23.12	13.48	22.70
Technology	13.16	21.73	12.53	21.11	12.04	20.96

Table 3: LSV measure for stock j in period t is computed as $LSV_{j,t} = |p_{j,t} - p_t| - AF_{j,t}$, where $p_{j,t}$ is the purchase intensity for the stock j, p_t is the purchase intensity across all stocks and $AF_{j,t}$ is an adjustment factor. With the same notations, FHW measure is computed as, $FHW^2_{j,t} = \left((p_{j,t} - p_t)^2 - E\left[(p_{j,t} - p_t)^2\right]\right) \frac{I_{j,t}}{I_{j,t-1}}$, where $I_{j,t}$ is the number of active traders. We consider a minimum number of 10 active traders per stock. Stocks with less than 10 active traders in period t are excluded of the analysis for this period. Quarterly measures are calulated for all stocks over the period 1990-2006. The LSV and FHW measures are calculated for 3 levels of stock capitalization ("Market capitalization"). Large (small) capitalizations correspond to the 30 % top (bottom) capitalizations. The medium category contains the remaining observations. The LSV and FHW measures are computed for 3 levels of trading volume in euros ("Volume of trading"). High trading volume (low trading volume) corresponds to the 30 % top (bottom) volume. The medium category contains the remaining observations.

		All stocks		Market Capitalization						Volume of trading					
				Small		Medium		Large		Low		Medium		High	
		LSV	FHW	LSV	FHW	LSV	FHW	LSV	FHW	LSV	FHW	LSV	FHW	LSV	FHW
1999	Q1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Q2	14.16	22.89	12.81	22.54	12.01	20.03	13.55	19.65	12.81	22.67	11.91	19.82	13.74	19.75
	Q3	13.33	21.94	9.04	18.61	13.02	22.03	16.42	23.34	9.12	18.86	13.08	22.18	16.90	23.55
	Q4	13.34	21.81	8.63	17.13	11.53	20.12	15.87	21.99	9.15	18.76	10.87	18.95	16.35	22.07
2000	Q1	15.74	24.78	12.71	22.32	16.72	25.91	17.00	25.69	13.92	24.62	15.95	24.91	16.94	25.01
	Q2	14.60	23.93	10.32	20.84	12.25	20.57	17.43	24.39	12.40	23.54	10.97	18.99	17.13	23.60
	Q3	15.15	24.44	11.70	22.60	14.05	23.26	17.75	25.38	11.92	23.39	14.19	23.43	17.27	24.32
	Q4	14.08	22.84	9.46	19.16	12.08	20.78	18.83	25.45	10.65	21.05	10.98	18.92	19.01	25.79
2001	Q1	13.35	22.01	9.80	20.04	11.40	19.62	17.33	24.13	9.65	20.17	11.63	19.85	17.18	23.71
	Q2	13.89	22.84	11.95	21.65	12.89	22.06	13.99	21.22	12.23	22.83	12.83	21.54	13.77	20.59
	Q3	13.61	22.41	15.28	25.42	11.84	20.76	13.76	20.73	14.73	25.37	11.79	20.28	14.32	21.36
	Q4	13.03	22.09	13.26	22.77	10.91	19.42	15.01	23.46	14.12	24.11	10.73	19.43	14.46	22.11
2002	Q1	13.71	23.19	11.48	22.04	11.83	21.50	17.38	24.91	12.54	23.91	11.14	20.30	17.16	24.46
	Q2	16.10	25.02	18.11	28.47	13.44	22.44	17.59	24.64	18.08	29.12	13.49	22.32	17.40	23.87
	Q3	14.83	24.03	20.92	30.03	11.71	21.08	13.42	20.84	21.21	30.75	11.55	20.70	13.25	20.19
	Q4	13.64	23.92	18.02	28.51	10.02	19.94	14.13	24.82	18.48	30.02	10.99	21.13	12.49	21.51
2003	Q1	13.82	23.00	17.19	27.44	10.66	19.91	14.13	21.14	17.30	27.75	12.33	21.66	11.61	18.03
	Q2	11.64	20.93	13.59	23.44	11.25	21.04	10.72	18.93	14.98	25.50	11.42	21.25	9.01	15.51
	Q3	11.90	21.16	12.41	23.31	11.31	20.66	11.91	19.03	13.68	25.17	10.97	20.12	10.99	17.24
	Q4	11.48	20.39	11.16	21.42	10.50	19.50	12.18	19.51	10.88	21.38	11.50	20.51	11.34	18.54
2004	Q1	10.81	19.23	8.93	18.19	10.33	19.25	13.10	20.07	9.65	19.78	10.68	19.40	11.86	18.26
	Q2	12.36	20.99	13.72	24.22	10.46	18.46	12.30	18.89	13.44	24.34	11.19	19.33	11.68	17.93
	Q3	12.11	20.88	13.11	23.90	8.89	16.87	13.79	19.83	13.76	24.47	8.77	16.90	13.28	19.05
	Q4	12.52	21.46	14.30	24.39	11.48	19.87	10.86	18.14	13.67	24.12	11.95	20.61	10.87	17.38
2005	Q1	11.58	19.92	9.98	18.95	11.61	20.21	11.46	18.31	12.24	22.34	9.11	16.80	12.54	19.06
	Q2	12.27	21.06	11.86	21.69	11.54	19.95	11.41	18.10	12.74	23.12	9.86	17.50	12.81	19.67
	Q3	11.82	20.33	10.55	20.45	10.80	18.83	12.02	18.22	10.71	20.71	10.36	18.49	12.65	18.61
	Q4	12.37	21.26	9.67	19.81	11.06	19.22	13.25	20.51	10.16	20.73	10.37	18.55	13.68	20.42
2006	Q1	12.12	20.37	9.76	18.54	10.32	17.63	13.15	19.29	10.50	20.14	10.13	16.99	12.66	18.44
	Q2	11.97	20.86	10.10	19.83	11.78	20.18	11.40	18.44	12.15	22.77	9.83	17.41	11.97	18.82
	Q3	12.25	20.61	9.20	18.77	12.06	19.99	14.07	21.09	9.83	19.94	11.21	19.00	14.58	21.24
	Q4	12.18	21.28	10.28	20.11	10.66	18.93	11.58	17.99	10.43	20.16	10.84	19.23	11.19	17.50

Table 4: Mean contemporaneous and time-series correlation of percentage buys by individual investors

Results are based on trades data from a major European broker house (01/1999-12/2006). For each stock in each month, we compute the proportion of all trades that are purchases. The second column of the table represents the correlations between percentage buys at time t and time t+τ where τ=1,...,24. The third column gives the correlation between the percentage buys by group 1 at time t and the percentage buys by group 2 at time t+τ. The first element of this column is the mean contemporaneous correlation across groups. T-statistics are based on the mean and standard deviation of the calculated correlations.

Horizon (τ)	*Correlation of % buys in month t with % buys in months $t+\tau$*		*t-Statistics*	
	Whole set of investors	Group 1 with group 2	Whole set of investors	Group 1 with group 2
0	100.00	85.06	n.a.	218.63***
1	32.80	33.69	23.33***	22.16***
2	18.92	19.46	16.23***	15.37***
3	15.14	14.09	13.42***	10.73***
4	11.34	11.30	10.32***	9.39***
5	11.48	10.36	11.65***	8.44***
6	9.46	9.11	8.83***	7.29***
7	6.72	6.17	6.30***	4.65***
8	5.83	5.91	6.20***	5.84***
9	4.17	4.30	4.26***	3.28***
10	3.20	1.79	3.11***	1.47
11	4.09	3.69	3.63***	2.77***
12	5.32	5.04	4.65***	3.99***
13	3.65	1.50	3.31***	1.27
14	1.99	1.60	2.04**	1.40
15	2.32	0.62	2.15**	0.55
16	1.93	1.06	1.71*	0.77
17	2.29	0.39	2.29**	0.30
18	2.26	2.08	1.97*	1.69*
19	1.99	1.58	2.06**	1.28
20	1.14	2.71	1.00	2.15**
21	0.57	−0.78	0.50	−0.59
22	1.98	1.21	1.76*	0.89
23	2.24	3.05	2.11**	2.40**
24	3.53	3.28	3.28***	2.63**
25	2.01	1.89	1.95*	1.63
26	0.26	−1.72	0.25	−1.39
27	−0.36	−0.73	−0.34	−0.56
28	−2.75	−2.79	−2.33**	−1.91*
29	−3.78	−4.52	−3.59***	−3.41***
30	−1.20	−1.52	−1.07	−1.09
31	0.65	0.62	0.58	0.44
32	0.38	−2.04	0.37	−1.25
33	−0.09	−1.76	−0.08	−1.38
34	−1.71	−2.07	−2.02**	−1.62
35	0.07	0.10	0.07	0.07
36	−0.32	1.36	−0.26	1.00

Table 5: Mean contemporaneous and time-series correlation of individual investors herding measure

Results are based on IHM values computed from trades data from a large European broker house (01/1999-12/2006). The second column of the table represents the correlations between IHM values at quarter t and quarter $t + \tau$ where τ=0,...,30. T-statistics are based on the mean and standard deviation of the calculated correlations.

Horizon (τ)	*Correlation of % buys in month t with % buys in months $t + \tau$*	*t-Statistics*
	Whole set of investors	Whole set of investors
0	100.00	n.a.
1	12.51	12.53***
2	11.63	13.76***
3	10.32	13.44***
4	10.93	12.00***
5	9.62	15.33***
6	9.30	14.92***
7	7.85	11.39***
8	7.43	10.04***
9	7.12	10.12***
10	6.81	10.36***
11	6.86	9.39***
12	5.69	7.97***
13	5.15	8.31***
14	4.60	6.51***
15	4.92	8.27***
16	5.78	5.36***

Table 6: This table reports average IHM values using various subsamples of investors. Four characteristics are considered: the gender, whether the investor trades warrants during the sample period, the total number of transactions and the average portfolio value. For each characteristic and each quarter, we compare the average IHM values of the two subsamples of investors. Reported p-values (computed with Monte-Carlo simulations) correspond to the test of no difference between the average IHM values of the two subsamples of investors.

		Gender			Warrants			Number of transactions			Average Portfolio Value		
		Male	Female	p-value	Yes	No	p-value	< 100	> 200	p-value	< 5000	> 100000	p-value
1999	Q1	0.1567	0.1636	0.0200	0.1407	0.1625	0.0000	0.1716	0.1325	0.0000	0.1844	0.1410	0.0000
	Q2	0.1259	0.1195	0.0280	0.1174	0.1268	0.0010	0.1323	0.1083	0.0000	0.1347	0.1201	0.0000
	Q3	0.1570	0.1583	0.6120	0.1403	0.1617	0.0000	0.1714	0.1309	0.0000	0.1769	0.1346	0.0000
	Q4	0.1315	0.1353	0.0750	0.1248	0.1341	0.0000	0.1373	0.1217	0.0000	0.1327	0.1432	0.0000
2000	Q1	0.1016	0.0986	0.0950	0.1047	0.1001	0.0170	0.1037	0.0985	0.0000	0.1044	0.1019	0.1530
	Q2	0.1288	0.1332	0.0420	0.1251	0.1308	0.0100	0.1299	0.1276	0.2060	0.1168	0.1474	0.0000
	Q3	0.1497	0.1558	0.0120	0.1235	0.1574	0.0000	0.1663	0.1183	0.0000	0.1520	0.1564	0.0270
	Q4	0.1257	0.1324	0.0000	0.1198	0.1287	0.0000	0.1328	0.1109	0.0000	0.1216	0.1258	0.0340
2001	Q1	0.1144	0.1211	0.0020	0.1084	0.1174	0.0000	0.1199	0.1025	0.0000	0.1002	0.1149	0.0000
	Q2	0.1070	0.1112	0.0870	0.0988	0.1100	0.0000	0.1151	0.0873	0.0000	0.0921	0.1217	0.0000
	Q3	0.0926	0.0947	0.3840	0.0902	0.0937	0.1380	0.0989	0.0760	0.0000	0.0815	0.0915	0.0000
	Q4	0.1091	0.1141	0.0590	0.1015	0.1121	0.0000	0.1187	0.0870	0.0000	0.0922	0.1082	0.0000
2002	Q1	0.1562	0.1563	0.9620	0.1459	0.1586	0.0000	0.1609	0.1400	0.0000	0.1221	0.1663	0.0000
	Q2	0.1136	0.1179	0.0940	0.1164	0.1140	0.3980	0.1151	0.1033	0.0000	0.0785	0.1209	0.0000
	Q3	0.0785	0.0812	0.2250	0.0757	0.0798	0.0750	0.0830	0.0636	0.0000	0.0619	0.0751	0.0000
	Q4	0.0963	0.1005	0.1070	0.0887	0.0990	0.0000	0.1071	0.0733	0.0000	0.0872	0.0790	0.0000
2003	Q1	0.1040	0.1117	0.0130	0.0988	0.1071	0.0080	0.1112	0.0896	0.0000	0.0738	0.1076	0.0000
	Q2	0.1188	0.1407	0.0000	0.0957	0.1294	0.0000	0.1477	0.0752	0.0000	0.1113	0.0899	0.0000
	Q3	0.0801	0.0902	0.0000	0.0713	0.0845	0.0000	0.0922	0.0634	0.0000	0.0663	0.0772	0.0000
	Q4	0.0619	0.0641	0.3330	0.0526	0.0647	0.0000	0.0692	0.0514	0.0000	0.0529	0.0567	0.0870
2004	Q1	0.0609	0.0645	0.1250	0.0529	0.0637	0.0000	0.0697	0.0484	0.0000	0.0582	0.0541	0.0540
	Q2	0.0930	0.1043	0.0000	0.0851	0.0976	0.0000	0.1044	0.0769	0.0000	0.0870	0.0834	0.1750
	Q3	0.1340	0.1419	0.0190	0.1029	0.1431	0.0000	0.1666	0.0864	0.0000	0.1523	0.0836	0.0000
	Q4	0.0972	0.0984	0.6510	0.0873	0.0999	0.0000	0.1099	0.0751	0.0000	0.0964	0.0670	0.0000
2005	Q1	0.0732	0.0791	0.0170	0.0735	0.0745	0.6360	0.0780	0.0659	0.0000	0.0561	0.0686	0.0000
	Q2	0.1128	0.1159	0.2770	0.1090	0.1145	0.0670	0.1295	0.0837	0.0000	0.1095	0.0924	0.0000
	Q3	0.1086	0.1130	0.1380	0.0929	0.1131	0.0000	0.1236	0.0849	0.0000	0.0988	0.0888	0.0000
	Q4	0.0995	0.1039	0.1050	0.0898	0.1027	0.0000	0.1139	0.0753	0.0000	0.1121	0.0787	0.0000
2006	Q1	0.0764	0.0787	0.2560	0.0682	0.0787	0.0000	0.0849	0.0591	0.0000	0.0759	0.0592	0.0000
	Q2	0.0792	0.0858	0.0080	0.0673	0.0834	0.0000	0.0970	0.0517	0.0000	0.0871	0.0611	0.0000
	Q3	0.0846	0.0859	0.6630	0.0773	0.0866	0.0010	0.0890	0.0742	0.0000	0.0629	0.0842	0.0000
	Q4	0.0789	0.0783	0.7860	0.0758	0.0794	0.1260	0.0854	0.0654	0.0000	0.0743	0.0686	0.0080

Table 7: The left part of this table (Backward) presents, for 32 quarters from January 1999 to December 2006, the coefficients (and associated p-values) of the Spearman correlation between the investors' IHM (computed for $[t : t+3]$) and, respectively, the previous quarter ($[t-3 : t]$) portfolios' average return, standard deviation and skewness. The right part of the Table (Forward) presents the same statistics but for the portfolios' subsequent returns, computed on $[t+3 : t+6]$

		Backward						Forward					
		Average Return		Standard Deviation		Skewness		Average Return		Standard Deviation		Skewness	
		Correlation (%)	*P-value* (%)	*Correlation* (%)	*P-value* (%)	*Correlation* (%)	*p-value* (%)	*Correlation* (%)	*P-value* (%)	*Correlation* (%)	*P-value* (%)	*Correlation* (%)	*P-value* (%)
2000	**Q1**	−11.79	0.00	−4.45	0.00	3.98	0.00	−12.11	0.00	−5.74	0.00	3.88	0.00
	Q2	−3.47	0.00	−21.49	0.00	14.06	0.00	7.43	0.00	−16.99	0.00	−9.93	0.00
	Q3	4.22	0.00	−11.54	0.00	−3.33	0.00	−5.29	0.00	−8.25	0.00	4.34	0.00
	Q4	1.25	3.53	−12.39	0.00	−2.08	0.04	2.21	0.01	−8.45	0.00	−3.66	0.00
2001	**Q1**	7.14	0.00	−16.51	0.00	−2.69	0.00	−5.74	0.00	−8.79	0.00	−2.80	0.00
	Q2	3.53	0.00	−13.50	0.00	−2.99	0.00	−5.38	0.00	−9.89	0.00	−8.19	0.00
	Q3	−0.09	88.51	−7.02	0.00	−1.58	1.27	−9.19	0.00	−5.67	0.00	−3.96	0.00
	Q4	−9.45	0.00	−10.41	0.00	−3.48	0.00	−5.61	0.00	−10.26	0.00	−1.62	1.47
2002	**Q1**	4.67	0.00	−8.21	0.00	0.81	22.83	2.27	0.07	−7.19	0.00	−7.70	0.00
	Q2	−0.82	19.98	−3.34	0.00	2.71	0.00	−8.16	0.00	1.11	8.43	−5.20	0.00
	Q3	−5.39	0.00	−6.89	0.00	0.20	75.17	−4.52	0.00	−4.80	0.00	0.10	87.52
	Q4	−8.49	0.00	−10.04	0.00	−1.33	4.63	−2.02	0.29	−6.17	0.00	−0.72	28.93
2003	**Q1**	2.91	0.01	−14.54	0.00	−3.00	0.00	−7.05	0.00	−10.00	0.00	−0.94	19.28
	Q2	−1.72	1.38	−27.32	0.00	3.58	0.00	−4.18	0.00	−19.25	0.00	−7.30	0.00
	Q3	−4.90	0.00	−19.21	0.00	−12.46	0.00	1.87	0.91	−16.58	0.00	2.04	0.44
	Q4	5.26	0.00	−8.50	0.00	1.27	7.53	5.03	0.00	−7.68	0.00	−0.45	53.14
2004	**Q1**	6.60	0.00	−13.17	0.00	1.08	12.18	3.09	0.00	−9.25	0.00	−6.43	0.00
	Q2	0.43	55.44	−7.49	0.00	1.84	1.15	4.75	0.00	−6.32	0.00	−1.27	8.09
	Q3	6.12	0.00	−12.01	0.00	−13.76	0.00	7.44	0.00	−14.76	0.00	−0.39	59.71
	Q4	2.56	0.04	−9.58	0.00	−6.03	0.00	2.06	0.49	−9.85	0.00	−1.32	7.33
2005	**Q1**	−0.28	69.58	−6.88	0.00	−3.03	0.00	1.32	6.81	−4.94	0.00	−2.11	0.37
	Q2	4.99	0.00	−10.13	0.00	5.18	0.00	−0.12	87.33	−11.72	0.00	0.99	17.65
	Q3	−5.54	0.00	−13.47	0.00	−4.70	0.00	−4.66	0.00	−14.31	0.00	0.84	22.90
	Q4	−4.01	0.00	−7.40	0.00	−4.54	0.00	4.93	0.00	−11.20	0.00	2.02	0.50
2006	**Q1**	0.13	84.91	−12.49	0.00	−3.58	0.00	1.45	3.41	−8.21	0.00	0.39	57.02
	Q2	−4.92	0.00	−14.12	0.00	−2.20	0.22	2.68	0.02	−14.39	0.00	−3.94	0.00
	Q3	5.52	0.00	−16.53	0.00	−5.83	0.00	8.66	0.00	−11.38	0.00	−1.88	1.85
	Q4	8.93	0.00	−10.42	0.00	−0.42	55.60	-	-	-	-	-	-

Table 8: The effect of past performance on herding

This table reports the results of the panel regression introduced in the last section of the article. The independent variable is the Investor Herding Measure (IHM) for quarter t. The explanatory variables are the lagged values of IHM (on quarter $t-1$ and $t-2$), 4 lagged values of the normed return, and,the Herfindahl index and the portfolio value at the beginning of quarter t. Investor and time fixed effects are not reported here.

Explanatory variable	Coefficients	Standard error	t-statistics	P-value
$IHM_{i,t-1}$	−0.068270	0.002030	−33.63	0.0000
$IHM_{i,t-2}$	−0.032374	0.002008	−16.13	0.0000
$\frac{\mu_{i,t}}{\sigma i,t}$	−0.000356	0.000152	−2.35	0.0188
$\frac{\mu_{i,t-1}}{\sigma i,t-1}$	−0.000873	0.000133	−6.55	0.0000
$\frac{\mu_{i,t-2}}{\sigma i,t-2}$	−0.000235	0.000132	−1.77	0.0761
$\frac{\mu_{i,t-3}}{\sigma i,t-3}$	0.000091	0.000134	0.69	0.4927
$DIV_{i,t}$	0.013749	0.002216	6.21	0.0000
$PV_{i,t}$	0.000000	0.000000	−0.44	0.6611
R^2	0.260021			
Nb. of observations	287738			