Exploring style herding by mutual funds*

Caterina Santi[†] Remco C.J. Zwinkels[‡]

March 2023

Abstract

We study intentional herding in investment styles by mutual funds, and its consequences. We find that style herding is significant and persistent. Herding tends to increase after periods of high market volatility and decrease with sentiment, consistent with the intentional character of herding. Furthermore, we find that herding is related to changes in market dynamics. Finally, we find that herding in certain styles tends to temporarily increase mutual funds' performance, whereas it reduces flows. Overall, the results illustrate that intentional herding in styles is prevalent and has important consequences for market dynamics, fund managers, and investors.

Keywords: herding, mutual funds, asset pricing.

JEL classification: C32; G12; G23.

*The authors gratefully acknowledge the comments and suggestions from the editor, two anonymous reviewers, Fulvio Corsi, Teodor Dyakov, Juri Marcucci, Roberto Renò, and Mathijs van Dijk. We would like also to thank the organizers of the 2018 Asian meeting of the Econometric Society, the participants at the 2018 China meeting of the Econometric Society, and the 5th Italian Econometric Society Ph.D. Workshop in Econometrics and Empirical Economics. The usual disclaimer applies.

[†]HEC Liège - Management School of the University of Liège – caterina.santi@uliege.be

[‡]Vrije Universiteit Amsterdam and Tinbergen Institute – r.zwinkels@vu.nl

1 Introduction

The share of U.S. equity held by mutual funds in the United States has reached 32% at the end of 2021 (Investment Company Institute, 2022). A common perception sustains that mutual fund herding may drive stock prices away from fundamentals (Dasgupta et al., 2011a; Brown et al., 2014; Di Guilmi et al., 2014) leading to excess volatility and market fragility. In particular, career-concerned portfolio managers tend to imitate past trades. This tendency is exploited by security dealers and thus affects prices (Dasgupta et al., 2011b). Moreover, Gong et al. (2016) show that correlated fund flows lead to skewed distributions, and Deng et al. (2018) have shown that mutual fund herding amplifies stock price crash risk afterward.

The main objective of our paper is to address two fundamental subsequent questions: Do mutual funds herd across investment styles? And: What are the consequences of herding for market stability and fund management? Different from previous research, which focuses on industry styles (see e.g. Choi and Sias, 2009; Celiker et al., 2015), we explore herding in investment styles, such as market, size, value, and momentum. This is a relevant extension to the literature because funds are typically classified by their investment style since the identification of a fund in styles simplifies problems of choice of individual investors, and it allows one to easily evaluate the performance of the fund through comparison with a benchmark specific to the style. There is evidence that both individual investors and mutual funds allocate more at the style level than at the stock level (see e.g., Froot and Teo, 2008). Moreover, funds tend to increase their exposure to styles that are expected to outperform; (Frijns et al., 2016). Related to this, Cooper et al. (2005) find evidence of mutual funds changing their names to take advantage of the most popular investment styles, not necessarily paired with changing their actual holdings. Furthermore, Fricke (2019) documents that mutual funds show significant portfolio overlap. Given this context, it is relevant to understand the reasons that lead mutual fund managers to relatively change the focus from one style to another.

This paper contributes to the empirical literature on mutual fund herding in styles. We extend the method to detect herding developed by Hwang and Salmon (2004) to the herding behavior of mutual fund managers. In particular, Hwang and Salmon (2004) use the cross-sectional standard deviation of individual assets' betas to measure herding towards certain industries or styles in the market including the market index itself. In this paper, we instead consider mutual fund portfolios as individual assets and hence the cross-sectional standard deviation of their betas could naturally measure the herding behavior of mutual funds in the sense of Hwang and

Salmon (2004). This method is different from some popular measures of fund herding (see e.g. Lakonishok et al., 1992; Sias, 2004), and it provides a better understanding of fund herding. Specifically, the Hwang and Salmon (2004) herding measure can disentangle intentional herding from common movements in asset returns induced by common fundamental information. Intentional herding occurs when investors intentionally imitate the actions of others because of fads (Friedman, 1984), reputation concerns (Scharfstein and Stein, 1990; Trueman, 1994; Popescu and Xu, 2018), preference for certain asset characteristics (Falkenstein, 1996), and because other funds may have more or more accurate information (Banerjee, 1992; Bikhchandani et al., 1992; Welch, 1992). Differently, the Lakonishok-Shleifer-Vishny (hereafter LSV) measure is not able to distinguish pure imitation behavior from switches in investment decisions justified by fundamentals.[1] Choi and Skiba (2015), using the Sias (2004) measure, find that institutional investors' herding is more prevalent in markets characterized by a high level of information transparency. This suggests that institutional investors' herding behavior detected with the Sias method may be driven by fundamental information. Another advantage of our approach is that it allows us to better understand mutual fund herding at a higher frequency.[2] Specifically, herding measures based on portfolio holdings (see e.g. Lakonishok et al., 1992; Sias, 2004) can detect mutual fund herding at a quarterly frequency as these data are published only quarterly in the U.S. Since our herding measure uses mutual funds returns data, we can detect style herding at a monthly frequency.[3] Moreover, unlike previous studies which adopt the LSV and Sias herding measures, we can study the evolution of mutual fund style herding over time rather than averaged over the full sample period. This allows for a more in-depth analysis of the drivers and consequences of herding. Furthermore, by studying style herding rather than overall herding we better exploit the cross-sectional variation in investment strategies.

Our paper is also related to the literature on factor crowding. Indeed, because of the increasing level of mutual fund ownership of stocks, style herding by these investors may contribute to

[1] Bikhchandani and Sharma (2000) argue that the LSV measure is not linked to theoretical models, and therefore lacks an inter-temporal dimension of herding. It also uses a binary measure of buys versus sells. As such, the measure may produce biased results (Wylie, 2005; Frey et al., 2014).

[2] Kremer and Nautz (2013) have shown that herding measures are severely affected by data frequency, in particular, they find that institutions herd daily by using daily investor-specific data that directly identify institutional transactions in the German stock market.

[3] Other herding measures that employ returns data are presented in Christie and Huang (1995); Chang et al. (2000); Chiang and Zheng (2010); Duygun et al. (2021); Hwang et al. (2021). However, they are not directly comparable to the herding measure by Hwang and Salmon (2004). Recently, Bohl et al. (2017) has challenged the herding measure by Chang et al. (2000). They show that their test fails to accept the hypothesis of herding too often. They propose a modified test that models the data-generating process under the alternative hypothesis suggested in Hwang and Salmon (2004) that herding arises if the betas shrink towards one. The empirical analysis confirms the misleading implications of Chang et al. (2000).

a significant increase in the inflows of a certain factor. The literature has widely documented the value added by factor investing (Clarke et al., 2005; Dimson et al., 2017; Kim et al., 2017; Bergeron et al., 2018). However, there is an increasing concern that there are only a few factors available and too many investors to chase them (Giamouridis, 2017). Recently, Dichtl et al. (2019) have suggested avoiding factors that exhibit crowding. Instead, Baltas (2019) has found that crowding in divergence factors, such as momentum, is likely to decrease performance. However, crowding in convergence factors, such as value, may increase performance.

We employ the CRSP Survivor-Bias-Free U.S. mutual fund database over the period September 1998 – September 2022 to study style herding in U.S. domestic equity funds. We use four benchmark portfolios from Kenneth French's Web site for the style returns of the market (r_m), size (SMB), value (HML), and momentum (UMD) factors.

We observe several interesting results. First, we find that mutual fund herding in styles is significant, persistent, and independent from fundamentals (Fama and French, 2015). Herding behavior of mutual funds has also been confirmed using the Sias measure by Celiker et al. (2015), and Popescu and Xu (2018) among others. In contrast, studies based on the LSV measure find limited evidence of institutional herding (see e.g. Grinblatt et al., 1995; Wermers, 1999, and others). We also observe that mutual fund herding towards the size factor tends to be less volatile than herding towards the other styles. A potential explanation might be that information asymmetry is larger for small firms than for large firms, this may lead mutual funds to imitate other funds when choosing their exposure to the size risk factor (Wermers, 1999; Sias, 2004). This result is consistent with Deng et al. (2018), who show that mutual fund herding is associated with a poor information environment and low disclosure quality.

Second, we show that an increase in market uncertainty leads mutual fund managers to rely more on sources of information not related to fundamentals, increasing herding behavior, consistent with Di Guilmi et al. (2014). In addition, we find that mutual fund herding in styles is negatively associated with market sentiment. This is explained by the fact that sentiment induces investors to rely more on private information (Keshk and Wang, 2018). Further, we report that mutual fund herding towards the market portfolio significantly increases when economic conditions worsen. We also find that mutual fund herding in styles increases when the ratio of actively versus passively managed funds decreases. Differently from Hwang and Salmon (2004) and D'Arcangelis and Rotundo (2019), who show that crises are turning points in investors' herding behavior, we do not find any significant difference in mutual fund style herding during the burst of the Dotcom bubble and the global financial crises. Only mutual fund herding towards value was significantly lower during the first wave of the COVID-19 pandemic.

Third, we report that mutual fund herding in styles affects factor returns. Specifically, we find a direct effect of mutual fund herding towards size on factor returns. Moreover, we observe that herding impacts the effect of mutual fund flows on factor returns, and it may also impact their autocorrelation structure. Furthermore, the results indicate that mutual fund herding towards the market can predict the returns of the market portfolio out-of-sample. Thus, the presence or absence of mutual fund herding in styles is valuable information for individual investors as well as practitioners since it can predict factor returns. These findings are consistent with the hypothesis that our measure captures intentional herding, unrelated to fundamentals since information-based herding should not lead to predictability.

Finally, and most importantly for the manager, mutual fund herding towards size and value factors generally increases the value-weighted performance of the mutual fund industry, while adverse herding towards these styles decreases it. This suggests that it may be beneficial for mutual fund managers to engage in herding towards size and value since it temporarily increases performance. Differently, mutual fund herding towards the other factors does not affect performance. These results are consistent with Koch (2017), who does not find evidence of outperformance of herding funds, and Jiang and Verardo (2018) who show that herding behavior is prominent among unskilled funds. Moreover, mutual fund-style herding tends to reduce flows to the mutual fund industry, suggesting that investors do not fully recognize the information embedded in herding.

All in all, our results have important practical implications as mutual fund style herding produces consequences for both market stability and fund management. These findings are of interest to fund managers, as they indicate that style herding is not necessarily beneficial to the fund because performance is at best improved temporarily, whereas inflows tend to decrease. As such, the results might be of interest also to investors attempting to select an optimal fund manager. From a practitioner's perspective, our results can inform the construction of multifactor portfolios indeed the presence or absence of mutual fund herding can be used to predict factor returns. Furthermore, we show that herding is more likely to occur after particular states of the market hence the tendency of mutual funds to herd in styles is predictable. Finally, results are of interest to regulators interested in market stability because we find evidence of herding-induced demand pressure from mutual funds.

2 Data and methodology

2.1 Data and descriptive statistics

We employ the CRSP Survivor-Bias-Free U.S. Mutual Fund Database. The database includes funds of all investment objectives, principally equity funds, taxable and municipal bond funds, international funds, and money market funds. The current paper focuses on U.S. domestic equity mutual funds for three main reasons: first, daily data for these funds are available since the end of the 1990s. Second, equity funds constitute the largest market share of the mutual fund market in the United States (55 percent end-2021, Investment Company Institute, 2022). Third, focusing on funds active only in a specific market allows a clearer identification of the benchmark portfolios for the style returns and it allows us to investigate mutual fund herding in styles. We use daily return data from September 1998 to September 2022, and we require only that a fund has no missing daily observations on a given month to be included in the sample. We obtain an unbalanced panel of 18,212 U.S. domestic equity mutual funds. We also employ monthly return, turnover ratio ($turn$), management fees ($fees$), and total net asset (TNA) data. We define net flows to fund i at time t ($Flow_{i,t}$) as the net growth in fund assets beyond reinvested dividends (Sirri and Tufano, 1998). Formally, it is computed as follows:

$$Flow_{i,t} = \frac{TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1}}{TNA_{i,t-1}}, \tag{1}$$

where $TNA_{i,t}$ is fund i's total net assets at time t, and $R_{i,t}$ is the fund's return over the period from $t-1$ to t. In Table 1, we report summary statistics for the variables of interest for all sampled funds. We have an average of 8,217 sampled funds per month. The average monthly fund return is 0.73%, the average Sharpe ratio is equal to 0.47%, and the average fund flows are 0.17% of the previous month's TNA. The average size of the funds is \$518.67 million while the median size is lower with a value of \$43.7 million. Management fees are computed as the ratio of the management fees in \$ and average net assets in \$, and average 0.43.

The return on the market portfolio (r_m), size factor (SMB), value factor (HML), and momentum (UMD) are obtained from Kenneth French's data library.[4] In Panel A of Table 2, we present descriptive statistics for the factor returns. In line with Fama and French (1993), we observe that small-cap portfolios outperform their large-cap counterpart, and value portfolios tend

[4] See the Kenneth French' Web site: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, for a description of the procedure used to construct the portfolios.

Table 1: Fund characteristics

	Mean	St.Dev	p05	p50	p95
Return (in percent)	0.727	4.465	-7.713	1.224	6.840
Sharpe ratio (in percent)	0.471	2.123	-3.059	0.636	3.673
TNA (in millions)	518.668	2,805.658	0.100	43.700	1,816.800
Flows (in percent)	0.173	0.508	-0.565	0.112	0.946
Turnover ratio	0.433	0.141	0.147	0.429	0.657
Management fees	0.431	0.095	0.172	0.443	0.558
Number of funds (per month)	8,217	1,246	6,021	8,655	10,089

Notes: This table presents summary statistics on value-weighted characteristics of the sampled mutual funds. We present values for the mean, the standard deviation, and the 5th (p05), 50th (p50), and 95th (p95) percentiles along the sampled period. We report statistics for the monthly return, Sharpe ratio, total net assets (TNA), fund flows, turnover ratio, and management fees. The bottom row reports the summary statistics of the number of funds per month.

to have higher returns than growth portfolios. We also report that the momentum factor has an average monthly return of 0.30% in the period under study. Panel B presents the pair-wise correlation matrix of the factor returns; the correlations are relatively low.

2.2 Measuring herding

We extend the approach by Hwang and Salmon (2004) to detect herding of mutual fund managers towards investment styles. In particular, the original contribution by Hwang and Salmon (2004) uses the cross-sectional standard deviation (CSSD) of individual assets' betas to measure herding towards certain industries or styles in the market including the market index itself. In this paper, we treat mutual fund portfolios as individual assets. Thus, the CSSD of their betas can naturally be used to measure the herding behavior of mutual funds in the sense of Hwang and Salmon (2004).

Specifically, style investing generates common factors in the return of the assets that belong to the same class regardless of fundamentals (Barberis and Shleifer, 2003). Hence, the expected excess return of a fund i in period t may be written as:

$$E_t[r_{i,t}] = \alpha_{i,t} + \sum_{k=1}^{K} \beta_{i,k,t} E_t[r_{k,t}], \tag{2}$$

in which $E_t[r_{i,t}]$ is the expected excess return of fund i at time t, the term $\alpha_{i,t}$ represents the out- or underperformance of fund i, $\beta_{i,k,t}$ and $E_t[r_{k,t}]$ capture the exposure to and the expected excess return of investment style k, respectively. We argue that mutual fund herding within a certain style biases the funds' betas, such that the CSSD of the betas will be smaller than it would be in the absence of herding. Similarly to Hwang and Salmon (2004), we assume that the funds' biased betas satisfy the following relationship in the presence of mutual fund herding towards a certain style k:

Table 2: Benchmark portfolios

	r_m	SMB	HML	UMD
Panel A: Descriptive Statistics				
Mean	0.613	0.264	0.065	0.301
St.Dev	4.616	3.141	3.492	5.229
Min	-17.230	-15.350	-13.970	-34.300
p05	-8.126	-4.194	-4.736	-8.198
p50	1.190	0.190	-0.170	0.490
p95	7.594	5.088	6.612	7.590
Max	13.650	18.340	12.750	18.200
Panel B: Correlation matrix				
r_m	1.000			
SMB	0.250**	1.000		
HML	−0.073	0.016	1.000	
UMD	−0.339***	−0.003	−0.250***	1.000

Notes: This table presents summary statistics and correlations for the monthly returns (in percent) of the benchmark portfolios, namely market portfolio (r_m), size factor (SMB), value factor (HML), and momentum factor (UMD) retrieved from Kenneth French's data library.

$$\beta^b_{i,k,t} = \beta_{i,k,t} - h_{k,t}(\beta_{i,k,t} - \mathbf{E}_c[\beta_{i,k,t}]), \tag{3}$$

where $\beta^b_{i,k,t}$ represents the biased beta of fund i, $\mathbf{E}_c[\beta_{i,k,t}]$ captures the cross-sectional expected beta, and $h_{k,t}$ is mutual fund herding towards style k, $h_{k,t} \leq 1$. When $h_{k,t} = 0$ there is no mutual fund herding towards style k and no bias in the betas. When $h_{k,t} = 1$ there is perfect herding towards the style, meaning that all the funds move in the same direction with the same magnitude as the style portfolio, $\beta^b_{i,k,t} = \mathbf{E}_c[\beta_{i,k,t}]$. Instead, positive values of $0 < h_{k,t} < 1$ suggest the presence of some degree of herding toward style k hence the factor loading is too high (low) relative to its value in the absence of herding. When mutual funds are not confident in the reliability of their signals, they tend to imitate the investment behavior of other funds which might have more or more accurate information. Therefore, the CSSD of the betas of the mutual fund industry will be smaller with herding than in the absence of herding. In general, when the performance of a style increases significantly, investors will try to buy underperforming assets (relative to the style portfolio) and sell overperforming assets. Now consider for instance a mutual fund with a beta less than $\mathbf{E}_c[\beta_{i,k,t}]$. When mutual fund managers herd towards the style the fund's beta will be biased and in particular it will be higher. Therefore, the fund will look riskier than it should. Differently, a mutual fund with a beta greater than $\mathbf{E}_c[\beta_{i,k,t}]$ would have a lower beta in case of herding, therefore the fund will look less risky than it should. From this discussion, it is evident how mutual fund herding may lead to misallocation of resources. The existence of mutual fund style herding implies the existence of adverse herding, which is obtained in this setting by allowing $h_{k,t} < 0$. Adverse herding allows the market to return to equilibrium. When mutual funds are overconfident in their signals, the CSSD of the betas will be higher than in equilibrium. In particular, betas larger than $\mathbf{E}_c[\beta_{i,k,t}]$ will become higher and betas less than $\mathbf{E}_c[\beta_{i,k,t}]$ will become lower.

From Eq. (3) we can compute the CSSD of the biased betas as:

$$\sigma(\beta^b_{i,k,t}) = \sigma(\beta_{i,k,t})(1 - h_{k,t}). \tag{4}$$

Hwang and Salmon (2004) point out that "the empirical evidence of time-varying betas may derive from behavioral anomalies such as herding, rather than from fundamental changes". Indeed, $\sigma(\beta_{i,k,t})$ should not vary substantially at least in the short run unless the capital structure of firms within the market changed dramatically. However, the assumption of a stable $\sigma(\beta_{i,k,t})$ might not be realistic for mutual funds. For this reason, we estimate the transition matrix of the funds Lipper classification.[5] We observe that the elements on the main diagonal of the transition matrix, which represent the probability of a fund being classified as investing in a specific style next quarter given that it has received the same style classification this quarter, have an average of 0.9583. The first quartile is 0.9446 while the third quartile is 0.9978. Thus, it is safe to conclude that $\sigma(\beta_{i,k,t})$ varies very slowly in the long run, also for mutual funds.

Given that $\sigma(\beta_{i,k,t})$ is allowed to vary only slowly and the unobservable herding component is assumed to follow a zero mean dynamic process such as an $AR(1)$, after a log transformation we can write the model in the following state-space form:

$$log(\sigma(\beta^b_{i,k,t})) = \mu_k + H_{k,t} + \vec{\theta} X_t + \epsilon_{k,t}, \qquad \epsilon_{k,t} \sim iid(0, \sigma^2_{k,\epsilon}), \tag{5}$$

with

$$H_{k,t} = \phi_k H_{k,t-1} + \eta_{k,t}, \qquad \eta_{k,t} \sim iid(0, \sigma^2_{k,\eta}), \tag{6}$$

and

$$H_{k,0} \sim N\left(0, \frac{\sigma^2_{k,\eta}}{1 - \phi^2_k}\right), \tag{7}$$

where $\mu_k = \mathbf{E}[log(\sigma(\beta_{i,k,t}))]$, $H_{k,t} = log(1 - h_{k,t})$, and X is a matrix of control variables. A

[5] Lipper classifications are available quarterly and they are based on how the fund invests. Specifically, Lipper runs the actual holdings of the fund through an internal model to determine market cap and style versus a benchmark. Classifications are based on scores for a specific set of portfolio characteristics (P/E, P/B, etc).

further requirement of this analysis is the stationarity of the herding process, hence we impose $|\phi_k| \leq 1$. Equations (5) and (6) are respectively known as measurement and transition equations. Eq. (7) represents the initial condition.

2.3 Estimation details

We focus on herding by mutual fund managers towards the market, size, value, and momentum factors. As a first step, we use daily data over monthly intervals to estimate the Carhart (1997) model for each mutual fund:

$$r_{i,t} = \alpha_{i,t} + \beta_{i,m,t} r_{m,t} + \beta_{i,smb,t} SMB_t + \beta_{i,hml,t} HML_t + \beta_{i,umd,t} UMD_t + \epsilon_{i,t}, \tag{8}$$

where $r_{i,t}$ denotes the excess return of fund i in day t, and $r_{m,t}$, SMB_t (Small minus Big), HML_t (High minus Low) and UMD_t (Up minus Down) represent the daily market excess return, size, value, and momentum factors in day t. The returns of these benchmark portfolios are used as style returns.

We follow Hwang and Salmon (2004) and use 1-month of data to estimate the betas using OLS.[6] The period of 1 month allows us to capture reasonably rapid changes in herding and at the same time obtain reliable estimates.[7] Previous research on institutional herding generally adopts data at a quarterly frequency; as such, they may underestimate herding if it occurs within shorter periods. To analyze the differences in the characteristics of funds with very high or low factor betas, in each month we test the difference in returns, flows, and total net assets of funds with above 90th percentile and below 10th percentile exposure to the risk factors in Eq. (8). In general, we observe that funds with factor betas higher than the 90th percentile have significantly higher returns (on average 55% of the sampled months) and flows (on average 82.3% of the sampled months) than funds with factor betas lower than the 10th percentile. Moreover, funds with high exposure to the size and value factors have significantly lower total net assets than funds with low exposure to these factors, respectively 98.30% and 54.2% of the sampled months. The opposite is observed for funds with high exposure to market and momentum risk

[6] Hwang and Salmon (2004) demonstrate that as long as the estimation error of the betas in Eq. (8) is not correlated with the error term in the measurement equation ($\epsilon_{k,t}$) and $H_{k,t}$, the herding measure will not be affected. The estimation error will simply increase the noise in the state space model making it more difficult to find significant estimates of ϕ. However, relative movements in the herding process should not be affected by the presence of the estimation error.

[7] We refrain from using non-parametric methods to estimate the betas at a daily frequency (Ang and Kristensen, 2012; Li and Yang, 2011) because beta estimates are data frequency dependent (Gilbert et al., 2014). In particular, estimates of the betas at a high frequency are subject to the effect of firms' opacity and to a greater noise which may make difficult the identification of intentional herding.

factors.

Table 3: Log-cross-sectional standard deviation of the betas

	Mean	St.Dev.	Skewness	Kurtosis	Jarque Bera Test (p-value)
Betas on r_m	-1.740	0.315	0.372**	0.332	0.020
Betas on SMB	-0.964	0.134	1.086***	1.888***	0.000
Betas on HML	-1.124	0.262	0.498***	0.025	0.003
Betas on UMD	-1.505	0.360	0.061	-0.625***	0.090

Notes: This table presents the descriptive statistics of the log cross-sectional standard deviation of the estimated exposures to the four benchmark portfolios. We use daily fund return data from 1 September 1998 to 30 September 2022. For each month, daily returns of the market excess return (r_m), size factor (SMB), value factor (HML), and momentum factor (UMD) factors are employed to estimate betas of the factors on each sampled fund. We estimate the monthly betas with OLS. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Subsequently, we compute the CSSD of the estimated betas as:

$$\sigma(\hat{\beta^b_{i,k,t}}) = \sqrt{\frac{\sum_{i=1}^{N_t}\left(\hat{\beta^b_{i,k,t}} - \frac{\sum_{i=1}^{N_t}\hat{\beta^b_{i,k,t}}}{N_t}\right)^2}{N_t - 1}}, \quad k = \{m,\ smb,\ hml,\ umd\}, \tag{9}$$

where N_t is the number of funds in month t.[8]

Table 3 reports some statistical properties of the estimated log-CSSDs of the betas. The series of log-CSSDs have a negative average, meaning that the average CSSD of the betas is below one. The series of log-CSSDs are positively skewed except for the log-CSSD of the betas of the UMD factor. Moreover, the series of log-CSSD of the betas on SMB (UMD) shows positive (negative) kurtosis. The log-CSSD of the market and momentum betas do not deviate significantly from Gaussianity (1% confidence level).[9]

After having computed the CSSD of the betas, we estimate the parameters of the following state space model with conditional maximum likelihood:

$$\begin{cases} log(\sigma(\beta^b_{i,k,t})) = \mu_k + H_{k,t} + \theta_1 r_{m,t} + \theta_2 SMB_t + \theta_3 HML_t + \theta_4 UMD_t + \theta_5 log(\sigma_{m,t}) + \epsilon_{k,t}, \\ H_{k,t} = \phi_k H_{k,t-1} + \eta_{k,t}. \end{cases} \tag{10}$$

According to Barberis and Shleifer (2003), fund managers engage in style-based feedback trading, that is they tend to allocate more capital to styles that performed relatively better in the past while they tend to reduce exposure to styles with poor past performance. Thus, adding factor returns and market volatility to the model allows us to control for style-based feedback trading, and hence to detect intentional style herding. Indeed, if $H_{k,t}$ becomes insignificant when

[8] As a robustness check we also performed the analysis using the value-weighted CSSD. Results are qualitatively the same and reported in Appendix A.

[9] Note that the residuals of the Kalman filter estimates of style herding are normally distributed for all the four styles.

fundamental variables are included in the model, changes in $\sigma(\beta^b_{i,k,t})$ are driven by changes in fundamentals rather than herding. In particular, in Eq. (10) we control for the Carhart (1997) factor returns plus market volatility (σ_m) which is calculated as the sum of squared daily returns (Schwert, 1989).[10]

As a last step, we apply the Kalman filter and smoothing to estimate the latent herding component $H_{k,t}$.

3 Style herding by mutual funds

Table 4 reports the maximum likelihood estimates of the model parameters in Eq. (10) for the four styles, namely market, size, value, and momentum.

The most important result is that the estimates of the standard deviation of the disturbance term in the herding equation (σ_η) are highly significant. Therefore, we can conclude that there is mutual fund herding towards all the styles considered during our sample period. Furthermore, we find that the estimates of the parameter ϕ are large and significant meaning that mutual fund herding is highly persistent. Mutual fund herding towards the market portfolio factor shows the highest level of persistence. Instead, mutual fund herding towards the size factor shows the lowest volatility as compared to herding towards the other styles. This evidence is consistent with Wermers (1999), who finds higher levels of herding in small stocks. Fund managers probably have less accurate information on earnings from these companies and they are more prone to ignore their own information if it is different than the consensus opinion. Differently, mutual fund herding towards the momentum factor shows the lowest level of persistence and the highest level of volatility as compared to herding towards the other styles. This might be due to the high volatility in the UMD factor itself, in combination with the relatively high costs associated with momentum trading.

We also observe that the signal-to-noise proportions vary for mutual fund herding towards different styles. In particular, it is higher for herding towards the momentum factor. Herding towards the size factor explains around 16.3% of the total variability in $log(\sigma(\beta^b_{i,k,t}))$, while mutual fund herding towards the momentum factor 23.8%.

As we mentioned above, we include in our state-space model some fundamental variables. Indeed, we want to estimate a herding measure that is not affected by the fact that mutual

[10] Experiments with a version of the model without the SMB, HML, and UMD factors, or including the RMW and CMA factors, yields highly similar results; available on request.

Table 4: Maximum likelihood estimates

	h_m	h_{smb}	h_{hml}	h_{umd}
μ	-2.412^{***}	-0.968^{***}	-0.976^{***}	-1.499^{***}
	(0.212)	(0.074)	(0.128)	(0.168)
σ_ϵ	0.147^{***}	0.089^{***}	0.147^{***}	0.177^{***}
	(0.008)	(0.004)	(0.008)	(0.010)
ϕ	0.986^{***}	0.979^{***}	0.967^{***}	0.960^{***}
	(0.011)	(0.015)	(0.018)	(0.019)
σ_η	0.053^{***}	0.022^{***}	0.055^{***}	0.086^{***}
	(0.010)	(0.004)	(0.009)	(0.013)
r_m	0.571^{**}	-0.041	0.327	-1.045^{***}
	(0.255)	(0.150)	(0.255)	(0.317)
SMB	-0.380	-0.357^{*}	0.027	0.211
	(0.317)	(0.186)	(0.318)	(0.394)
HML	0.039	0.226	0.056	0.034
	(0.305)	(0.177)	(0.306)	(0.383)
UMD	-0.454^{**}	0.018	0.378^{*}	0.397
	(0.209)	(0.123)	(0.210)	(0.261)
$log(\sigma_m)$	-0.113^{***}	-0.006	0.021	0.002
	(0.015)	(0.008)	(0.015)	(0.019)
Signal to noise	0.169	0.163	0.211	0.238
ML values	358.150	520.899	358.997	290.495

Notes: This table presents the maximum likelihood estimates of the state space model parameters in Eq. (10). The series of 289 monthly CSSDs of betas is used to estimate the state-space model and then extract the latent herding component. Standard errors are reported in parenthesis. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5: Correlation of herding towards different styles

Variable	h_m	h_{smb}	h_{hml}	h_{umd}
h_m	1.000			
h_{smb}	0.812***	1.000		
h_{hml}	0.677***	0.729***	1.000	
h_{umd}	0.751***	0.617***	0.641***	1.000

This table presents the pairwise correlation matrix of the estimated series of mutual fund herding towards the market, size, value, and momentum styles. Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

funds might have similar trading behavior because they are following the same signals. We find that market uncertainty is significantly associated with the CSSD of the market betas. This implies that mutual fund managers tend to adjust their exposure to market risk in response to information embedded in the level of market uncertainty. Moreover, we find that also changes in UMD and market returns tend to lead mutual fund managers to adjust their exposure to market risk. The CSSD of the betas of SMB are negatively associated with SMB returns, and the CSSD of momentum betas are negatively associated with market returns. This is consistent with Cooper et al. (2004), who show that momentum profits depend on the state of the market. In line with Asness et al. (2013), who document a negative correlation between value and momentum returns, we observe that mutual fund managers adjust their exposure to the value risk factor following changes in momentum returns.

We report the correlation matrix of the four herding measures in Table 5. We can see that mutual fund herding towards the four styles is strongly positively and significantly correlated.

The highest level of correlation is reported for mutual fund herding towards the market portfolio and size factor (0.812). The lowest level of correlation is observed between h_{smb} and h_{umd} (0.617), however, it is still quite high. Thus, an increase in mutual fund herding towards a certain style is likely to be accompanied by an increase in herding towards the other styles, suggesting herding is a state variable.

Figures 1 to 4 show the evolution of mutual fund herding in styles $h_{k,t}$ ($h_{k,t} = 1 - e^{H_{k,t}}$) in the U.S. equity market. First, we observe several cycles of mutual fund herding and adverse herding in styles as h_k oscillates around its long-term average of zero. Second, h_k is always far less than unity indicating that there was never an extreme degree of herding towards any of the styles during our sample period. Third, we find evidence of mutual fund herding in styles when both the market is rising and when it is falling.

Interestingly, the four style herding measures show rather comparable trends. In the late 1990s the then-Federal Reserve Board chairman, Alan Greenspan warned about the irrational exuberance of the financial markets. The burst of the dot-com bubble occurred in March 2000 and it was followed by a recession in the United States from March to November 2001. In those years, we register significant adverse herding towards all the styles under study. This might be explained by the fact that the IT bubble was largely confined to a particular industry. Hence, we could expect industry herding in that period rather than factor herding.

Later, the 2000s were the decade of subprime borrowers. From 2004 to 2007 the worst loans were generated. In that period, we document adverse herding towards value and adverse herding towards momentum from January 2006 to August 2007. Adverse herding implies that mutual funds are overconfident in their own source of information which will cause the CSSD of the betas to be higher than in equilibrium. We conjecture that mutual funds, although uncertain about the overall performance of the market, may have been overconfident in their ability to pick up the "right" stocks for momentum and value investing.

Moreover, we register significant herding towards the market portfolio from February to September 2007, right before the Dow Jones Industrial Average hit its peak closing price on 9th October 2007. Also, existing home sales peaked in October and began to decline. In September 2008, the crisis hit its most critical stage. In that month, Lehman Brothers went bankrupt, Fannie Mae and Freddie Mac were taken over by the federal government, and Goldman Sachs and Morgan Stanley converted themselves into bank holding companies to increase their protection by Federal Reserve. From July to December 2008 we find significant adverse herding towards the

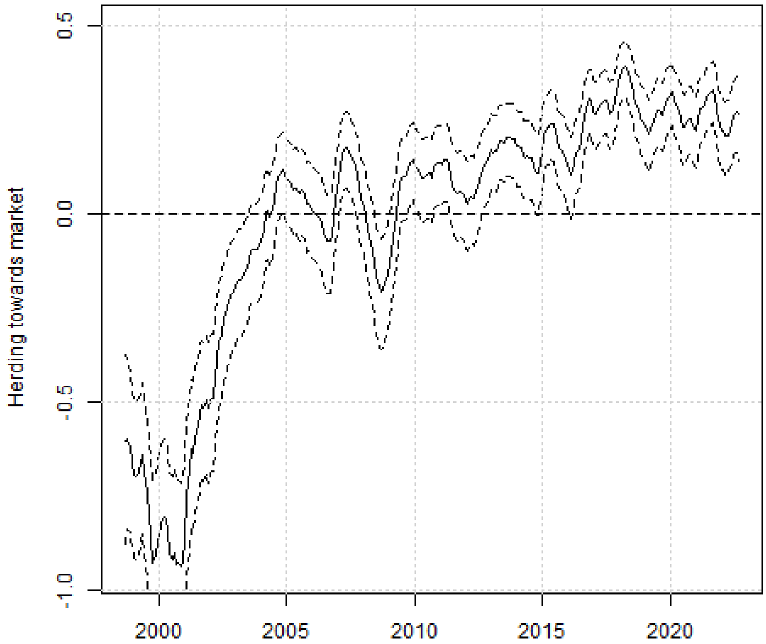

Figure 1: Herding towards the market portfolio. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

Figure 2: Herding towards the size factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

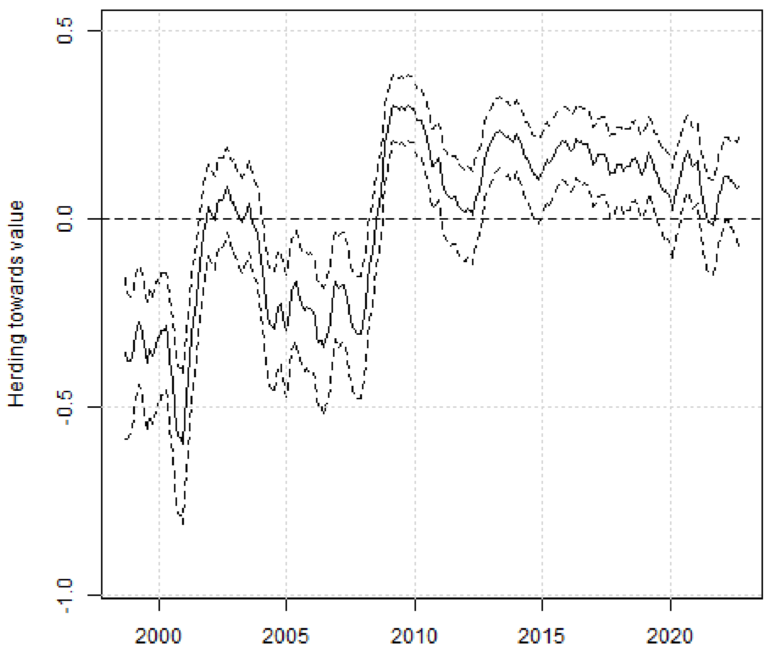

Figure 3: Herding towards the value factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

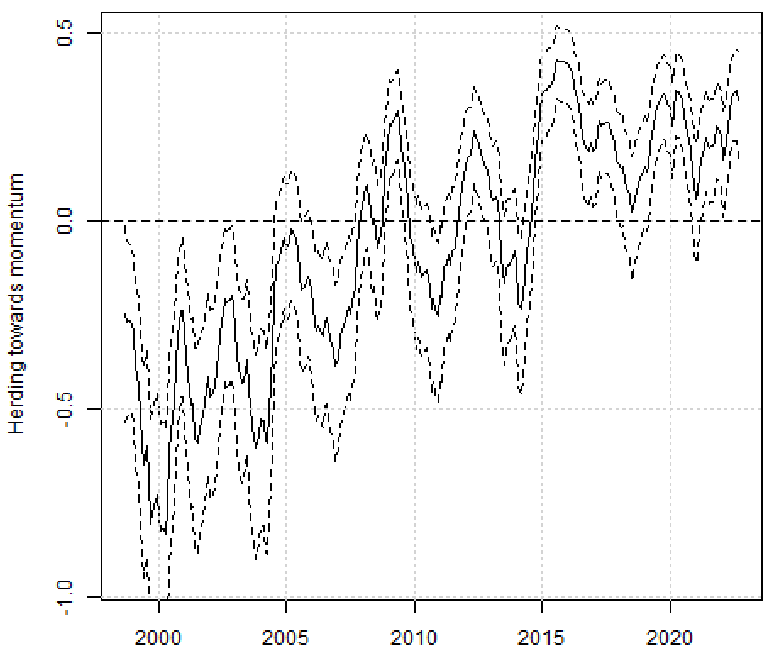

Figure 4: Herding towards the momentum factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

market portfolio which can be interpreted as a return towards the equilibrium. A few months later, from December 2008 to July 2009, instead, we register significant herding towards the momentum factor. In those months we also observe herding towards the size factor which lasts until November 2015. Fund managers are likely to have less reliable information on small firms which will make them more prone to follow the herd. This is especially true when the financial markets are going through a period of high uncertainty. Mutual fund herding towards the value factor is significantly positive from October 2008 to January 2011.

From 2009 until the end of the sampled period, mutual fund herding towards all the styles alternate periods in which it is significantly positive to periods in which it is not significant. We do not record significant adverse herding except for adverse herding towards momentum for a few months in 2011, and adverse herding towards size in 2022. Interestingly, we do not find any considerable change in mutual fund style herding during the COVID-19 pandemic.[11]

4 Determinants of style herding by mutual funds

Now that we have determined that style herding is significant and persistent, we examine the determinants of style herding by mutual funds. Specifically, we study whether market conditions, such as the overall market-wide investor sentiment as well as macroeconomic conditions, affect the intensity of mutual fund herding in styles. Given the intentional character of our style herding measure, we expect herding to decrease in sentiment and increase in uncertainty. We estimate the following model:

$$\begin{aligned} h_{k,t} =& \beta_0 + \beta_1 R_{m,t-6} + \beta_2 R_{smb,t-6} + \beta_3 R_{hml,t-6} + \beta_4 R_{umd,t-6} + \beta_5 \log(\sigma_{m,t-1}) + \\ & \beta_6 Sent_{t-1} + \beta_7 ADS_{t-1} + \beta_8 Active_{t-1} + \beta_9 Dotcom_t + \\ & \beta_{10} FinCrisis_t + \beta_{11} Covid_t + \epsilon_{k,t}, \end{aligned} \tag{11}$$

in which $k = \{m, smb, hml, umd\}$, $R_{m,t-6}$, $R_{smb,t-6}$, $R_{hml,t-6}$, and $R_{umd,t-6}$ are the cumulative returns of the previous six months of the market portfolio, the SMB, the HML, and the UMD factors. Market volatility is denoted by σ_m, we expect volatility to be positively related to herding. Indeed, uncertainty may lead mutual fund managers to rely more on other sources

[11] In Appendix B, we perform the analysis for the sub-samples of actively and passively managed funds. We find that the parameters' estimates of style herding in the two sub-samples are comparable. We observe that herding towards size and value is more volatile for active funds than passively managed funds. However, herding towards these styles by passively managed funds is more persistent. Moreover, mutual fund style herding follows a similar pattern independently from the fund management style (active or passive).

of information which may increase herding behavior. Furthermore, since intentional herding is an irrational behavior, we include investor sentiment ($Sent$) in Eq. (11). $Sent$ is the monthly investor sentiment index developed by Baker and Wurgler (2006, 2007) and it can be defined as beliefs that are not based on the facts at hand.[12] In particular, in periods of high (low) sentiment, people tend to make overly optimistic (pessimistic) judgments and choices. Therefore, we expect sentiment to be negatively related to herding, as it may cause investors to rely more on their own judgments. To investigate the effect of economic activity on herding, we include the ADS index (Aruoba et al., 2009). The index measures economic activity at high frequency using a dynamic factor model that includes several economic variables.[13] The data for the index is obtained from the Philadelphia Fed's website. The average value of the ADS index is zero. Positive values indicate better-than-average conditions, whereas negative values indicate worse-than-average conditions. We expect ADS to be negatively related to herding, as uncertainty is higher in a negative economy. $Active_t$ is the ratio between the number of actively and passively managed funds included in the sample in month t. To classify a fund as either actively or passively managed we follow Appel et al. (2016), Busse and Tong (2012), and Iliev and Lowry (2014). Specifically, we identify a fund as passively managed if either the CRSP mutual fund database classifies it as an index fund or if its fund name contains a text string that identifies it as an index fund.[14] We classify all other funds that do not satisfy the above criteria as actively managed, and we leave the funds with missing information on the CRSP index fund identifier and fund name as unclassified. We expect $Active_t$ to be negatively related to herding, as a high ratio of passive funds may lower the CSSD of the betas. $Dotcom_t$, $FinCrisis_t$, and $Covid_t$ are three dummies representing the burst of the dot-com bubble (March 2000 - December 2001), the global financial crisis (August 2007 - December 2008), and the first wave of the COVID-19 pandemic (March 2020 - August 2020), respectively. Again, we expect herding to increase in periods of crisis due to increased uncertainty. To mitigate endogeneity issues, the explanatory variables enter the regression with one lag.

Table 6 reports the OLS estimates of Eq. (11) for mutual fund herding in the four styles. Newey and West (1987) standard errors are reported in parenthesis. We observe that mutual

[12]We thank Jeffrey Wurgler for making the data available on his website.

[13]The ADS index (Aruoba et al., 2009) considers the following variables: weekly initial jobless claims; monthly payroll employment, industrial production, personal income, fewer transfer payments, manufacturing and trade sales; and quarterly real GDP.

[14]We use the following text strings to identify a fund as an index fund: Index, Ind, Idx, Indx, Mkt, Market, Russell, S & P, S and P, S&P, SandP, SP, DOW, DJ, MSCI, Bloomberg, KBW, NASDAQ, NYSE, STOXX, ETF, iShares, FTSE, Wilshire, Morningstar, 100, 400, 500, 600, 900, 1000, 1500, 2000, 3000, and 5000.

Table 6: Determinants of style herding by mutual funds

Dep.Var.	h_m	h_{smb}	h_{hml}	h_{umd}
Constant	0.699***	0.316***	0.913***	1.047***
	(0.157)	(0.106)	(0.138)	(0.227)
$R_{m,t-6}$	-0.078	-0.153**	-0.178	-0.289
	(0.100)	(0.060)	(0.168)	(0.177)
$R_{smb,t-6}$	-0.167	0.105	0.466**	-0.998***
	(0.284)	(0.106)	(0.228)	(0.272)
$R_{hml,t-6}$	0.122	0.170	-0.135	0.162
	(0.229)	(0.121)	(0.258)	(0.280)
$R_{umd,t-6}$	-0.171*	-0.094*	-0.081	-0.170
	(0.103)	(0.054)	(0.122)	(0.253)
$\log(\sigma_{m,t-1})$	-0.043***	0.003	0.043***	0.024
	(0.011)	(0.006)	(0.015)	(0.016)
$Sent_{t-1}$	-0.004	-0.036**	-0.115***	0.007
	(0.055)	(0.017)	(0.043)	(0.047)
ADS_{t-1}	-0.003**	-0.001	-0.001	-0.002
	(0.002)	(0.001)	(0.005)	(0.005)
$Active_{t-1}$	-0.122***	-0.032***	-0.081***	-0.120***
	(0.016)	(0.011)	(0.012)	(0.019)
$Dotcom_t$	-0.350	-0.086	0.052	-0.107
	(0.254)	(0.087)	(0.151)	(0.095)
$FinCrises_t$	0.034	0.024	-0.130	0.056
	(0.079)	(0.023)	(0.084)	(0.067)
$Covid_t$	0.042	0.032	-0.139*	0.007
	(0.061)	(0.036)	(0.075)	(0.137)
R^2	0.907	0.773	0.714	0.762
Adj. R^2	0.903	0.764	0.703	0.752
Num. obs.	279	279	279	279

Notes: This table presents the OLS estimates of Eq. (11), Newey and West (1987) standard errors are reported in parenthesis. Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

fund herding towards the market portfolio is negatively associated with past cumulative returns of the momentum factor. Moreover, mutual fund herding towards the size factor is negatively associated with past cumulative returns of the market portfolio and UMD factor. We also observe that mutual fund herding towards the momentum (value) factor tends to increase (decrease) after periods of low cumulative returns of the SMB factor.

In line with the theoretical findings in Di Guilmi et al. (2014), we show that in general uncertainty in the market performance (σ_m) is positively associated with mutual fund herding in styles. However, the coefficient of σ_m is significantly positive only for herding towards the value factor. Interestingly, mutual fund herding towards the market portfolio tends to decrease when σ_m increases. In Table 4, we have shown that market volatility is negatively associated with the CSSD of the market betas. Hence, during periods of high uncertainty mutual fund managers may invest more in the market portfolio and this behavior is driven by fundamentals rather than herding.

In addition, we find that market sentiment is negatively related to mutual fund herding in styles. Hence, an increase in sentiment causes investors to rely more on their private information. Further, we report that mutual fund herding towards the market portfolio tends to increase when economic conditions (ADS) worsen. We also find that mutual fund herding in styles increases when the ratio of actively versus passively managed funds decreases. This result is quite intuitive, indeed passively managed funds replicate the performance of an index. Hence, when the number of passively managed funds included in the sample is higher, we will observe a decrease in the CSSD of the factor betas which, if unjustified by fundamentals, will be interpreted as an increase in herding. Furthermore, we find that mutual fund herding towards the value factor was lower during the first wave of the COVID-19 pandemic. We do not find any significant difference in mutual fund style herding during the Dotcom bubble and the global financial crises. All in all, these results confirm the intentional non-information character of the herding measure.

5 Style herding and return predictability

The previous sections have laid bare that style herding by mutual funds is an important characteristic of their behavior. Furthermore, we find that our measure of intentional herding is positively associated with uncertainty, and negatively with sentiment. The next question we want to answer is whether mutual fund managers' herding has implications for asset pricing. Because of its sheer size, the mutual fund industry can affect prices and drive them (temporarily) away from fundamentals. Indeed, Ben-Rephael et al. (2011) show that mutual fund flows can cause temporary price pressure. Further, Gong et al. (2016) find that correlated fund flows contribute to skewed stock return distributions. By the same token, we conjecture that stock returns might be affected by mutual fund herding, as intentional herding causes financial markets to be more one-sided. As such, herding should be negatively related to future returns. We study the effects of herding both in-sample and out-of-sample.

5.1 In-sample predictability of returns

In this section we investigate the in-sample predictive ability of mutual fund style herding by estimating the following regression equation:

$$\begin{aligned} r_{k,t} =& \beta_0 + \beta_1 h_{k,t-1} + \beta_2 Pos \times h_{k,t-1} + \beta_3 r_{k,t-1} + \beta_4 r_{k,t-1} \times h_{k,t-1} + \\ & \beta_5 r_{k,t-1} \times Pos \times h_{k,t-1} + \beta_6 Flow_{t-1} + \beta_7 Flow_{t-1} \times h_{k,t-1} + \\ & \beta_8 Flow_{t-1} \times Pos \times h_{k,t-1} + \epsilon_t, \end{aligned} \tag{12}$$

Table 7: In-sample Predictive Ability of Style Herding

Dep.Var.	r_m	SMB	HML	UMD
Constant	0.007	-0.007**	-0.006	0.001
	(0.006)	(0.004)	(0.006)	(0.005)
h_{t-1}	-0.003	-0.161***	0.010	-0.024
	(0.017)	(0.049)	(0.031)	(0.031)
$Pos \times h_{t-1}$	0.019	0.261***	0.023	0.030
	(0.043)	(0.078)	(0.060)	(0.050)
r_{t-1}	0.324**	-0.228	0.004	-0.150
	(0.134)	(0.151)	(0.119)	(0.143)
$r_{t-1} \times h_{t-1}$	0.275	-0.690	-0.499	-0.270
	(0.261)	(0.677)	(0.498)	(0.217)
$r_{t-1} \times Pos \times h_{t-1}$	-2.082**	2.673	1.748*	1.657*
	(0.819)	(2.394)	(0.979)	(0.948)
$Flow_{t-1}$	-1.803	2.941***	1.532	0.342
	(1.208)	(1.107)	(0.967)	(0.956)
$Flow_{t-1} \times h_{t-1}$	-0.008	22.010***	-2.472	1.309
	(2.980)	(7.611)	(2.735)	(2.862)
$Flow_{t-1} \times Pos \times h_{t-1}$	8.163	-52.169***	-8.656	13.210
	(11.599)	(19.632)	(9.098)	(14.949)
R^2	0.052	0.061	0.092	0.044
Adj. R^2	0.024	0.034	0.065	0.017
Num. obs.	287	287	287	287

Notes: This table presents the OLS estimates of different forms of Eq. (12). Newey and West (1987) standard errors are reported in parenthesis. Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

in which r_k is the return of factor k, which can either be the excess market return, SMB, HML or UMD factor ($k = \{m, SMB, HML, UMD\}$), h_k is herding towards that style, and the dichotomous variable Pos is equal to one when the herding estimate is positive and zero otherwise. $Flow_t$ is aggregate value-weighted flow in the mutual fund industry at time t (fund flows are computed as in Eq. (1)). The coefficients β_1 and β_2 measure the predictive ability of herding and adverse herding. We also include the interaction of herding with lagged returns and mutual fund flows. The coefficients β_4 and β_5 capture whether herding affects the autocorrelation structure of markets, which is a measure of market efficiency (Rösch et al., 2016). If our measure of intentional herding affects returns, we expect the autocorrelation to be negatively related to herding as it introduces mean-reversion dynamics, i.e., we expect $\beta_5 < 0$. When fund flows push prices above fundamental, we expect β_6 to be negative due to mean reversion. If this price pressure effect is aggravated by herding, β_7 is negative as well. β_8 captures the possible asymmetry in this relationship driven by the fact that there is more need for immediacy in sells than in buys (Gong et al., 2016).

Table 7 reports the OLS estimates of Eq. (12). Newey and West (1987) standard errors are reported in parenthesis. We document that herding towards the size factor has a direct effect on next-month SMB returns. In particular, the coefficient is positive for herding and negative for adverse herding. Furthermore, we find that an increase in herding towards the market portfolio decreases the autocorrelation in r_m, while adverse herding does not have any significant impact. Differently, mutual fund herding towards value and momentum increases the autocorrelation in the respective risk factors, while adverse herding does not have any significant impact. Moreover, the interaction of fund flows with herding towards the size factor predicts higher next-month returns of the SMB factor, while the interaction of fund flows with adverse herding towards the size factor predicts lower next-month returns of the SMB factor.[15] All in all, the results suggest that herding has important implications on asset prices.

5.2 Out-of-sample predictability of returns

In the previous section, we observed that mutual fund style herding can help predict future factor returns. To evaluate the robustness of the in-sample results, Table 8 reports results for the out-of-sample test of return predictability. We use the first 100 observations to estimate the model parameters. Then, we compute a predictive regression forecast of model i ($\hat{y}_{i,t+1}$) by substituting the parameter estimates in Equation (12) and using the time series of the model variables in the predictive period. To assess the predictive ability of our models we compare them to parsimonious null models which are nested in our larger models. Specifically, we compare the model in Eq. (12) with the prevailing mean forecast model, and the baseline model which does not include style herding. The prevailing mean forecast corresponds to the constant expected excess return model and implies that returns are not predictable. The baseline model includes only lagged returns and fund flows as predictor variables. So, once we have assessed that the predictive ability of our model is higher than that of the prevailing mean forecast model we can check whether the gain in predictive ability is due to style herding or the other predictors in the equations.

The second through fifth columns of Table 8 report the test statistic of Clark and West (2007) for mutual fund herding towards different styles. Let $\hat{e}_{n,t+1}$ and $\hat{e}_{a,t+1}$ be the one-step-ahead forecast errors of the null and alternative model respectively. Let the adjusted mean squared predictive error be $\hat{f}_{t+1} = \hat{e}^2_{n,t+1} - \hat{e}^2_{a,t+1} + (\hat{y}_{n,t+1} - \hat{y}_{a,t+1})^2$. Let $\bar{f}$ and σ_f be the corresponding sample average and standard deviation of $\hat{f}$. The Clark and West (2007) test statistic is $\sqrt{P}\bar{f}/\sigma_f$, where P is the number of observations in the predictive period. A positive

[15] We also investigate whether herding contributes to factor timing. To do so, we compare the risk-adjusted performance of two portfolios composed of the market, SMB, HML, and UMD factors. In the first portfolio, the four factors are equally weighted; in the second portfolio, we use time-varying weights that depend negatively on our herding measure. We find that the Sharpe ratio of the second portfolio (21.2%) is significantly higher than the Sharpe ratio of the first portfolio (11.8%). Hence, mutual fund style herding contributes to factor timing.

sign of the test statistic should be interpreted as a failure of the null model to outperform our larger models which include style herding.

Table 8: Out-of-Sample predictability of returns

Null Model	r_m	SMB	HML	UMD
Panel A: All sample				
Prevailing mean forecast	0.367	0.149	-0.931	0.755
Baseline model	0.682	0.506	-1.398	0.619
Panel B: September 1998 - August 2010				
Prevailing mean forecast	1.153	-0.970	-0.594	0.695
Baseline model	1.466*	-0.333	-1.242	0.604
Panel C: September 2010 - September 2022				
Prevailing mean forecast	1.456*	-1.324	-0.855	0.937
Baseline model	1.490*	-1.774	-1.319	0.971

Notes: This table presents the test statistic of Clark and West (2007). In the first row, we use the prevailing mean forecast as a null model. In the second row, the null model includes only lagged returns and fund flows as predictor variables. Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

In Panel A of Table 8, we report the results for the entire sample period, in Panel B for the period from September 1998 to August 2010, and in Panel C for the period from September 2010 to September 2022. We observe that, in the most recent period, herding towards the market portfolio can significantly predict the returns of the market portfolio out-of-sample. Mutual fund herding towards the other styles cannot predict out-of-sample factor returns.

6 Style herding, flow, and performance

Having established that herding is a prominent feature of mutual fund manager behavior and that it predicts market returns out-of-sample, in this final section we proceed to study managers' incentives to herd. Given that fund managers' compensation is typically a function of fund size, we test whether herding leads to extra capital inflow and/or to extra performance. To do so, we estimate the following regression equation:[16]

$$\begin{aligned} Flow_t = &\beta_0 + \beta_1 h_{k,t-1} + \beta_2 Pos \times h_{k,t-1} + \beta_3 Flow_{t-1} + \beta_4 \alpha_{t-1} + \beta_5 r_{m,t-1} + \\ &\beta_6 turn_{t-1} + \beta_7 fees_{t-1} + \beta_8 Dotcom_t + \beta_9 FinCrisis_t + \beta_{10} Covid_t + \epsilon_t, \end{aligned} \tag{13}$$

where $Flow_t$ represents the value-weighted aggregate flows to the mutual fund industry in month t, h_k is mutual fund herding towards style k, Pos is a dichotomous variable which equals one when herding is positive and zero otherwise. The interaction term $Pos \times h_{k,t}$ allows us to study

[16] Note that this model captures aggregate flows into equity mutual funds. The results can therefore not be directly compared to cross-sectional studies.

Table 9: Mutual fund herding and fund flows

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	-0.005*	-0.004**	-0.003***	-0.002	-0.003
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
h_m	-0.012***				
	(0.002)				
$Pos \times h_m$	0.009*				
	(0.005)				
h_{smb}		-0.025***			
		(0.008)			
$Pos \times h_{smb}$		-0.000			
		(0.012)			
h_{hml}			-0.009***		
			(0.003)		
$Pos \times h_{hml}$			0.004		
			(0.006)		
h_{umd}				-0.011***	
				(0.003)	
$Pos \times h_{umd}$				0.008*	
				(0.004)	
h_{pca1}					-0.003***
					(0.001)
$Pos \times h_{pca1}$					0.002*
					(0.001)
h_{pca2}					0.000
					(0.001)
$Pos \times h_{pca2}$					0.002
					(0.002)
$Flow_{t-1}$	0.217**	0.187**	0.223**	0.167*	0.106
	(0.093)	(0.093)	(0.101)	(0.086)	(0.086)
α_{t-1}	1.568	1.152	1.173	1.899*	1.186
	(1.061)	(1.028)	(1.008)	(1.147)	(1.019)
$r_{m,t-1}$	0.013*	0.013*	0.012	0.012	0.014*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$turn_{t-1}$	-0.014**	0.000	0.011**	0.000	-0.012*
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
$fees_{t-1}$	0.025***	0.014**	-0.002	0.004	0.020***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
$Dotcom_t$	-0.004**	-0.003**	-0.002*	-0.001	-0.004**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
$FinCrises_t$	-0.000	-0.000	-0.002*	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Covid_t$	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.429	0.433	0.405	0.428	0.475
Adj. R^2	0.408	0.412	0.383	0.406	0.451
Num. obs.	279	279	279	279	279

Notes: This table reports the OLS estimates of Eq. (13). The dependent variable is the value-weighted aggregate flows to the mutual fund industry, where fund flows are computed as in Eq. (1). The last column reports the estimates when we use the principal component of the estimated herding parameters (h_{pca}) as a regressor. Newey and West (1987) standard errors are reported in parenthesis. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

whether herding and adverse herding have a different impact on fund flows. The variable α_t represents the value-weighted risk-adjusted performance (OLS estimate of the Carhart (1997) model intercept computed with daily data over monthly intervals) of the mutual fund industry, $turn$ is the value-weighted aggregate turnover ratio of the sampled mutual funds, and $fees$ represents value-weighted management fees of the sampled mutual funds. Again, $Dotcom_t$, $FinCrisis_t$, and $Covid_t$ are three dummies representing the burst of the dot-com bubble (March 2000 - December 2001), the global financial crisis (August 2007 - December 2008), and the first wave of the COVID-19 pandemic (March 2020 - August 2020), respectively.

Table 10: Mutual fund herding and fund performance

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.008	-0.000	0.004	0.007	-0.002
	(0.009)	(0.007)	(0.006)	(0.007)	(0.011)
h_m	-0.020				
	(0.013)				
$Pos \times h_m$	0.009				
	(0.022)				
h_{smb}		-0.095**			
		(0.042)			
$Pos \times h_{smb}$		0.113*			
		(0.061)			
h_{hml}			-0.042***		
			(0.015)		
$Pos \times h_{hml}$			0.064**		
			(0.027)		
h_{umd}				-0.010	
				(0.012)	
$Pos \times h_{umd}$				0.013	
				(0.021)	
h_{pca1}					-0.005*
					(0.003)
$Pos \times h_{pca1}$					0.006
					(0.005)
h_{pca2}					-0.007
					(0.007)
$Pos \times h_{pca2}$					0.009
					(0.010)
α_{t-1}	-0.034	-0.053	-0.062	-0.028	-0.054
	(0.071)	(0.076)	(0.075)	(0.076)	(0.077)
$Flow_{t-1}$	0.122	0.082	0.079	0.203	0.011
	(0.395)	(0.375)	(0.378)	(0.395)	(0.403)
$turn_{t-1}$	-0.058	-0.055*	-0.017	-0.021	-0.048
	(0.036)	(0.030)	(0.020)	(0.022)	(0.035)
$fees_{t-1}$	0.037	0.048	-0.004	-0.001	0.038
	(0.043)	(0.041)	(0.027)	(0.027)	(0.042)
$Dotcom_t$	0.004	0.004	0.003	0.008	0.000
	(0.010)	(0.009)	(0.009)	(0.010)	(0.009)
$FinCrises_t$	-0.008	-0.006	-0.010**	-0.006	-0.007
	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)
$Covid_t$	-0.003	-0.004	-0.001	-0.004	-0.005
	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)
R^2	0.042	0.056	0.061	0.030	0.056
Adj. R^2	0.010	0.025	0.029	-0.002	0.018
Num. obs.	279	279	279	279	279

Notes: This table reports the OLS estimates of Eq. (14). The dependent variable is the aggregate value-weighted risk-adjusted performance (in percent) of the mutual fund industry measured as the alpha of the Carhart (1997) four-factor model computed with daily data over monthly intervals for each fund in the sample. The last column reports the estimates when we use the principal component of the estimated herding parameters (h_{pca}) as a regressor. Newey and West (1987) standard errors are reported in parenthesis. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 9 reports the OLS estimates of Eq. (13). Newey and West (1987) standard errors are reported in parenthesis. Since mutual fund herding towards different styles is highly correlated (see Table 5), we study the effect of mutual fund herding on flows separately for each style. To investigate the overall effect of style herding on aggregate mutual fund flows, we consider the first two principal components of the estimated herding parameters which account for 88.1% of the total variance. The results are reported in the last column of Table 9.

We show that mutual fund herding towards all the styles predicts lower next-month fund flows to the mutual fund industry.

Concerning the control variables, we observe that mutual fund flows are significantly positively autocorrelated. Consistently from Del Guercio and Tkac (2002) who point out that fund performance is one of the main drivers of fund flows, we report that the coefficient of α is always positive although it is significant only in Model 4. Consistent with the fact that investors in the mutual fund industry withdraw money from the funds during bad periods and increase investments during good periods, we find that the return of the market portfolio (r_m) is positively associated with mutual fund flows. Furthermore, previous period management fees ($fees$) tend to be positively associated with fund flows. We also observe that fund flows were significantly lower during the burst of the dot-com bubble, while they were not significantly different during the global financial crisis and the first wave of the COVID-19 pandemic.

Subsequently, we investigate the relationship between mutual fund herding in styles and the outperformance of the mutual fund industry. We estimate the following regression equation:

$$\begin{aligned}\alpha_t = \beta_0 + \beta_1 h_{k,t-1} + \beta_2 Pos \times h_{k,t-1} + \beta_3 \alpha_{t-1} + \beta_4 Flow_{t-1} + \\ + \beta_5 turn_{t-1} + \beta_6 fees_{t-1} + \beta_7 Dotcom + \beta_8 FinCrisis + \beta_9 Covid + \epsilon_t,\end{aligned} \tag{14}$$

where α_t is the value-weighted risk-adjusted performance of the mutual fund industry.[17] Specifically, α_t is the OLS estimate of the Carhart (1997) model intercept computed with daily data over monthly intervals. As in the previous equation, h_k is mutual fund herding towards style k, Pos is a dichotomous variable that equals one when herding is positive and zero otherwise. The term $Pos \times h_{k,t}$ allows us to study whether herding and adverse herding have a different impact on fund performance. $Flow_t$ represents the value-weighted aggregate flows to the mutual fund industry in month t, $turn$ is the value-weighted aggregate turnover ratio of the sampled mutual funds, and $fees$ represents value-weighted management fees of the sampled mutual funds. $Dotcom_t$, $FinCrisis_t$, and $Covid_t$ are three dummies representing the burst of the dot-com bubble (March 2000 - December 2001), the global financial crisis (August 2007 - December 2008), and the first wave of the COVID-19 pandemic (March 2020 - August 2020), respectively.

Table 10 reports the OLS estimates of Eq. (14). Newey and West (1987) standard errors are

[17] Since the estimates of α_t are sensitive to model specification, we have estimated Eq. (14) using the value-weighted CAPM, Fama-French three and five factors model alphas. The main results are qualitatively the same and they are available upon request. An alternative approach to measure mutual fund performance that does not rely on any asset pricing benchmark is the stochastic dominance approach (Joy and Porter, 1974; Cho et al., 2007; Chui et al., 2020).

reported in parenthesis. We study the effect of mutual fund herding on performance separately for each style as mutual fund herding towards different styles is highly correlated (see Table 5). To investigate the overall effect of style herding on aggregate mutual fund performance, we consider again the first two principal components of the estimated herding parameters. The results are reported in the last column of Table 10.

Table 10 shows that mutual fund herding towards size and value factors (see Models 2 and 3) temporarily increase performance while adverse herding in these styles decreases performance. Herding toward the other factors does not affect mutual fund performance. Fricke (2019) find that even modest levels of fund's portfolio overlap can imply substantial return correlations. This evidence may explain why we do not find any significant effect of herding towards some styles on funds' performance. From model 5, we can see that overall style herding contributes to decreasing fund performance. This result is consistent with Koch (2017) which does not find evidence of outperformance of herding funds, and Jiang and Verardo (2018) which show that herding behavior is prominent among unskilled funds.

We observe that mutual fund performance is negatively autocorrelated, pointing towards mean reversion of outperformance (Carhart, 1997). Moreover, the coefficients of the remaining control variables are generally not significant at the standard level.

In sum, mutual fund managers may have an incentive to herd towards size and value factors as they temporarily increase fund performance. However, mutual fund style herding leads to a decrease in fund flows. Hence, there does not appear to be a clear incentive for managers to engage in intentional herding.

7 Conclusions

Whereas the herding literature is already quite sizeable, it has not yet uncovered how it interacts with the most important investment strategy of mutual funds, namely style investing. As such, we contribute to the literature by exploring the extent, drivers, and consequences of style herding towards four investment styles (market, size, value, and momentum) for U.S. domestic equity mutual funds.

We find that mutual fund herding in styles is significant and persistent during our sample period. We also report that mutual fund herding tends to increase after periods of high volatility of the market and increasing pessimism. Moreover, we report that mutual fund herding towards the market portfolio significantly increases when economic conditions worsen. We also find that mutual fund herding in styles increases when the ratio of actively versus passively managed funds decreases. Interestingly, we do not find any considerable change in mutual fund style herding during the COVID-19 pandemic. Specifically, only mutual fund herding towards value was significantly lower during the first wave of the COVID-19 pandemic.

The paper also sheds light on the effect of mutual fund herding in styles on factor returns. Specifically, we find a direct effect of only mutual fund herding towards size on factor returns. Moreover, we observe that herding impacts the effect of mutual fund flows on factor returns, and it may also impact their autocorrelation structure. Furthermore, the results indicate that mutual fund herding towards the market can predict the returns of the market portfolio out-of-sample.

Finally, we examine the relationship between mutual fund herding in styles and the value-weighted flows and risk-adjusted performance of the mutual fund industry. We show that, in general, aggregate fund flows decrease when mutual funds herd in styles. Moreover, mutual fund style herding tends to reduce the value-weighted risk-adjusted performance of the mutual fund industry. However, mutual fund herding towards size and value temporarily increases performance.

These results can inform practitioners on the construction of multi-factor portfolios, indeed the presence or absence of mutual fund herding can predict factor returns. Furthermore, the tendency of mutual funds to herd in styles is predictable as herding is more likely to occur after particular states of the market. For mutual fund managers, our results suggest that style herding is not necessarily beneficial to the fund. Moreover, these results may guide investors in the choice of an optimal fund manager.

Given the large share of corporate equity held by mutual funds in the U.S., style herding by these investors may contribute to a significant increase in the inflows of a certain factor. This phenomenon is known as factor crowding and it increases drawdown risk. Specifically, a sudden change in market sentiment may result in many investors selling the popular factor causing drawdowns. Future research could explore the contribution of mutual fund style herding on factor crowding and its effect on the performance and risk of multi-factor portfolio strategies.

References

Ang, A. and D. Kristensen (2012). Testing conditional factor models. *Journal of Financial Economics 106*(1), 132–156.

Appel, I. R., T. A. Gormley, and D. B. Keim (2016). Passive investors, not passive owners. *Journal of Financial Economics 121*(1), 111–141.

Aruoba, S. B., F. X. Diebold, and C. Scotti (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics 27*(4), 417–427.

Asness, C. S., T. J. Moskowitz, and L. H. Pedersen (2013). Value and momentum everywhere. *The Journal of Finance 68*(3), 929–985.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance 61*(4), 1645–1680.

Baker, M. and J. Wurgler (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives 21*(2), 129–151.

Baltas, N. (2019). The impact of crowding in alternative risk premia investing. *Financial Analysts Journal 75*(3), 89–104.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics 107*(3), 797–817.

Barberis, N. and A. Shleifer (2003). Style investing. *Journal of Financial Economics 68*(2), 161–199.

Ben-Rephael, A., S. Kandel, and A. Wohl (2011). The price pressure of aggregate mutual fund flows. *Journal of Financial and Quantitative Analysis 46*(02), 585–603.

Bergeron, A., M. Kritzman, and G. Sivitsky (2018). Asset allocation and factor investing: An integrated approach. *The Journal of Portfolio Management 44*(4), 32–38.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy 100*(5), 992–1026.

Bikhchandani, S. and S. Sharma (2000). Herd behavior in financial markets. *IMF Economic Review 47*(3), 279–310.

Bohl, M. T., N. Branger, and M. Trede (2017). The case for herding is stronger than you think. *Journal of Banking & Finance 85*, 30–40.

Brown, N. C., K. D. Wei, and R. Wermers (2014). Analyst recommendations, mutual fund herding, and overreaction in stock prices. *Management Science 60*(1), 1–20.

Busse, J. A. and Q. Tong (2012). Mutual fund industry selection and persistence. *The Review of Asset Pricing Studies 2*(2), 245–274.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance 52*(1), 57–82.

Celiker, U., J. Chowdhury, and G. Sonaer (2015). Do mutual funds herd in industries? *Journal of Banking & Finance 52*, 1–16.

Chang, E. C., J. W. Cheng, and A. Khorana (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance 24*(10), 1651–1679.

Chiang, T. C. and D. Zheng (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance 34*(8), 1911–1921.

Cho, Y.-H., O. Linton, and Y.-J. Whang (2007). Are there monday effects in stock returns: A stochastic dominance approach. *Journal of Empirical Finance 14*(5), 736–755.

Choi, N. and R. W. Sias (2009). Institutional industry herding. *Journal of Financial Economics 94*(3), 469–491.

Choi, N. and H. Skiba (2015). Institutional herding in international markets. *Journal of Banking & Finance 55*, 246–259.

Christie, W. G. and R. D. Huang (1995). Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal 51*(4), 31–37.

Chui, D., W. W. Cheng, S. C. Chow, and L. Ya (2020). Eastern halloween effect: A stochastic dominance approach. *Journal of International Financial Markets, Institutions and Money 68*, 101241.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(1), 291–311.

Clarke, R. G., H. de Silva, and R. Murdock (2005). A factor approach to asset allocation. *The Journal of Portfolio Management 32*(1), 10–21.

Cooper, M. J., H. Gulen, and P. R. Rau (2005). Changing names with style: Mutual fund name changes and their effects on fund flows. *The Journal of Finance 60*(6), 2825–2858.

Cooper, M. J., R. C. Gutierrez Jr, and A. Hameed (2004). Market states and momentum. *The Journal of Finance 59*(3), 1345–1365.

D'Arcangelis, A. M. and G. Rotundo (2019). Herding in mutual funds: A complex network approach. *Journal of Business Research* (129), 679–686.

Dasgupta, A., A. Prat, and M. Verardo (2011a). Institutional trade persistence and long-term equity returns. *The Journal of Finance 66*(2), 635–653.

Dasgupta, A., A. Prat, and M. Verardo (2011b). The price impact of institutional herding. *The Review of Financial Studies 24*(3), 892–925.

Del Guercio, D. and P. A. Tkac (2002). The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds. *Journal of Financial and Quantitative Analysis 37*(04), 523–557.

Deng, X., S. Hung, and Z. Qiao (2018). Mutual fund herding and stock price crashes. *Journal of Banking & Finance 94*, 166–184.

Di Guilmi, C., X.-Z. He, and K. Li (2014). Herding, trend chasing and market volatility. *Journal of Economic Dynamics and Control 48*, 349–373.

Dichtl, H., W. Drobetz, H. Lohre, C. Rother, and P. Vosskamp (2019). Optimal timing and tilting of equity factors. *Financial Analysts Journal 75*(4), 84–102.

Dimson, E., P. Marsh, and M. Staunton (2017). Factor-based investing: The long-term evidence. *The Journal of Portfolio Management 43*(5), 15–37.

Duygun, M., R. Tunaru, and D. Vioto (2021). Herding by corporates in the us and the eurozone through different market conditions. *Journal of International Money and Finance 110*, 102311.

Falkenstein, E. G. (1996). Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *The Journal of Finance 51*(1), 111–135.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Frey, S., P. Herbst, and A. Walter (2014). Measuring mutual fund herding–a structural approach. *Journal of International Financial Markets, Institutions and Money 32*, 219–239.

Fricke, D. (2019). Are specialist funds "special"? *Financial Management 48*(2), 441–472.

Friedman, B. M. (1984). A comment: stock prices and social dynamics. *Brookings Papers on Economic Activity 2*, 504–508.

Frijns, B., A. B. Gilbert, and R. C. Zwinkels (2016). On the style-based feedback trading of mutual fund managers. *Journal of Financial and Quantitative Analysis 51*(3), 771–800.

Froot, K. and M. Teo (2008). Style investing and institutional investors. *Journal of Financial and Quantitative Analysis 43*(04), 883–906.

Giamouridis, D. (2017). Systematic investment strategies. *Financial Analysts Journal 73*(4), 10–14.

Gilbert, T., C. Hrdlicka, J. Kalodimos, and S. Siegel (2014). Daily data is bad for beta: Opacity and frequency-dependent betas. *The Review of Asset Pricing Studies 4*(1), 78–117.

Gong, M., M. Lin, and R. Zwinkels (2016). Forecasting crashes: Correlated fund flows and skewness in stock returns. *Journal of Financial Econometrics 15*(1), 36–61.

Grinblatt, M., S. Titman, and R. Wermers (1995). Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *American Economic Review 85*(5), 1088–1105.

Hwang, S., A. Rubesam, and M. Salmon (2021). Beta herding through overconfidence: A behavioral explanation of the low-beta anomaly. *Journal of International Money and Finance 111*, 102318.

Hwang, S. and M. Salmon (2004). Market stress and herding. *Journal of Empirical Finance 11*(4), 585–616.

Iliev, P. and M. Lowry (2014). Are mutual funds active voters? *The Review of Financial Studies 28*(2), 446–485.

Investment Company Institute (2022). 2022 investment company fact book. Technical report, Investment Company Institute.

Jiang, H. and M. Verardo (2018). Does herding behavior reveal skill? an analysis of mutual fund performance. *The Journal of Finance 73*(5), 2229–2269.

Joy, O. M. and R. B. Porter (1974). Stochastic dominance and mutual fund performance. *Journal of Financial and Quantitative Analysis 9*(1), 25–31.

Keshk, W. and J. J. Wang (2018). Determinants of the relationship between investor sentiment and analysts' private information production. *Journal of Business Finance & Accounting 45*(9-10), 1082–1099.

Kim, J. H., W. C. Kim, and F. J. Fabozzi (2017). Robust factor-based investing. *The Journal of Portfolio Management 43*(5), 157–164.

Koch, A. (2017). Herd behavior and mutual fund performance. *Management Science 63*(11), 3849–3873.

Kremer, S. and D. Nautz (2013). Causes and consequences of short-term institutional herding. *Journal of Banking & Finance 37*(5), 1676–1686.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics 32*(1), 23–43.

Li, Y. and L. Yang (2011). Testing conditional factor models: A nonparametric approach. *Journal of Empirical Finance 18*(5), 972–992.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation-consistent covariance matrix. *Econometrica 55*, 703–708.

Popescu, M. and Z. Xu (2018). Mutual fund herding and reputational concerns. *Journal of Economics and Finance 42*(3), 550–565.

Rösch, D. M., A. Subrahmanyam, and M. A. van Dijk (2016, 10). The Dynamics of Market Efficiency. *The Review of Financial Studies 30*(4), 1151–1187.

Scharfstein, D. S. and J. C. Stein (1990). Herd behavior and investment. *The American Economic Review 80*(3), 465–479.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance 44*(5), 1115–1153.

Sias, R. W. (2004). Institutional herding. *The Review of Financial Studies 17*(1), 165–206.

Sirri, E. R. and P. Tufano (1998). Costly search and mutual fund flows. *The Journal of Finance 53*(5), 1589–1622.

Trueman, B. (1994). Analyst forecasts and herding behavior. *The Review of Financial Studies 7*(1), 97–124.

Welch, I. (1992). Sequential sales, learning, and cascades. *The Journal of Finance 47*(2), 695–732.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *The Journal of Finance 54*(2), 581–622.

Wylie, S. (2005). Fund manager herding: A test of the accuracy of empirical results using uk data. *The Journal of Business 78*(1), 381–403.

A Robustness checks

To check the robustness of the results to a value-weighted cross-sectional expectation, we adopt the value-weighted CSSD of the estimated betas in the estimation of the state-space model:

$$\sigma^w(\hat{\beta}^b_{i,k,t}) = \sqrt{\sum_{i=1}^{N_t} w_{i,t}\left(\hat{\beta^b_{i,k,t}} - \sum_{i=1}^{N_t} w_{i,t}\hat{\beta^b_{i,k,t}}\right)^2}, \quad k = \{m,\, smb,\, hml,\, umd\}, \tag{15}$$

where N_t is the number of funds in month t, and $w_{i,t}$ is the relative size of fund i (TNA_i) to the mutual fund industry at time t.

Results are reported in Table 11, we find only limited difference compared to the results shown with the equally-weighted CSSD. Mutual fund herding in the four styles is still significant and persistent. Moreover, the correlation with the herding measure estimated using the equally-weighted CSSD of the betas is always close to 1.

Table 11: Maximum likelihood estimates (Value-weighted CSSD)

	h^w_m	h^w_{smb}	h^w_{hml}	h^w_{umd}
μ	-2.490^{***}	-1.220^{***}	-1.131^{***}	-1.590^{***}
	(0.166)	(0.072)	(0.120)	(0.169)
σ_ϵ	0.163^{***}	0.114^{***}	0.169^{***}	0.199^{***}
	(0.010)	(0.005)	(0.009)	(0.012)
ϕ	0.975^{***}	0.964^{***}	0.941^{***}	0.948^{***}
	(0.016)	(0.023)	(0.027)	(0.023)
σ_η	0.062^{***}	0.023^{***}	0.062^{***}	0.095^{***}
	(0.012)	(0.005)	(0.011)	(0.015)
r_m	0.488^{*}	-0.052	0.297	-1.225^{***}
	(0.283)	(0.190)	(0.293)	(0.357)
SMB	-0.461	-0.512^{**}	0.130	0.253
	(0.353)	(0.235)	(0.365)	(0.443)
HML	-0.003	0.315	0.056	-0.097
	(0.341)	(0.223)	(0.350)	(0.431)
UMD	-0.502^{**}	0.036	0.457^{*}	0.386
	(0.233)	(0.156)	(0.241)	(0.293)
$log(\sigma_m)$	-0.107^{***}	-0.011	0.020	0.009
	(0.017)	(0.010)	(0.017)	(0.022)
Signal to noise	0.207	0.161	0.246	0.256
ML values	327.778	457.881	323.096	258.453
Correlation with h_k extracted from the model in Eq. (10)	0.964^{***}	0.895^{***}	0.977^{***}	0.986^{***}

Notes: This table presents the maximum likelihood estimates of the state space model parameters in Eq. (10). The series of 289 monthly cross-sectional weighted standard deviations of betas is used. Newey and West (1987) standard errors are reported in parenthesis. Note: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

B Style herding by actively and passively managed funds

We estimate style herding for the sub-samples of actively and passively managed funds. To classify a fund as either actively or passively managed we follow Appel et al. (2016), Busse and Tong (2012), and Iliev and Lowry (2014). Specifically, we identify a fund as passively managed if either the CRSP mutual fund database classifies it as an index fund or if its fund name contains a text string that identifies it as an index fund.[18] We classify all other funds that do not satisfy the above criteria as actively managed, and we leave the funds with missing information on the CRSP index fund identifier and fund name as unclassified.

Table 12 reports the maximum likelihood estimates of the model parameters in Eq. (10) for actively managed funds in Panel A, and for passively managed funds in Panel B. The results in the two sub-samples are comparable. We observe that herding towards size and value is more volatile for active funds than passively managed funds. However, herding towards these styles by passively managed funds is more persistent.

When we look at Figures 5 to 8, we can see that mutual fund style herding follows a similar pattern independently from the fund management style (active or passive).

Summing up, although passively managed funds should follow their benchmark index no matter the state of the market, our results suggest that they tend to mimic more style returns in certain periods than others.

[18]We use the following text strings to identify a fund as an index fund: Index, Ind, Idx, Indx, Mkt, Market, Russell, S & P, S and P, S&P, SandP, SP, DOW, DJ, MSCI, Bloomberg, KBW, NASDAQ, NYSE, STOXX, ETF, iShares, FTSE, Wilshire, Morningstar, 100, 400, 500, 600, 900, 1000, 1500, 2000, 3000, and 5000.

Table 12: Maximum likelihood estimates

	h_m	h_{smb}	h_{hml}	h_{umd}
Panel A: Active Funds				
μ	-2.447^{***}	-1.021^{***}	-0.952^{***}	-1.500^{***}
	(0.225)	(0.078)	(0.132)	(0.169)
σ_ϵ	0.147^{***}	0.093^{***}	0.150^{***}	0.173^{***}
	(0.008)	(0.004)	(0.008)	(0.010)
ϕ	0.987^{***}	0.980^{***}	0.966^{***}	0.960^{***}
	(0.011)	(0.014)	(0.018)	(0.019)
σ_η	0.054^{***}	0.021^{***}	0.057^{***}	0.086^{***}
	(0.010)	(0.004)	(0.009)	(0.013)
r_m	0.635^{**}	0.003	0.302	-1.152^{***}
	(0.264)	(0.161)	(0.270)	(0.324)
SMB	-0.314	-0.402^{**}	0.060	0.209
	(0.319)	(0.194)	(0.327)	(0.391)
HML	-0.170	0.180	0.031	0.139
	(0.322)	(0.193)	(0.329)	(0.397)
UMD	-0.480^{**}	0.006	0.377^{*}	0.344
	(0.211)	(0.128)	(0.216)	(0.259)
$log(\sigma_m)$	-0.118^{***}	-0.009	0.021	-0.002
	(0.015)	(0.009)	(0.015)	(0.019)
Signal to noise	0.171	0.152	0.213	0.237
ML values	346.525	495.949	342.251	285.485
Panel B: Passive Funds				
μ	-2.406^{***}	-0.836^{***}	-1.033^{***}	-1.537^{***}
	(0.200)	(0.073)	(0.136)	(0.154)
σ_ϵ	0.150^{***}	0.086^{***}	0.154^{***}	0.191^{***}
	(0.008)	(0.004)	(0.008)	(0.012)
ϕ	0.985^{***}	0.988^{***}	0.977^{***}	0.941^{***}
	(0.012)	(0.012)	(0.015)	(0.027)
σ_η	0.053^{***}	0.013^{***}	0.044^{***}	0.086^{***}
	(0.010)	(0.003)	(0.009)	(0.015)
r_m	0.756^{***}	0.110	0.256	-1.107^{***}
	(0.269)	(0.147)	(0.272)	(0.353)
SMB	-0.656^{*}	-0.607^{***}	-0.191	0.095
	(0.326)	(0.177)	(0.330)	(0.426)
HML	0.038	0.274	-0.193	-0.154
	(0.327)	(0.175)	(0.328)	(0.431)
UMD	-0.411^{*}	0.019	0.343	0.404
	(0.215)	(0.117)	(0.218)	(0.282)
$log(\sigma_m)$	-0.104^{***}	-0.001	0.026^{*}	0.018
	(0.016)	(0.008)	(0.015)	(0.021)
Signal to noise	0.177	0.116	0.170	0.258
ML values	342.551	525.263	345.860	266.512

Notes: This table presents the maximum likelihood estimates of the state space model parameters in Eq. (10). Panel A reports the results for the subsample of actively managed funds, and Panel B reports the results for the subsample of passively managed funds. The series of monthly CSSDs of betas is used to estimate the state-space model and then extract the latent herding component. Standard errors are reported in parenthesis. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

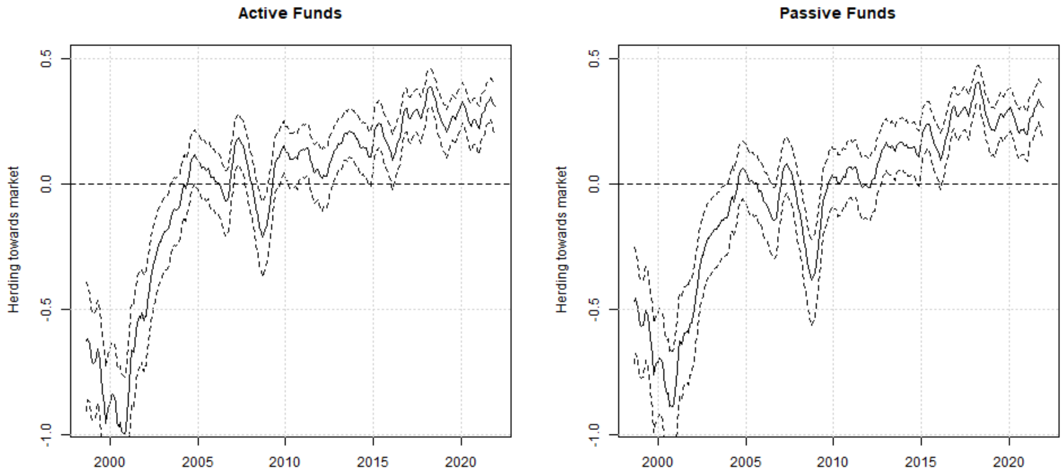

Figure 5: Herding towards the market portfolio. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

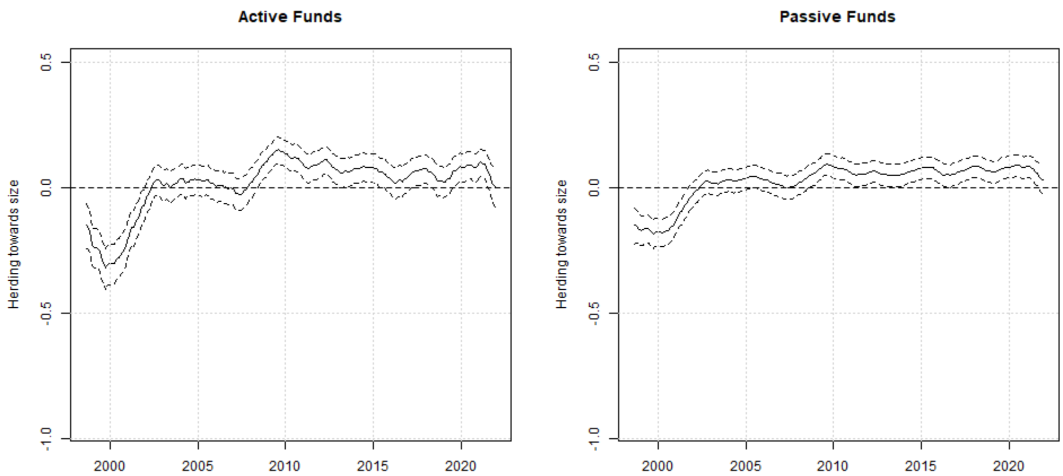

Figure 6: Herding towards the size factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

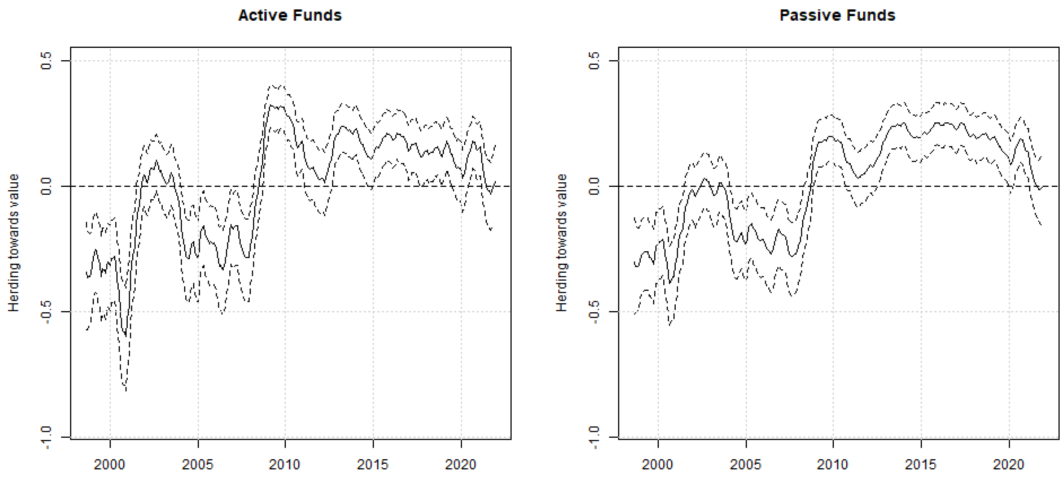

Figure 7: Herding towards the value factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.

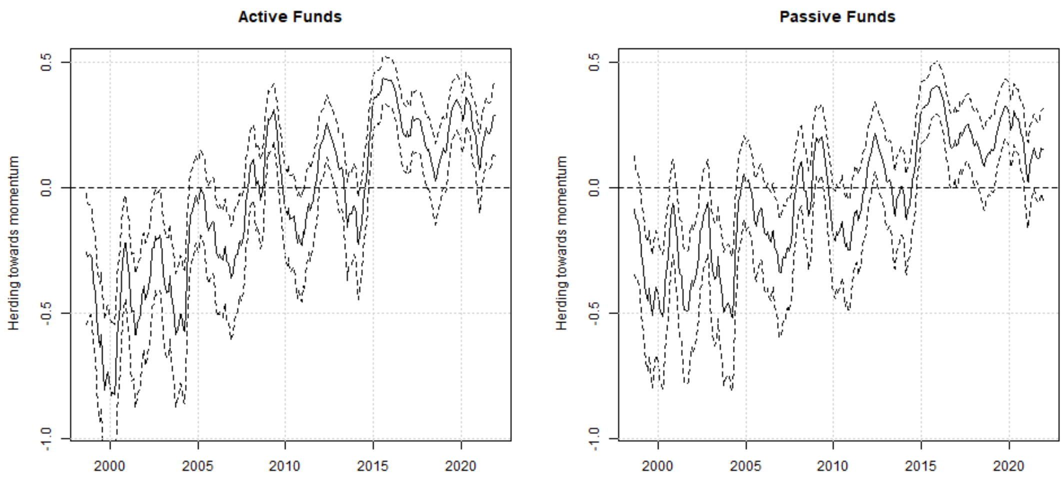

Figure 8: Herding towards the momentum factor. The black straight line is the smoothed series of herding, and the black dashed lines are the 95% confidence intervals.