Research in International Business and Finance 62 (2022) 101740

Contents lists available at ScienceDirect

Research in International Business and Finance

journal homepage: www.elsevier.com/locate/ribaf

Analysis of herding behavior in individual investor portfolios using machine learning algorithms

Taylan Mavruk

University of Gothenburg, Centre for Finance/Department of Business Administration, Box 610, SE-405 30 Gothenburg, Sweden

ARTICLE INFO

JEL classifications:
G11
G14

Keywords:
Herding
Portfolio performance
Machine learning algorithms

ABSTRACT

This paper examines the determinants of herding at both stock and individual investor levels and studies the portfolio performance of herd vs. non-herd portfolios using machine learning algorithms. The disposition effect and the attention effect seem to explain herding behavior at the stock level. At the individual investor level, the cumulative number of buys and portfolio values reduce the prediction of herding behavior, while high values of portfolio return lead to a small increase in herding. Individuals who herd do not outperform either market or non-herd portfolios, suggesting that herding is a behavioral bias. Thus, such behavior seems to destabilize stock markets, creating temporary discrepancies in stock prices followed by reversals back to fundamentals. The most predictive factor in the performance tests of individual portfolios is the market risk premium and using equally-weighted factors rather than value-weighted factors seem to provide more consistent results in the portfolio performance analyses.

1. Introduction

This paper first examines the determinants of herding at both stock and individual investor levels. It then explores whether investors who herd outperform the market and those who do not exhibit herding behavior in their portfolios. Determinants of herding, choice of a performance model, and predictability of Carhart (1997) 4 factors are evaluated using machine learning algorithms. Investors may engage in trading behavior that could be a consequence of various behavioral biases. Herding behavior is one example suggesting that investors systematically correlate their trades with other investors' trades in a group (see e.g., Lakonishok et al., 1992; Sias, 2004). Anti-herding or contrarian behavior would suggest the opposite; investors deviate from the trade behavior of the crowd.

The explanations behind herding behavior are based on either information asymmetry or behavioral motivations such as fads, fear, greed, noise, (Shiller, 1984; Barber et al., 2009; Merli and Roger, 2013) or reputational concerns (Scharfstein and Stein, 1990; Trueman, 1994). It is more likely to find these behavioral explanations in (a limited number of) studies examining herding behavior in individual investor portfolio choice. These studies suggest that herding is irrational and has a destabilizing effect on financial markets by pushing asset prices away from their fundamental values followed by subsequent return reversals.[1] However, the information asymmetry explanation of herding behavior would suggest that such trading behavior is rational if it improves risk-adjusted returns for

E-mail address: taylan.mavruk@handels.gu.se.

[1] The potential explanations for herding documented in the previous literature are principal-agent concerns (Scharfstein and Stein, 1990), informational cascades (Feng and Seasholes, 2005), correlated information (Sias, 2004), passive responses of individual investors to institutional herding (Dorn et al., 2008), common responses to tax law changes (Barber et al., 2009), systematic changes to individual investor risk aversion (Kumar and Lee, 2006), psychological biases; representativeness heuristic (Tversky and Kahnemann, 1974), limited attention (Barber and Odean, 2008), and disposition effect (Shefrin and Statman, 1985).

https://doi.org/10.1016/j.ribaf.2022.101740
Received 4 August 2021; Received in revised form 14 June 2022; Accepted 12 August 2022
Available online 18 August 2022

investors. The empirical literature studying herding behavior in institutional investor portfolio choice often supports information-based trading for herding behavior and separates the information story into cascading and investigative herding (see e.g., Grinblatt et al., 1995; Sias, 2004).

The explanation based on informational cascades suggests that investors ignore their private information signals when they decide which stocks to trade and whether to buy or sell the stocks (Hirshleifer and Teo, 2009). This means investors favor observational learning, observing others' actions, payoffs, or conversations because they value observational learning more than their private signals. Hirshleifer and Teo (2009) relate cascades to information blockages that are an aspect of an informational externality. Individuals make a trading choice to maximize their own utility, and they disregard whether their trades provide potential information benefit to other investors. An investigative herding contrasts with the notion of cascading; once investors cascade, they would not have a reason to investigate. Sias (2004) observes that institutional investors analyze the same underlying information and draw the same conclusions on the expected value of assets and thus explain their trade behavior with investigative herding. The consequences of such information-based herding are that asset prices adjust faster to fundamental information, increasing market efficiency rather than destabilizing the markets.

This paper contributes to the above-mentioned studies in a number of ways. First, the main contribution is that the paper provides a detailed examination of the determinants and performance of herding in the portfolios of individual investors. The determinants of herding at the stock level have been widely examined (see e.g., Lakonishok et al., 1992; Sias, 2004; Hirshleifer and Teo, 2009), and the studies draw conclusions in line with the theories of information asymmetry, attention effect, and disposition effect.

The previous empirical evidence also documents herding by institutional investors across markets, asset classes, and different types of investors. The majority of these studies document evidence of herding by examining the clustering behavior of market experts such as analysts or institutional investors (see Welch, 2000; Sias, 2004; Barber et al., 2009; Choi and Sias, 2009; Hirshleifer and Teo, 2009). For instance, Sias (2004) finds that institutional investors herd in the US, follow their own lagged trades, and are momentum traders. As to the main reason why institutional investors herd, Sias (2004) suggests information-based trading. Similarly, Grinblatt et al. (1995) document herding behavior among money managers at US-based mutual funds, suggesting an information-based explanation for herding behavior. Consistent with these studies, Cipriani and Guarino (2014) show that informed traders can herd to a significant degree on a daily basis. There is also evidence that investors herd in Japan (Kim and Nofsinger, 2005), albeit to a lesser extent than in the US markets. The authors suggest that herding in Japan is most likely investigative, with a positive price impact since it speeds up price adjustments. The overall results found in the herding literature on the trade behavior of institutional investors suggest that herding helps to create faster price adjustment and greater market efficiency.

Regarding the determinants of herding at the individual investor level, a detailed examination is still scarce aside from some exceptions (see e.g., Merli and Roger, 2013), and existing literature does not provide a reduced form of regression model for herding behavior in trades of individual investors. For instance, even if Barber et al. (2009) have detailed individual-level data, their analyses of determinants of herding are at the stock level, in which many of the trade theories can be tested in the aggregate. The determinants of such trade behavior at the individual investor level are important to explore in detail, especially if the trade bias is behavioral and the aim is to eliminate that bias.

Moreover, the detailed investor-level determinants are difficult to examine at the stock level analysis, which would provide in and out capital flows at the aggregate level from certain groups. At the individual investor level, one may need to study a large number of predictors or covariates that are highly correlated, have non-linear relationships to herding, and have complex variable interactions. Thus, we may need to rely on various advanced statistical tools to minimize noise and redundant data. Machine learning (ML) techniques play a crucial role, particularly in large datasets on investor portfolios, in order to reduce the number of parameters, reduce overfitting by enhancing generalization, and to avoid the curse of dimensionality (see Chen et al., 2020). The techniques provide computational power and functional flexibility, particularly in high-dimensional datasets.

The recent developments in ML models also allow us to test theories and examine the relations among the variables for further exploration and prediction. The ML models can also help with detecting outliers and extracting features as well as performing regressions of high dimensional data by partitioning a complex dataset into subsets that may identify linear patterns. Particularly, the decision tree algorithms can identify patterns and produce plausible outcomes. To explore the benefits of ML (and Artificial Intelligence) methods, Goodell et al. (2021) review the extant literature on ML application in finance. The authors use bibliometric methodologies to examine the financial research in the area including portfolio construction, valuation, and investor behavior. Their analysis shows that the use of ML improves trading and investment decisions. The methods seem also to improve the predictive ability of traditional econometric modelling by extracting effective information from large and macro datasets (Duan et al., 2022). For instance, examining the nonlinear relationship between systemic risk and the macroeconomy in China, Duan et al. (2022) suggest that ML techniques can be used for data learning ability and exploring nonlinearity for the extraction of systemic risk information.

Using ML methods, this paper examines various features both in determinants and performance tests of herd vs. non-herd portfolios to explore which features are important for the model. The analyses present both predictive power and the wall time to run the regressions. As the main model, the study implements random forest regressions (Breiman, 2001) and thereafter uses recursive feature elimination (RFE) in the training dataset to select the top three important features in predicting the herding measures. The predictions from random forest regressions are explained with SHAP (SHapley Additive exPlanations) values, a method from coalitional game theory (Shapley, 1953), which measures the importance of a feature by comparing the random forest prediction results with and without the pertinent feature. In other words, SHAP values measure the effect of a feature by taking into account its interaction with other features (see Strumbelj and Kononenko, 2010; Lundberg and Lee, 2017). The study further implements two other machine learning methods as robustness, feature selection with gradient boosting and least absolute shrinkage and selection operator (LASSO).

The paper closely follows the study by Merli and Roger (2013), which provides an analysis of the determinants and performance of herd portfolios. The main contribution to their study is the examination of whether more detailed individual, portfolio, firm, and market-level characteristics explain herding behavior. Moreover, since herding is defined as systematic trade correlation within a group, this paper also relates other types of trade biases, such as local bias and birthplace bias, observed within a herding group by using locality as a social network. (Lindblom et al., 2018). Regardless of the type of herding, investors may compare their performance with that of their communication peers in a local community or in their birthplace and decide whether they are doing relatively better or worse. To examine such details, this study considers networks and examines the herding behavior of individuals who share a common birthplace and 5-digit residential zip code to define birthplace and local biases, respectively.

Second, this paper also examines herding behavior at the stock level by replicating the well-known herding measures used in the literature to compare their findings with those in data from Sweden from a more recent period. It applies two measures for stock level analyses, starting with the most used measure introduced by Lakonishok et al. (1992) and by Frey et al. (2014), as well as two measures for individual investor level analyses introduced by Grinblatt et al. (1995) and Merli and Roger (2013). Stock level herding measures are examined and compared in small vs. large firms, low vs. high book-to-market value firms, and firms that are headquartered in urban vs. rural areas. The analysis also covers stock level herding measures by industry.

There are a number of advantages in examining herding behavior in stock markets in Sweden. The ownership structure in Sweden is known as concentrated, and the long-term family ownership structure that uses differentiated voting rights offers fruitful research on investor behavior. Euroclear Sweden, the Swedish Central Securities Depository (CSD), offers detailed regular records on individual ownership, and the book series, “Owners and Power in Sweden’s Listed Companies,” routinely identifies family owners of firms based in Sweden. The Nasdaq OMX and the other stock markets in Sweden are the largest markets in terms of size and number of investors in Scandinavia. Sweden also has a long history of being an attractive stock market for foreign investors because of its highly sophisticated innovation activities. Taken together, these features of Sweden and ownership data have enabled researchers to make a significant contribution to research on investor behavior (see e.g., Giannetti and Simonov, 2006; Hamberg et al., 2013; Mavruk, 2021).

Finally, this paper also explores and compares the performance of individual portfolios that exhibit herding vs. non-herding against the market using widely recognized market models. On the one hand, if investors infer information from each other’s trades, we would expect that herding improves investors’ portfolio performance. This would be consistent with the notion of “wisdom of crowds” – groups of investors communicating with each other enable market prediction more so than isolated individuals (see Hirshleifer and Hong, 2003). If true, risk-averse investors would be better off frequently interacting and exchanging information with each other. On the other hand, if herding occurs because of behavioral reasons such as noise trading, we would expect negative or no abnormal performance in herd portfolios since noise traders as a group tend to lose money trading (Barber et al., 2009).

Merli and Roger (2013) provide detailed analysis on the performance of herd vs. non-herd portfolios of individual investors using market models such as CAPM and Carhart 4 factor models. Although we know about the theoretical structure of these market models, we still have some doubts whether, for example, the CAPM (one factor of market risk premium) or Fama-French 3 factor model, the latter of which adds small minus big (SMB) and high minus low (HML) factors (so-called anomalies), would explain the portfolio returns of investors in small markets such as Sweden. Additional examination could determine if Carhart’s 4th factor, Momentum (MOM), should be included in the models. There are also other factors/anomalies that have been suggested, but those are beyond this study. In addition, there are no clear guidelines on whether the portfolio of stocks generating these factors should be value-weighted or equally-weighted. It seems as if the choice of the model depends on the data (geography and time) and research question, however, some reverse engineering is also involved, meaning the researcher chooses the model based on the outcome. Such an approach leads the researcher to engage in a specification search, selecting sample criteria and test procedures until insignificant parameters become significant, a process known as *p-hacking* (see Hou et al., 2020). The second part of the paper aims to examine these questions using ML algorithms, which may also assist with comparing the predictive power and consistency of the most important features using equally-weighted and value-weighted factors. The results from ML algorithms are also compared with the results from ordinary least squares of market model regressions, as previous research shows that ordinary least square regressions applied to cross-sectional regressions of returns on anomaly variables are highly sensitive to microcap outliers (see Hou et al., 2020).

Using quarterly investor data from Euroclear and data on firm/stock characteristics from Compustat and FinBas between 2006 and 2014, the results suggest a number of predictive determinants at both the stock and individual investor levels. The disposition and attention effects seem to explain herding behavior at the stock level. At the investor level, portfolio value, portfolio return, firm size, liquidity, and cumulative number of buys have the most predictive power in explaining widely used herding measures. SHAP values indicate that high values of cumulative number of buys and portfolio values consistently reduce the prediction of herding behavior by approximately 0.2 percentage points (ppt.), while the high values of portfolio return lead to a small increase of 0.005 ppt. of the prediction. Regarding the firm characteristics, the high values of the average liquidity (-0.09 ppt.) and MTB (-0.01 ppt.) in investor portfolios are consistent in reducing the prediction of herding behavior.

The results also suggest that individuals who herd in stock markets do not outperform the market or those who do not herd, a result indicating that herding is behavioral. These results are consistent with those obtained from the stock level analysis suggesting that herding is greater in large firms, in firms with low book-to-market value, and in urban firms. If information asymmetry was the reason behind such behavior, one should have observed the opposite results. Moreover, the outcome of ML algorithms suggests that the market risk premium is the most predictive determinant of returns in herd vs. anti-herd portfolios. SHAP values consistently indicate that high values of market risk premiums increase the prediction of excess returns on portfolios by approximately 0.5 ppt. These results are more robust in models using equally weighted factors, while the SMB factor in some value-weighted factor models has more predictive power than the market risk premium. This finding indicates that using Carhart 4 factor models with equally-weighted factors would provide more robust results.

The result that herding is behavioral has implications for investor wealth and asset prices. If herding behavior is systematic, we expect stocks with more herding to be overpriced since this behavior may generate investor sentiment that leads to more net purchases (more than one would expect from chance). This creates buying pressure on these stocks. These effects can be observed even if only a small group of individual investors exhibit herding behavior, particularly if investing in illiquid stocks. This would concur with prior research showing that individual investors are net purchasers of attention-grabbing stocks because they hold fewer stocks in their portfolios and do not seem to sell short (Barber and Odean, 2008). Herding trading behavior seems to be a result of reasons other than private information that lead individuals to deviate from a pure portfolio diversification strategy. It is therefore likely that such herding behavior creates temporary discrepancies in stock prices followed by reversals back to fundamentals. The fact that the impact of individuals' herding behavior is not permanent suggests that such behavior is destabilizing for the stock market. It is not driven by fundamentals and thus cannot facilitate price discovery; rather, it creates temporary mispricing that is exploited, perhaps by institutional investors.

The detailed examination of determinants of herding also has practical implications. Knowledge about the characteristics of individuals who show herding biases will allow us to design better financial services and education programs to improve individual investment choices and alleviate such trade biases. Particularly, if the motive behind the biases is just behavioral and leads deviations of stock prices away from the fundamental value of stocks, eliminating such biases would decrease return predictability, improving financial market stability and investor welfare.

The rest of the paper proceeds as follows. Section 2 introduces the data and method, which includes measuring herding at stock and investor levels, portfolio performance measures, and model specification. Section 3 provides results from stock level herding measures, Section 4 presents results from investor level herding measures, Section 5 shows results from the portfolio performance of individual investors, and finally, Section 6 concludes the paper.

2. Data and method

2.1. Data

The proprietary investor dataset is provided by Euroclear Sweden AB and covers the quarterly holdings of all shareholders of firms listed mainly on the OMX (large, mid, and small cap) exchange. The dataset covers 36 waves of unbalanced panels, ranging from March 2006 to December 2014, and includes the number of stocks held in each firm, personal identification number of investors, five-digit postal zip code of residential address, age, gender, and investor nationality, as well as organization identification data (ISIN, and company VAT number). This dataset is matched with information on market prices and firm characteristics obtained from Compustat and the Swedish House of Finance Research Data Center, FinBas, databases.

The descriptive statistics on the investor data are presented in Table 1. Price information is available for 932 different stocks, averaging approximately 517 stocks over 36 quarters. The dataset contains 2,451,508 individual investors and their holding positions of 206,275,967 over 36 quarters. Approximately 70 % of the investors made at least one trade. The investors' number of purchases is more than the number of sells, which is consistent with findings in Barber and Odean (2008) and Barber et al. (2009). This suggests that individual investors, compared with institutional investors, are net buyers of the stocks. Additionally, consistent with their data, the mean value of buys of investors is greater than the mean value of sells (approximately 1.16 times *t value*: 144.40). These analyses support the focus of this study on purchase intensity rather than sell intensity. In addition, the mean transaction values (both buys and sells) are greater than the median transaction values, indicating the right skew in individual trades.

2.2. Measuring herding at stock level

2.2.1. Lakonishok et al. (1992), LSV, measure of herding

The study measures herding as the tendency for traders to accumulate on the same side (buy or sell) of the market for stock j in quarter t. To determine buy or sell, the difference in the number of shares held in stock j between quarters t and $t-1$ for individual i is calculated, more formally, $n_{i,j,t} - n_{i,j,t-1}$. The number of shares held is adjusted for corporate actions such as splits, reverse splits, dividends, equity issuance, and bonuses in stocks. If the difference in the number of shares held in stock j between quarters t and $t-1$ is positive (negative), then investor i increased (decreased) her holdings in stock j. The standard $LSV_{j,t}$ measure is given by Eq. (1):

Table 1
Descriptive statistics on investor data. The table provides information on investors, the number of transactions they made and the transaction values for the stocks with price information. Price information is available for 932 different stocks, averaging about 517 stocks over 36 quarters, which amounts to 18,588 stock-quarter observations.

Period	January 2006 to December 2014
Number of investors	2451,508
Number of positions	206,275,967
No. of investors who made at least one trade	1705,757
Number of buys	12,009,973
Mean (median) buy value in SEK	76,969 (21,138)
Number of sells	6823,280
Mean (median) sell value in SEK	66,172 (33,370)

$$LSV_{j,t} = |p_{j,t} - p_t| - E[|p_{j,t} - p_t|], \tag{1}$$

where p_t is the purchase intensity across all stocks to account for liquidity shocks, that is, aggregate shifts in and out of the stock market. $E\left[\left|p_{j,t} - p_t\right|\right]$ is the adjustment factor ($AF_{j,t}$), which allows measure $LSV_{j,t}$ to be unbiased in the case of no herding, the null hypothesis. The purchase intensity $p_{j,t}$ is the number of investors that increased their holdings in stock j in quarter t in relation to the number of investors that traded stock j in quarter t and is given by Eq. (2):

$$p_{j,t} = \frac{\sum_{i=1}^{I_{j,t}} b_{i,j,t}}{\sum_{i=1}^{I_{j,t}} (b_{i,j,t} + s_{i,j,t})}, \tag{2}$$

The binary variable $b_{i,j,t}$ ($s_{i,j,t}$) takes the value of 1 if investor i increased (decreased) her holdings in stock j in quarter t and it is 0 otherwise. $I_{j,t}$ is the number of active traders over quarter t. Moreover, the adjustment factor, $AF_{j,t}$, is given by Eq. (3).

$$AF_{j,t} = \sum_{k=0}^{I_{j,t}} \binom{I_{j,t}}{k} (p)^k (1 - p_t)^{I_{j,t}-k} \left| \frac{k}{I_{j,t}} - p_t \right|, \tag{3}$$

where the number of purchases, k, is binomially distributed with the probability of the purchase intensity across all stocks, p_t, and the number of active traders, $I_{j,t}$.

One alternative to this widely used standard measure of herding at the stock level, $LSV_{j,t}$, is the cross-sectional absolute deviation model ($CSAD_t$), which measures the dispersion between the stock return, $R_{i,t}$, and the return on the market, $R_{m,t}$, $CSAD_t = \frac{1}{N}\sum_{i=1}^{N} |R_{i,t} - R_{m,t}|$ (see Chang et al., 2000). There are two reasons the measure of $CSAD_t$ is not used in this paper. First, $CSAD_t$ is not a measure of herding; instead, $CSAD_t$ is in turn regressed on $R_{m,t}$, and its response rate is used to detect herd behavior in the markets (Chang et al., 2000). Second, $CSAD_t$ is often used to compare herd behavior across stock markets as the measure begins with stock level return dispersion, and then the average over the cross section of stocks in period t is taken to represent herd behavior at the stock market level. However, the herding measures chosen in this paper begin with a stock level measure, $LSV_{j,t}$, which is then used to measure herding at the individual investor level that satisfies the main aim of the paper, which is the examination of both determinants and the performance of herd behavior in individual investor portfolios.

2.2.2. Frey et al. (2014), FHW, measure of herding

The *LSV* measure has been criticized over time; for instance, Bikhchandani and Sharma (2001), Oehler (1998), and Wermers (1999) argue that the *LSV* measure may capture unintentional herding by counting the number of traders and does not account for the amount of trade bought and sold. Frey et al. (2014) show that the *LSV* measure is downward biased, as the adjustment factor used does account for the level of herding. However, Frey et al. (2014) also show that the bias declines with the number of active traders. Thus, this bias would have less impact on the study's sample, as it includes approximately 1758,693 investors on average in each period. Nevertheless, to compare the results to Frey et al. (2014), the study also computes their measure of herding. More formally, the *FHW* herding measure for stock j in quarter t is defined as in Eq. (4):

$$FHW_{j,t}^2 = \left((p_{j,t} - p_t)^2 - E\left[(p_{j,t} - p_t)^2 \right] \right) \frac{I_{j,t}}{I_{j,t-1}}, \tag{4}$$

where the notations are the same as in Eq. (1) but in squared terms. The average *FHW* measure in quarter t can be computed as in Eq. (5):

$$\overline{FHW}_t = \sqrt{\frac{1}{J} \sum_{j=1}^{J} FHW_{j,t}^2}. \tag{5}$$

The above measure should theoretically provide a higher value than the *LSV* measure, and Bellando (2010) and Merli and Roger (2013) argue that the two measures together provide lower and upper bounds of the true value of herding.

2.3. Measuring herding at investor level

2.3.1. Active investor herding measure

Using the *LSV* measure as a starting point, similar to the model in Merli and Roger (2013), first investor level herding is measured by discriminating sell herding ($p_{j,t} < p_t$) and buy herding ($p_{j,t} > p_t$). The signed herding (*SLSV*) follows Grinblatt et al. (1995), and Wermers (1999):

$$SLSV_{j,t} = \left\{ \begin{array}{c} LSV_{j,t} | p_{j,t} > p_t | \\ -LSV_{j,t} | p_{j,t} < p_t | \end{array} \right\}, \tag{6}$$

the investor level herding (*IHM*) is then given by Eq. (7):

$$IHM_{i,t} = \frac{\sum_{j=1}^{J}\left(n_{i,j,t} - n_{i,j,t-1}\right)\overline{P}_{j,t}SLSV_{j,t}}{\sum_{j=1}^{J}\left|n_{i,j,t} - n_{i,j,t-1}\right|\overline{P}_{j,t}}, \tag{7}$$

where $\overline{P}_{j,t}$ is the average price of the stock between quarter t and $t-1$, $n_{i,j,t} - n_{i,j,t-1}$ is the change in the number of adjusted shares between quarter t and $t-1$, and $\left(n_{i,j,t} - n_{i,j,t-1}\right)\overline{P}_{j,t}$ is the average value of the transaction made between quarter t and $t-1$. The measure is defined only for *active trades* that are non-zero change in adjusted share. The average price of the stock during the quarter is used because the data do not allow one to observe the exact day of purchase or sales within a quarter. Furthermore, $SLSV_{j,t}$ is the signed *LSV* herding measure as described in Eq. (6) and the denominator in Eq. (7) is the total value of all transactions made by individual i between quarters t and $t-1$. This measure is positive if investor i is herding and negative if the investor does not show herding behavior.

2.3.2. Investor herding measure including passive positions

The second investor level herding measure follows Grinblatt et al. (1995), who compute the herding behavior of fund managers. This study applies this measure to investor data and names it, *GTW*, more formally:

$$GTW_{i,t} = \sum_{j=1}^{J}\left(w_{i,j,t} - w_{i,j,t-1}\right)SLSV_{j,t}, \tag{8}$$

where $w_{i,j,t}$ is the weight of stock j in investor i's portfolio in quarter t and $SLSV_{j,t}$ is the signed *LSV* as described in Eq. (6). One potential caveat with this measure is that a price variation in a stock may cause a change in weight of the stock; thus, investors might be classified as herding in stock, but they may not be actively trading. Additionally, the measure is defined only for non-missing (since the equation includes lagged weight) and non-zero values of change in weight. If the individual includes a new stock in her portfolio in period $t+1$, the change in weight for that stock will have a missing observation in the period, which is also true for the *IHM* measure.

2.4. Determinants of herding

2.4.1. Determinants of herding at stock level

For the stock level determinants, first, the mean contemporaneous and time-series correlation of percentage buys by all individual investors across all stocks are analyzed. Following Barber et al. (2009), the data are divided into randomly chosen equal groups of investors. For each stock in each quarter, the percentage of all purchase trades is calculated. Then, the mean contemporaneous correlation across groups and the mean temporal correlation from one quarter to eight quarters covering two years of data (to match the period to their data) are examined. The correlation of group one with group two represents the temporal correlation of percentage buys by group one in quarter t with the percentage buys by group two in quarter $t+L$, where $L = 1$ to 8.

Second, for each quarter, the study regresses the percentage buys on each stock on lagged quarterly log returns over three years (Ret. Q_1 through Q_12), log abnormal volume in the stock during the quarter, and one period lagged percentage buys. The average coefficients and *t*-values from quarterly regressions are presented. Abnormal volume is calculated as the number of trades made in stock *j* in quarter *t* divided by its average over the previous year. The proportion of buys tends to be estimated more precisely for stocks with many trades, and thus, the study uses a weighted least squares regression in each quarter, where the weights are equal to the square root of the number of trades in stock *j*. Stocks with fewer than ten trades are dropped. Since the sum of percentage of buys and percentage of sells is 1, the results from percentage of sells regressions yield the same coefficients with a reversed sign (untabulated).

2.4.2. Determinants of herding at investor level

For investor level determinants, random forest regression algorithms are applied to examine the predictive power of a wide range of determinants (features) of the *IHM* or *GTW* measure of herding (target variable). The determinants are commonly used in the portfolio choice and individual trading literature (e.g., Giannetti and Simonov, 2006; Barber et al., 2009; Giannetti and Laeven, 2016; Barber et al., 2009; Merli and Roger, 2013). Appendix Table 1 presents the definition of these variables, and Appendix Table 2 presents descriptive statistics.

Investor age (mean: 52) is the difference in years between birth year and current year in the data, *PortValue* is the value-weighted average of the value of stocks in MSEK (mean: 0.23), *CumNoTrades* (mean: 4.36) represents the number of trades an individual made up until quarter *t*, proxying for investor's trade experience. Gender is coded as 1 for males (72%) and 0 for females. The portfolio beta (*PortBeta,* mean: 0.92) is estimated by daily CAPM regressions during a quarter using daily stock returns and the SIXRX return index as the return of the market portfolio. Closing daily stock prices are obtained from FinBas, and daily risk-free rate of returns and three-month treasury bill rates (SSVX3M) are obtained from the Central Bank of Sweden (Riksbanken). The average risk premium on the portfolio (*Rp-rf*) is 1.5%. The variables *Size, MTB, Liquidity, Leverage, and Firm Age* represent the average firm characteristics in an individual portfolio. *Size* is the natural logarithm of the firm's total assets at year end (mean: 23.58), *MTB* is the market to book ratio of the stock at quarter end (mean: 3.89), *Liquidity* is measured as the average trading volume of the stock during the quarter divided by the stock market valuation at quarter end (mean: 3.49), *Leverage* is the ratio of debt to total assets of the firm at year end (mean: 0.45), and *Firm Age* is measured as the logarithm of the number of quarters from the IPO (mean: 3.08). The remaining variables in the table are the Carhart 4 factors obtained from the Swedish House of Finance Research Data Center.

2.4.2.1. Machine learning algorithms. The study uses a random forest model as a means to identify informative determinants of herding at the investor level in a supervised learning setting (see Breiman, 2001). Previous comparative studies using machine learning techniques for feature selection on large datasets (see e.g., Grömping, 2009) suggest that random forest has a high precision and is considered a more reliable method in both regression and classification settings than other dimensionality reduction techniques; classifiers, such as K-nearest neighbor, linear discriminant analysis, and support vector machines (Chen et al., 2020); embedded methods, such as least absolute shrinkage and selection operator (LASSO) (Tibshirani, 1996); multivariate filter methods, e.g., correlation-based feature selector (Hall, 1999); and multivariate adaptive regression splines (Hastie et al., 2001).

The classification models as well as logistic regression models that may have either nominal or binary output are not appropriate to apply in this study, as dependent variables, herding measures, and portfolio returns are continuous variables. Moreover, despite the convenient prediction results, the relation between the dependent variable and the explanatory variables is difficult to understand in methods using neural nets because the variables are transformed by the nets at each layer (Grömping, 2009). Additionally, non-parametric linear regression models may have difficulties providing good prediction results for determinants of herding when faced with more complex non-linear decision boundaries. The random forest model is applicable to the analyses in this study since the explanatory variables are both numerical and category variables. The technique is also useful in modelling complex non-linear relations without feature engineering. Most importantly, the embedded feature ranking technique of variable importance measures in random forests would provide the most important determinants and performance predictors of herding for future research.

Hence, in this study, the main analyses have been implemented with random forest (Breiman, 2001). The evidence from previous research shows that ordinary least squares regressions applied to cross-sectional regressions of returns on anomaly variables are highly sensitive to microcap outliers (see Hou et al., 2020). In the portfolio performance tests, the study compares the results from ordinary least squares of market model regressions with those from random forest regressions. The model maps the determinants onto the target variable of herding with overall accuracy and uses a feature importance measure to identify a subset of determinants that are predictive of herding. Although the random forest model chooses bootstrap samples, correlation among the variables is important for accuracy; therefore, a pairwise correlation analysis to examine the multicollinearity between the pairs of variables is performed before running random forest regressions in this study. As a result, three variables – district, cumulative number of trades, and population are dropped from the model. Appendix Table 3 presents the final pairwise correlation matrix, showing that none of the pairwise correlations exceed 0.7.

Moreover, the recursive feature elimination (RFE) technique in the training dataset is used to select the top three important features in predicting the herding measures and the top two important features in performance tests. To compare the feature importance values and hence the regression accuracy, four metrics are used: R^2, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Median Absolute Error. However, because of its simplicity in the interpretation, R^2 is used as the main metric for the features' importance values for the model. It is defined as $R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{\widehat{y}}_i)^2}$, where the numerator is the mean squared error of the predictions against the actual values in the current model. The denominator is the mean squared error of the mean prediction against the actual values in the baseline model. Thus, R^2 in these analyses presents the explanatory power of the regression model compared to a basic model that predicts the mean value of the target from the training set.

To determine the predictive power of each variable, the random forest model computes the feature (variable) importance score (relative to scores of all variables) in each model by estimating a set of coefficients based on the reduction in the criterion used to select split points at each node. The high values of the feature importance score indicate that the feature/variable is important (predictive) for the model. The values of these scores for each variable are presented in a bar chart in this study. These coefficients are also used to calculate the weighted sum of the input values, which is the prediction of the model. The analyses also present the wall time, which shows how long it takes for the algorithm to run each model in the study.

The feature importance can be calculated in different ways – using built-in feature importance, permutation-based importance, and SHAP values. To interpret the predictions from random forest, the study uses SHAP values, which are calculated using a method from coalitional game theory that measures the importance of a feature by comparing the results that random forest predicts with and without that feature (see Lundberg and Lee, 2017). Random Forest built-in feature computes the feature importance during the random forest training but the technique has a tendency to select important features with high cardinality. Particularly, when the features are correlated, it may select one feature as important but neglect the second important feature. This critique also counts for permutation-based importance technique, although it randomly selects each feature and computes the change in a model's performance. There is a risk that permutation-based importance technique reports important features as unimportant if the features are highly correlated. Also, the technique is computationally expensive. Although computing SHAP values can also be computationally expensive, resulting decision plots and summary plots are very informative for determining the most important features. Lundberg and Lee (2017) argue that SHAP values can help interpret the output of any machine learning model.

Besides this, SHAP values can be calculated for any tree model; they are also superior in terms of global interpretability and local interpretability. The analysis in this paper focuses on the global interpretability as the main aim is to explore how much each feature contributes (either positively or negatively) to herding measures and return in herd vs. anti-herd portfolios. The focus is to examine the direction of the relationship for each feature with the target variable. For this purpose, the SHAP summary plots are very informative.

Although it provides useful information, the focus is not the local interpretability where each observation gets its own, individual, SHAP values. This may improve transparency by showing why each individual observation receives a prediction and contributes to the prediction of a target variable. However, such predictions at individual observation level are more interesting if they are used in further analysis. In contrast, the focus in this paper is to examine predictability at feature level not at observation level.

Since the algorithm or the evaluation procedure is stochastic, the results may vary each time one runs the regressions. To alleviate such concerns, this study follows two approaches. First, the models are run several times, and similar rankings of the variables are observed, although the values of the scores are slightly altered. Since random forest regressions have a long runtime in large datasets, a randomly chosen sample of 30% of the data is used in each trial. A randomly chosen 70% of this sample covers the training dataset, and the remaining 30% constitutes the test dataset.

Second, for robustness analyses, two other estimation machine learning methods are applied. Unlike random forest, which combines a large number of decision trees at the end of the process using the average values of the decision tree values, the technique of feature selection with gradient boosting combines decision trees at the beginning of the process and thus combines the trees one by one. Thus, the trees in gradient boosting are not built independently as in random forest. If the data do not suffer from noise, gradient boosting may perform better than random forest. Although it is unlikely that data at hand are noise free, the gradient boosting technique is also implemented in this study to compare the results from those obtained from random forest regressions. In addition, it is likely that further studies will implement a generalized linear model when examining the determinants of herding. To present and compare the results from a linear feature selection model, LASSO (although it has non-linear versions) is also implemented. Cross-validation is applied to all the models to automatically select the best performing parameters. Finally, it is likely that model learning depends on hyperparameters; to tune hyperparameters, the grid search technique is applied for optimization of the parameters.

2.5. *Measure of portfolio performance and tests*

The portfolio performance measure follows the method proposed by Feng and Seasholes (2005) and Barber et al. (2009) using a time series of net purchases and sales made by investors in a particular stock. For the number of shares purchased in quarter $t-n$ that are still held by investor i in quarter t, the current price of a stock, $P_{j,t}$ is compared to the one-period lagged stock price.[2] Once the returns on the number of shares purchased still held by investors are computed, the portfolio performance is calculated as the value weighted average of stock returns in the portfolio and used in the market model regressions. To test portfolio performance while ensuring that any potential outperformance does not result from risk, the Carhart 4 factor model regressions are run. More formally,

$$R_{p,t} - R_{ft} = \alpha_i + \beta_{mkt}(R_{m,t} - R_{f,t}) + \beta_{smb}SMB_t + \beta_{hml}HML_t + \beta_{mom}MOM_t + e_{i,t}, \tag{9}$$

where the portfolio return of individuals, $R_{p,t}$ is calculated using the last quarter price as the reference price in stock return calculations.[3] R_{ft} is the risk-free rate (one-month Swedish T-bill rate), $R_{m,t}$ is the return on the stock market in Sweden (SIXRX index), and SMB_t, HML_t, and MOM_t are the returns on the size, value, and momentum factors. Two regressions are run for each group. While $R_{p,t}$ and market risk premium are always value-weighted, other factors are either equally-weighted or value-weighted in the regressions.

To examine and compare CAPM, Fama-French 3 factors, and Carhart 4 factors, the study applies random forest regressions, examines the variable importance scores along with RFE for the top two features and explains the results with SHAP values. The study also examines the predictability of these variables in separate models for value-weighted and equally-weighted factors. As in the determinants of herding analyses, gradient boosting and LASSO techniques of feature selection are also implemented and compared. Cross validation is applied to all models and grid search is used to tune hyperparameters.

3. Stock level herding

The stock level analyses in Panel A of Table 2 show that the average *LSV* (10.3%, t-value:10.02) and *FHW* (18.3% (sqrt(0.033), t-value: 16.32) measures are significantly different from zero.

Compared with findings in Lakonishok et al. (1992), who examine herding behavior in pension funds, the result on herding is approximately 7.6 percentage points higher (their average is 2.7 %). Barber et al. (2009) examine herding behavior in individual investors' trades and find results similar to this study. They show 6.8 % herding in individual trades at a large discount broker and 12.8 % herding at a large retail broker; thus, their simple average of 9.8 % is comparable to the results in this paper. Moreover, examining herding behavior in mutual funds, Frey et al. (2014) show that *FHW* is 2.8 times higher than the traditional measure of *LSV*. In regard to herding in individual investors, the results in this study show that *FHW* is 1.6 times higher than *LSV*. Frey et al. (2014) show approximately 17.3 % of *FHW*, which is about 1 percentage point lower than the finding in this study, *FHW*, 18.3 %. Moreover, the average purchase intensity in stocks is approximately 62.7 % across all stocks, which is similar to the purchase intensity of individual

[2] The reference price can be calculated as first purchase price, highest purchase price, average purchase price, and most recent purchase price. However, in the main analyses, the one period lagged stock price is used as the reference price, although all other measures are tested to ensure the robustness of the results.

[3] Using the last period price as reference price is the most commonly used performance measure (see e.g., Feng and Seasholes, 2005; Frazzini, 2006; Lindblom et al., 2017, 2018; and Barber et al., 2009).

Table 2
Descriptive analysis of stock level herding measures. Panel A shows the descriptive statistics on the stock level herding measures. Panel B splits the data by median market cap as proxy for firm size. Panel C splits the data by book-to-market value. In Panel D firm location is classified as urban if the firm headquarter is located in Stockholm, Gothenburg, or Malmö. Rural firms are those with headquarters located in other areas. Finally, Panel D presents stock level herding measures by industry following ICB classification. *FHW* values are the square root of the statistics in Panel D. *T*-values of the difference tests for the split samples based on unequal variance, using Satterthwaite method, are presented in the panels.

Panel A. Descriptive statistics on stock level herding measures									
Variable	N	N Miss	Minimum	25th Pctl	Mean	Median	75th Pctl	Maximum	Std Dev
Purchase intensity	14,054	0	0.002	0.536	0.627	0.625	0.725	1.000	0.168
LSV	14,054	0	-0.241	0.012	0.103	0.072	0.160	0.873	0.135
FHW	14,054	0	-0.249	0.000	0.033	0.009	0.038	0.772	0.075
Panel B. Stock level herding measures by firm size									
	N	Mean	Std Dev	t-value					
LSV									
Small firms (1)	7030	0.099	0.143						
Large firms (2)	7024	0.107	0.127						
Diff (2–1)		0.009	0.135	3.87					
FHW									
Small firms (1)	7030	0.031	0.067						
Large firms (2)	7024	0.036	0.083						
Diff (2–1)		0.005	0.075	3.88					
Panel C. Stock level herding measures by book to market ratio									
	N	Mean	Std Dev	t-value					
LSV									
Low BTM (1)	7027	0.108	0.129						
High BTM (2)	7027	0.098	0.141						
Diff (2–1)		-0.009	0.135	-4.12					
FHW									
Low BTM (1)	7027	0.033	0.068						
High BTM (2)	7027	0.034	0.081						
Diff (2–1)		0.001	0.075	0.45					
Panel D. Stock level herding measures by firm location									
	N	Mean	Std Dev	t-value					
LSV									
Rural (1)	2340	0.098	0.133						
Urban (2)	11,699	0.104	0.135						
Diff (2–1)		0.006	0.135	1.99					
FHW									
Rural (1)	2340	0.032	0.073						
Urban (2)	11,699	0.034	0.075						
Diff (2–1)		0.001	0.075	0.89					
Panel E. Stock level herding measures by industry									
ICBname		LSV		FHW					
	N	Mean	Std Dev	Mean	Std Dev				
Basic Materials	932	0.106	0.137	0.185	0.279				
Consumer Goods	1300	0.100	0.133	0.178	0.275				
Consumer Services	1458	0.098	0.141	0.183	0.283				
Financials	1984	0.106	0.139	0.184	0.280				
Health Care	1782	0.099	0.127	0.173	0.257				
Industrials	3769	0.105	0.135	0.186	0.276				
Oil & Gas	455	0.118	0.144	0.200	0.286				
Technology	2119	0.101	0.131	0.180	0.264				
Telecommunications	172	0.111	0.143	0.209	0.300				
Utilities	83	0.083	0.102	0.154	0.208				

investors reported by Barber et al. (2009). Table 2 also adds new results to these previous findings and shows that herding is more pronounced in large firms (Panel B), firms with a low book-to-market ratio (Panel C), and firms that are headquartered in urban areas (Panel D). Panel E shows that the distribution of both herding measures is similar across the industries, although it is slightly more pronounced in telecommunications and oil and gas industries.

Taken together, the results suggest that there exist significant herding tendencies in individual trading in Swedish stock markets, rejecting the null hypothesis that the contemporaneous buy/sell decisions of individual investors are uncorrelated. In terms of economic interpretation, for example, using the 10.3 % *LSV* measure, the result suggests that if the average purchase intensity across all stocks in Eq. (1) is 0.5, then 60.3 % of the individual investors would be changing their holdings of an average stock in one direction and 39.7% in the opposite direction. These results in Panels B-C also indicate that herding is a behavioral bias. If information asymmetry was the reason behind individual investors herding, the opposite results should have been observed: more herding in small firms, young firms, and firms located in rural areas.

3.1. Contemporaneous and time-series correlations

Table 3 explores the stock level purchase intensity and examines the mean contemporaneous and time-series correlation of percentage buys by all individual investors across all stocks.

The first row of the table shows that the mean contemporaneous correlation of purchase intensity across groups is 71.5 %, which is qualitatively the same as the result reported in Barber et al. (2009), which is approximately 71.5 % for retail investors and approximately 73.4 % for discount investors. This result indicates that in a given quarter, both groups of investors concentrate their purchases in the same stocks. It also means that the purchase intensity of one group can explain approximately 51 % of the variation in purchase intensity of group 2 as the square of the correlation coefficient, 0.715^2, reflects R-squared from a regression of the purchase intensity of group one on the purchase intensity of group 2. The remaining rows in the table show the time series correlations between the purchase intensity in quarter t and quarter $t+1$ both within and across groups 1 and 2. The correlation across the groups ranges between 9% and 15% and seems to decrease over time, although there is some non-linearity; that is, it increases after one year and then decreases in quarter 8 to the level below what is observed in the first quarter. Although the time series mean correlations tend to be lower in Sweden than what is reported in the US, the overall results suggest strong persistence in purchase intensity over time.

3.2. Results from determinants of herding at stock level

Table 4 explores determinants of herding at stock level. Barber et al. (2009) show that behavioral explanations are more plausible for the herding behavior of individual investors, while representativeness heuristics explain individual buying behavior, the disposition effect seems to explain individual selling behavior. To examine whether their results hold for this study and to gain better insights from the determinants of trading in a different and smaller market (Sweden) and in a more recent period (between 2006 and 2014), the analyses in Table 4 revisit their model.

The table shows that there is a significant and negative relationship between percentage buys and return at Q-1, − 12.8% (t-value: −11.20). The effects seem to dissipate after Q-2, but some evidence of a negative relationship at lag 6 still exists. The results indicate that individuals buy fewer (and hence sell more) stocks with strong recent returns, which means that when recent returns are low, they buy more, a result that can be explained by the disposition effect. This result concurs with the findings of Heimer (2016) and suggests that social interaction is associated with increases in the disposition effect. In general, more distant past returns do not seem to matter for individuals. Additionally, the coefficients for the abnormal trading volume, which captures the attention effect, are significant and positive. Compared to the Barber et al. (2009) result showing a 4.9% attention effect for discount investors and a 3.4 % attention effect for retail investors, the results in this study show a higher attention effect, 8.1%, in individual trading. The results also show persistency of trading in stocks, as past purchase intensity significantly increases current purchase intensity. Compared to these current results, the persistency rate reported in Barber et al. (2009) is higher – 46.2%, for discount investors and 60.5 % for retail investors.

The overall results are consistent with the disposition effect and that individuals are net buyers of stocks with high abnormal trading volume, i.e., attention grabbing stocks. These stock level analyses indicate, thus far, that herding is behavioral, implying that we expect a higher degree of short-term mispricing in stocks with a higher level of herding. Thus, suboptimal (biased) decisions or, more precisely, under-diversification, may cause short-term price discrepancies in these stocks. The next section analyzes the determinants of herding at the individual investor level.

4. Investor level herding

Appendix Table 1 shows that the average quarterly *IHM* value is significantly different from zero (t-value: 13.94) and indicates that 16.3 % of the investors are trading more on the same side than what would be predicted if decisions were randomly made. The average

Table 3

Mean contemporaneous and time-series correlation of percentage buys by all individuals. The data are split into randomly chosen equal group of investors. For each-quarter, the percentage of all trades that are purchases are calculated. The table presents the mean contemporaneous correlation across groups in the first row. The remaining rows represent the mean temporal correlation from one quarter to eight quarters covering two years of data. The correlation of group one with group two represents the temporal correlation of percentage buys by group one in quarter t with the percentage buys by group two in quarter $t + L$, where $L = 1,8$. T-statistics are based on the mean and standard deviation of the calculated correlations. Stars * , * *, * ** indicate the significance at the 10%, 5%, and 1% level, respectively.

	Correlation of % buys in quarter t with % buys in quarter $t + L$			*t*-statistics		
Lag	Group 1 with group 1	Group 2 with group 2	Group 1 with group 2	Group 1 with group 1	Group 2 with group 2	Group 1 with group 2
0	1.000	1.000	0.715			92.96 ***
1	0.163	0.162	0.139	21.58 ***	21.43 ***	18.11 ***
2	0.116	0.111	0.095	14.98 ***	14.36 ***	12.37 ***
3	0.106	0.100	0.091	13.48 ***	12.74 ***	11.82 ***
4	0.185	0.178	0.155	23.33 ***	22.49 ***	20.21 ***
5	0.089	0.084	0.080	10.92 ***	10.40 ***	10.39 ***
6	0.077	0.061	0.068	9.42 ***	7.53 ***	8.79 ***
7	0.076	0.071	0.061	9.21 ***	8.64 ***	7.97 ***
8	0.109	0.110	0.092	13.13 ***	13.37 ***	11.96 ***

Table 4
Stock level cross-sectional regressions of buying intensity. For each quarter, the percentage buys are regressed on each stock on lagged quarterly log returns over three years (Ret. Q_1 through Q_12), log abnormal volume in the stock during the quarter, and one period lagged percentage buys. Abnormal volume is calculated as the number of trades made in stock *j* in quarter *t* divided by its average over the previous year. The lagged dependent variable is included to account for the previously documented time-series dependence in the proportion of buys. The proportion of buys tends to be estimated more precisely for stocks with many trades and thus a weighted least squares regression is estimated in each quarter, where the weights are equal to the square root of the number of trades in stock *j*. Stocks with fewer than ten trades are excluded. Since the sum of % Buys and % Sells is 1, the results from % Sells regressions yield the same coefficients with a reversed sign (untabulated). The table reports the mean coefficient estimates across 23 quarters. *T*-values (in parentheses) are based on the time series of coefficient estimates and the stars indicates that the coefficient is significant at the 5% or lower levels.

	All	
Variable	% Buys	t-value
Intercept	0.499	(14.48)**
Ret. *Q*-1	-0.128	(−11.20)**
Ret. *Q*-2	0.012	(1.04)
Ret. *Q*-3	-0.015	(−1.69)
Ret. *Q*-4	-0.009	(−1.06)
Ret. *Q*-5	0.001	(0.06)
Ret. *Q*-6	-0.027	(−2.92)**
Ret. *Q*-7	-0.013	(−1.70)
Ret. *Q*-8	-0.003	(−0.44)
Ret. *Q*-9	0.001	(0.07)
Ret. *Q*-10	0.000	(−0.03)
Ret. *Q*-11	-0.001	(−0.12)
Ret. *Q*-12	0.007	(0.73)
Abn. Vol.	0.081	(6.31)**
Lagged dep. var.	0.268	(5.73)* *

value of the *IHM* measure is approximately 4% points larger than the result shown by Merli and Roger (2013) using French data. When the passive positions of individuals considered using the measure of *GTW*, there is on average no herding, which is similar to the low level of herding of mutual funds (approximately 2.5 %) reported by Grinblatt et al. (1995). One reason for this result is related to the fact that most of the individuals in the data are passive investors. Since all investors cannot be buying and selling as a herd, as the buys must equal sells in the aggregate, it is more likely to observe herding in a subset of investors using the *GTW* measure but not in the aggregate.

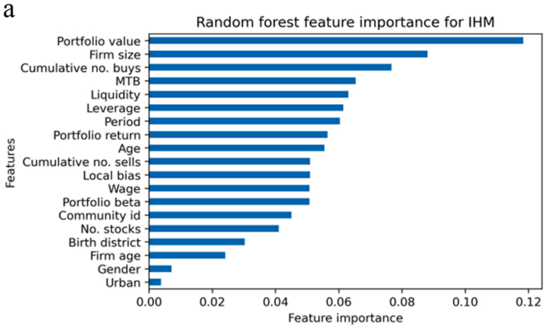
4.1. *Results from determinants of herding at the investor level*

Fig. 1a presents the feature importance of all variables for *IHM*, along with the results from recursive feature elimination in the training dataset for the top 5 features, to select the three most important variables. The reason for applying RFE to the top 5 features is that it takes a tremendous amount of time if it is applied to all features. The results from the regression predictability and accuracy metrics are also shown for R^2, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Median Absolute Error (Med. AR). The wall time showing how long it takes to run each algorithm is presented in the figure as well. In particular, the wall time increases exponentially with the number of features when computing SHAP values. Thus, similar to RFE, SHAP values are calculated for the top 5 features and presented in the figure. Fig. 1b revises these analyses for *GTW*.

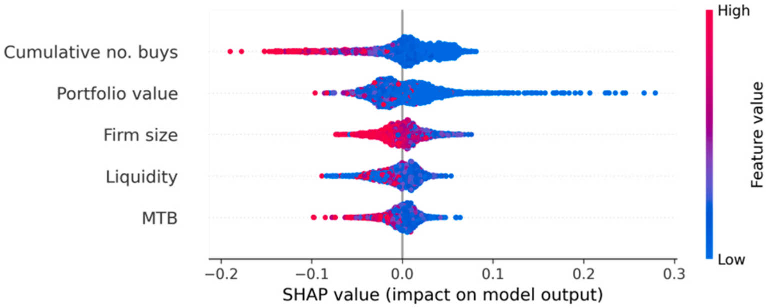
Fig. 1a shows that the IHM model has a slightly low but still reasonable predictability with an R^2 of 43 %, and this final regression takes approximately 1 h 11 min 4 s of wall time. RMSE of the model is 0.117. The top three most important features are portfolio value (12 %), followed by firm size (9 %) and cumulative number of buys (8 %). The variable local bias measuring how much individuals overweigh firm/stocks headquartered in the district that they live in is fairly important (5 %) for *IHM*. The variables of whether the individual lives in an urban area, their gender, and firm age are ranked as the bottom three and do not seem to matter as much for *IHM*. The RFE method ranks portfolio value, firm size, and liquidity on top, followed by MTB and cumulative number of buys. The RFE method increases R^2 from 0.43 to 0.803, indicating that these top 5 features are highly predictive for *IHM*.

SHAP values show the sign and magnitude of feature importance and hence its contribution across all possible permutations of the top 5 features to the model's prediction for *IHM*. The red (blue) color in the SHAP summary plot indicates high (low) values of the feature. The vertical axis shows the feature importance among the top 5 features, and the horizontal axis shows whether high (low) values of the feature lead to lower or higher prediction within the model. The results show that high values of cumulative number of buys reduce the prediction of *IHM* by approximately 0.2 ppt. The lower the portfolio values, the higher the prediction of IHM (an increase of 0.3 ppt.). Low values of firm size, liquidity, and MTB tend to increase the prediction of IHM.

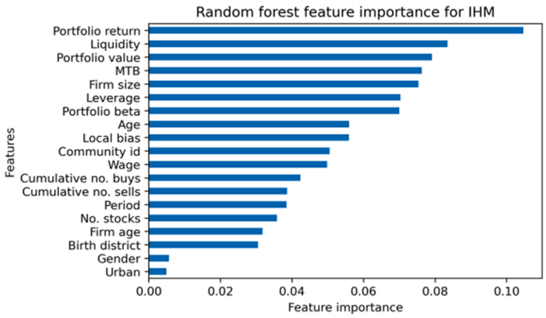
Fig. 1b plots the feature importance of all features for *GTW*. Compared with the *IHM* model, *GTW* seems to have less predictive a



R2	0.425
MAE	0.078
RMSE	0.117
Med. AE	0.05
Wall time	1h 7min 11s
RFE ranks for top 5 features	
Portfolio value	1
Firm size	1
Liquidity	1
MTB	2
Cumulative no. buys	3
RFE-R2	0.803



b



R2	0.129
MAE	0.006
RMSE	0.014
Med. AE	0.002
Wall time	1h 11min 4s
RFE ranks for top 5 features	
Portfolio value	1
Firm size	3
Liquidity	1
MTB	2
Portfolio return	1
RFE-R2	0.803

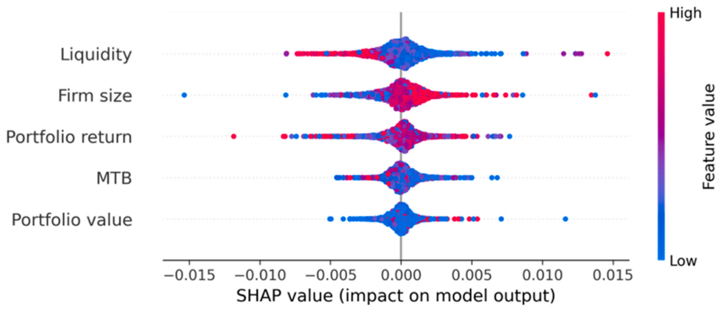


Fig. 1. Determinants of herding at individual investor level. Figures present the results from random forest regressions, feature importance, the recursive feature elimination (RFE), and SHAP values for *IHM* and *GTW* measures of herding. The feature importance is implemented by the random forest algorithm in scikit-learn in Python. In random forest regressions, a randomly chosen 70% of the sample is used as training dataset and the remaining 30 % is used as test dataset. The model predictability as R^2, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Median Absolute Error (Med. AR) as well as the wall time showing how long it takes to run each algorithm are shown in the figures. To save time, RFE method is applied to top 5 features among which top three features are selected in training dataset for each model and the results are presented along with R^2 after RFE. Fig. 1a presents the results from random forest regressions and RFE and the SHAP values for the top 5 features for *IHM*. Fig. 1b shows the results from random forest regression and RFE for GTW and the SHAP values for the top 5 features for *GTW*.

ability, with an R^2 of 13 %, and it takes a wall time of 1 h 11 min 4 s to run the regression. RMSE in *GTW* model is 0.014, lower than that of the *IHM* model, but it is difficult to compare RMSEs as the number of observations differs significantly in these models. Thus, the relative (to total variance) measure of R^2 is easier to interpret when comparing the models. One reason for the poorer performance of *GTW* is that the measure does not exclude the passive positions and most of the individuals are passive in the data; that is to say *GTW* is zero, on average. This measure seems to be doing much better when applied to institutional investor herding behavior (see Grinblatt et al., 1995). The top three important features in this model are portfolio return (11 %), liquidity (8 %), and portfolio value (8 %). The results are similar but slightly better from local bias (6 %) in comparison to the *IHM* model. The bottom three ranked variables are urban, gender, and birth district of individuals. The RFE method ranks the top 5 features as portfolio return, liquidity, and portfolio value as the highest, followed by MTB and firm size. The R^2 metric from the recursive feature elimination method increases from 0.13 to 0.803. SHAP values in Fig. 1b show that the high values of portfolio return lead to a small increase (0005 ppt.) in the prediction of *GTW*. Similar results are observed from the portfolio value. While high values of liquidity and MTB decrease the prediction, high values of firm size increase the prediction of *GTW*.

Appendix Table 4 presents the results from variable rankings using gradient boosting and LASSO models for the top three determinants. Gradient boosting ranks portfolio value, cumulative number of buys, and cumulative number sells on top in the *IHM* regressions, which yield the highest R^2 among other models. All the other model accuracy measures, Mean Square Error (MSE), Mean Absolute Error (MAE) and Median Absolute Error (Median AE), are in favor of gradient boosting as the errors tend to be lower in this model than those in LASSO. LASSO ranks cumulative number of buys, cumulative number sells, and wage on top, but the performance of this model is poorer than that of gradient boosting in *IHM* regressions (R^2: 0.038). In *GTW* regressions, gradient boosting provides slightly better results than LASSO model, the R^2 values are 0.064 vs. 0.0002, respectively but gradient boosting has a slightly higher in Median AE (0.002). Gradient boosting ranks portfolio return, number of stocks in a portfolio, and cumulative number of buys as the top three features. Finally, LASSO ranks cumulative number of buys, cumulative number sells, and wage on top. The ranking results from gradient boosting are more consistent with random forest regressions, while the results from LASSO deviate slightly from those former results. This indicates that there are some non-linear effects and interactions among the predictors that are not captured by LASSO, while they are accounted for in random forest and gradient boosting models.

In summary, *IHM*, which measures active herding behavior, seems to perform better in individual investor data. The important features are slightly different in the *GTW* measure. The most important variables that explain both herding measures can be ordered as portfolio value, portfolio return, firm size, liquidity, and cumulative number of buys.

5. Performance of individual investors' herd portfolios

Table 5 provides the results from the OLS of Carhart 4 factor model regressions where the factors other than market risk premium are either value-weighted or equally-weighed. Panel A splits the data on the herding measures and shows results from herd vs. non-herd portfolios.

Both herding measures in Panel A show that individuals who do not herd earn significantly higher abnormal returns than those who herd. For example, the value-weighted *IHM* regressions show a difference of 1.52 % (*chi2*: 152.04), and the value-weighted *GTW* regressions show a difference of 0.61 % (*chi2*: 162.96) for those who do not herd vs. those who herd, respectively. The corresponding differences in alpha in equally-weighted measures are 1.1 % (*chi2*: 92.34) in *IHM* and 0.8 % (*chi2*: 156.98) in *GTW* regressions. These

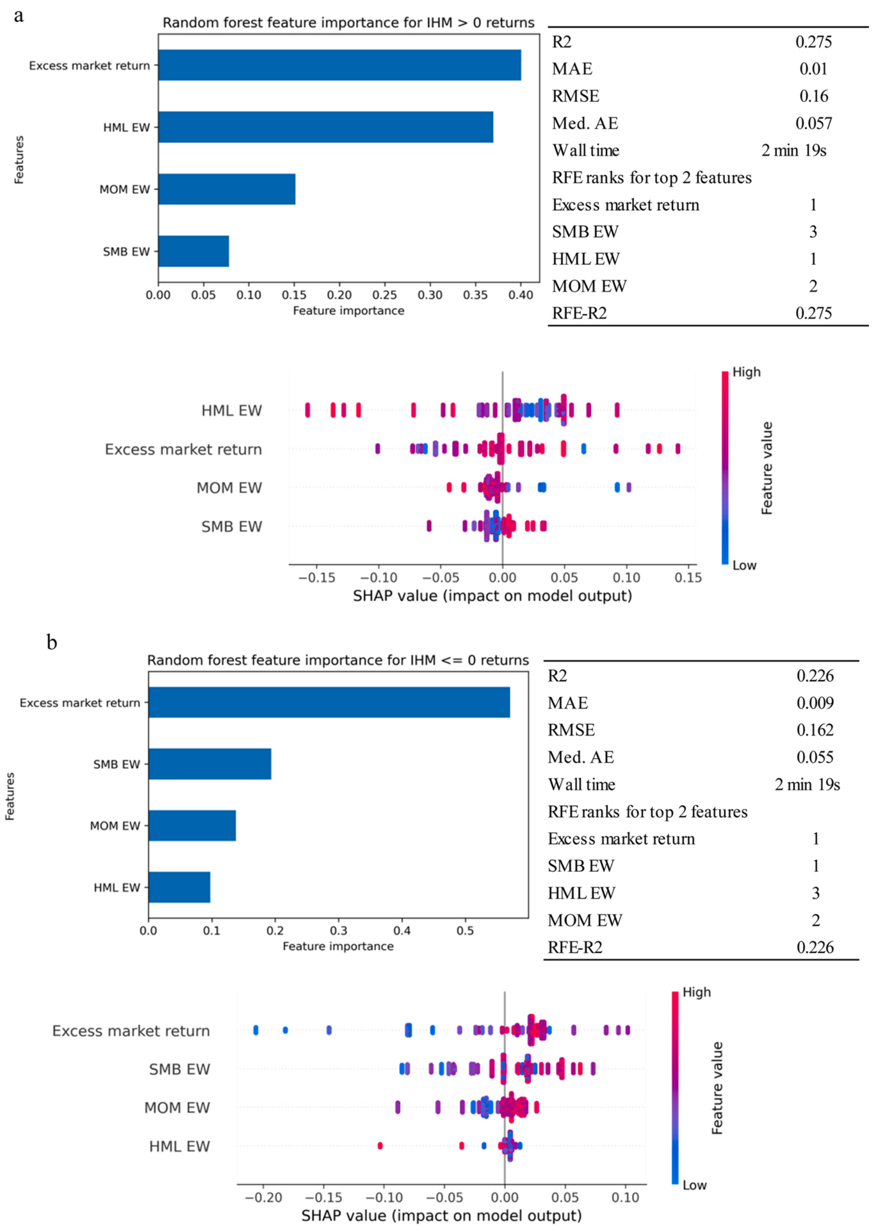
Table 5

Regression results from performance tests. The Carhart (1997) 4 factor model regressions are run in a classical regression model setting to test portfolio performance. More formally: $R_{p,t} - R_{ft} = \alpha_i + \beta_{mkt}(R_{m,t} - R_{f,t}) + \beta_{smb}SMB_t + \beta_{hml}HML_t + \beta_{mom}MOM_t + e_{i,t}$, where, $R_{p,t}$ is the value-weighted portfolio return R_{ft} is risk-free rate (one-month Swedish T-bill rate), $R_{m,t}$ is the stock market return in Sweden (SIXRX index), SMB_t, HML_t, and MOM_t are the returns on the size, value, momentum factors and obtained from the Swedish House of Finance Research Data Center. Two regressions are run for each group. While $R_{p,t}$ and market risk premium are always value weighted, other factors are either equally-weighted (denoted as EW) or value-weighted (denoted as VW). The panel splits the data on the herding measures. Standard errors are clustered on district and period.

Panel A. Herding vs. no-herding								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	IHM< =0	IHM< =0	IHM> 0	IHM> 0	GTW< =0	GTW< =0	GTW> 0	GTW> 0
	EW	VW	EW	VW	EW	VW	EW	VW
VARIABLES	Rp-rf	Rp-rf	Rp-rf	Rp-rf	Rp-rf	Rp-rf	Rp-rf	Rp-rf
Alpha	0.0270	0.0370	0.0147	0.0218	0.0158	0.0261	0.0083	0.0200
	(3.76)***	(5.95)***	(1.62)	(3.11)***	(2.11)**	(4.24)***	(1.38)	(4.36)***
Beta	0.7262	0.6629	0.6554	0.7738	0.5056	0.5130	1.1021	1.0666
	(4.27)***	(4.28)***	(5.84)***	(8.15)***	(2.91)***	(3.30)***	(14.00)***	(13.22)***
SMB	0.2545	0.4871	0.1622	0.5525	0.2934	0.5477	0.1589	0.3290
	(6.41)***	(4.73)***	(2.33)**	(3.94)***	(6.48)***	(4.80)***	(3.98)***	(3.54)***
HML	-0.3923	-0.1493	-0.4443	-0.0563	-0.3923	-0.3026	-0.6448	-0.4408
	(−2.30)**	(−0.70)	(−1.94)*	(−0.22)	(−2.23)**	(−1.46)	(−3.68)***	(−2.38)**
MOM	-0.5870	-0.3431	-0.2708	0.2461	-0.5210	-0.2184	-0.5380	-0.5273
	(−5.59)***	(−1.92)*	(−1.80)*	(1.07)	(−4.17)***	(−1.21)	(−5.26)***	(−3.35)***
Observations	469,320	469,320	1,486,488	1,486,488	6,876,569	6,876,569	2,552,882	2,552,882
R-squared	0.056	0.064	0.052	0.061	0.045	0.058	0.146	0.150
F-value	110.21	27.05	34.06	38.64	92.19	28.39	134.35	62.85

results suggest that herding does not yield abnormal returns for the average investor, which is consistent with the findings in Barber et al. (2009) that herding is not information but behavioral-based. However, it is also evident that the results from equally-weighted factors differ from those obtained using value-weighted factors, particularly in herd portfolios. Next, the study explores the consistency of the results from equally-weighted and value-weighted factors by implementing machine learning algorithms. The analyses yield the factors that provide the most predictability in market models for herding measures.

Fig. 2 (*IHM*) and Fig. 3 (*GTW*) show the results from random forest regressions, feature importance, RFE, and SHAP values, using

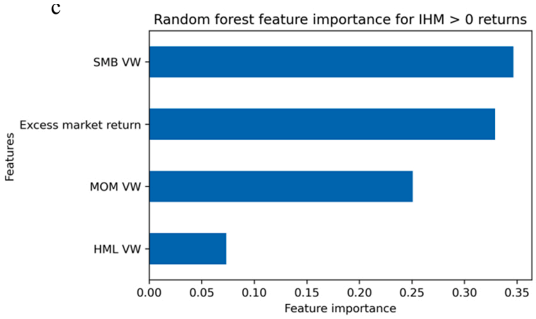


R2	0.275
MAE	0.01
RMSE	0.16
Med. AE	0.057
Wall time	2 min 19s
RFE ranks for top 2 features	
Excess market return	1
SMB EW	3
HML EW	1
MOM EW	2
RFE-R2	0.275

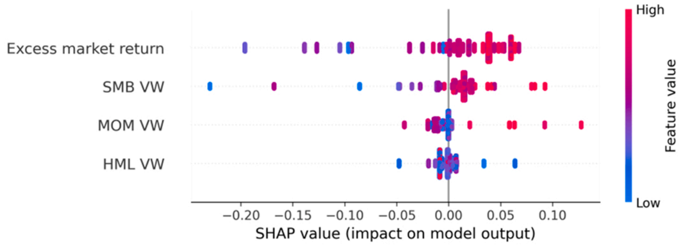
R2	0.226
MAE	0.009
RMSE	0.162
Med. AE	0.055
Wall time	2 min 19s
RFE ranks for top 2 features	
Excess market return	1
SMB EW	1
HML EW	3
MOM EW	2
RFE-R2	0.226

Fig. 2. Results from performance tests for *IHM* measure in herd vs. non-herd portfolios. Figures present the results from random forest regressions, feature importance, the recursive feature elimination (RFE), and SHAP values for *IHM* measure of herding in Carhart 4 factor model. The results from equally-weighted and value-weighted factors are presented for both herd and non-herd portfolio performance regressions. The feature importance is implemented by the random forest algorithm in scikit-learn in Python. In random forest regressions, a randomly chosen 70% of the sample is used as training dataset and the remaining 30% is used as test dataset. The model predictability as R^2, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Median Absolute Error (Med. AR) as well as the wall time showing how long it takes to run each algorithm are shown in the figures. RFE method selects the top two features in training dataset for each model and the results are presented along with R^2 after RFE. Fig. 2a presents the results from herd portfolios, *IHM* > *0,* using equally weighted factors, Fig. 2b shows the results from non-herd portfolios, *IHM* < =*0,* using equally-weighted factors. Fig. 2c presents the results from herd portfolios, *IHM* > *0,* using value-weighted factors, and finally, Fig. 2d shows the results from non-herd portfolios, *IHM* < =*0,* using value-weighted factors.

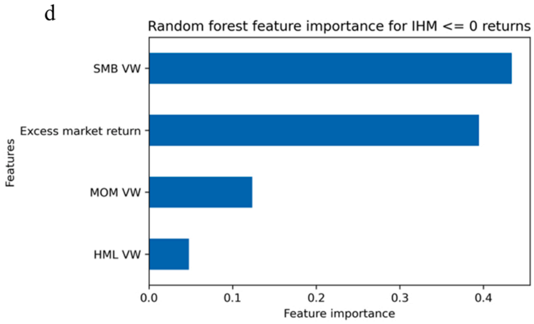
c



R2	0.275
MAE	0.01
RMSE	0.16
Med. AE	0.058
Wall time	2 min 31s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	3
MOM VW	2
RFE-R2	0.275



d



R2	0.23
MAE	0.094
RMSE	0.16
Med. AE	0.055
Wall time	8 min 13s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	3
MOM VW	2
RFE-R2	0.225

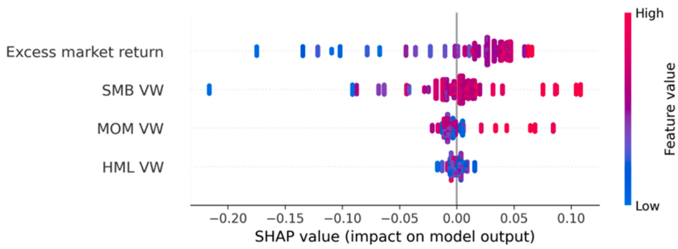
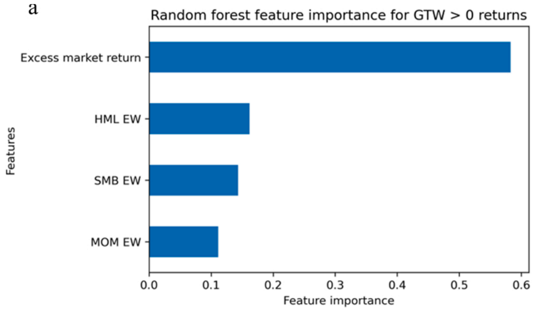


Fig. 2. (*continued*).

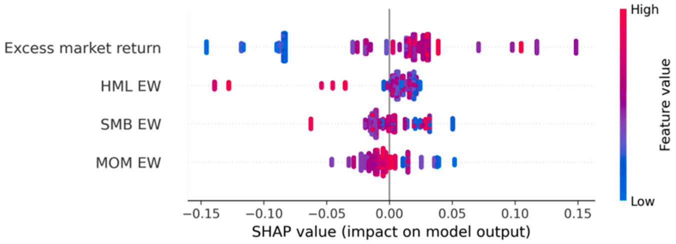
both equally-weighed factors and value-weighted factors in herd vs. non-herd portfolios for *IHM* and for *GTW* measures, respectively.

In herd portfolios *(IHM >0)* using equally-weighted factors, in Fig. 2a, the most predictive variable is the excess market return (with feature importance of 0.4), followed by HML (0.37), MOM (0.15), and SMB (0.07). The random forest regression has an R^2 of 28 %, RMSE is 0.16, and the wall time to run these regressions is 2 min and 19 s. Although the number of observations is the same as in Table 5, the number of variables is much lower in these portfolio performance regressions, which seems to reduce the wall time tremendously. The RFE method ranks the top 2 predictors among all four factors as excess market return and HML. SHAP values are calculated based on a randomly selected 30 % of the data for computation feasibility. The results show that high values of excess market return lead to a positive prediction (0.15) of excess returns in herd portfolios. This interpretation also holds for SMB, albeit the effect is lower, 0.04. High HML values negatively predict excess returns. Finally, the high values of MOM lead to a negative prediction on excess returns in herd portfolios.

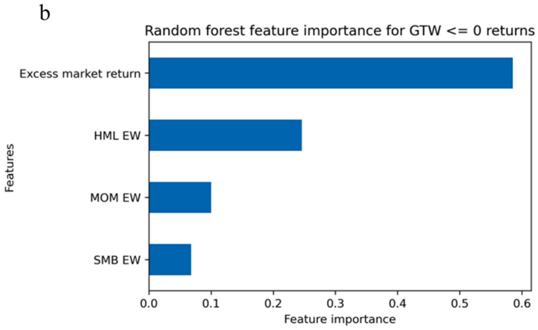
In non-herd portfolios *(IHM $\leq$ 0)* using equally-weighted factors, random forest regressions, in Fig. 2b, also rank excess market return on top (with a feature importance score of 0.6). However, SMB (0.2) is ranked as the second most important predictor, followed by MOM (0.13) and HML (0.1). R^2 is slightly lower than that in herd portfolios, 0.23, and RMSE is 0.162. RFE ranks excess market a



R2	0.361
MAE	0.074
RMSE	0.122
Med. AE	0.047
Wall time	39.1s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	2
MOM VW	3
RFE-R2	0.358



b

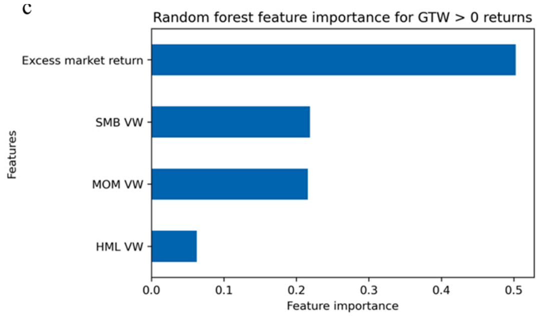


R2	0.267
MAE	0.103
RMSE	0.172
Med. AE	0.061
Wall time	1 min 39.1s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	2
MOM VW	3
RFE-R2	0.267

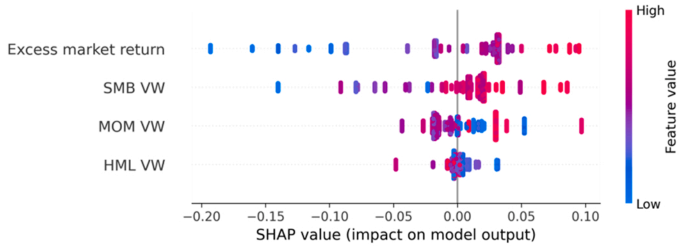
Fig. 3. Results from performance tests for *GTW* measure in herd vs. non-herd portfolios. Figures present the results from random forest regressions, feature importance, the recursive feature elimination (RFE), and SHAP values for *GTW* measure of herding in Carhart 4 factor model. The results from equally-weighted and value-weighted factors are presented for both herd and non-herd portfolio performance regressions. The feature importance is implemented by the random forest algorithm in scikit-learn in Python. In random forest regressions, a randomly chosen 70% of the sample is used as training dataset and the remaining 30 % is used as test dataset. The model predictability as R^2, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Median Absolute Error (Med. AR) as well as the wall time showing how long it takes to run each algorithm are shown in the figures. RFE method selects the top two features in training dataset for each model and the results are presented along with R^2 after RFE. Fig. 3a presents the results from herd portfolios, $GTW > 0$, using equally weighted factors, Fig. 3b shows the results from non-herd portfolios, $GTW < =0$, using equally-weighted factors. Fig. 3c presents the results from herd portfolios, $GTW > 0$, using value-weighted factors, and finally, Fig. 3d shows the results from non-herd portfolios, $GTW < =0$, using value-weighted factors.

return and SMB as the top two predictors of excess return, and there is not much change in R^2 according to the RFE method. The SHAP values suggest that high values of excess market return positively predict excess returns in non-herd portfolios. The SHAP value is positive for high values of SMB and MOM, but high values of HML decrease the predictability of excess returns.

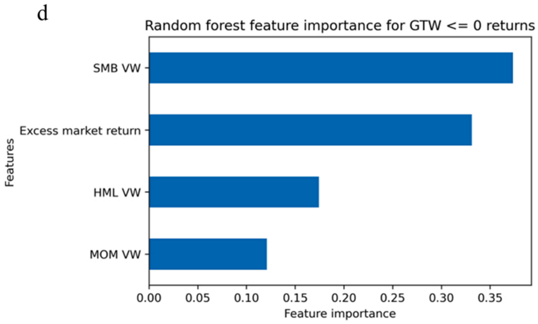
Fig. 2c and d present the results for feature importance, RFE ranks, and SHAP values in herd vs. non-herd portfolios using value-c



R2	0.363
MAE	0.073
RMSE	0.122
Med. AE	0.047
Wall time	37.7.1s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	3
MOM VW	2
RFE-R2	0.267



d



R2	0.267
MAE	0.073
RMSE	0.103
Med. AE	0.062
Wall time	1 min 59s
RFE ranks for top 2 features	
Excess market return	1
SMB VW	1
HML VW	2
MOM VW	3
RFE-R2	0.267

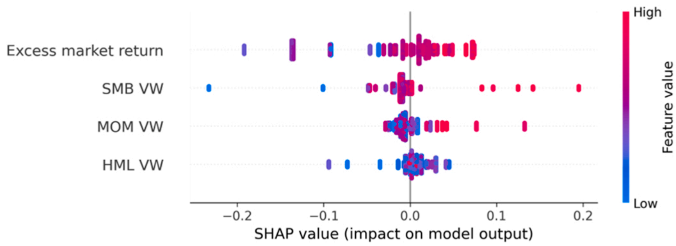


Fig. 3. (*continued*).

weighted factors. Unlike in the analyses with equally-weighted factors, the random forest regressions rank SMB as the highest, and excess market return is ranked second in both herd and non-herd portfolios. Both R^2 (0.275) and other model evaluation metrics are similar to those analyses with equally-weighted factors. Using a random sample of 30 %, SHAP values indicate that high values of excess market return increase the prediction of excess returns in both herd and non-herd portfolios. High values of SMB and MOM also improve the prediction of excess returns, while HML does not seem to contribute to the prediction of excess returns on the portfolios. Fig. 3 revises all these analyses for the *GTW* measure of herding.

The results from random forest regressions in Fig. 3a and b are consistent with those obtained from Fig. 2. Using equally-weighted factors, random forest regressions rank excess market return as the highest feature, with an importance score of 0.58 (0.59), R^2 of 0.361 (0.267) and RMSE of 0.122 (0.172) in herd (non-herd) portfolios. SHAP values indicate that excess market return positively predicts excess returns in both herd and non-herd portfolios. Using value-weighted factors in herd portfolios, Fig. 3c also ranks excess market return as the highest feature (0.5), with an R^2 of 3.63 and RMSE of 0.122. SHAP values indicate that high excess market return increases the prediction of excess returns in herd portfolios. However, using value-weighted factors in Fig. 3d, random forest regressions rank SMB as the most important feature (0.36), with an R^2 of 0.267 and RMSE of 0.103. Although this result is consistent with the results from the *IHM* measure in Fig. 2d, it does not concur with all other random forest regression analyses that rank excess market returns at the top.

In summary, using equally-weighted factors in the OLS of Carhart 4 factor (classical) regression models, the results show that herd portfolios outperform neither market nor non-herd portfolios. The results from random forest regressions in both herd and non-herd portfolios are consistent. In the equally-weighted portfolios, the market risk premium has the most predictive power. It is followed by HML, MOM, and SMB in herd portfolios, and SMB takes the second highest rank after the market risk premium in non-herd portfolios, particularly in predicting the *IHM* measure of herding. The results from random forest regressions using value-weighted portfolios are different in non-herd portfolios, regardless of which measure of herding has been used. SMB seems to have the highest degree of feature importance. It is followed by market risk premium, MOM, and HML. However, in herd portfolios, the market risk premium still has the most predictive power regardless of the weighting scheme and the herding measure. The results from SHAP values are consistent with the fact that high values of the market risk premium positively predict excess returns in both herd and non-herd portfolios. Taken as a whole, the highest rank order of feature importance alternates between the market risk premium and SMB using value weighted-factors in non-herd portfolios. The market risk premium consistently has the highest rank order using equally-weighted factors in both herd and non-herd portfolios. Thus, using equally-weighted factors appears to provide more robust and consistent results in portfolio performance tests.

Appendix Table 5 shows the ranking results from gradient boosting and LASSO for the top two factors in the performance tests. Panel A shows the results from *IHM* regressions for both herd and non-herd portfolios, and Panel B shows the results from *GTW* regressions for both herd and non-herd portfolios. In Panel A, in the *IHM* regressions gradient boosting seems to rank market risk premiums on top in both herd and non-herd portfolios, regardless of whether the factors are equally-weighted or value-weighted. Gradient boosting seems to provide better results in terms of R2 (22.9% vs 6.3%), MSE (0.026 vs. 0.032), MAE (0.095 vs. 0.112), and Median AE (0.056 vs. 0.072). However, using value-weighted factors, LASSO ranks market risk premium as the second feature in non-herd portfolios. In other regressions, LASSO ranks market risk premium on top and concurs with gradient boosting regressions, although gradient boosting seems to perform much better than LASSO as the average R^2 is 0.25 vs. 0.05, and the model accuracy measures concerning errors are lower in gradient boosting.

In Panel B, using equally-weighted factors in *GTW* regressions, both gradient boosting and LASSO rank market risk premium on top in both herd and non-herd portfolios. However, using value-weighted factors, these results do not hold in all the regressions. In non-herd portfolios, using value-weighted factors in performance tests, the market risk premium is ranked as the second most important feature by both gradient boosting and LASSO. This result concurs with those obtained from the *IHM* measure in Panel A and those from random forest regressions implemented in the main analysis, suggesting that using equally-weighted factors in these performance tests yields more consistent results than using value-weighted factors. Consistent with the results from Panel A, gradient boosting seems to be a better choice of model as it has a higher R2 and lower average errors than observed in LASSO model.

Appendix Table 6 provides a summary of these model accuracy statistics including the results from random forest regressions. The table also reports mean and standard deviation of MAE from cross validation of each machine learning model. For the cross validation, the scikit-learn library in Python is used; the program makes the MAE negative since it is maximized instead of minimized in the analysis. This means that the smaller values indicate a better model fit – the model is perfect if MAE is 0. In Panel A, among the machine learning models examining the determinants of herding, random forest is superior as it has a higher R2 and a lower Median AE than gradient boosting and LASSO models, but MAE seems to be lower in gradient boosting. The mean negative MAE from the cross validation is similar in random forest and gradient boosting models, although random forest is doing slightly better. We observe less negative mean MAE in the regression using *GTW* as a herding measure.

In Panel B, examining the models testing portfolio performance of *IHM* measures, random forest and gradient boosting produce similar results for all accuracy statistics including cross validation, but MAE is lower in random forest (0.009 vs. 0.026). LASSO seems to perform less than the other models. The results are similar in the other panels. Random forest and gradient boosting tend to produce similar results, although random forest performs slightly better. LASSO is ranked as the third model according to the accuracy statistics and cross validation. The R2 is around 22.6% and mean negative MAE is around − 0.102 across random forest and gradient boosting models.

In Appendix Table 7, grid search method is applied to optimize hyperparameters and hence to find the best parameters in the regression models. The cross-validation technique is used to account for overfitting, when the model performs well in training dataset but not in test dataset. The analysis shows results from hyper parameter tuning with 5-Fold CV implemented in scikit-learn. For the determinants of herding (Panel A) and portfolio performance analysis (Panels B and C), Appendix Table 1 shows the result for the best performing model with i) the max number of levels in each decision tree (max depth), ii) minimum number of data points placed in a node before the node is split (minimum sample split), iii) the number of trees, (n estimators), iv) grid search score, and v) average error for the best performing model. In both panels, max depth of 11 in the decision tree is consistent. The minimum sample split of 2 seem to be a valid choice. The number of estimators vary between 400 and 550 depending on the herding measure (*IHM* or *GTW*), whether the investors exhibit herding, and whether the portfolios are equally-weighted or value-weighted in the performance tests. The average grid score for the best model is 0.310 and the average error is around 0.090 in the models.

6. Conclusions and implications

Investors seem to have relative wealth concerns; they are concerned with the correlation of their portfolio returns with the returns of other investors, so-called herding, which biases their portfolio choice. This paper examines the determinants of herding at both the stock and individual investor levels. Particularly, determinants of herding at the individual investor level are scarce, and this paper contributes to the existing studies by exploring the predictive power of widely used portfolio, firm, and investor characteristics by implementing machine learning algorithms. At the stock level, the disposition effect and the attention effect seem to explain herding behavior. Herding seems to be greater in large firms, in firms with low book-to-market value, and in urban firms, indicating that information asymmetry is not the reason behind such behavior. An information-based explanation would suggest greater herding, particularly in small and rural firms.

At the investor level, the results from SHAP analysis show that high values of the cumulative number of buys and portfolio values reduce the prediction of herding behavior, while high values of portfolio return lead to a small increase in herding. Regarding the firm characteristics in the portfolios of individuals, the high values of liquidity and MTB seem to reduce the prediction of herding behavior.

The results also show that herd portfolios do not outperform market or non-herd portfolios, suggesting that herding is a behavioral bias. The most predictive factor in these performance regressions is market risk premium, and SHAP values consistently indicate that high values of market risk premium increase the prediction of excess returns on individual portfolios. Using equally-weighted factors, these results are robust and consistent; however, the SMB factor in some value-weighted factor models has more predictive power than the market risk premium.

Taken together, the results suggest that herding is a systematic behavioral bias, leading to short-term mispricing in stocks. Thus, any potential decrease in herding behavior would decrease the degree to which stock prices deviate from their fundamental values and subsequent return reversals. Further research could examine herding behavior within a social network and explore how individual investors within the network react to firm-specific events. Moreover, using machine learning algorithms can help explore the determinants of herding with the most predictive power among many candidates and can guide a more predictive model. However, the credibility of the results may improve substantially by a closer connection with economic theory, although theory is not immune to its own problems. Thus, in further research, the outcome of ML algorithms can be combined with economic theory that explains the fundamental properties of herding to draw strong conclusions about the general principles of such trading behavior.

CRediT authorship contribution statement

Taylan Mavruk is responsible for Conceptualization, Methodology, Software, Data curation, Formal analysis, Writing – original draft preparation, Visualization, Investigation, Software, Validation, Writing – review & editing.

Data Availability

The authors do not have permission to share data.

Acknowledgement

I am thankful to Saif Ullah and Turki Alshammari for encouraging comments at Vietnam Symposium in Banking and Finance 2021. I thank Xindi He and other seminar participants at Southwestern Finance Association 2021 Annual Meeting for valuable comments and suggestions. I would also thank the editor, John W. Goodell, the guest-editor, Sabri Boubaker, and two anonymous reviewers for the Research in International Business and Finance for insightful comments, which have helped developing the paper. Financial support by VINNOVA (grant 2010-02449) is gratefully acknowledged.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ribaf.2022.101740.

References

Barber, B., Odean, T., 2008. All that glitters: the effect of attention on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21, 785–818.
Barber, B.M., Odean, T., Zhu, N., 2009. Do retail trades move markets? Rev. Financ. Stud. 22, 151–186.
Bellando, R., 2010. Measuring Herding Intensity: A Hard Task., HAL Working Paper, 26 pages.
Bikhchandani S., Sharma, S., 2001. Herd Behaviour in Financial Markets. IMF Staff Pap. 47, 279–310.
Breiman, L., 2001. Random forests. Mach. Learn 45, 5–32.
Carhart, M.M., 1997. On persistence in mutual fund performance. J. Financ. 52, 57–82.
Chang, E.C., Cheng, J.W., Khorana, A., 2000. An examination of herd behavior in equity markets: an international perspective. J. Bank. Financ. 24, 1651–1679.
Chen, R.-C., Dewil, C., Huang, S.-W., Caraka, R.E., 2020. Selecting critical features for data classification based on machine learning methods. J. Big Data 7, 1–26.
Choi, N., R. Sias, W., 2009. Institutional industry herding. J. Financ. Econ. 94, 469–491.
Cipriani, M., Guarino, A., 2014. Estimating a structural model of herd behavior in financial markets. Am. Econ. Rev. 104, 224–251.
Dorn, D., Huberman, G., Sengmueller, P., 2008. Correlated trading and returns. J. Financ. 63, 885–920.
Duan, Y., Goodell, J.W., Li, H., Li, X., 2022. Assessing machine learning for forecasting economic risk: evidence from an expanded Chinese financial information set. Financ. Res. Lett. 46, 102273.
Feng, L., Seasholes, M., 2005. Do investor sophistication and trading experience eliminate behavioral biases in financial markets? Rev. Financ 9, 305–351.
Frazzini, A., 2006. The disposition effect and underreaction to news. J. Financ. 61, 2017–2046.
Frey, S., Herbst, P., Walter, A., 2014. Measuring mutual fund herding – a structural approach. J. Int. Financ. Mark. Inst. Money 32, 219–239.
Giannetti, M., Laeven, L., 2016. Local ownership, crises, and asset prices: evidence from US mutual funds. Rev. Financ 20, 947–978.
Giannetti, M., Simonov, A., 2006. Which investors fear expropriation? Evidence from investors' portfolio choices. J. Financ. 61, 1507–1547.

Goodell, J.W., Kumar, S., Lim, W.M., Pattnaik, D., 2021. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. J. Behav. Exp. Financ. 32, 100577.
Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. Am. Econ. Rev. 85, 1088–1105.
Grömping, U., 2009. Variable importance assessment in regression: linear regression versus random forest. Am. Stat. 63, 308–319.
Hall, M.A., 1999. Correlation-based feature selection for machine learning. PhD Thesis. The University of Waikato, New Zealand.
Hamberg, M., Mavruk, T., Sjögren, S., 2013. Investment allocation decisions, home bias and the mandatory IFRS adoption. J. Int. Money Financ. 36, 107–130.
Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer Series in Statistics. Springer New York Inc., New York, NY, USA.
Heimer, R.Z., 2016. Peer pressure: social interaction and the disposition effect. Rev. Financ. Stud. 29, 3177–3209.
Hirshleifer, D., Hong, S.H., 2003. Herd behaviour and cascading in capital markets: a review and synthesis. Eur. Financ. Manag. 9, 25–66.
Hirshleifer, D., Teo, S.H., 2009. Thought and behavior contagion in capital markets. In: Reiner Schenk-Hoppé, K., Hens, T. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. Elsevier, pp. 1–56.
Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. Rev. Financ. Stud. 33, 2019–2133.
Kim, K.A., Nofsinger, J.R., 2005. Institutional herding, business groups, and economic regimes: evidence from Japan. J. Bus. 78, 213–242.
Kumar, A., Lee, C., 2006. Retail investor sentiment and return comovements. J. Financ. 61, 2451–2486.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1992. The IMPACT OF INSTITUTIONAL TRADING ON STOCK Prices. J. Financ. Econ. 32, 23–43.
Lindblom, T., Mavruk, T., Sjögren, S., 2017. Proximity Bias in Investors' Portfolio Choice. Palgrave Macmillan Publishers Ltd., UK. ISBN: 978-3-319-54761-9.
Lindblom, T., Mavruk, T., Sjögren, S., 2018. East or west, home is best: the birthplace bias of individual investors. J. Bank. Financ. 92, 323–339.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30, 4768–4777.
Mavruk, T., 2021. Cross-regional differences in news tone and local stock ownership. Manag. Sci. 67, 2657–3320.
Merli, M., Roger, T., 2013. What drives the herding behavior of individual investors? Finance 34, 67–104.
Oehler, A., 1998. Do mutual funds specializing in German stocks herd? Finanz. PFM 12, 452–465.
Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. Am. Econ. Rev. 80, 465–479.
Shapley, L.S., 1953. A value for n-person games. In: Kuhn, Harold W., Tucker, Albert W. (Eds.), Contributions to the Theory of Games Ii. Princeton: Princeton University Press, USA, pp. 307–317.
Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long: theory and evidence. J. Financ. 40, 777–790.
Shiller, R., 1984. Stock market and social dynamics Brook. Pap. Econ. Act., 2, pp. 457–510.
Sias, R.W., 2004. Institutional herding. Rev. Financ Stud. 17, 165–206.
Strumbelj, E., Kononenko, I., 2010. An efficient explanation of individual classifications using game theory. J. Mach. Learn. Res. 11, 1–18.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. 58, 267–288.
Trueman, B., 1994. Analyst forecasts and herding behavior. Rev. Financ. Stud. 7, 97–124.
Tversky, A., Kahnemann, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.
Welch, I., 2000. Herding among security analysts. J. Financ. Econ. 58, 369–396.
Wermers, R., 1999. Mutual fund herding and the impact on stock prices. J. Financ. 54, 581–622.