Accepted Manuscript

Mutual Fund Herding and Stock Price Crashes

Xin Deng , Shengmin Hung , Zheng Qiao

PII: S0378-4266(18)30163-8
DOI: 10.1016/j.jbankfin.2018.07.014
Reference: JBF 5390

To appear in: *Journal of Banking and Finance*

Received date: 17 September 2016
Revised date: 29 June 2018
Accepted date: 25 July 2018

Please cite this article as: Xin Deng , Shengmin Hung , Zheng Qiao , Mutual Fund Herding and Stock Price Crashes, *Journal of Banking and Finance* (2018), doi: 10.1016/j.jbankfin.2018.07.014

Mutual Fund Herding and Stock Price Crashes

Xin Deng Shengmin Hung Zheng Qiao*

June 2018

Abstract

We investigate the impact of mutual fund herding behaviours on stock price crashes. There are competing hypotheses with respect to how investors' herding behaviours are associated with information processing. Our empirical evidence shows that mutual fund herding is associated with a poor information environment and low disclosure quality. More importantly, mutual fund herding amplifies stock price crash risk afterwards. The main finding is concentrated on buy-herding rather than sell-herding. To mitigate the endogeneity concern, we use the 2004 SEC mutual fund disclosure regulation change as an exogenous shock and the results hold. We further use propensity score matching to alleviate the impact of information asymmetry. Finally, additional analysis reveals that our results are not driven by the price impact hypothesis.

Keywords: Stock Price Crashes; Corporate Disclosure; Mutual Fund; Herding

---

*Xin Deng, School of Finance, Shanghai University of Finance and Economics, Shanghai 200433, China, Email: deng.xin@mail.shufe.edu.cn. Shengmin Hung, College of Business, Soochow University, Taipei 100, Taiwan, Email: shung@scu.edu.tw. Corresponding Author: Zheng Qiao, School of Management, Xiamen University, Xiamen 361005, China, Email: zqiao@xmu.edu.cn. We would like to thank the editor Carol Alexander and two anonymous referees for their invaluable suggestions. We are grateful to Michelle Hanlon, Chuan Yang Hwang, Jiang Luo, Donghui Wu, Huai Zhang, Lei Zhang, and participants of 2016 American Accounting Association Annual Meeting and 2017 European Accounting Association Annual Congress for their invaluable comments. Deng and Qiao acknowledge the support from the National Natural Science Foundation of China (71790601, 71402087, 71772111) and the MOE Project of Humanities and Social Sciences (16YJC790082). We also thank Yao Ge for her excellent research assistance. All errors are ours.

1. Introduction

As a prevalent phenomenon in financial markets, the herding behaviour of institutional investors has attracted considerable research interest. There are competing hypotheses with respect to how herding behaviour is associated with information processing. On the one hand, herding weakens information collection activities when institutional investors herd in the face of information cascades (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992; Welch, 1992), relative-performance induced agency motives (Roll, 1992; Brennan, 1993; Admati and Pfleiderer, 1997), or reputation-based mimicking motives (Scharfstein and Stein, 1990; Trueman, 1994; Zweibel, 1995; Prendergast and Stole, 1996; Graham, 1999). On the other hand, fund managers simultaneously and independently move when they receive correlated information shocks (Froot, Scharfstein, and Stein, 1992; Hirshleifer, Subrahmanyam, and Titman, 1994), which speeds up information incorporation. Thus, it is an empirical question as to how mutual fund herding is associated with information processing. We use stock price crash as a consequential outcome to further examine this question. When herding fund managers become less active in collecting and processing information, their monitoring becomes less effective. As a result, corporate managers can withhold bad news more easily, increasing the likelihood of a stock price crash. However, if herding is associated with more active information collection, fund managers become more effective monitors, thus reducing information hoarding and stock price crash risk.

This paper investigates the variation in the information acquisition activities of mutual funds and focuses specifically on herding behaviour. We use stock price crashes as the outcome variable to examine how varying mutual fund herding behaviours are associated with information disclosure. Mutual funds, whose portfolio holdings are subject to mandatory disclosure regulations by the U.S. Securities and Exchange Commission (SEC), provide us with a great setting to study how the herding behaviours of institutional investors affect the disclosure strategy and consequently the stock price crash risk of their holding companies. Our main results show that mutual fund herding (especially buy-herding) is positively associated with stock price crashes. This positive relationship is consistent with the information blockage hypothesis. We also find supporting evidence that mutual fund (buy) herding is associated with deteriorated disclosure quality.

We obtain accounting data from COMPUSTAT and stock return data from CRSP for the sample period from 1989 to 2013. We first find that mutual fund herding (*HM*) is associated with deteriorated corporate disclosure quality. We use four proxies to measure the disclosure quality. The first measure is idiosyncratic volatility, which is based on the price synchronicity measure R^2. Higher idiosyncratic volatility results in better disclosure quality. We find that mutual fund herding is negatively associated with idiosyncratic volatility, which suggests that less firm-specific information is revealed. Additionally, we find that mutual fund herding is negatively associated with the firm-specific earnings transparency measure (*ET*) from Barth et al. (2013), the accounting misstatement likelihood *F-SCORE* from Dechow et al. (2011), and the accounting conservatism *C-SCORE* from Khan and Watts (2009).[2] Overall, our findings suggest that higher mutual fund herding intensity is associated with deteriorated corporate disclosure quality.

Second, we show that higher mutual fund herding intensity leads to higher stock price crash risk. Jin and Myers (2006) and Hutton et al. (2009) find that the private information availability or disclosure opaqueness predict future stock price crashes. As mutual fund herding

[2] We put a negative sign in front of the *F-SCORE* in Dechow et al. (2013) so that higher values proxy for higher disclosure quality.

deteriorates corporate disclosure quality, we propose that the stock prices of firms with intensive mutual fund herding are more likely to crash. Following Hutton et al. (2009), Kim et al. (2011a), and Chen, Hong and Stein (2001), we construct the crash likelihood ($CRASH$) and negative skewness ($NCSKEW$) to measure the stock price crash risk. Consistent with our prediction, the results show that mutual fund herding is positively related to the stock price crash risk.

Finally, the predicting power of herding on stock price crash risk is mostly concentrated in buy-herding rather than sell-herding. Buy- and sell-herding can be different for the following reasons. First, selling pressure usually exerts a disciplining effect on corporate managers and improves disclosure quality. For example, prior studies show that short-selling can restrain earnings management (Fang et al., 2015; Massa et al., 2015) and speed up the discovery of financial misconduct (Karpoff and Lou, 2010). Similarly, sell-herding is predicted to constrain management team through its downward pressure on stock price as well. Therefore, managers are less incentivized to exploit disclosure discretion under strengthened scrutiny induced by sell-herding. In addition, in face of an information cascade, the information processing is asymmetric between buy and sell decisions. Fund managers tend to overweigh good (bad) news over bad (good) in buy (sell)-herding. Consequently, the negligence of bad news in buy-herding can induce bad news hoarding and stock price crashes afterwards, while sell-herding is unlikely to do so. In sum, given such counteractive effects which weaken the relation between sell-herding and stock price crashes, our main findings are concentrated in buy-herding.

Our results could suffer from endogeneity problems of both reverse causality and omitted variables. For example, firms that experience stock price crashes are more likely to be subject to mutual fund herding, so that firms with higher mutual fund herding is positively associated with stock price crash risk. It is also possible that both mutual fund herding and future stock price crashes are simultaneously determined by omitted variables. To mitigate such endogenous concerns, we conduct three tests. We first use the 2004 regulation change as an exogenous shock to the visibility of peers' holding details. This regulation requires mutual funds to increase the disclosure frequency from a semi-annual basis to a quarterly basis. The increase in reporting frequency of mutual funds' holding is predicted to increase herding intensity as the mimicking strategy is more easily implemented, while it is not likely to directly affect the stock price crash risk. We find that compared with firms that have already experienced herding from mutual funds with quarterly disclosures, those that mainly have herding from mutual funds with semi-annual disclosures are associated with higher stock price crash risk after the 2004 regulation. In addition, to deal with the problem of omitted variables in a general term, we use firm fixed effects to control for firm-specific and time-invariant factors that potentially correlate with mutual fund herding and affect stock price crash risk. Our results hold after controlling for firm fixed effects. Finally, Chan, Hwang, and Mian (2005) document that information asymmetry proxies (such as analyst coverage and stock volatility) are likely to be the determinants of fund herding. To specifically rule out the possibility that information asymmetry is the omitted variable that drives our results, we adopt the approach of propensity score matching (PSM) to explicitly control for multiple information asymmetry measures. Our main results consistently hold over these tests.

It is worth noting that the price impact might explain our results. To rule out this possibility, we further control for two sets of variables. First, we add price impact measure (*MFFLOWIN)* by Edmans, Goldstein, and Jiang (2012) as additional control in the quarterly baseline regression and the results still hold. Second, we add future stock returns to rule out the alternative story of stock return mean reversion. Although future stock returns display strong explanatory power on stock price crashes, the coefficients on herding measures stay negative and statistically significant, suggesting its distinctive role in predicting future stock price crashes.

To investigate whether disclosure quality is the potential mechanism through which herding affects stock price crashes, we illustrate a stronger effect of mutual fund herding on stock prices crashes among firms that are of low disclosure quality and are subject to more severe information asymmetry. To some extent, it confirms our conjecture that mutual fund herding is likely to induce stock price crashes through deteriorated information environment.

This paper contributes to the literature in two ways. On the one hand, we investigate stock price crash as a specific consequence of institutional herding behaviours in financial markets. Prior empirical studies on mutual fund herding can be categorized into two main strands. The first strand of literature focuses on the existence and determinants of mutual fund herding behaviours (see, e.g., Sias, 2004; Chan et al., 2005; Wylie, 2005; Choi and Sias, 2009; Chiang and Zheng, 2010). The second strand of literature examines how herding is associated with future stock returns (see, e.g., Wermers, 1999; Dasgupta, 2011; Brown et al., 2014). Rather than the overall stock returns, our study focuses on extreme stock price changes—stock price crashes. Extreme negative return on the left tail of return distribution, i.e. huge loss with small probability, is a critical concern for asset management. To our best knowledge, this is the first paper explaining stock price crashes using a specific institutional investor trading behaviour.[3] Our study provides useful implications for investment practices, especially for risk management.

On the other hand, we explore the possible underlying mechanism of the impact of mutual fund herding on stock price crashes. We find partial evidence that corporate disclosure quality deterioration can be one potential explanation for the relation between mutual fund

[3] An and Zhang (2013) examine the effect of institutional ownership on crash risk.

herding and stock price crashes. Prior studies on herding and stock returns are mostly about asset pricing, while we try to identify a new underlying mechanism from the perspective of the corporate side through which the mutual fund herding leads to stock price crashes.

The remainder of this paper is organized as follows. Section 2 reviews the literature and develops the main hypotheses. Section 3 describes the data and empirical methodology of this paper. Section 4 presents a series of empirical results. Section 5 provides conclusions.

2. Literature and Hypothesis Development

According to the information blockage argument, mutual fund managers can herd out of different motives. First, information cascade is one reason why fund managers ignore their private information or stop acquiring private information and instead turn to the public information pool (see, e.g., Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992; Welch, 1992). Second, mutual fund managers are normally evaluated by their relative performance with respect to certain benchmarks, which incentivizes fund managers to adjust their investments according to peers' strategies. The inefficient contracting in the fund industry constitutes another reason for herding (see, e.g., Roll, 1992; Brennan, 1993; Admati and Pfleiderer, 1997). Third, fund managers (especially those facing uncertainties in their careers or those of lower ability) choose to mimic star fund managers out of reputation concerns (see, e.g., Scharfstein and Stein, 1990; Trueman, 1994; Zweibel, 1995; Prendergast and Stole, 1996; Chevalier and Ellison, 1999; Graham, 1999). These three motives induce fund managers to exert less effort in collecting and processing information.

On the other hand, information incorporation argues that fund managers can make simultaneous decisions when they receive similar information shocks (see, e.g., Froot, Scharfstein, and Stein, 1992; Hirshleifer, Subrahmanyam, and Titman, 1994). In other words, the herding phenomenon may actually be driven by independent and informed decisions.

There is a consensus that institutional investors are aggressive claimers for corporate information, and this demand for information plays an important role in determining corporate disclosure strategy. For example, Ajinkya, Bhojraj, and Sengupta (2005) find that firms with greater institutional ownership on average issue forecasts more frequently and in a more specific, accurate, and objective way. Bushee and Noe (2000) and Healy, Hutton, and Palepu (1999) show that corporate disclosure practices are positively correlated with institutional ownership. Bird and Karolyi (2016) use the discontinuity in the Russell 1000 index and Russell 2000 index as identification to document the causality between the institutional ownership and the amount and quality of corporate disclosure.

While institutional investors on average require more information, the degree of demand for information varies between different institutional investors. Based on the theoretical arguments of information cascade, agency-induced herding, and reputation-based herding, herding mutual funds passively acquire or process private information. When herding fund managers become less active in collecting and processing information, their monitoring becomes less effective. As a result, corporate managers can withhold bad news more easily. Corporate disclosure quality is expected to decrease in equilibrium. Meanwhile, if herding fund managers are more active in information processing and monitoring, it is more difficult for managers to withhold bad news. We expect to observe the improvement of disclosure quality in this case.

H1. *Firms with high mutual fund herding levels are associated with deteriorated disclosure quality.*

There are extensive asset pricing studies on asymmetric volatility and stock price crashes that are based on different theoretical models and hypotheses. For example, leverage effect attributes it to the financial leverage in a firm (Black, 1976; Christie, 1982). Volatility feedback turns to a negative casual relation between volatility risk and required rate of return (Pindyck, 1984; French et al., 1987; Campbell and Hentschel, 1992). Chen, Hong, and Stein (2001) propose that investor heterogeneity, together with short-selling constraint, facilitate more optimistic information over pessimistic information, leading to stock price crashes afterwards (see, e.g., Hong and Stein, 2003; Yuan, 2005).

Meanwhile, corporate finance researchers also investigate the determinants of stock price crashes from the information hoarding perspective. Jin and Myers (2006) find that the R^2 (the stock price synchronicity measure that implies a lack of private information availability) is related to stock price crashes by using cross-country data. Hutton et al. (2009) further examine the firm level and use earnings management as a disclosure opaqueness measure to predict future stock price crashes. In addition, there are many follow-up studies that explore new determinants of information hoarding, including corporate tax avoidance, managerial compensation incentives, corporate social responsibility, short interest, and accounting conservatism, and use stock price crashes as the outcome variable in their empirical designs (see, e.g., Kim et al., 2011a; Kim et al., 2011b; Kim et al., 2014; Callen and Fang, 2015; Kim and Zhang, 2016).

Based on hypothesis 1, more active monitoring from herding mutual fund managers can facilitate information disclosure and hence reduce information hoarding. Stock price crash risk would be reduced. Alternatively, less effective monitoring from herding mutual fund managers results in less information supplied in the corporate disclosure. As a result, stock price crash risk tends to be amplified, which is a direct outcome of information hoarding.

H2. *Firms with higher mutual fund herding levels are more likely to experience future stock price crashes.*

In addition to the general herding level, we predict that the impacts of buy-herding and sell-herding can be different. Compared with buy-herding, sell-herding has some distinctive characteristics. First and most importantly, selling pressure usually has a disciplinary effect on corporate managers, which prevents information hoarding and improves disclosure quality. Prior studies document short-selling can restrain managerial discretion in disclosure.[4] Similar to short-selling, sell-herding also imposes selling pressure and, as a result, is predicted to enhance governance effectiveness. Second, according to the information cascade theory, when fund managers herd to sell (buy), they tend to overweigh bad (good) news over good (bad) news. This is unlikely to reduce disclosure quality and may even generate more pressure on corporate disclosures. Third, after sell-herding occurs, its negative price impact has been incorporated in stock prices. It is thus less likely to further result in extreme negative returns. As a result, we expect that the results of buy-herding and sell-herding will be asymmetric and that only buy-herding can induce future stock price crashes.

H3. *Compared with sell-herding of mutual funds, buy-herding of mutual funds has stronger predictive power for stock price crashes.*

3. Variable Construction, Empirical Methodology and the Sample

3.1. Measurement of Mutual Fund Herding

[4] Fang et al. (2015) and Massa et al. (2015) document how short-selling weakens earnings management. Karpoff and Lou (2010) find that short-selling help reveal financial misconduct at earlier dates.

Following Lakonishok et al. (1992), Wermers (1999), and Brown et al. (2014) [5], we define $HM_{i,q}$ as the measure of herding by funds into (or out of) firm *i* during quarter q:

$$HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$$

where $p_{i,q}$ is the proportion of mutual fund buyers relative to the total number of mutual funds trading firm *i* during quarter q. $E|p_{i,q} - E[p_{i,q}]|$ is subtracted to adjust for random fluctuations of the expected proportion of buyers (Wermers, 1999). This measure captures the extent to which the proportion of buying (selling) of mutual funds exceeds the expected random proportion of buying (selling) funds. With the absolute value, $HM_{i,q}$ is unsigned in that either excessive buy-herding or sell-herding would both increase $HM_{i,q}$. To further distinguish between buy-herding and sell-herding from mutual funds, we follow Wermers (1999) and Brown et al. (2014) to respectively define buy-herding and sell-herding as follows:

$$BHM_{i,q} = HM_{i,q} | \, p_{i,q} > E[p_{i,q}]$$

$$SHM_{i,q} = HM_{i,q} | \, p_{i,q} < E[p_{i,q}]$$

3.2. Measurement of Stock Price Crash

Following prior literature, we use two measures for stock price crashes, $CRASH_{i,t}$ and $NCSKEW_{i,t}$. We define $CRASH_{i,t}$ as the outlier in firm-specific weekly returns. We first run the following regression to obtain the residual returns for each firm-week:

$$r_{i,n} = \beta_0 + \beta_{1i} r_{m,n-2} + \beta_{2i} r_{m,n-1} + \beta_{3i} r_{m,n} + \beta_{4i} r_{m,n+1} + \beta_{5i} r_{m,n+2} + \varepsilon_{i,n}$$

where $r_{i,n}$ is the weekly stock return and $r_{m,n}$ is the weekly market return. Leading and lagging

[5] Lakonishok et al. (1992) proposed a measure for institutional investors' herding using the abnormal proportion of buying (selling) investors. Wermers (1999) improved the measure by incorporating the correlation between mutual fund herding and the momentum effect. Brown et al. (2014) use the same mutual fund herding measure and interact it with analysts' recommendations. Meanwhile, Nofsinger and Sias (1999) and Choi and Sias (2009) document the dynamic change of the prevalence of institutional herding behaviors in time series.

weekly market returns are included to control for nonsynchronous trading. Firm-specific weekly return $W_{i,n}$ is the natural logarithm of one plus the residual $\hat{\varepsilon}_{i,n}$

$$W_{i,n} = ln\left(1 + \hat{\varepsilon}_{i,n}\right)$$

When $W_{i,n}$ falls below 3.2 standard deviations of the mean value of $W_{i,n}$, it is taken as a crash week.[6] $CRASH_{i,t}$ is a dummy variable that equals 1 if firm *i* experiences at least 1 crash week in fiscal year *t* and 0 otherwise. The calculation of our second measure, $NCSKEW_{i,t}$, is the negative conditional residual return skewness at firm-year level;

$$NCSKEW_{i,t} = [-N(N-1)^{\frac{3}{2}} \sum W_{i,n}^{3}]/[(N-1)(N-2)\left(\sum W_{i,n}^{2}\right)^{\frac{3}{2}}]$$

where *N* is the number of firm-specific weekly return observations for firm *i* in year *t*.

3.3. Measurement of disclosure quality

3.3.1. Measurement of R^2

We follow Morck et al. (2000) and Hutton et al. (2009) and adopt the logistic transformation of $R_{i,t}^2$ to measure firm-specific volatility $IDIOSYN_{i,t}$, which captures the lack of market synchronicity. The $R_{i,t}^2$ for each firm-year is obtained from the weekly return regression model in Kim et al. (2011a). A lower $IDIOSYN_{i,t}$ implies the lack of private information available for a specific firm. Based on the firm-year regression using weekly returns, $IDIOSYN_{i,t}$ for firm *i* in fiscal year *t* is defined as follows:

$$IDIOSYN_{i,t} = \ln[(1 - R_{i,t}^2)/R_{i,t}^2]$$

3.3.2. Measurement of Earnings Transparency

[6] We also tried alternative thresholds, namely 3.09, 3.15, 3.25, and 3.30 standard deviations, in defining crash weeks. Our results are robust and similar.

We strictly follow the two-step method in Barth et al. (2013) to construct the measure of earnings transparency (ET) at the firm-level in order to capture both intertemporal and cross-sectional differences in the return-earnings relationship. In the first step, we run regressions at the industry-year level to obtain a residual for each firm-year. In the second step, we sort the sample into 4 portfolios according to the residual values from the first step in each year and run regressions at the portfolio-year level. The earnings transparency measure $ET_{i,t}$ is equal to the sum of the two adjusted R^2 in the two steps.[7]

$$ET_{i,t} = ETI_{j,t} + ETP_{p,t}$$

where firm *i* belongs to industry *j* and portfolio *p*. $ET_{i,t}$ represents the earnings transparency for firm *i* in year *t*. $ETI_{j,t}$ represents the earnings transparency for industry *j* in year *t*. $ETP_{p,t}$ represents earnings transparency for portfolio *p* in year *t*.

3.3.3 Measurement of C-SCORE

We follow Khan and Watts (2009) to obtain the $C-SCORE_{i,t}$ for each firm-year. The

[7] To compute the first component $ETI_{j,t}$, the first step regression is ran at the Fama-French 17 industry-year level:

$$RET_{i,j,t} = \alpha_0^I + \alpha_1^I E_{i,j,t}/P_{i,j,t-1} + \alpha_2^I \Delta E_{i,j,t}/P_{i,j,t-1} + \varepsilon_{i,j,t}$$

where $RET_{i,j,t}$ is annual stock return of firm *i* of industry *j* in year *t*; $E_{i,j,t}$ is income before extraordinary items and discontinued operations of firm *i* of industry *j* in year *t*; $P_{i,j,t-1}$ is the stock price of firm *i* of industry *j* in year *t*-1; $\Delta E_{i,j,t}$ is change of income before extraordinary items and discontinued operations firm *i* of industry *j* from year *t*-1 to year *t*. $ETI_{j,t}$ is the adjusted R^2 obtained from each regression. Different firms of the same industry in a given year share the same industry earnings transparency value $ETI_{j,t}$.

Each NYSE, AMEX, and NASDAQ stock is assigned to an industry portfolio at the end of June of year t according to its four-digit SIC code. Industry classification data is obtained from Kenneth R. French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_17_ind_port.html.

We further calculate the firm-year residual estimates from the first step and sort all firms within the same industry into 4 portfolios every year by the absolute value of the residuals. The second stage regression is ran at portfolio-year level as follows:

$$RET_{i,p,t} = \alpha_0^{IN} + \alpha_1^{IN} E_{i,p,t}/P_{i,p,t-1} + \alpha_2^{IN} \Delta E_{i,p,t}/P_{i,p,t-1} + \varepsilon_{i,p,t}$$

where $RET_{i,p,t}$ is annual stock return of firm *i* of portfolio *p* in year *t*; $E_{i,p,t}$ is income before extraordinary items and discontinued operations of firm *i* of portfolio *p* in year *t*; $P_{i,p,t-1}$ is the stock price of firm *i* of portfolio *p* in year *t*-1; $\Delta E_{i,p,t}$ is change of income before extraordinary items and discontinued operations firm *i* of portfolio *p* from year *t*-1 to year *t*. $ETP_{p,t}$ is the adjusted R^2 obtained from each portfolio-year regression. Different firms of the same portfolio in a given year share the same portfolio earnings transparency value $ETP_{p,t}$. The firm-specific earnings transparency measure $ETP_{i,t}$ is the sum of $ETI_{j,t}$ and $ETP_{p,t}$.

$C-SCORE_{i,t}$ is a composite measure of accounting conservatism for firm *i* in fiscal year *t*. The key idea is to calculate a predicted coefficient of the interaction term in Basu (1997). In other words, the regression coefficients in the Basu (1997) model are cross-sectionally calculated using relevant firm characteristics including market equity value $LNMV_i$, market to book ratio MB_i, and the debt ratio $DEBT_i$ for each year. For a given year, firm *i*'s $C-SCORE$ is the predicted value of the interaction term, where $\hat{\lambda}_{1,t}$, $\hat{\lambda}_{2,t}$, $\hat{\lambda}_{3,t}$, and $\hat{\lambda}_{4,t}$ are the estimated coefficients on firm characteristics to obtain the predicted $C-SCORE_{i,t}$ for firm *i* in year *t*. Higher $C-SCORE_{i,t}$ stands for higher level of accounting conservatism for firm *i* in year *t*.

$$C-SCORE_{i,t} = \hat{\lambda}_{1,t} + \hat{\lambda}_{2,t}LNMV_{i,t} + \hat{\lambda}_{3,t}MB_{i,t} + \hat{\lambda}_{4,t}DEBT_{i,t}$$

3.3.4 Measurement of F-SCORE

We follow Dechow et al. (2011) in calculating the $F-SCORE_{i,t}$ for each firm-year. $F-SCORE_{i,t}$ is a predicted value from a battery of accounting items. In Dechow et al. (2011), a higher $F-SCORE_{i,t}$ stands for a higher likelihood of financial misstatement, which is a signal of worsened financial disclosure quality. To be consistent with the other disclosure quality measures used in this paper, we put a negative sign before the conventional $F-SCORE_{i,t}$ so that a higher $F-SCORE_{i,t}$ in this paper implies higher disclosure quality.

$$F-SCORE_{i,t} = -\frac{e^{predicted\ value_{i,t}}}{1+e^{predicted\ value_{i,t}}} * \frac{1}{0.0037}$$

where

$$\begin{aligned} Predicted\ Value_{i,t} &= -7.90 + 0.790 * RSST\_ACC_{i,t} + 2.518 * CH\_REC_{i,t} + 1.191 * CH\_INV_{i,t} \\ &+1.979 * SOFF\_ASSETS_{i,t} + 0.171 * CH\_CS_{i,t} - 0.932 * CH\_ROA_{i,t} \\ &+1.029 * ISSUE_{i,t} \end{aligned}$$

$RSST\_ACC_{i,t}$ is the change in non-cash net operating assets scaled by average total assets. $CH\_REC_{i,t}$ is the change in receivables scaled by average total assets. $CH\_INV_{i,t}$ is the change in inventory scaled by average total assets. $SOFT\_ASSETS_{i,t}$ is the total assets minus PPE (property, plant, and equipment) and cash holdings scaled by total assets. $CH\_CS_{i,t}$ is the percentage change in cash sales excluding account receivable changes. $CH\_ROA_{i,t}$ is the change in the return on assets. $ISSUE_{i,t}$ is an indicator variable that is equal to 1 if the firm issued securities during the year and 0 otherwise. We drop the extreme observations with an $F-SCORE_{i,t}$ smaller than -10.

3.4. Empirical Design

We mainly use multivariate regressions to investigate the relationship between mutual fund herding and crash risk. First, to examine the relationship between mutual fund herding and the corporate information environment, we use $IDIOSYN_{i,t}$, $ET_{i,t}$, $C-SCORE_{i,t}$, and $F-SCORE_{i,t}$ as the dependent variables to proxy for disclosure quality. All four measures are positively associated with disclosure quality. The regression is specified below.

$$DISCLOSURE\_MEASURE_{i,t} = \beta_0 + \beta_1 HM_{i,t} + \beta_2 ROA_{i,t} + \beta_3 LEV_{i,t} + \beta_4 MB_{i,t} + \beta_5 SIZE_{i,t} + \beta_6 IO_{i,t} + INDUSTRY\ FE + YEAR\ FE + \varepsilon_{i,t}$$

The expected signs of all four disclosure quality measures are negative when the information blockage hypothesis holds and herding causes information deterioration in firms. $ROA_{i,t}$ is the income before extraordinary items divided by lagged total assets for firm *i* in year *t*. $LEV_{i,t}$ is the long-term debts divided by total assets for firm *i* in year *t*. $MB_{i,t}$ is the market value of equity divided by the book value of equity for firm *i* in year *t*. $SIZE_{i,t}$ is the natural log value of the total assets for firm *i* in year *t*. $IO_{i,t}$ is the average institutional ownership percentage from 13F across four quarters for firm *i* in year *t*.[8]

Second, to investigate the predictive relationship between mutual fund herding and crash

[8] We calculate the correlation between the overall institutional ownership percentage (*IO*) and buy/sell-herding measures (*BHM/SHM*). The Pearson correlations of *(IO, BHM*) and *(IO, SHM*) are 0.255 and 0.149, respectively. Both correlations are significant at the 1% level.

risks, we specify the regression below.

$$\begin{aligned} CRASH\_MEASURE_{i,t} = {} & \beta_0 + \beta_1 HM_{i,t-1} + \beta_2 DTURN_{i,t-1} + \beta_3 SIGMA_{i,t-1} + \beta_4\, RETURN_{i,t-1} \\ & + \beta_5 ROA_{i,t-1} + \beta_6 LEV_{i,t-1} + \beta_7 MB_{i,t-1} + \beta_8 SIZE_{i,t-1} + \beta_9 ACCM_{i,t-1} \\ & + INDUSTRY\ FE + YEAR\ FE + \varepsilon_{i,t} \end{aligned}$$

When the dummy variable $CRASH_{i,t}$ is set to be the dependent variable, we use the Probit model in our main regressions. In annual regressions, to be consistent with the prior literature, we use the 1-year lagged average annual herding measure $HM_{i,t-1}$ in year *t*-1 along with other 1-year lagged control variables to explain future crash variables in year *t*. The lagged explanatory herding variable can also help reduce the potential endogeneity problem by estimating using non-overlapping windows. According to prior studies, we control for the variables related to crash risk. $DTURN_{i,t-1}$ is the average monthly share turnover (in hundreds) divided by the total number of shares outstanding (in thousands) over the previous fiscal-year period for firm *i* in year *t*-1, where the monthly share turnover is calculated as the monthly trading volume divided by the total number of shares outstanding during the month. $SIGMA_{i,t-1}$ is the standard deviation of weekly returns over the fiscal-year period for firm *i* in year *t*-1. $RETURN_{i,t-1}$ is the buy-and-hold returns over the fiscal-year period for firm *i* in year *t*-1. $ACCM_{i,t-1}$ is the abnormal accrual calculated as in Ball and Shivakumar (2006, 2008).

3.5. The Sample

We obtain accounting data from COMPUSTAT and stock return data from the CRSP. Our sample period spans from 1989 to 2013. We strictly follow Wermers (1999) and impose the same hurdle rate of 5 trades per stock-quarter, which means that at least 5 funds are buying (selling) the stock in a given stock-quarter. As argued in Wermers (1999), it does not make economic sense to call it herding when there is barely any trade going on.[9] In our regression analysis, we follow the extant crash risk literature that defines crash measures and specifies regressions on an annual basis. As a result, we transform quarterly herding numbers into an annualized number by taking the average of the four quarters. To capture the herding intensity at the firm-year level in a meticulously way, each firm-year observation is required to have at least one herding quarter in our sample. When the firm-quarter herding measure is missing due to the number of trades in the quarter being less than 5, we use the minimum herding value in that quarter.[10] Firms in the financial industry and observations with extreme values are excluded from the sample.[11] Finally, we end up with 59,094 firm-year observations as our main sample.

[Insert Table 1 in Here]

The distribution of stock price crashes across years is shown in Panel A of Table 1. On average, 17.76 percent of firm-year observations contain at least one weekly crash across our sample period from 1989 to 2013. By definition, stock price crash is a rare event at individual firm level. At aggregate level, however, the percentage of firms experiencing crashes is nonnegligible. Panel B of Table 1 lists the industry distribution of our sample. The durable

[9] For robustness, we also remove the hurdle filter, and the results are almost the same.

[10] Intuitively, if there is no active trading from mutual funds for a firm in a specific quarter, the mutual fund herding level is low. To obtain a more accurate average annualized measure, we fill the missing *HM* quarters with quarterly minimum herding numbers similar to the winsorizing method. There are two alternative methods that replace the missing data with 0s or only use the available quarter-level data. However, these two methods suffer from serious bias in measuring herding at the annual level. For example, stock A has an *HM* of 2% in Q2 and Q4 only, while stock B has an *HM* of 2% in each quarter from Q1 to Q4. If we take the average without considering the missing quarters in stock A, we would come to a wrong conclusion that the herding levels of A and B are the same. However, stock A is barely traded by mutual funds in Q1 and Q3. Therefore, it is improper to assign the same annual herding levels to stock A and B. Similarly, it would also be incorrect to replace the missing quarters with 0s, since 0 has an economic meaning in herding calculation.(in the average herding level). By filling the two quarter gaps with their respective minimum herding numbers (i.e. winsorizing extremely low herding levels to minimum herding numbers), we obtain a lower annualized average herding number for stock A compared to that of stock B. Compared with alternatives, this method can better depict the cross-sectional herding differences at the annual level.

[11] We drop firm-year observations with negative total assets, negative book value of common equity, ROA smaller than -1, LEV greater than 1, $ACCM$ greater than 0.5, or M/B ratio greater than 20.

manufacturing industry and the computer industry are the two largest industries, accounting for 25.3% and 15.9% of the whole sample, respectively.

[Insert Table 2 in Here]

Table 2 presents the summary statistics of the variables. Both the mean and median values of $HM_{i,t-1}$ are positive, suggesting that the proportion of mutual funds traded in the same direction is higher than the benchmark percentage of independent and random trading choices. This confirms the existence of a herding phenomenon among mutual funds. Although the means (medians) of herding measures are around zero, what we are interested in is the cross-sectional differences among firms of different herding levels. Summary statistics of control variables are also shown in Table 2. The correlation matrix of key variables is listed in Table 3. We observe that $HM_{i,t-1}$ is positively correlated with both crash risk measures $CRASH_{i,t}$ and $NCSKEW_{i,t}$ in the next period. Consistent with Hutton et al. (2009), we find that the discretionary accrual measure $ACCM_{i,t-1}$ (financial opaqueness) is positively correlated with both crash measures. We also find that firms with higher cumulative returns and those with a high M/B ratio are more likely to experience stock price crashes in the next year.

[Insert Table 3 in Here]

4. Empirical Results

4.1. Mutual Fund Herding and Information Environment

To test H1, we use multivariate regressions to examine the association between mutual fund herding and corporate disclosure quality using a battery of alternative measures in Table 4. We find that overall mutual fund herding is associated with a deteriorated disclosure environment in Panel A of Table 4. To be specific, a higher $HM_{i,t-1}$ is associated with a lower $IDIOSYN_{i,t}$, which implies a lack of private information available for a specific firm. For accounting disclosure, a higher $HM_{i,t-1}$ is associated with lower accounting conservatism $C-SCORE_{i,t}$ (Khan and Watts, 2009), lower earnings transparency $ET_{i,t}$ (Barth et al., 2013), and a worsened overall earnings quality $F-SCORE_{i,t}$ (Dechow et al., 2011). In conclusion, in the face of high levels of mutual fund herding, a firm displays worsened disclosure quality. This confirms our conjecture regarding the information supply and demand equilibrium. A firm provides less information when external herding institutional investors exhibit lower information demand. We also separately examine buy-herding and sell-herding in Panel B and Panel C of Table 4. Consistent with our prediction, buy-herding and sell-herding generally have asymmetric impacts on corporate disclosure. The overall information deterioration effect is concentrated in buy-herding, illustrated by the significantly negative coefficient estimates reported through column (1) to (8) in Panel B. On the other hand, due to counteractive effects such as disciplinary impact of selling-herding, managers are generally willing to provide more information when firms experience selling-herding. We find that the sell-herding is positively correlated with all disclosure measures through column (1) to (6) except for $C-SCORE_{i,t}$ in column (7) and (8). While $IDIOSYN_{i,t}$, $ET_{i,t}$, and $F-SCORE_{i,t}$ are constructed in a way to measure the overall transparency of a firm, $C-SCORE_{i,t}$ is an accounting conservatism measure that mainly captures whether a firm can timely recognize all probable losses and expenditures. In other words, accounting conservatism requires a firm to disclose bad news more actively over good news in earnings (Basu, 1997). Even though sell-herding can exert disciplining effect and encourage managers to improve the disclosure quality, the counteractive effect might not be big enough to incentivize the firm to voluntarily disclose bad news more actively given the downward price pressure caused by sell-herding. In addition, the coefficient is only significant in column (7), while it becomes insignificant after we control for firm fixed effects in column (8).

[Insert Table 4 in Here]

4.2. Mutual Fund Herding and Crash Risk

Panel A of Table 5 presents the baseline regression results of the crash measures with respect to mutual fund herding. With either crash probability $CRASH_{i,t}$ or $NCSKEW_{i,t}$ as the dependent variable, we find that mutual fund herding in the lagged period $HM_{i,t-1}$ is always positively associated with future stock price crash risk, which supports H2 that unsigned mutual fund herding leads to elevated crash risk in the next period. Year fixed effects are included in all columns. We use both industry and firm fixed effects in our baseline regression. The economic impact is also significant. Take column (2) of Table 5 for example, one standard deviation increase of $HM_{i,t-1}$ from its mean value is associated with an increase of $NCSKEW_{i,t}$ from 0.045 to 0.074, i.e. a 64.4% increase. In column (3) of Table 6 with firm fixed effect, one standard deviation increase of $HM_{i,t-1}$ from its mean value is associated with an increase of $NCSKEW_{i,t}$ from 0.045 to 0.061, i.e. a 35.6% increase.

Furthermore, to verify our findings using a more accurate window, we adopt quarterly frequency regressions between $CRASH\_MEASURE_{i,q}$ and $HM_{i,q-1}$ in Panel B of Table 5. While crash measures and herding measures are both set at quarterly frequency, firm-level control variables are lagged for one year. Quarter fixed effects are included in all columns. Both industry and firm fixed effects are used across different columns of Panel B.

[Insert Table 5 in Here]

In Table 6, we further study buy-herding and sell-herding, respectively. According to H3, our findings are mostly concentrated in buy-herding. Consistent with this prediction, Panel A of

Table 6 reports that only the coefficient estimates of buy-herding $BHM_{i,t-1}$ in columns (1) and (2) are significant and positive, while those on $SHM_{i,t-1}$ in columns (3) and (4) are not significantly different from zero. In column (2) of Table 6, one standard deviation increase of $BHM_{i,t-1}$ from its mean value is associated with an increase of $NCSKEW_{i,t}$ from 0.045 to 0.084, i.e. a 86.7% increase. In column (3) of Table 6 with firm fixed effect, one standard deviation increase of $BHM_{i,t-1}$ from its mean value is associated with an increase of $NCSKEW_{i,t}$ from 0.045 to 0.068, i.e. a 51.1% increase. Since the coefficients of $SHM_{i,t-1}$ are not significantly from zero, we no longer interpret their economic significance. Again, we adopt the quarterly frequency regression in Panel B of Table 6, and the results are robust and similar to Panel A.

Compared with buy-herding, sell-herding can have different impacts on stock price crashes for the following reasons. First, selling pressure from mutual funds has a disciplining impact on corporate managers, similar to the argument that short sellers can help prevent firms from manipulating earnings (Massa et al., 2015; Fang et al., 2016). As a result, rather than hoard more bad news, corporate managers are under heightened scrutiny and thus may have to disclose more information. We examine the relationship between herding and disclosures for buy-herding and sell-herding separately. Consistent with the disciplining effect, we find that sell-herding (contrary to buy-herding) is actually associated with improved disclosure quality. This finding also supports our main results that buy-herding (rather than sell-herding) is more associated with stock price crashes. Second, there exists an asymmetry of information processing in buy and sell decisions in face of an information cascade. When fund managers herd to sell (buy), they tend to overweigh bad (good) news and ignore good (bad) news for a firm. This is consistent with information cascade theory in that later arrivers choose to rely on earlier arrivers' information, which means relying on bad news in sell-herding and ignoring potential good news. As a result, the information processing is asymmetric with respect to buy-herding and sell-herding. Buy-herding fund managers tend to ignore bad news for a firm, while sell-herding fund managers tend to ignore good news for a firm. Third, sell-herding trades from mutual fund managers should have already driven down stock prices ex-ante, maybe even below its intrinsic value. Therefore, it is unlikely to further observe extreme negative stock returns. The realized sell trades have already been incorporated into stock prices. Technically, it can also explain why we cannot find stock price crashes after observing strong mutual fund sell-herding.

[Insert Table 6 in Here]

From Table 7 to Table 8, we use alternative herding and crash measures to examine the robustness of our results. In Table 7, we use the modified second moment herding measure proposed by Frey et al. (2014). This measure is designed to correct the downward bias in the expected benchmark and inflated herding measure in Lakonishok et al. (1992). In Table 7, we replace the measure $HM_{i,q}$ from Lakonishok et al. (1992) with $HM2_{i,q}$ from Frey et al. (2014). The coefficients on the annualized alternative herding measure $HM2_{i,t-1}$ are significant and positive and of similar magnitude as $HM_{i,t-1}$. This suggests that our main findings are robust to herding measure construction bias.

[Insert Table 7 in Here]

In Table 8, we adopt another continuous crash risk measure $DUVOL_{i,t}$ from Chen, Hong and Stein (2001). This measure calculates the ratio of down-week returns' standard deviation over up-week returns' standard deviation in order to capture the left-skewness in the return distribution. From column (1) to column (3), our findings are consistent with the results from previous tables that unsigned mutual fund herding and buy-herding predict higher crash risk.

[Insert Table 8 in Here]

4.3. Endogeneity Tests

4.3.1. Reverse Causality

Econometrically, all endogeneity problems are caused by the violation of conditional mean independence between independent variables and residuals. As a result, an exogenous variation in independent variables can technically mitigate the endogenous concern. We use the 2004 mutual fund holding disclosure regulation change as an exogenous shock on herding. The new regulation requires that all mutual funds increase their holding portfolio disclosure frequency from a semi-annual to a quarterly basis (Agarwal et al., 2015). The rise in disclosure frequency increases the visibility of fund managers' investment strategies with respect to one another. With an investment portfolio more frequently exposed to one's peers, a mutual fund manager becomes more reluctant to collect information and develop unique investment strategies since the relative advantage of exerting more efforts on private information is lower after the regulation shock. Instead, with the decreasing marginal benefits of proprietary investment strategies, fund managers find it better to simply herd with others.[12] Agarwal et al. (2015) document evidence that supports the statement above. We exploit this one-dimensional shock on mutual fund herding around 2004 to examine whether it leads to stock price crash risk.

We specify our model in a standard difference-in-difference regression model with a treatment dummy and a time dummy. *TREAT* equals to one when a firm in the first quarter of 2004 (i.e., one quarter before the regulation becomes effective in May 2004) is held by at least one fund with a semi-annual disclosure frequency for four consecutive quarters in 2003 right before our event window starts, and zero otherwise. This leaves us with approximately 13% of

[12] We observe a significant jump in the overall mutual fund herding level right after the new regulation was enforced in May 2004. Moreover, within a longer window, the ex-post mean herding level is on average higher than the ex-ante mean herding level.

the full sample exposed to stronger ex-post disclosure frequency shocks. As shown in Table 9, the interaction terms between the treatment dummy and post-event dummy are significant and positive in both columns. This verifies our conjecture that exogenous increases in herding around the 2004 SEC regulation change leads to higher stock price crash risk. The results from this natural experiment largely rule out the possibility of reverse causality.

[Insert Table 9 in Here]

4.3.2. Omitted Variable Bias

Another alternative story is that information asymmetry is the driving force for both mutual fund herding and future stock price crashes. On the one hand, Chan, Hwang, and Mian (2005) find that the level of herding in individual stocks is positively related to the measures of information asymmetry. In other words, in the face of a noisy information environment, fund managers are more likely to give up private information seeking and turn to herding with others. On the other hand, firms with a worse information environment and hidden news are more likely to be exposed to stock price crashes afterwards. Given the arguments above, the spurious correlation between mutual fund herding and stock price crashes can be driven by the omitted information asymmetry variables. Although the natural experiment results can technically resolve this bias, we further conduct two tests to mitigate the omitted variable bias (OVB) in a more rigorous way. First, in our main table, we complement industry-fixed effects with firm-fixed effects that control for firm-specific and time-invariant factors that are potentially correlated with both mutual fund herding and stock price crash risk. Second, to rule out this specific explanation, we use propensity score matching (PSM) to control for information asymmetry factors that may both attract fund herding and cause stock price crashes. Following Chan, Hwang, and Mian (2005), we use both stock return volatility and analyst coverage as matching parameters in the first stage. In addition, we also put earnings transparency and the Big-N auditor dummy variable in the matching procedure to further control for information asymmetry between corporate managers and public investors. We divide our full sample into 3 terciles. Firms with $HM_{i,t-1}$ falling into the top tercile in a given year are set to be the treatment sample.[13] We conduct the logistic regression in the first stage. To ensure the matching accuracy, we conduct a 1-to-1 matching with the maximum caliper distance of 3%. The results in Panel A of Table 10 are consistent with Chan, Hwang, and Mian (2005) that higher analyst coverage is associated with lower herding, and higher stock return volatility is associated with higher herding. The Big-N auditor indicator and earnings transparency are also associated with lower herding. In addition, firms with larger size, higher growth, lower leverage, and higher earnings management are more likely to have more mutual fund herding. The propensity score matching enables us to explicitly control the impacts from these factors.

Panel B of Table 10 presents the characteristic comparison between the treatment sample and matching sample. The 1-on-1 matching outcome is satisfactory. By default, the treatment sample's average herding level is higher than that of the matching sample at the 1% significance level, while all other characteristics display no significant difference, including analyst coverage, stock return volatility, earnings transparency, and the Big-N indicator.

Using the matched sample, we re-do the baseline regressions. The results are reported in Panel C of Table 10. All three columns display positive and significant relationships between $HM_{i,t-1}$ and crash measures. After controlling for the impact of information asymmetry and other firm characteristics, our main findings still hold. For robustness, we also conduct the same propensity matching for $BHM_{i,t-1}$ and $SHM_{i,t-1}$, respectively, and the results are consistent

[13] We also tested alternative subsample numbers such as 4 quartiles or 5 quintiles, and the results were similar.

with our main findings in Panel B of Table 6.[14] The results using the PSM largely rule out the possibility that our results are driven by information asymmetry.

[Insert Table 10 in Here]

4.4. Additional Analyses

4.4.1. Price Impact

Another alternative explanation for our finding is that herding funds' flow-induced price impact leads to stock price crashes. To be specific, buy-herding funds' money flows into a stock in a short time period and pushes the stock price away from its fundamentals. Thus, the stock price is expected to revert back in the future given the uninformative price impact.[15] To ensure that our results are not purely driven by flow-induced price impacts (such as liquidity shocks from mutual funds), we conduct regressions that explicitly control for price impact measures using quarterly frequency regressions.

We first construct the mutual fund flow-induced price impact measure according to Edmans, Goldstein, and Jiang (2012), who examined how uninformative exogenous stock price deviations affect takeover.[16] In columns (1) to (6) of Table 11, we include the price impact measure $MFFLOWIN_{i,q-1}$ as an additional control and our main results still hold. Second, we further control for the future one or two quarters' cumulative stock returns with both industry

[14] To save space, we do not report the propensity score matching results for $BHM_{i,t-1}$ and $SHM_{i,t-1}$. We follow the same procedure as for $HM_{i,t-1}$. The results are consistent with our baseline regressions.

[15] Actually, the price impact story and our information demand story are not mutually exclusive. We argue that herding fund managers tend to make less informative moves. This can end up leading to uninformative price impacts and future price reversals.

[16] The flow-induced price impact measures in Edmans, Goldstein, and Jiang (2012) are built on the idea of previous measures, including Coval and Stanford (2007). It is also quite similar to the price impact measure in Mozaffar, Kogan, and George (2012). We choose to follow Edmans, Goldstein, and Jiang (2012) because it has the lowest sample loss.

and firm fixed effects along with time fixed effects in columns (7) to (12).[17]

As shown in the Table 11, the significant coefficient of future returns does imply its explanatory power on future crashes. In the meantime, the herding measures remain positive and significant, with a slight reduction in the magnitudes of the regression coefficients from those in Table 5. This suggests that although price impacts do impact future stock price crash risk, mutual fund herding alone has its own predictive power on future stock price crashes even after accounting for the price impact issue.[18]

[Insert Table 11 in Here]

4.4.2. Different Empirical Models

In addition, in unreported tables, we use the VaR method to re-define the crash thresholds under the generalized extreme value distribution (GEV), which is a generalized form of the Fréchet distribution, the Gumbel distribution, and the Weibull distribution. We also adjusted the crash dummy cutoffs of 3.09, 3.15, 3.25, and 3.30 standard deviations. Our results are robust to these sensitivity tests. We further examine the sensitivity of our results to model specifications. We use dynamic panel regressions for both balanced and unbalanced panels to take into account the impact of lagged dependent variables. We strictly follow the rigorous methodology for short panels in Bun and Kiviet (2003). The main results still hold for both balanced and unbalanced

[17] However, when the crash dummy is the dependent variable, the non-linear Probit regression with future returns on the right-hand side fails to generate a converged solution in our programming process. Thus, we present the results for *NCSKEW* only in this table.

[18] Furthermore, in unreported tables, we double sort our sample by both $MFFLOWIN_{i,t-1}$ and $HM_{i,t-1}$ to examine how our main findings are affected by marginal changes in price impacts. The 5×3 double-sorting results for $CRASH_{i,t}$ and $NCSKEW_{i,t}$, have two main findings. First, as mutual fund inflow-induced price impacts increase from quintile 1 to quintile 5, we observe higher crash risks in the *High_HM* group over the *Low_HM* group, except for quintile 4. Second, price impacts from mutual fund flows do seem to explain certain parts of the association between fund herding and crash risks. When we examine the difference-of-difference at the lower right corner of the table, the relationship between $HM_{i,t-1}$ and crash risks (although it remains) is significantly weakened.

dynamic panel regressions. In sum, our main results are not sensitive to different model specifications.

4.4.3. Triple Sorting by Herding, Size, and Disclosure Quality

In order to explore whether corporate disclosure quality acts as one of the potential channeling mechanisms. Table 12 reports the 10×2×2 triple sorting results by herding, size, and various corporate disclosure quality measures. We find robust and positive relation between herding and future crash likelihood in various subsamples. However, the focus of the triple sorting results is the difference-in-differences tests (DID); we intend to test whether the positive relation between herding and stock price crashes is stronger in low disclosure quality firms. We find partial support for this conjecture. Only in small firms in which information asymmetry is more severe, low disclosure quality firms seem to display stronger relation between herding and future stock price crashes. This serves as supporting evidence for corporate disclosure quality as one potential channeling mechanism between herding and stock price crashes. It is worth noting that the results should be interpreted with caution because this partial evidence cannot rule out other alternative parallel mechanisms through which herding leads to stock price crashes.

[Insert Table 12 in Here]

5. Conclusion

This paper investigates the impacts of mutual fund herding behaviours on corporate disclosure and consequently on stock price crash risk. First, we find that mutual fund herding deteriorates corporate disclosure quality. Firms with high mutual fund herding have less private information available, lower earnings transparency, higher likelihoods of accounting errors, and lower accounting conservatism. Second, we find a strong predictive relationship between mutual fund herding and stock price crashes. We use one natural experiment using the 2004 SEC regulation change on mutual fund portfolio disclosure to address the reverse causality issue. To mitigate the omitted variable bias that our findings are driven by other factors (such as information asymmetry between outside investors and managers), we further control for firm fixed effects and adopt propensity score matching to explicitly control for information asymmetry. Our main findings are also robust to alternative herding and crash measures and different regression model specifications. Third, we find that the predictive relationship between mutual fund herding and stock price crashes is concentrated in buy-herding rather than sell-herding. We further take into account the price impact factor and discover that our findings cannot be fully explained by it. Finally, after triple sorting the sample by firm size, disclosure quality, and herding measures, we find that the relation between mutual fund herding and stock price crashes is much stronger in the subsample of firms with smaller size and with lower disclosure quality. This test partially supports our conjecture that disclosure quality acts as one of the channels through which mutual fund herding leads to stock price crashes, when firms suffer severe information asymmetry.

Previous literature focuses on either the impact of institutional investor herding on stock returns or the relation between institutional ownership and crash risk. Our study complements the literature by examining herding as a specific trading behaviour of institutional investors on crash risk. Such predictive relation between mutual fund herding intensity and stock price crash provides useful implications for investment practice; strong mutual fund buy-herding signal can be an alert for holding investors.

References

Admati, A. and P. Pfleiderer. 1997. Does it all add up? Benchmarks and the compensation of active portfolio managers. *Journal of Business* 70: 323-350.

Agarwal, V., K. Mullally, Y. Tang, and B. Yang. 2015. Mandatory portfolio disclosure, stock liquidity, and mutual fund performance. *Journal of Finance* 70:2732-2776.

Ajinkya, B., S. Bhojraj, and P. Sengupta. 2005. The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of Accounting Research* 43:343–76.

An, H. and T. Zhang. 2013. Stock price synchronicity, crash risk, and institutional investors. *Journal of Corporate Finance* 21: 1-15

Banerjee, A. 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107: 797-817.

Ball, R. and L. Shivakumar. 2006. Role of accruals in asymmetrically timely gain and loss recognition. *Journal of Accounting Research* 44: 207-242.

Ball, R. and L. Shivakumar. 2008. Earnings quality at initial public offerings. *Journal of Accounting and Economics* 45: 324-349.

Barth, M., Y. Konchitchki, and W. Landsman. 2013. Cost of capital and earnings transparency. *Journal of Accounting and Economics* 55: 206-224.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24: 3-37.

Beyer, A., D. Cohen, T. Lys and B. Walther, 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50: 296-343.

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: 992-1026.

Bird, A. and S. Karolyi. 2016. Do institutional investors demand public disclosure? *Review of Financial Studies* 29: 3245-3277.

Black, F. 1976. Studies of stock price volatility changes. *Proceeding of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*: 177-181.

Bleck, A. and X. Liu. 2007. Market transparency and the accounting regime. *Journal of Accounting Research* 45: 229-256.

Brennan, M. 1993. Agency and asset pricing. *Working Paper*. (Anderson School, UCLA #6–93, 1993)

Bun, M. and J. Kiviet. 2003. On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters* 79:145-152.

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

Bushee, B. and C. Noe. 2000. Corporate disclosure practices, institutional investors, and stock return volatility. *Journal of Accounting Research* 38: 171-202.

Callen, J. and X. Fang. 2015. Short interest and stock price crash risk. *Journal of Banking and Finance* 60: 181-194.

Campbell, J., and L. Hentschel. 1992. No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* 31: 281-318.

Chiang, T. and D. Zheng. 2010. An empirical analysis of herd behavior in global stock markets. *Journal of Banking and Finance* 34:1911-1921.

Chevalier, J. and G. Ellison. 1999. Are some mutual fund managers better than others?

Cross-sectional patterns in behavior and performance. *Journal of Finance* 54: 875-899.
Chan, K., C.-Y. Hwang, and M.G. Mian. 2005. Mutual fund herding and dispersion of analysts' earnings forecasts. *Working Paper*.
Chen, J., H. Hong, and J. Stein. 2001. Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics* 61: 345-381.
Choi, N. and R. Sias. 2009. Institutional industry herding. *Journal of Financial Economics* 94: 469-491.
Christie, A. 1982. The stochastic behavior of common stock variances- Value, leverage, and interest rate effects. *Journal of Financial Economics* 10:407-432.
Coval, J. and E. Stafford. 2007. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics* 86: 479-512.
Dasgupta, A., A. Prat, and M. Verardo. 2011a. Institutional trade persistence and long-term equity returns. *Journal of Finance* 66: 635-653.
Dasgupta, A., A. Prat, and M. Verardo. 2011b. The price impact of institutional herding. *Review of Financial Studies* 24:892-925.
Dechow, P., W. Ge, C. Larson, and R. Sloan. 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28: 17-82.
Edmans, A., I. Goldstein, and W. Jiang. 2012. The real effects of financial markets: The impact of prices on takeovers. *Journal of Finance* 67: 933-971.
Fang, V., A. Huang, and J. Karpoff. 2016. Short selling and earnings management: A controlled experiment. *Journal of Finance* 71: 1251-1294.
Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32: 137-164.
French, K., G. Schwert, and R. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19: 3-29.
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.
Froot, K., D. Scharfstein, and J. Stein. 1992. Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance* 47: 1461-1484.
Graham, J. 1999. Herding among investment newsletters: Theory and evidence. *Journal of Finance* 54: 237-268.
Healy, P., A. Hutton, and K. Palepu.1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. *Contemporary Accounting Research* 16:485-520.
Hirshleifer, D., A. Subrahmanyam, and S. Titman. 1994. Security analysis and trading patterns when some investors receive information before others. *Journal of Finance* 49: 1665-1698.
Hirshleifer, D. and S. Teoh. 2003. Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management* 9: 25-66.
Hong, H. and J. Stein. 2003. Differences of opinion, short-sales constraints, and market crashes. *Review of Financial Studies* 16: 487-525.
Hutton, A., A. Marcus, and H. Tehranian. 2009. Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics* 94: 67-86.
Jin, L. and S. Myers. 2006. R^2 around the world: New theory and new tests. *Journal of Financial Economics* 79: 257-292.
Karpoff, J. M. and X. Lou. 2010. Short sellers and financial misconduct. *Journal of Finance* 65: 1879–1913.

Khan, M. and R. Watts. 2009. Estimation and empirical properties of a firm-year measure of accounting conservatism. *Journal of Accounting and Economics* 48: 132-150.

Khan, M., L. Kogan, and G. Serafeim. 2012. Mutual fund trading pressure: Firm-level stock price impact and timing of SEOs. *Journal of Finance* 67: 1371-1395.

Kim, Y., H. Li, and S. Li. 2014. Corporate social responsibility and stock price crash risk. *Journal of Banking and Finance* 43: 1-13.

Kim, J.-B., Y. Li, and L. Zhang. 2011a. Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics* 100: 639-662.

Kim, J.-B., Y. Li, and L. Zhang. 2011b. CFOs versus CEOs: Equity incentives and crashes. *Journal of Financial Economics* 101: 713-730.

Kim, J.-B. and L. Zhang. 2016. Accounting conservatism and stock price crash risk: Firm-level evidence. *Contemporary Accounting Research* 33: 412-441.

Lakonishok, J., A. Shleifer, and R. Vishny. 1992. The impact of institutional trading on stock Prices. *Journal of Financial Economics* 32: 23-43.

Massa, M., B. Zhang, and H. Zhang. 2015. The invisible hand of short-selling: Does short-selling discipline earnings manipulation? *Review of Financial Studies* 28: 1701-1736

Morck, R., B. Yeung, and W. Yu. 2000. The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58: 215-260.

Pindyck, R.1984. Risk, inflation, and the stock market. *American Economic Review* 74: 334-351.

Prendergast, C. and L. Stole. 1996. Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning. *Journal of Political Economy* 104: 1105-1134.

Roll, R. 1992. A mean/variance analysis of tracking error. *Journal of Portfolio Management* 18: 13-22.

Scharfstein, D. and J. Stein. 1990. Herd behavior and investment. *American Economic Review* 80: 465-479.

Sias, R. 2004. Institutional herding. *Review of Financial Studies* 17: 165-206.

Trueman, B. 1994. Analyst forecasts and herding behavior. *Review of Financial Studies* 7:97-124.

Welch, I. 1992. Sequential sales, leaning, and cascades. *Journal of Finance* 47: 695-732.

Wermers, R.1999. Mutual fund herding and the impact on stock prices. *Journal of Finance* 54: 581-622.

Wylie, S. 2005. Fund manager herding: A test of the accuracy of empirical results using UK data. *Journal of Business* 78: 381-403.

Yuan, K. 2005. Asymmetric price movements and borrowing constraints: A REE model of crisis, contagion, and confusion. *Journal of Finance* 60: 379-411.

Zwiebel, J. 1995. Corporate conservatism and relative compensation. *Journal of Political Economy* 103: 1-25.

Table 1: Sample Description

Panel A: Crash Distribution by Fiscal Year

Fiscal Year	No. of firms	No. of Firms without crash	No. of firms with crash	Crash Percentage
1989	1,175	1,012	163	13.87%
1990	1,331	1,153	178	13.37%
1991	1,480	1,288	192	12.97%
1992	1,619	1,373	246	15.19%
1993	1,815	1,505	310	17.08%
1994	2,080	1,783	297	14.28%
1995	2,367	2,030	337	14.24%
1996	2,479	2,144	335	13.51%
1997	2,665	2,257	408	15.31%
1998	2,838	2,367	471	16.60%
1999	2,766	2,346	420	15.18%
2000	2,554	2,119	435	17.03%
2001	2,579	2,141	438	16.98%
2002	2,823	2,277	546	19.34%
2003	2,797	2,313	484	17.30%
2004	2,745	2,166	579	21.09%
2005	2,715	2,126	589	21.69%
2006	2,701	2,119	582	21.55%
2007	2,641	2,107	534	20.22%
2008	2,641	2,022	619	23.44%
2009	2,638	2,208	430	16.30%
2010	2,536	2,116	420	16.56%
2011	2,422	2,009	413	17.05%
2012	2,384	1,841	543	22.78%
2013	2,303	1,778	525	22.80%
Total	59,094	48,600	10,494	17.76%

Panel B: Sample Firm Distribution by SIC Code

	SIC codes	No.of firms
Agriculture	0-999	184
Mining and construction	1000-1299, 1400-1999	1,670
Food	2000-2111	1,675
Textile, printing, and publishing	2200-2790	3,490
Chemicals	2800-2824, 2840-2899	1,928
Pharmaceuticals	2830-2836	3,327
Extractive industries	2900-2999, 1300-1399	2,664
Durable manufacturers	3000-3569, 3580-3669, 3680-3999	14,971
Computers	7370-7379, 3570-3579, 3670-3679	9,410
Transportation	4000-4899	3,891
Utilities	4900-4999	3,292
Retail	5000-5999	6,794
Service	7000-7369, 7380-9999	5,798
Total		59,094

Table 2: Descriptive Statistics

This table presents the summary statistics of variables used in this paper. $CRASH_{i,t}$ is the indicator variable which equals to 1 if a firm experiences at least 1 crash week, i.e. firm-specific weekly return falls below 3.2 standard deviations of the mean value in a given year, and 0 otherwise. $NCSKEW_{i,t}$ is defined as the negative value of sample skewness of the logarithm of residual returns, where sample skewness is obtained by dividing the unique symmetric unbiased third moment estimator by the standard deviation of $W_{i,t}$ raised to third power. $HM_{i,t-1}$ is the unsigned measure of mutual fund herding intensity calculated as the average of quarterly herding measure $HM_{i,q}$ in year t-1, where $HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$. $BHM_{i,t-1}$ is the buy-herding intensity for firm *i* in year t-1 calculated as the average of quarterly buy-herding measure $BHM_{i,q}$ in year t-1, where $BHM_{i,q} = HM_{i,q}|\ p_{i,q} > E[p_{i,q}]$. $SHM_{i,t-1}$ is the sell-herding intensity for firm *i* in year *t*-1 calculated as the average of quarterly sell-herding measure $SHM_{i,q}$ in year t-1, where $SHM_{i,q} = HM_{i,q}|\ p_{i,q} < E[p_{i,q}]$. $DTURN_{i,t-1}$ is the average monthly share turnover difference for firm *i* between fiscal-year *t-1* and *t-2*. $SIGMA_{i,t-1}$ is the standard deviation of weekly returns over the fiscal-year period for firm *i* in year *t*-1. $RETURN_{i,t-1}$ is the buy-and-hold returns over the fiscal-year period for firm *i* in year *t*-1. $ROA_{i,t-1}$ is the income before extraordinary items divided by lagged total assets for firm *i* in year *t*-1. $LEV_{i,t-1}$ is the long-term debts divided by total assets for firm *i* in year *t*-1. $MB_{i,t-1}$ is the market value of equity divided by the book value of equity for firm *i* in year *t*-1. $SIZE_{i,t-1}$ is the natural log value of total assets for firm *i* in year *t*-1. $ACCM_{i,t-1}$ is the absolute value of abnormal accruals, where abnormal accruals are estimated from the model suggested by Ball and Shivakumar (2006, 2008). $IO_{i,t}$ is the total institutional ownership for firm *i* in year *t*. $IDIOSYN_{i,t}$ is the idiosyncratic volatility for firm *i* in year *t*. $ET_{i,t}$ is the firm-specific earnings transparency measure calculated as the sum of explanatory power of earnings on returns from both industry level and portfolio level as in Barth et al. (2013). $FSCORE_{i,t}$ is the negative value of predicted likelihood of misstatement proposed by Dechow et al. (2011). $CSCORE_{i,t}$ is a composite measure of accounting conservatism for firm *i* in year *t* by Khan and Watts (2009).

	N	MEAN	SD	P5	P25	P50	P75	P95
$CRASH_{i,t}$	59,094	0.178	0.382	0.000	0.000	0.000	0.000	1.000
$NCSKEW_{i,t}$	59,094	0.045	0.819	-1.129	-0.389	0.000	0.409	1.412
$DUVOL_{i,t}$	59,094	-0.111	0.683	-1.207	-0.554	-0.119	0.321	1.012
$HM_{i,t-1}$	59,094	0.012	0.059	-0.093	-0.014	0.011	0.038	0.104
$BHM_{i,t-1}$	59,094	-0.071	0.042	-0.133	-0.103	-0.074	-0.044	0.002
$SHM_{i,t-1}$	59,094	-0.086	0.046	-0.145	-0.118	-0.091	-0.061	-0.010
$DTURN_{i,t-1}$	59,094	-0.023	2.015	-1.384	-0.259	-0.004	0.242	1.356
$SIGMA_{i,t-1}$	59,094	0.066	0.036	0.026	0.041	0.058	0.082	0.134
$RETURN_{i,t-1}$	59,094	0.170	0.766	-0.611	-0.205	0.069	0.359	1.217
$ROA_{i,t-1}$	59,094	0.026	0.146	-0.255	0.003	0.047	0.093	0.194
$LEV_{i,t-1}$	59,094	0.167	0.165	0.000	0.005	0.132	0.279	0.480
$MB_{i,t-1}$	59,094	2.718	2.438	0.641	1.273	1.977	3.245	7.381
$SIZE_{i,t-1}$	59,094	6.246	1.912	3.504	4.818	6.008	7.477	9.790
$ACCM_{i,t-1}$	59,094	0.062	0.068	0.003	0.018	0.041	0.081	0.199
$IO_{i,t}$	63,619	0.509	0.261	0.103	0.290	0.506	0.725	0.932
$IDIOSYN_{i,t}$	63,619	1.301	0.987	-0.274	0.650	1.295	1.945	2.910
$ET_{i,t}$	55,713	0.469	0.188	0.168	0.324	0.481	0.610	0.763
$FSCORE_{i,t}$	55,428	-1.029	0.740	-2.240	-1.315	-0.875	-0.537	-0.283
$CSCORE_{i,t}$	54,918	0.102	0.095	-0.049	0.045	0.102	0.159	0.251

Table 3: Coefficients of Correlation

This table presents the correlation matrix of the key variables in this paper. $CRASH_{i,t}$ is the indicator variable which equals to 1 if a firm experiences at least 1 crash week, i.e. firm-specific weekly return falls below 3.2 standard deviations of the mean value in a given year, and 0 otherwise. $NCSKEW_{i,t}$ is defined as the negative value of sample skewness of the logarithm of residual returns, where sample skewness is obtained by dividing the unique symmetric unbiased third moment estimator by the standard deviation of $W_{i,t}$ raised to third power. $HM_{i,t-1}$ is the unsigned measure of mutual fund herding intensity calculated as the average of quarterly herding measure $HM_{i,q}$ in year *t*-1, where $HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$. $BHM_{i,t-1}$ is the buy-herding intensity for firm *i* in year *t*-1 calculated as the average of quarterly buy-herding measure $BHM_{i,q}$ in year *t*-1, where $BHM_{i,q} = HM_{i,q}|\ p_{i,q} > E[p_{i,q}]$. $SHM_{i,t-1}$ is the sell-herding intensity for firm *i* in year *t*-1 calculated as the average of quarterly sell-herding measure $SHM_{i,q}$ in year *t*-1, where $SHM_{i,q} = HM_{i,q}|\ p_{i,q} < E[p_{i,q}]$. $DTURN_{i,t-1}$ is the average monthly share turnover difference for firm *i* between fiscal-year *t-1* and *t-2*. $SIGMA_{i,t-1}$ is the standard deviation of weekly returns over the fiscal-year period for firm *i* in year *t*-1. $RETURN_{i,t-1}$ is the buy-and-hold returns over the fiscal-year period for firm *i* in year *t*-1. $ROA_{i,t-1}$ is the income before extraordinary items divided by lagged total assets for firm *i* in year *t*-1. $LEV_{i,t-1}$ is the long-term debts divided by total assets for firm *i* in year *t*-1. $MB_{i,t-1}$ is the market value of equity divided by the book value of equity for firm *i* in year *t*-1. $SIZE_{i,t-1}$ is the natural log value of total asset for firm *i* in year *t*-1. $ACCM_{i,t-1}$ is the absolute value of abnormal accruals, where abnormal accruals are estimated from the model suggested by Ball and Shivakumar (2006, 2008). $IO_{i,t-1}$ is the institutional ownership for firm *i* in year *t*-1. *p*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

	$CRASH_{i,t}$	$NCSKEW_{i,t}$	$DUVOL_{i,t}$	$HM_{i,t-1}$	$BHM_{i,t-1}$	$SHM_{i,t-1}$	$DTURN_{i,t-1}$	$SIGMA_{i,t-1}$	$RETURN_{i,t-1}$	$ROA_{i,t-1}$	$LEV_{i,t-1}$	$MB_{i,t-1}$	$SIZE_{i,t-1}$	$ACCM_{i,t-1}$
$CRASH_{i,t}$	1.000													
$NCSKEW_{i,t}$	0.623***	1.000												
	(0.00)													
$DUVOL_{i,t}$	0.437***	0.815***	1.000											
	(0.00)	(0.00)												
$HM_{i,t-1}$	0.034***	0.051***	0.045***	1.000										
	(0.00)	(0.00)	(0.00)											
$BHM_{i,t-1}$	0.048***	0.077***	0.083***	0.488***	1.000									
	(0.00)	(0.00)	(0.00)	(0.00)										
$SHM_{i,t-1}$	0.007*	0.007*	0.004	0.505***	-0.340***	1.000								
	(0.08)	(0.09)	(0.35)	(0.00)	(0.00)									
$DTURN_{i,t-1}$	0.005	0.024***	0.021***	0.027***	0.059***	-0.023***	1.000							
	(0.19)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)								
$SIGMA_{i,t-1}$	-0.005	-0.013***	-0.083***	0.075***	-0.035***	0.048***	0.058***	1.000						
	(0.24)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)							
$RETURN_{i,t-1}$	0.031***	0.075***	0.083***	0.053***	0.173***	-0.133***	0.101***	0.057***	1.000					
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)						
$ROA_{i,t-1}$	0.033***	0.069***	0.109***	-0.031***	0.056***	-0.028***	0.041***	-0.400***	0.157***	1.000				
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)					
$LEV_{i,t-1}$	-0.030***	-0.018***	-0.001	-0.017***	0.011***	-0.005	0.008**	-0.123***	-0.033***	-0.024***	1.000			
	(0.00)	(0.00)	(0.77)	(0.00)	(0.01)	(0.23)	(0.05)	(0.00)	(0.00)	(0.00)				
$MB_{i,t-1}$	0.055***	0.094***	0.085***	0.077***	0.114***	0.014***	0.046***	0.035***	0.283***	0.039***	-0.043***	1.000		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)			
$SIZE_{i,t-1}$	-0.003	0.052***	0.100***	0.145***	0.150***	0.222***	0.007*	-0.379***	-0.012***	0.224***	0.339***	-0.004	1.000	
	(0.52)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.09)	(0.00)	(0.00)	(0.00)	(0.00)	(0.36)		
$ACCM_{i,t-1}$	0.017***	0.007*	-0.028***	0.037***	-0.014***	0.008**	0.006	0.303***	-0.015***	-0.253***	-0.160***	0.138***	-0.244***	1.000
	(0.00)	(0.09)	(0.00)	(0.00)	(0.00)	(0.04)	(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	

Table 4: Mutual Fund Herding and Corporate Disclosure Quality

This table presents the regression results of corporate disclosure quality on mutual fund herding. $IDIOSYN_{i,t}$ is the idiosyncratic volatility for firm *i* in year *t*. $ET_{i,t}$ is the firm-specific earnings transparency measure calculated as the sum of explanatory power of earnings on returns from both industry level and portfolio level, as in Barth et al. (2013). $F-SCORE_{i,t}$ is the negative value of predicted likelihood of misstatement proposed by Dechow et al. (2011). $C-SCORE_{i,t}$ by Khan and Watts (2009) is a composite measure of accounting conservatism for firm *i* in year *t*. $HM_{i,t}$ is the unsigned measure of mutual fund herding intensity calculated as the average of quarterly herding measure $HM_{i,q}$ in year *t*, where $HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$. $BHM_{i,t}$ is the buy-herding intensity for firm *i* in year *t* calculated as the average of quarterly buy-herding measure $BHM_{i,q}$ in year *t*-1, where $BHM_{i,q} = HM_{i,q}|\, p_{i,q} > E[p_{i,q}]$. $SHM_{i,t}$ is the sell-herding intensity for firm *i* in year *t* calculated as the average of quarterly sell-herding measure $SHM_{i,q}$ in year *t*, where $SHM_{i,q} = HM_{i,q}|\, p_{i,q} < E[p_{i,q}]$. Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *t*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Panel A: Mutual Fund Herding and Disclosure Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$IDIOSYN_{i,t}$	$IDIOSYN_{i,t}$	$ET_{i,t}$	$ET_{i,t}$	$FSCORE_{i,t}$	$FSCORE_{i,t}$	$CSCORE_{i,t}$	$CSCORE_{i,t}$
Constant	2.716***	2.384***	0.665***	0.584***	-0.794***	-0.112**	0.282***	0.285***
	(33.82)	(44.44)	(46.50)	(56.48)	(-6.56)	(-1.98)	(36.14)	(69.27)
$HM_{i,t}$	-1.364***	-1.298***	-0.122***	-0.077***	-0.870***	-0.653***	-0.031***	-0.008*
	(-17.79)	(-15.25)	(-8.94)	(-5.01)	(-13.09)	(-10.91)	(-6.21)	(-1.72)
$ROA_{i,t}$	0.005	-0.128***	0.025***	-0.003	-0.401***	-0.356***	-0.049***	-0.059***
	(0.17)	(-3.55)	(4.11)	(-0.31)	(-11.43)	(-8.08)	(-16.12)	(-16.20)
$LEV_{i,t}$	0.373***	0.378***	-0.009*	0.005	-0.193***	-0.022	0.127***	0.102***
	(13.27)	(8.98)	(-1.72)	(0.59)	(-4.83)	(-0.54)	(35.38)	(26.39)
$MB_{i,t}$	-0.029***	-0.027***	-0.009***	-0.012***	-0.007***	-0.026***	-0.014***	-0.012***
	(-16.84)	(-12.70)	(-24.44)	(-24.03)	(-3.38)	(-11.93)	(-43.53)	(-28.63)
$SIZE_{i,t}$	-0.161***	-0.144***	0.010***	0.017***	-0.001	-0.187***	-0.032***	-0.032***
	(-40.74)	(-14.23)	(17.30)	(8.95)	(-0.12)	(-15.92)	(-123.49)	(-39.37)
$IO_{i,t}$	-0.418***	-0.285***	0.013***	0.064***	-0.037	0.126***	-0.022***	-0.029***
	(-18.25)	(-7.69)	(3.31)	(9.26)	(-1.30)	(3.67)	(-13.79)	(-10.60)
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	63,619	63,619	55,713	55,713	55,428	55,428	54,918	54,918
adj. R^2	0.327		0.274		0.090		0.693	
Overall R^2		0.315		0.253		0.003		0.683

Panel B: Mutual Fund Buy-Herding and Disclosure Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$IDIOSYN_{i,t}$	$IDIOSYN_{i,t}$	$ET_{i,t}$	$ET_{i,t}$	$FSCORE_{i,t}$	$FSCORE_{i,t}$	$CSCORE_{i,t}$	$CSCORE_{i,t}$
Constant	2.514***	2.168***	0.648***	0.566***	-0.919***	-0.230***	0.278***	0.283***
	(31.25)	(40.14)	(44.89)	(54.08)	(-7.64)	(-4.06)	(35.62)	(66.49)
$BHM_{i,t}$	-2.524***	-2.497***	-0.206***	-0.194***	-1.517***	-1.241***	-0.048***	-0.019***
	(-25.33)	(-23.65)	(-11.47)	(-10.06)	(-17.94)	(-16.54)	(-7.51)	(-2.88)
$ROA_{i,t}$	0.057**	-0.073**	0.031***	0.003	-0.366***	-0.330***	-0.048***	-0.059***
	(2.10)	(-2.06)	(5.04)	(0.32)	(-10.47)	(-7.53)	(-15.77)	(-16.12)
$LEV_{i,t}$	0.377***	0.325***	-0.008	0.003	-0.188***	-0.048	0.128***	0.102***
	(13.55)	(7.81)	(-1.51)	(0.37)	(-4.73)	(-1.16)	(35.54)	(26.33)
$MB_{i,t}$	-0.024***	-0.021***	-0.009***	-0.011***	-0.004**	-0.023***	-0.014***	-0.012***
	(-14.65)	(-9.68)	(-24.10)	(-23.27)	(-2.20)	(-10.57)	(-43.37)	(-28.34)
$SIZE_{i,t}$	-0.163***	-0.141***	0.010***	0.017***	-0.001	-0.184***	-0.032***	-0.032***
	(-42.36)	(-14.19)	(16.97)	(8.98)	(-0.26)	(-15.79)	(-123.60)	(-39.31)
$IO_{i,t}$	-0.393***	-0.269***	0.016***	0.068***	-0.021	0.135***	-0.021***	-0.029***
	(-17.42)	(-7.32)	(4.11)	(9.94)	(-0.73)	(3.97)	(-13.42)	(-10.46)
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	63,619	63,619	55,713	55,713	55,428	55,428	54,918	54,918
adj. R^2	0.337		0.274		0.096		0.693	
Overall R^2		0.323		0.254		0.004		0.683

Panel C: Mutual Fund Sell-Herding and Disclosure Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$IDIOSYN_{i,t}$	$IDIOSYN_{i,t}$	$ET_{i,t}$	$ET_{i,t}$	$FSCORE_{i,t}$	$FSCORE_{i,t}$	$CSCORE_{i,t}$	$CSCORE_{i,t}$
Constant	2.855***	2.520***	0.684***	0.607***	-0.639***	-0.001	0.280***	0.284***
	(35.70)	(45.77)	(47.44)	(57.14)	(-5.19)	(-0.02)	(35.33)	(69.06)
$SHM_{i,t}$	0.699***	0.924***	0.100***	0.155***	0.834***	0.709***	-0.023***	-0.006
	(8.05)	(9.97)	(5.83)	(8.28)	(10.40)	(10.31)	(-3.70)	(-0.92)
$ROA_{i,t}$	0.050*	-0.123***	0.030***	0.000	-0.355***	-0.343***	-0.049***	-0.059***
	(1.81)	(-3.34)	(4.90)	(0.04)	(-10.07)	(-7.76)	(-16.08)	(-16.30)
$LEV_{i,t}$	0.406***	0.413***	-0.005	0.007	-0.164***	-0.009	0.127***	0.102***
	(14.41)	(9.76)	(-0.98)	(0.78)	(-4.11)	(-0.22)	(35.10)	(26.46)
$MB_{i,t}$	-0.032***	-0.029***	-0.010***	-0.012***	-0.008***	-0.027***	-0.014***	-0.012***
	(-18.42)	(-13.10)	(-24.90)	(-23.85)	(-4.20)	(-11.96)	(-43.73)	(-28.63)
$SIZE_{i,t}$	-0.171***	-0.155***	0.009***	0.016***	-0.010**	-0.194***	-0.032***	-0.032***
	(-43.02)	(-15.21)	(15.09)	(8.30)	(-2.23)	(-16.50)	(-121.84)	(-39.46)
$IO_{i,t}$	-0.410***	-0.225***	0.012***	0.065***	-0.033	0.156***	-0.022***	-0.029***
	(-17.72)	(-5.95)	(3.01)	(9.44)	(-1.14)	(4.57)	(-13.77)	(-10.56)
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	63,619	63,619	55,713	55,713	55,428	55,428	54,918	54,918
adj. R^2	0.320		0.273		0.087		0.693	
Overall R^2		0.306		0.252		0.002		0.683

Table 5: Mutual Fund Herding and Crash Risk

This table presents the regression results of crash risk on mutual fund herding. $CRASH_{i,t}$ is an indicator variable which equals to 1 if a firm experiences at least 1 crash week, i.e. firm-specific weekly return falls below 3.2 standard deviations of the mean value in a given year, and 0 otherwise. $NCSKEW_{i,t}$ is defined as the negative value of sample skewness of the logarithm of residual returns for firm *i* in year *t*, where sample skewness is obtained by dividing the unique symmetric unbiased third moment estimator by the standard deviation raised to third power. $HM_{i,t-1}$ is the unsigned measure of mutual fund herding intensity calculated as the average of quarterly herding measure $HM_{i,q}$ in year *t*-1, where $HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$. $CRASH_{i,q}$ is an indicator variable which equals to 1 if firm *i* experiences at least 1 crash week in quarter *q*, i.e. firm-specific weekly return falls below 3.2 standard deviations of the mean value, in a given quarter, and 0 otherwise. $NCSKEW_{i,q}$ is defined as the negative value of sample skewness of the logarithm of residual returns for firm *i* in quarter *q*, where sample skewness is obtained by dividing the unique symmetric unbiased third moment estimator by the standard deviation raised to third power. $DTURN_{i,q-1}$ is the average monthly share turnover difference for firm *i* between quarter *q-1* and *q-2*. $SIGMA_{i,q-1}$ is the standard deviation of daily returns for firm *i* in quarter *q*-1. $RETURN_{i,q-1}$ is the cumulative buy and hold returns for firm *i* over quarter *t*-1. Column (1) in Panel A and Column (1) in Panel B are Probit regressions. Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *z(t)*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Panel A: Panel Regression at Annual Frequency

	(1) $CRASH_{i,t}$	(2) $NCSKEW_{i,t}$	(3) $NCSKEW_{i,t}$
Constant	-1.191***	-0.511***	-0.920***
	(-6.45)	(-4.97)	(-17.17)
$HM_{i,t-1}$	0.438***	0.488***	0.264***
	(4.15)	(7.92)	(3.83)
$DTURN_{i,t-1}$	0.001	0.005**	0.002
	(0.27)	(2.26)	(1.07)
$SIGMA_{i,t-1}$	-0.550**	0.054	-0.757***
	(-2.28)	(0.38)	(-4.29)
$RETURN_{i,t-1}$	0.027***	0.054***	0.052***
	(3.25)	(9.14)	(8.32)
$ROA_{i,t-1}$	0.474***	0.370***	0.467***
	(9.07)	(11.70)	(10.46)
$LEV_{i,t-1}$	0.012	-0.072***	-0.217***
	(0.26)	(-2.85)	(-5.10)
$MB_{i,t-1}$	0.018***	0.021***	0.040***
	(6.56)	(12.05)	(15.61)
$SIZE_{i,t-1}$	-0.007	0.021***	0.146***
	(-1.56)	(8.90)	(16.20)
$ACCM_{i,t-1}$	0.271***	0.160***	0.144**
	(2.77)	(2.79)	(2.23)
Industry FE	Yes	Yes	No
Firm FE	No	No	Yes
Year FE	Yes	Yes	Yes
N	59,094	59,094	59,094
pseudo R^2	0.018		
adj. R^2		0.030	
Overall R^2			0.013

Panel B: Panel Regression at Quarterly Frequency

	(1)	(2)	(3)
	$CRASH_{i,q}$	$NCSKEW_{i,q}$	$NCSKEW_{i,q}$
Constant	-1.924***	-0.104	-0.501***
	(-6.72)	(-1.02)	(-5.15)
$HM_{i,q-1}$	0.171***	0.112***	0.100***
	(3.64)	(4.39)	(3.71)
$DTURN_{i,q-1}$	0.012***	-0.001	0.000
	(3.81)	(-0.52)	(0.24)
$SIGMA_{i,q-1}$	-1.104***	0.470***	0.166**
	(-7.42)	(6.92)	(2.02)
$RETURN_{i,q-1}$	0.023	0.560***	0.608***
	(1.33)	(51.92)	(53.02)
$ROA_{i,t-1}$	0.265***	0.094***	0.176***
	(8.36)	(7.01)	(8.85)
$LEV_{i,t-1}$	0.016	-0.032***	-0.101***
	(0.53)	(-2.67)	(-4.92)
$MB_{i,t-1}$	0.013***	0.006***	0.015***
	(7.75)	(7.63)	(13.43)
$SIZE_{i,t-1}$	-0.011***	0.003***	0.083***
	(-3.66)	(2.75)	(19.38)
$ACCM_{i,t-1}$	0.236***	0.103***	0.106***
	(3.68)	(3.74)	(3.35)
Industry FE	Yes	Yes	No
Firm FE	No	No	Yes
Quarter FE	Yes	Yes	Yes
N	252,665	252,665	252,665
pseudo R^2	0.023		
adj. R^2		0.034	
Overall R^2			0.023

Table 6: Mutual Fund Buy (Sell) Herding and Crash Risk

This table presents the regression results of crash risk on mutual fund buying and selling herding, respectively. $BHM_{i,t-1}$ is the buying herding intensity for firm *i* in year *t*-1 calculated as the average of quarterly buy-herding measure $BHM_{i,q}$ in year *t*-1, where $BHM_{i,q} = HM_{i,q} | \, p_{i,q} > E[p_{i,q}]$. $SHM_{i,t-1}$ is the selling herding intensity for firm *i* in year *t*-1 calculated as the average of quarterly selling herding measure $SHM_{i,q}$ in year *t*-1, where $SHM_{i,q} = HM_{i,q} | \, p_{i,q} < E[p_{i,q}]$. Column (1) and (4) in Panel A and Panel B are Probit regressions. Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *z(t)*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Panel A: Panel Regression at Annual Frequency

	(1) $CRASH_{i,t}$	(2) $NCSKEW_{i,t}$	(3) $NCSKEW_{i,t}$	(4) $CRASH_{i,t}$	(5) $NCSKEW_{i,t}$	(6) $NCSKEW_{i,t}$
Constant	-1.115***	-0.444***	-0.869***	-1.209***	-0.529***	-0.935***
	(-6.01)	(-4.29)	(-15.99)	(-6.52)	(-5.12)	(-17.04)
$BHM_{i,t-1}$	1.007***	0.927***	0.543***			
	(6.78)	(10.84)	(5.94)			
$SHM_{i,t-1}$				0.020	0.042	-0.049
				(0.14)	(0.50)	(-0.53)
$DTURN_{i,t-1}$	0.001	0.004**	0.002	0.001	0.005**	0.002
	(0.16)	(2.16)	(1.00)	(0.31)	(2.32)	(1.11)
$SIGMA_{i,t-1}$	-0.417*	0.201	-0.694***	-0.464*	0.142	-0.715***
	(-1.74)	(1.41)	(-3.94)	(-1.92)	(0.99)	(-4.04)
$RETURN_{i,t-1}$	0.018**	0.045***	0.047***	0.028***	0.055***	0.052***
	(2.11)	(7.78)	(7.57)	(3.33)	(9.17)	(8.21)
$ROA_{i,t-1}$	0.468***	0.365***	0.462***	0.472***	0.368***	0.470***
	(8.96)	(11.55)	(10.37)	(9.03)	(11.61)	(10.53)
$LEV_{i,t-1}$	0.008	-0.077***	-0.212***	0.006	-0.078***	-0.220***
	(0.18)	(-3.05)	(-4.98)	(0.12)	(-3.08)	(-5.18)
$MB_{i,t-1}$	0.017***	0.020***	0.039***	0.018***	0.021***	0.040***
	(6.31)	(11.83)	(15.41)	(6.72)	(12.29)	(15.62)
$SIZE_{i,t-1}$	-0.007	0.022***	0.145***	-0.005	0.023***	0.147***
	(-1.49)	(9.35)	(16.07)	(-1.04)	(9.65)	(16.31)
$ACCM_{i,t-1}$	0.277***	0.168***	0.144**	0.279***	0.169***	0.147**
	(2.83)	(2.93)	(2.25)	(2.85)	(2.95)	(2.28)
Industry FE	Yes	Yes	No	Yes	Yes	No
Firm FE	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	59,094	59,094	59,094	59,094	59,094	59,094
pseudo R^2	0.018			0.017		
adj. R^2		0.031			0.029	
Overall R^2			0.014			0.013

Panel B: Panel Regression at Quarterly Frequency

	(1)	(2)	(3)	(4)	(5)	(6)
	$CRASH_{i,q}$	$NCSKEW_{i,q}$	$NCSKEW_{i,q}$	$CRASH_{i,q}$	$NCSKEW_{i,q}$	$NCSKEW_{i,q}$
Constant	-1.920***	-0.082	-0.478***	-1.927***	-0.136	-0.535***
	(-6.72)	(-0.80)	(-4.91)	(-6.74)	(-1.33)	(-5.48)
$BHM_{i,q-1}$	0.184***	0.330***	0.340***			
	(3.30)	(10.66)	(10.56)			
$SHM_{i,q-1}$				0.064	-0.156***	-0.180***
				(1.20)	(-5.36)	(-5.99)
$DTURN_{i,q-1}$	0.012***	-0.001	0.000	0.012***	-0.001	0.000
	(3.78)	(-0.73)	(0.04)	(3.89)	(-0.54)	(0.23)
$SIGMA_{i,q-1}$	-1.059***	0.523***	0.215***	-1.092***	0.518***	0.208**
	(-7.12)	(7.73)	(2.62)	(-7.34)	(7.64)	(2.53)
$RETURN_{i,q-1}$	0.016	0.548***	0.596***	0.027	0.554***	0.602***
	(0.92)	(50.40)	(51.62)	(1.51)	(50.83)	(51.93)
$ROA_{i,t-1}$	0.265***	0.095***	0.176***	0.264***	0.095***	0.180***
	(8.35)	(7.06)	(8.85)	(8.31)	(7.09)	(9.01)
$LEV_{i,t-1}$	0.014	-0.034***	-0.101***	0.016	-0.036***	-0.105***
	(0.44)	(-2.85)	(-4.91)	(0.52)	(-3.05)	(-5.11)
$MB_{i,t-1}$	0.013***	0.005***	0.015***	0.013***	0.006***	0.015***
	(7.76)	(7.33)	(13.16)	(7.85)	(8.01)	(13.55)
$SIZE_{i,t-1}$	-0.010***	0.003***	0.084***	-0.010***	0.005***	0.085***
	(-3.45)	(2.91)	(19.52)	(-3.54)	(4.28)	(19.78)
$ACCM_{i,t-1}$	0.241***	0.106***	0.109***	0.240***	0.109***	0.111***
	(3.76)	(3.88)	(3.43)	(3.75)	(3.95)	(3.51)
Industry FE	Yes	Yes	No	Yes	Yes	No
Firm FE	No	No	Yes	No	No	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	252,665	252,665	252,665	252,665	252,665	252,665
pseudo R^2	0.023			0.023		
adj. R^2		0.034			0.034	
Overall R^2			0.023			0.023

Table 7: Mutual Fund Herding and Crash Risk: Alternative Herding Measure

This table presents the regression results of crash risk on mutual fund herding measure constructed by Frey et al (2014). $HM2_{i,t-1}$, is calculated as the average of quarterly herding measure $HM2_{i,q}$, the mutual fund herding measure constructed by Frey et al (2014), in year *t*-1. $HM2_{i,q}$, the measure of herding in stock of firm *i* during quarter *q*, is specified as

$$HM2_{i,q} = \{(p_{i,q} - E[p_{i,q}])^2 - E[(p_{i,q} - E[p_{i,q}])^2]\}\, n/(n-1).$$

Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *z(t)*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

	(1) $CRASH_{i,t}$	(2) $NCSKEW_{i,t}$	(3) $NCSKEW_{i,t}$
Constant	-1.208***	-0.530***	-0.928***
	(-6.55)	(-5.16)	(-17.33)
$HM2_{i,t-1}$	0.432**	0.580***	0.420***
	(2.14)	(4.81)	(3.02)
$DTURN_{i,t-1}$	0.001	0.005**	0.002
	(0.30)	(2.31)	(1.09)
$SIGMA_{i,t-1}$	-0.496**	0.105	-0.741***
	(-2.06)	(0.73)	(-4.21)
$RETURN_{i,t-1}$	0.027***	0.054***	0.052***
	(3.31)	(9.20)	(8.33)
$ROA_{i,t-1}$	0.474***	0.371***	0.470***
	(9.07)	(11.72)	(10.52)
$LEV_{i,t-1}$	0.007	-0.076***	-0.219***
	(0.16)	(-3.01)	(-5.15)
$MB_{i,t-1}$	0.018***	0.021***	0.040***
	(6.70)	(12.27)	(15.64)
$SIZE_{i,t-1}$	-0.005	0.023***	0.147***
	(-1.14)	(9.79)	(16.32)
$ACCM_{i,t-1}$	0.276***	0.165***	0.146**
	(2.82)	(2.88)	(2.27)
Industry FE	Yes	Yes	No
Firm FE	No	No	Yes
Year FE	Yes	Yes	Yes
N	59,094	59,094	59,094
pseudo R^2	0.017		
adj. R^2		0.029	
Overall R^2			0.013

Table 8: Mutual Fund Herding and Crash Risk: Alternative Crash Risk Measure

This table presents the regression results of an alternative crash risk measure on mutual fund herding. $DUVOL_{i,t}$ is calculated as the logarithm of the ratio of firm specific weekly return standard deviation in down weeks over firm specific weekly return standard deviation in up weeks for firm *i* in year *t* (Chen, Hong, and Stein, 2001).

$$DUVOL_{i,t} = \ln\{\frac{(\sum_{DOWN} W_{i,n}^2)/(N_d - 1)}{(\sum_{UP} W_{i,n}^2)/(N_u - 1)}\}$$

where $W_{i,n}$ is firm-specific weekly return; N_u and N_d are the number of up and down weeks, respectively. Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *t*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

	(1)	(2)	(3)	(4)	(5)	(6)
	$DUVOL_{i,t}$	$DUVOL_{i,t}$	$DUVOL_{i,t}$	$DUVOL_{i,t}$	$DUVOL_{i,t}$	$DUVOL_{i,t}$
Constant	-0.485***	-0.866***	-0.433***	-0.824***	-0.503***	-0.881***
	(-6.03)	(-19.76)	(-5.33)	(-18.44)	(-6.18)	(-19.65)
$HM_{i,t-1}$	0.333***	0.191***				
	(6.53)	(3.33)				
$BHM_{i,t-1}$			0.699***	0.433***		
			(9.83)	(5.67)		
$SHM_{i,t-1}$					-0.006	-0.067
					(-0.08)	(-0.87)
$DTURN_{i,t-1}$	0.004*	0.002	0.003	0.001	0.004*	0.002
	(1.74)	(0.87)	(1.62)	(0.80)	(1.78)	(0.90)
$SIGMA_{i,t-1}$	-0.974***	-1.137***	-0.871***	-1.089***	-0.907***	-1.102***
	(-8.56)	(-8.05)	(-7.70)	(-7.73)	(-7.92)	(-7.78)
$RETURN_{i,t-1}$	0.055***	0.051***	0.048***	0.047***	0.055***	0.051***
	(10.64)	(9.42)	(9.51)	(8.75)	(10.56)	(9.26)
$ROA_{i,t-1}$	0.343***	0.382***	0.339***	0.378***	0.341***	0.384***
	(14.04)	(11.32)	(13.92)	(11.22)	(13.95)	(11.37)
$LEV_{i,t-1}$	-0.075***	-0.217***	-0.078***	-0.213***	-0.080***	-0.220***
	(-3.60)	(-6.26)	(-3.77)	(-6.12)	(-3.83)	(-6.32)
$MB_{i,t-1}$	0.017***	0.034***	0.016***	0.033***	0.017***	0.034***
	(11.88)	(16.55)	(11.65)	(16.35)	(12.11)	(16.57)
$SIZE_{i,t-1}$	0.023***	0.118***	0.023***	0.117***	0.025***	0.119***
	(11.17)	(16.13)	(11.49)	(15.98)	(11.83)	(16.24)
$ACCM_{i,t-1}$	0.055	0.080	0.060	0.080	0.061	0.082
	(1.18)	(1.51)	(1.29)	(1.52)	(1.31)	(1.56)
Industry FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	59,094	59,094	59,094	59,094	59,094	59,094
adj. R^2	0.041		0.042		0.040	
Overall R^2		0.029		0.030		0.029

Table 9: Mutual Fund Herding and Crash Risk: The 2004 SEC Regulation Change

This table presents the difference-in-differences test results of crash measures on mutual fund herding. The SEC regulation change on mutual fund disclosure frequency from quarterly to semi-annual basis was effective since May 2004. We take the two-quarter sample right before and after the effective date of this regulation, i.e. first and second quarter-end herding values and crash measures in the subsequent quarters. $TREAT$ is an indicator which equals to 1 if a firm has at least one holding mutual fund with semi-annual reporting frequency for all four consecutive quarters through 2003. $POST$ is an indicator variable which equals to 1 if an firm-quarter observation falls in the second quarter of 2004, i.e. the first quarter after the regulation becomes effective, and 0 if an observation falls in the first quarter of 2004, i.e. the last quarter before the regulation becomes effective. Column (1) is a Probit regression. Heteroscedasticity-robust standard errors are estimated and clustered at firm level. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

	(1) $CRASH_{i,q}$	(2) $NCSKEW_{i,q}$
Constant	-1.509***	-0.079
	(-3.15)	(-0.27)
$TREAT$	-0.137	-0.115
	(-1.05)	(-1.59)
$POST$	0.151**	0.100**
	(2.50)	(2.53)
$TREAT \times POST$	0.293*	0.312***
	(1.79)	(2.76)
$DTURN_{i,q-1}$	0.033	0.013
	(1.61)	(0.97)
$SIGMA_{i,q-1}$	-2.987***	1.163*
	(-2.67)	(1.80)
$RETURN_{i,q-1}$	-0.086	0.891***
	(-0.65)	(10.81)
$ROA_{i,t-1}$	0.673***	-0.047
	(3.15)	(-0.36)
$LEV_{i,t-1}$	0.322*	0.237**
	(1.76)	(2.16)
$MB_{i,t-1}$	0.010	0.009
	(0.97)	(1.36)
$SIZE_{i,t-1}$	-0.044**	-0.028***
	(-2.52)	(-2.83)
$ACCM_{i,t-1}$	0.437	0.294
	(1.12)	(1.13)
Industry FE	Yes	Yes
N	5,478	5,478
pseudo R^2	0.030	
adj. R^2		0.042

Table 10: Mutual Fund Herding and Crash Risk: Propensity Score Matching

This table presents the regression results of crash on mutual fund herding using propensity score matching. Panel A reports the logistic regression result of the 1st stage regression. Treatment sample is defined as the firms, the herding values of which are in the top tercile for each year. A 1-on-1 matching is done with the maximum caliper distance of 3% by matching firms from the rest of the sample. $HERD\_TREAT_{i,t-1}$ is an indicator and is equal to 1 if the $HM_{i,t-1}$, the herding value of the firm *i* in year *t*-1 belongs to the top tercile, 0 otherwise. $COVERAGE_{i,t-1}$ is the log value of 1 plus the average number of analyst following for firm *i* in year *t*-1. $ET_{i,t-1}$ is the firm-specific earnings transparency measure calculated as the sum of explanatory power of earnings on returns from both industry level and portfolio level as in Barth et al. (2013). $BIGN_{i,t-1}$ is an indicator and is equal to 1 if the company uses a Big-N audit firm in year *t*-1, 0 otherwise. $SIGMA_{i,t-1}$ is the standard deviation of weekly returns for firm *i* in fiscal year *t*-1. $LITIGATION_{i-1}$ is the industry dummy which equals to 1 if SIC code of firm *i* in year *t*-1 is within 2833-2836, 3570-3577, 3600-3674, 5200-5961, and 7370-7374 and 0 otherwise (Francis et al. 1994; LaFond and Roychowdhury 2008). Panel B reports the summary statistics for treatment and matched samples, and the sample mean difference testing results are also provided. Panel C provides the 2nd stage regression results. Heteroscedasticity-robust standards errors are estimated and clustered at firm level. *z(t)*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Panel A: The First Stage Result of Logistic Regression

	$HERD\_TREAT_{i,t-1}$
$Constant$	-1.040***
	(-9.75)
$COVERAGE_{i,t-1}$	-0.133***
	(-6.18)
$ET_{i,t-1}$	-0.394***
	(-6.22)
$BIGN_{i,t-1}$	-0.261***
	(-7.03)
$SIGMA_{i,t-1}$	10.374***
	(23.65)
$LITIGATION_{i,t-1}$	0.109***
	(4.64)
$SIZE_{i,t-1}$	0.083***
	(8.88)
$MB_{i,t-1}$	0.022***
	(4.87)
$LEV_{i,t-1}$	-0.118*
	(-1.68)
$ROA_{i,t-1}$	-0.012
	(-0.13)
$ACCM_{i,t-1}$	0.762***
	(4.55)
Year FE	Yes
N	45,617
pseudo R^2	0.020

Panel B: Summary Statistics for Treatment and Matched Samples

	Herd_Treat=0					Herd_Treat=1				
Variable	**N**	**Mean**	**s.d**	**Min**	**Max**	**Mean**	**s.d**	**Min**	**Max**	**Mean Dif.**
$HM_{i,t-1}$	14,880	-0.01	0.04	-0.16	0.04	0.06	0.05	0.01	0.80	0.07***
$COVERAGE_{i,t-1}$	14,880	1.56	0.65	0.00	3.72	1.57	0.66	0.69	3.80	0.01
$ET_{i,t-1}$	14,880	0.47	0.19	0.02	1.42	0.47	0.19	0.02	1.42	-0.00
$BIGN_{i,t-1}$	14,880	0.90	0.30	0.00	1.00	0.90	0.30	0.00	1.00	0.00
$SIGMA_{i,t-1}$	14,880	0.07	0.03	0.01	0.38	0.07	0.03	0.01	0.37	0.00
$LITIGATION_{i,t-1}$	14,880	0.36	0.48	0.00	1.00	0.36	0.48	0.00	1.00	-0.00
$SIZE_{i,t-1}$	14,880	6.35	1.76	2.34	12.84	6.37	1.81	2.36	12.79	0.02
$MB_{i,t-1}$	14,880	2.89	2.57	0.08	20.00	2.88	2.47	0.08	19.99	-0.01
$LEV_{i,t-1}$	14,880	0.17	0.17	0.00	0.86	0.17	0.17	0.00	0.95	-0.00
$ROA_{i,t-1}$	14,880	0.04	0.13	-0.94	0.57	0.04	0.13	-0.96	0.57	-0.00
$ACCM_{i,t-1}$	14,880	0.06	0.07	0.00	0.50	0.06	0.07	0.00	0.50	-0.00

Panel C: The 2nd Stage Regression Results

	(1) $CRASH_{i,t}$	(2) $NCSKEW_{i,t}$	(3) $DUVOL_{i,t}$
Constant	-1.064***	-0.187*	-0.231**
	(-5.51)	(-1.90)	(-2.36)
$HM_{i,t-1}$	0.350**	0.352***	0.238***
	(2.28)	(3.98)	(3.19)
$DTURN_{i,t-1}$	0.016**	0.012***	0.009***
	(2.06)	(3.21)	(3.00)
$SIGMA_{i,t-1}$	-0.272	0.408**	-0.851***
	(-0.72)	(1.98)	(-4.86)
$RETURN_{i,t-1}$	0.035***	0.081***	0.078***
	(2.69)	(10.29)	(11.47)
$ROA_{i,t-1}$	0.535***	0.384***	0.329***
	(6.86)	(8.59)	(9.14)
$LEV_{i,t-1}$	-0.031	-0.105***	-0.085***
	(-0.50)	(-3.03)	(-3.00)
$MB_{i,t-1}$	0.017***	0.017***	0.014***
	(4.50)	(7.61)	(7.42)
$SIZE_{i,t-1}$	-0.008	0.014***	0.015***
	(-1.29)	(4.09)	(5.33)
$ACCM_{i,t-1}$	0.259*	0.247***	0.136**
	(1.91)	(3.20)	(2.18)
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	29,760	29,760	29,760
pseudo R^2	0.018		
adj. R^2		0.031	0.038

Table 11: Mutual Fund Herding and Crash Risk: Controlling for Price Impact

This table provides the regression results by further controlling for mutual fund inflow-induced price impact and future stock returns. $MFFLOWIN_{i,q-1}$ is the price impact caused by mutual fund inflow for firm *i* in quarter *q*-1 by Edmans, Goldstein, and Jiang (2012). $RETURN_{i,(q,q+1)}$ is the cumulative buy-and hold returns for firm *i* from quarter *q* to quarter *q*+1. $RETURN_{i,q}$ is the cumulative buy-and hold returns for firm *i* over quarter *q*. Column (1) to column (3) are Probit regressions. Heteroscedasticity-robust standard errors are estimated and clustered at firm level in regressions with industry and year fixed effects. *z(t)*-values are reported in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

	(1) $CRASH_{i,q}$	(2) $CRASH_{i,q}$	(3) $CRASH_{i,q}$	(4) $NCSKEW_{i,q}$	(5) $NCSKEW_{i,q}$	(6) $NCSKEW_{i,q}$	(7) $NCSKEW_{i,q}$	(8) $NCSKEW_{i,q}$	(9) $NCSKEW_{i,q}$	(10) $NCSKEW_{i,q}$	(11) $NCSKEW_{i,q}$	(12) $NCSKEW_{i,q}$
Constant	-2.006**	-2.004**	-2.015**	-0.539***	-0.515***	-0.578***	0.527***	0.285**	0.554***	0.314***	0.465***	0.216*
	(-6.65)	(-6.65)	(-6.68)	(-5.50)	(-5.25)	(-5.88)	(6.14)	(2.56)	(6.45)	(2.84)	(5.39)	(1.93)
$HM_{i,q-1}$	0.171***			0.128***			0.102***	0.085***				
	(3.13)			(4.05)			(4.24)	(3.22)				
$BHM_{i,q-1}$		0.174***			0.376***				0.464***	0.502***		
		(2.85)			(10.67)				(17.02)	(15.52)		
$SHM_{i,q-1}$			0.036			-0.199***					-0.331***	-0.372***
			(0.60)			(-5.81)					(-12.23)	(-12.10)
$MFFLOWIN_{i,q-1}$	-0.013	-0.013	-0.013	-0.003	-0.001	-0.002						
	(-1.17)	(-1.13)	(-1.18)	(-0.44)	(-0.24)	(-0.34)						
$RETURN_{i,q}$							-2.436***		-2.434***		-2.435***	
							(-41.77)		(-41.74)		(-41.75)	
$RETURN_{i,(q,q+1)}$								-0.804***		-0.802***		-0.803***
								(-9.91)		(-9.91)		(-9.92)
$RETURN_{i,q-1}$	0.031*	0.025	0.033*	0.612***	0.599***	0.605***						
	(1.69)	(1.34)	(1.80)	(49.95)	(48.52)	(48.74)						
$DTURN_{i,q-1}$	0.017***	0.017***	0.017***	0.002	0.002	0.002	0.005***	0.007***	0.004**	0.006***	0.005***	0.007***
	(3.82)	(3.80)	(3.90)	(0.98)	(0.81)	(1.01)	(2.96)	(3.78)	(2.52)	(3.37)	(2.79)	(3.60)
$SIGMA_{i,q-1}$	-1.124**	-1.081**	-1.108**	0.193**	0.250***	0.241***	1.615***	1.434***	1.657***	1.476***	1.669***	1.492***
	(-7.15)	(-6.88)	(-7.05)	(2.22)	(2.88)	(2.77)	(10.11)	(15.90)	(10.38)	(16.40)	(10.45)	(16.53)
$ROA_{i,t-1}$	0.286***	0.285***	0.285***	0.179***	0.180***	0.183***	0.026	0.049**	0.026	0.048**	0.031	0.054**
	(8.55)	(8.52)	(8.49)	(8.48)	(8.50)	(8.65)	(1.10)	(2.20)	(1.10)	(2.19)	(1.32)	(2.45)
$LEV_{i,t-1}$	-0.024	-0.026	-0.025	-0.124***	-0.124***	-0.129***	0.116***	0.058**	0.116***	0.058**	0.109***	0.050*
	(-0.72)	(-0.79)	(-0.75)	(-5.73)	(-5.72)	(-5.93)	(4.69)	(2.09)	(4.70)	(2.09)	(4.42)	(1.81)
$MB_{i,t-1}$	0.013***	0.013***	0.013***	0.016***	0.016***	0.016***	-0.014***	-0.006**	-0.015***	-0.006***	-0.014***	-0.006**
	(7.31)	(7.33)	(7.41)	(13.33)	(13.07)	(13.48)	(-9.78)	(-2.52)	(-10.07)	(-2.71)	(-9.61)	(-2.41)
$SIZE_{i,t-1}$	-0.002	-0.002	-0.002	0.085***	0.086***	0.087***	-0.113***	-0.062***	-0.111***	-0.060***	-0.109***	-0.057***
	(-0.74)	(-0.58)	(-0.59)	(18.53)	(18.65)	(18.94)	(-18.66)	(-5.22)	(-18.42)	(-5.06)	(-17.99)	(-4.82)
$ACCM_{i,t-1}$	0.297***	0.301***	0.301***	0.122***	0.125***	0.128***	0.035	0.038	0.039	0.043	0.044	0.048
	(4.43)	(4.49)	(4.50)	(3.59)	(3.67)	(3.75)	(0.96)	(1.14)	(1.08)	(1.27)	(1.20)	(1.41)
Industry FE	Yes	Yes	Yes	No	No	No	No	No	No	No	No	No
Firm FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	223,641	223,641	223,641	223,641	223,641	223,641	252,665	251,986	252,665	251,986	252,665	251,986
pseudo R^2	0.022	0.022	0.022									
Overall R^2				0.025	0.025	0.025	0.260	0.087	0.262	0.088	0.261	0.087

Table 12: Triple Sorted Portfolios by Herding, Size, and Corporate Disclosure Quality

This table reports the 10×2×2 triple sorting results by herding, size, and various corporate disclosure quality measures. $HM_{i,t}$ is the unsigned measure of mutual fund herding intensity calculated as the average of quarterly herding measure $HM_{i,q}$ in year *t*, where $HM_{i,q} = |p_{i,q} - E[p_{i,q}]| - E|p_{i,q} - E[p_{i,q}]|$. $IDIOSYN_{i,t}$ is the idiosyncratic volatility for firm *i* in year *t*. $ET_{i,t}$ is the firm-specific earnings transparency measure calculated as the sum of explanatory power of earnings on returns from both industry level and portfolio level, as in Barth et al. (2013). $F-SCORE_{i,t}$ is the negative value of predicted likelihood of misstatement proposed by Dechow et al. (2011). $C-SCORE_{i,t}$ by Khan and Watts (2009) is a composite measure of accounting conservatism for firm *i* in year *t*. Small and Large size firms are divided by the median of firm size, which is measured by the average of total asset of firm *i* in year *t*. Low and high disclosure quality subgroups are divided by the median of each of the four disclosure quality measures by the end of year *t-1*, respectively. We report the average future crash likelihood $CRASH_{i,t}$ for each of the 40 portfolios. We test the high-minus-low differences of $CRASH_{i,t}$ when herding increases within different subsamples along the dimensions of size and disclosure quality. We also report the difference-in-differences one-sided *t*-test with respect to whether high-minus-low is larger in low disclosure quality subgroup than that in high disclosure subgroup. *, ** and *** denote statistical significance at the 10%, 5% and 1% levels.

Panel A: Triple Sorted Portfolios by Herding, Size, and Idiosyncratic Volatility

	Small Size			Large Size		
	Low *IDIO*	High *IDIO*	DID	Low *IDIO*	High *IDIO*	DID
Low HM	0.127	0.127		0.114	0.111	
2	0.141	0.161		0.154	0.148	
3	0.205	0.192		0.180	0.166	
4	0.197	0.184		0.185	0.174	
5	0.193	0.166		0.205	0.193	
6	0.200	0.195		0.192	0.201	
7	0.220	0.171		0.201	0.192	
8	0.201	0.210		0.183	0.213	
9	0.187	0.188		0.165	0.185	
High HM	0.201	0.176		0.161	0.181	
High minus Low	0.075***	0.048***	-0.026*	0.047***	0.070***	0.024

Panel B: Triple Sorted Portfolios by Herding, Size, and Earnings Transparency

	Small Size			Large Size		
	Low *ET*	High *ET*	DID	Low *ET*	High *ET*	DID
Low HM	0.132	0.121		0.114	0.110	
2	0.154	0.149		0.161	0.144	
3	0.192	0.206		0.175	0.173	
4	0.208	0.180		0.187	0.182	
5	0.183	0.186		0.204	0.199	
6	0.207	0.198		0.198	0.188	
7	0.196	0.183		0.200	0.200	
8	0.215	0.189		0.194	0.191	

9	0.190	0.181		0.191	0.167	
High HM	0.176	0.162		0.185	0.159	
High minus Low	0.043***	0.041***	-0.002	0.070***	0.049***	-0.021

Panel C: Triple Sorted Portfolios by Herding, Size, and *F-SCORE*

	Small Size			Large Size		
	Low *F-SCORE*	High *F-SCORE*	DID	Low *F-SCORE*	High *F-SCORE*	DID
Low HM	0.129	0.124		0.126	0.106	
2	0.156	0.156		0.192	0.121	
3	0.194	0.194		0.203	0.156	
4	0.211	0.179		0.207	0.154	
5	0.211	0.156		0.237	0.168	
6	0.211	0.194		0.216	0.169	
7	0.220	0.156		0.222	0.181	
8	0.220	0.185		0.214	0.178	
9	0.198	0.169		0.205	0.153	
High HM	0.195	0.162		0.202	0.152	
High minus Low	0.066***	0.039***	-0.028*	0.076***	0.045***	−0.031

Panel D: Triple Sorted Portfolios by Herding, Size, and *C-SCORE*

	Small Size			Large Size		
	Low *C-SCORE*	High *C-SCORE*	DID	Low *C-SCORE*	High *C-SCORE*	DID
Low HM	0.125	0.125		0.121	0.097	
2	0.167	0.146		0.150	0.154	
3	0.203	0.194		0.168	0.192	
4	0.208	0.187		0.188	0.182	
5	0.207	0.178		0.195	0.217	
6	0.242	0.188		0.185	0.223	
7	0.182	0.191		0.197	0.213	
8	0.218	0.196		0.196	0.193	
9	0.199	0.180		0.174	0.196	
High HM	0.204	0.157		0.161	0.180	
High minus Low	0.079***	0.032***	-0.048**	0.040***	0.083***	0.043