参赛队员姓名：Clement Man Yiu Liu

中学：深圳国际交流学院

省份：广东省

国家/地区：中国，南方赛区

指导教师姓名：陈钊

指导教师单位：复旦大学

论文题目：Herding Behavior around "Smart Money"

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Herding Behavior around "Smart Money"

Clement Man Yiu Liu

Shenzhen College of International Education

September 14, 2022

Abstract

We study the consequences of boundedly rational investors chasing conspicuous signals in the information explosion era. In this paper, we show theoretically that investors would follow the exogenous smart money flow instead of conducting self-analysis, leading to investors' herding behavior, and the positive sentiment could exacerbate such effect. Comfortingly, the Shanghai-Hong Kong Stock Connect and Shenzhen-Hong Kong Stock Connect provide a cornerstone for our empirical research. The capital flow from Hong Kong to the mainland (Northbound Capital Flow, hereafter NCF) through this channel is widely praised as smart money. In this context, we find that the NCF would induce significant stock-level herding, which is more pronounced among retail investors, and the findings are consistent in several robustness tests. Further, investors' panic suppresses the herding behavior around smart money, while investors' mania stimulates investor herding. In addition, investors' perception of smart money flow also impacts the herding behavior, including shareholding ratio of northbound capital, market capitalization and smart money type.

Keywords: Smart Money; Herding; Evolutionary Game Model; Investor Sentiment

Classification: G11; G20; C73

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

"*But one of the remaining ten thousand hunters will surely make a choice to fire on that position, because for civilizations at a certain level of technological development, attacking may be safer and less of a hassle than probing.*"

***The Dark Forest* by Cixin Liu, 2015**

1 Introduction

Boundedly rational investors are unable to cope with the increasingly redundant information in the market (Elliott et al., 2015), so they are trying to find simplified signals to guide their investment. Like hunters, investors are committed to finding clear and obvious targets, such as mutual fund performance or ratings (Kaniel and Parham, 2017; Ben-David et al., 2022), alphabetical order (Doellman et al., 2019), and naïve diversification rules (Thaler et al., 2001). Although these studies have confirmed the fact that investors chase signals, very little is known about whether and to what extent these simple signals affect investors' herding behavior. We address this question by examining the influence of smart money flow on stock-level investors' herding theoretically and empirically. Our focus on smart money is motivated by investors' consensus on its importance (Akbas et al., 2015), while other signals mentioned above may only be accepted or noticed by some investors.

Herding behavior is defined as the convergence of investors' trading behavior. Prior theoretical studies differ in their explanation of what might trigger herd behavior (Demirer et al., 2010; Jegadeesh and Kim, 2010). The first set of theories attributes herding to investor psychology, including investors' sentiment and sense of security in following the crowd (Devenow and Welch, 1996). The second approach focuses on information-driven herding, that is, investors may not have access to valuable information and will follow others' decisions with private information (Galariotis et al., 2015; Celiker et al., 2015). The third branch of theories mainly focuses on the institutional herding caused by the principal-agent relationship (Demirer et al., 2010). In this paper, under the background of increasingly redundant information, we lay our research on the impact of conspicuous signals in the market, specifically smart money flow, on investors' herding.

In the popular press and academia, dumb money and smart money are opposite concepts. Lou (2012) argued that price pressure from dumb money is generally presupposed

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

to make prices depart from fundamentals, while Frazzini and Lamont (2008) found that arbitrage by smart money makes prices converge to fundamental values. In practice, researchers are more accustomed to delineating smart money as specific types of funds, for instance, experienced venture capital (Sørensen, 2007), hedge funds (Akbas et al., 2015), experienced stock issuers (Gibson et al., 2004), and even institutional investors. But different from previous studies, our paper does not focus on which funds or funds with certain characteristics could be called smart money. Instead, we only pay attention to whether the well-accepted smart money would stimulate herding behavior.

First, we build an evolutionary game-theoretic model to study heterogenous investors' herding in response to the exogenous trading behavior of smart money. The evolutionary game model (hereinafter referred to as EGM) is very suitable for this study. On the one hand, EGM incorporates players with bounded rationality in an environment of incomplete information, which is more practical than the classic game model. On the other hand, EGM could demonstrate the dynamics of strategy change (Wang et al., 2021).

Under the framework of the evolutionary game, investors are categorized into institutional investors and retail investors, who can observe exogenous smart money trading behavior directly. In our model, institutional and retail investors can choose *Follow* or *Self-Analysis* strategies, with different levels of costs and profitability, which will induce the stock return movement. Our model predicts that when the difference between the profitability of the two strategies exceeds the critical value, both types of investors will choose the *Follow* strategy, resulting in herding behavior. And the critical value is related to the difference between costs and stock return. Even if the complete condition of herding cannot be satisfied, our model also provides an analytical framework for the herding possibility. Ulteriorly, investors' sentiment plays a role in their mental accounts, which is directly reflected in their perception of costs. Our model also depicts the impact of sentiment on the evolutionarily stable state, that is, positive sentiment will stimulate people to herd around smart money, and vice versa. Finally, we also propose four typical cases and simulate them respectively, to illustrate herding behavior in the model and the heterogenous impact of sentiment on herding.

Second, the stock connect mechanism between mainland China's stock market and Hong Kong's stock market provides a unique environment for our empirical research, i.e.,

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Shanghai–Hong Kong Stock Connect and Shenzhen–Hong Kong Stock Connect (hereinafter referred to as Stock Connect Program). The Stock Connect Program allows investors in mainland China and Hong Kong to trade designated stocks listed on the other market through their local exchanges (Zhao et al., 2021). The capital flowing from Hong Kong to mainland China is known as Northbound Capital Flow (hereafter NCF). Lu et al. (2021) summarized that institutional investors contribute more than 90% of NCF, and NCF is even recognized as a key indicator that can predict the rise and fall of the A-share market. Huang et al. (2016) argued that these investors are considered more mature and professional, relying more on financial statements in their investment decisions. Thus, NCF invested by these investors is also known as "Smart Money" in academic papers and media, for instance Lu et al. (2021).[1]

Thus, we empirically examine the relationship between investors' herding behavior and smart money flow using a sample consisting of 2,330 Shanghai and Shenzhen A-share listed firms from Q1 2017 to Q1 2022. Following Lakonishok et al. (1992), Jiao and Ye (2014), Brown et al. (2014), and Cai et al. (2019), we construct the proxy variables of the stock-level herding behavior, institutional herding behavior, retail herding behavior, and buy/sell herding behavior. Our core independent variable is the ratio of stock-level NCF divided by the stock market value. In our main results, we find that herding behavior is positively related to the NCF, and retail investors respond more strongly than institutional investors. These findings are consistent with the theoretical model, suggesting that two types of investors all choose the *Follow* strategy, and are also robust with a battery of tests. To empirically verify our proposition regarding the impact of investors' heterogeneous sentiment, we incorporate two sets of empirical analyses, i.e., panic-sentiment and mania-sentiment tests. Our results indicate that investors' panic would suppress the herding behavior, especially that of retail investors. In contrast, the investors' mania would stimulate the herding behavior, all in line with the proposed theoretical model.

In additional analyses, we find that investors' perception of smart money also plays a key role in herding behavior formation. The empirical results first show that investors

[1] The media also widely use "Smart Money" to refer to NCF. For example, in the South China Morning Post on May 14, 2019, an article entitled *Smart money pulls out of China's equity market via Hong Kong as hopes of trade war's resolution fade to glory* wrote that "Smart money is accelerating its pullout from China's equities after riding Asia's largest stock market to its crest in the first three months of 2019, . . . , Capital pouring into the Shanghai and Shenzhen stock markets via Hong Kong, . . . ".

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

are not sensitive to stocks where northbound capital has held more positions or stocks with larger market value. Moreover, as disclosed by the Hong Kong Exchanges (HKEX), most brokers of the Stock Connect Program are securities companies or banks. The former prefers short-term trading, while the latter prefers long-term asset allocation. According to this difference, research institutions name NCF triggered by securities companies and banks as Trading-oriented NCF and Allocation-oriented NCF, respectively. We can find that the Trading-oriented NCF inspires herding, but not the Allocation-oriented NCF. This also shows that A-share investors seem to track NCF only from the appearance, but are not sensitive to its internal structure, which is consistent with the assumption that investors simply track conspicuous signals.

Our work yields significant contributions to the literature in the following ways. First, we contribute to the literature on the consequences of boundedly rational investors in the information explosion era. Bounded rationality means that individual investors possess limited attention and processing power (Daniel et al., 2002; Hirshleifer et al., 2009). A large body of literature in behavioral economics has revealed that boundedly rational agents tend to underreact to less vivid and salient sources of information (Peng, 2005; Peng and Xiong, 2006). Further research has verified that bounded rationality would affect market price efficiency and may generate mispricing (Dietrich et al., 2001; Giglio and Shue, 2014). During the information explosion era, information overload would exacerbate inefficiency caused by bounded rationality (Cardinaels et al., 2019). We contribute to this literature by exploring the herding consequence triggered by the boundedly rational investors, who chase the simplified and conspicuous signals in the market.

Second, our theoretical model brings a unique insight to the literature on the nature of herding behavior. Theorists have long been interested in the mechanism of herding behavior (Cai et al., 2019). Our theoretical model provides evidence for investors' herding behavior in the face of exogenous deterministic return information. Meanwhile, under the unified framework, our paper also proves the heterogeneous response of institutional and retail investors due to their endowments. Different from taking analysts and other subjects as the source of herding benefits (Brown et al., 2014), we provide a new perspective on the source of the herding effect. Correspondingly, we extend the literature on the smart money effect. Prior literature highlights that some specific investors can be called smart

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

money, such as experienced venture capital (Sørensen, 2007), hedge funds (Akbas et al., 2015), experienced stock issuers (Gibson et al., 2004), and smart money can be nurtured by education (Cole et al., 2014) and reduce market volatility (Hervé et al., 2019). We instead focus on the signaling effect of widely recognized smart money for investors.

Third, our empirical study is rooted in the unique environment brought by emerging capital markets. As stated initially, investors are committed to finding conspicuous targets. However, the simplified signals mentioned in relevant studies are not direct and obvious enough to potentially significantly induce investors' herding behavior, like the mutual fund rating (Kaniel and Parham, 2017) and the alphabetical order (Doellman et al., 2019). In China, the Stock Connect Program is a unique institutional arrangement. The consensus that NCF is widely recognized as smart money, provides a natural soil for us to identify the herding effect brought by smart money empirically. Based on this, we can also make up for the gap in the literature on whether smart money causes the herding behavior of investors.

For this program itself, researchers have been paying more attention to its impact on bilateral market change, such as volatility linkages (Lin, 2017), market quality (Xu et al., 2020), corporate financialization (Ying et al., 2021), investment efficiency (Peng et al., 2021) and asset pricing (Liu et al., 2021). Zhao et al. (2021) incorporated the difference-in-difference (DID) method to verify the negative relationship between stock market liberalization and institutional herding using the Stock Connect Program as an exogenous shock, indicating that a more liberal market will reduce the potential herding behavior of institutional investors. Our paper differentiates from Zhao et al. (2021) quite a lot in many aspects. On the one hand, we focus on the long-term effects of NCF after launching the Stock Connect Program, rather than the instantaneous impact. Accordingly, our results are just the opposite of Zhao et al. (2021), which is determined by investors' long-term understanding of NCF. On the other hand, the key point of this paper is that NCF is widely recognized as smart money, leading to investors' herding behavior, and has nothing to do with market liberalization. Thus, our research is a useful and indispensable supplement to Zhao et al. (2021).

Fourth, we also supplement the evidence of the impact of investor sentiment and perception on the degree of herding. The existing literature proves that the herding of analysts

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

and other subjects can be affected by sentiment. We enrich the conclusions of this literature from the two sentiment dimensions, i.e., panic and mania. More importantly, we verify how the perception of information affects the degree of herding. The existing literature mainly focuses on the impact of risk perception on herding (Kizys et al., 2021), but our results explore a deeper level of perception, that is, the perception of the herding source itself. The results show that this kind of perception significantly impacts the degree of herding, and there is a difference between retail and institutional investors.

The remainder of the paper proceeds as follows. In Section 2, we review the related literature and correspondingly propose our hypothesis based on the theoretical model. In Section 3, we introduce our sample construction and empirical methodology. In Section 4, we report our main empirical results, sentiment analysis and robustness tests. In Section 5, we incorporate the results of further analysis. Finally, we conclude this paper in Section 6.

2 The theoretical model

In this section, we propose the background of the theoretical model, including the model setup, the evolutionary game analysis, and the simulation results.

2.1 Model setup

We consider three types of participant agents: "smart money" with private information sets, institutional investors, and retail investors. For the three market participants, we have the following assumptions:

First, we assume that smart money will take the initiative after receiving private information. Suppose that institutional investors and retail investors can take two strategies {*Follow*, *Self-Analysis*} after seeing the actions of smart money.

Second, institutional and retail investors have different levels of capabilities and resources, so heterogeneity exists between the two participants. It is assumed that the costs of institutional and retail investors tracking smart money are C_1 and C_2 respectively, and the costs of choosing independent analysis instead of tracking are C_1' and C_2'. Naturally, we can have the following parameters' relationship, i.e., $C_2' > C_1' > C_2 > C_1$.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Third, suppose that the intrinsic stock value contained in the private information owned by smart money is P, but the behavior of smart money alone cannot promote the price to reach the intrinsic value level. The tracking or independent analysis behavior of different participants in the market has a significant impact on the securities price so that the stock can achieve the corresponding return. For simplification, if institutional and retail investors follow the behavior of smart money, the stock return will reach $\theta_1 r$ and $\theta_2 r$ accordingly and respectively. Instead, if the two types of investors choose self-analysis, the stock returns will reach $\theta_1' r$ and $\theta_2' r$. Since smart money has information advantages, we can get $\theta_2 > \theta_2'$ and $\theta_1 > \theta_1'$.

Fourth, although the stock price and return will change due to the market participants' behavior, investors are not able to fully benefit from the fluctuations due to their investment decision-making ability, which is strongly related to their risk exposure and investment timing. Thus, we introduce the investors' profit probability parameter. When investors choose to track smart money, the profit probabilities of institutional and retail investors are α_1 and α_2 respectively, and vice versa are α_1' and α_2' respectively. Meanwhile, since the profit probability of investors is correspondingly higher with independent analysis, we set $\alpha_1' > \alpha_1$ and $\alpha_2' > \alpha_2$.

According to the above hypothesis, after seeing the exogenous smart money behavior, in the game system of institutional investors and retail investors, there are four circumstances that occur, that is $\{$*Follow*, *Self-Analysis*$\}\times\{$*Follow*, *Self-Analysis*$\}$. Consequently, the payment matrix is shown in Table 1.

[PLEASE INSERT TABLE 1 HERE]

2.2 The evolutionary game

Assume that the ratio of dominant "Institutional Investor" adopting the "*Follow*" mixed strategy is $X(0 \le X \le 1)$ and the ratio of adopting the "*Self-Analysis*" mixed strategy is $1-X$. The ratio of "Retail Investor" adopting the "*Follow*" mixed strategy is $Y(0 \le Y \le 1)$, and the ratio of adopting the "*Self-Analysis*" mixed strategy is $1 - Y$.

Under the premise that institutional investors choose the "*Follow*" strategy when retail investors adopt the strategy of "*Follow*" and "*Self-Analysis*", the sum of institutional

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

investors' expected income value is as follows:

$$\pi_I^F = Y\alpha_1 r\left(\theta_2 - \theta_2'\right) + \alpha_1 r\left(\theta_2' + \theta_1\right) - C_1 \tag{1}$$

Under the premise that institutional investors choose the "*Self-Analysis*" strategy when retail investors adopt the strategy of "*Follow*" and "*Self-Analysis*", the sum of institutional investors' expected income value is as follows:

$$\pi_I^S = Y\alpha_1' r\left(\theta_2 - \theta_2'\right) + \alpha_1' r\left(\theta_2' + \theta_1'\right) - C_1' \tag{2}$$

Thus, we can get the institutional investors' expected payoff as follows:

$$\pi_I = X\pi_I^F + (1-X)\pi_I^S \tag{3}$$

According to the basic principles of the evolutionary game, from formulas (1) and (3), institutional investors choose the "*Follow*" strategy's duplicate replication dynamic equation as follows:

$$\frac{dX}{dt} = X\left(\pi_I^F - \pi_I\right) = X(1-X)\left(\pi_I^F - \pi_I^S\right) \tag{4}$$

$$\pi_I^F - \pi_I^S = \mathrm{Yr}\left(\theta_2 - \theta_2'\right)\left(\alpha_1 - \alpha_1'\right) + \left(\alpha_1 - \alpha_1'\right) r\theta_2' + r\left(\alpha_1\theta_1 - \alpha_1'\theta_1'\right) + \left(C_1' - C_1\right) \tag{5}$$

By repeating the above analysis steps for retail investors, we can obtain the following equation:

$$\pi_R^F = X\alpha_2 r\left(\theta_1 - \theta_1'\right) + \alpha_2 r\left(\theta_1' + \theta_2\right) - C_2 \tag{6}$$

$$\pi_R^S = X\alpha_2' r\left(\theta_1 - \theta_1'\right) + \alpha_2' r\left(\theta_2' + \theta_1'\right) - C_2' \tag{7}$$

$$\pi_R = Y\pi_R^F + (1-Y)\pi_R^S \tag{8}$$

$$\frac{dY}{dt} = Y\left(\pi_R^F - \pi_R\right) = Y(1-Y)\left(\pi_R^F - \pi_R^S\right) \tag{9}$$

$$\pi_R^F - \pi_R^S = Xr\left(\theta_1 - \theta_1'\right)\left(\alpha_2 - \alpha_2'\right) + \left(\alpha_2 - \alpha_2'\right) r\theta_1' + r\left(\alpha_2\theta_2 - \alpha_2'\theta_2'\right) + \left(C_2' - C_2\right) \tag{10}$$

When setting $dX/dt = 0$ and $dY/dt = 0$, there are five local equilibrium points on the plane $N = \{(X,Y); 0 \le X, Y \le 1\}$, i.e., $(0,0), (0,1), (1,0), (1,1), (X^*, Y^*)$, where

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

$$X^* = \frac{(\alpha_2 - \alpha_2')\, r\theta_1' + r\,(\alpha_2\theta_2 - \alpha_2'\theta_2') + (C_2' - C_2)}{r\,(\theta_1 - \theta_1')\,(\alpha_2' - \alpha_2)} \tag{11}$$

$$Y^* = \frac{(\alpha_1 - \alpha_1')\, r\theta_2' + r\,(\alpha_1\theta_1 - \alpha_1'\theta_1') + (C_1' - C_1)}{r\,(\theta_2 - \theta_2')\,(\alpha_1' - \alpha_1)} \tag{12}$$

However, only the $(0,0)$ and $(1,1)$ are able to be the Evolutionarily Stable Strategy (ESS). Denote $\Delta_1 = \alpha_1' - \alpha_1 > 0$, $\Delta_2 = \alpha_2' - \alpha_2 > 0$, and we can easily show that $dX^*/d\Delta_2 < 0$, $dY^*/d\Delta_1 < 0$. Thus, we can get our **Proposition 1**:

Proposition 1: The strategy $(1,1)$ is the unique Evolutionarily Stable Strategy (ESS) when $\Delta_1 > \Delta_1^*$ and $\Delta_2 > \Delta_2^*$, that is, both institutional and retail investors will always choose to herd around smart money, where

$$\Delta_1^* = \frac{(C_1' - C_1) + r\alpha_1\,(\theta_1 - \theta_1')}{r\,(\theta_1' + \theta_2')} \tag{13}$$

$$\Delta_2^* = \frac{(C_2' - C_2) + r\alpha_2\,(\theta_2 - \theta_2')}{r\,(\theta_1' + \theta_2')} \tag{14}$$

To elaborate it graphically, we depict the vector fields for the five possible stable equilibria. As demonstrated above, the local equilibrium points $(0,1)$ and $(1,0)$ are unstable, (X^*, Y^*) is a saddle point, and only $(0,0)$ or $(1,1)$ is the stable point. From Figure 1, we can find the dynamic evolution process of all initial states. When the initial probability set (X_0, Y_0) are in the two regions denoted $\boldsymbol{A}$ and $\boldsymbol{B}$, the final stable state is $(1,1)$. On the contrary, if the initial probability set (X_0, Y_0) are in the two regions denoted $\boldsymbol{C}$ and $\boldsymbol{D}$, the final stable state is $(0,0)$. In a sense, the area of each region in the figure determines the final evolutionarily stable strategy.

[PLEASE INSERT FIGURE 1 HERE]

Thus, if $\Delta_1 \to \Delta_1^{*-}$ and $\Delta_2 \to \Delta_2^{*-}$, although X^* and Y^* are positive, they will approach 0. Correspondingly, the areas of $\boldsymbol{C}$ and $\boldsymbol{D}$ will be very trivial. Based on this, we propose the first corollary of the paper:

Corollary 1: In practice, strategy $(1,1)$ is more likely to be the equilibrium than strategy $(0,0)$, when $\Delta_1 \to \Delta_1^{*-}$ and $\Delta_2 \to \Delta_2^{*-}$.

Under the same logic, we can further analyze different types of investors and get the second corollary of this paper, that is,

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Corollary 2: Retail investors are more likely to follow the smart money when $X^* < Y^*$.

In deeper research to consider the irrational herding behavior of investors, we naturally need to consider the impact of investors' mental accounting. The previous literature on mental accounting argues that people group their financial resources and expenditures into "mental accounts" and make decisions within the context of those narrowly defined accounts instead of integrating all decisions together in a single optimization problem (Grinblatt and Han, 2005; Milkman and Beshears, 2009). Any result of an investment alternative affects the evaluation of the result after being registered to a mental account (Thaler, 2008), especially including investors' sentiment.

To clarify the impact of investor sentiment on herding behavior around smart money, we can incorporate an additional parameter τ to modify the costs of investors' decisions. Investors' sentiment is divided into positive type and negative type, which will reduce or increase the costs of investors' mental accounts respectively, i.e., $0 < \tau < 1$ or $1 < \tau$. Then, the payment matrix will be fine-tuned as shown in Table 2.

[PLEASE INSERT TABLE 2 HERE]

Following the verification process of Eq.(1) to Eq.(10), we can also verify that there are five local equilibrium points, i.e., $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, (X^{**}, Y^{**}), where

$$X^{**} = \frac{(\alpha_2 - \alpha_2')\, r\theta_1' + r\,(\alpha_2\theta_2 - \alpha_2'\theta_2') + \tau\,(C_2' - C_2)}{r\,(\theta_1 - \theta_1')\,(\alpha_2' - \alpha_2)} \tag{15}$$

$$Y^{**} = \frac{(\alpha_1 - \alpha_1')\, r\theta_2' + r\,(\alpha_1\theta_1 - \alpha_1'\theta_1') + \tau\,(C_1' - C_1)}{r\,(\theta_2 - \theta_2')\,(\alpha_1' - \alpha_1)} \tag{16}$$

Consequently, the strategy $(1,1)$ is the Evolutionarily Stable Strategy (ESS) when $\Delta_1 > \Delta_1^{**}$ and $\Delta_2 > \Delta_2^{**}$,

$$\Delta_1^{**} = \frac{\tau\,(C_1' - C_1) + r\alpha_1\,(\theta_1 - \theta_1')}{r\,(\theta_1' + \theta_2')} \tag{17}$$

$$\Delta_2^{**} = \frac{\tau\,(C_2' - C_2) + r\alpha_2\,(\theta_2 - \theta_2')}{r\,(\theta_1' + \theta_2')} \tag{18}$$

According to the above formulas, we can tell that $dX^{**}/d\tau > 0$ and $dY^{**}/d\tau > 0$. Thus, our second proposition related to the impact of investor sentiment would fall into place rapidly.

Proposition 2: With more positive investor sentiment, the costs of mental accounts

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

would decrease, increasing the probability that the strategy $(1,1)$ is an evolutionarily stable strategy. In other words, both institutional and retail investors will be more likely to herd around smart money.

2.3 The simulation results

In this section, we utilize the MATLAB system to simulate the dynamic evolution, and then the evolution process of the investors' strategies would be observed more intuitively. Without loss of generality, we introduce four typical cases to elaborate our propositions and corollaries in the above section, and the parameter setting is shown in Table 3.

[PLEASE INSERT TABLE 3 HERE]

In case 1, the generalized and benchmark case is simulated. As seen from Panel A of Figure 2, since $\Delta_1 = \Delta_2 = 0.1000, \Delta_1^* = 0.1933, \Delta_2^* = 0.2400$, the precondition of Proposition 1 is not satisfied, so not all investors in the initial state will choose the Follow strategy. However, we can also find the saddle point is (0.3500,0.2333), which can determine the area of the region $\boldsymbol{A}$ and $\boldsymbol{B}$ plotted in Figure 1. In this simulation environment, the initial state of most investors will eventually move towards strategy (1,1), that is, forming herd behavior.

In case 2, we will show the evolutionary process when the condition of **Proposition 1** is satisfied. In Panel B of Figure 2, $\Delta_1 = \Delta_2 = 0.1000, \Delta_1^* = 0.0964, \Delta_2^* = 0.0945$. Hence, Δ_1 and Δ_2 are larger than Δ_1^* and Δ_2^*, respectively, which meets the requirements of **Proposition 1**. The figure indicates that all initial states of the investors would converge to the unique Evolutionarily Stable Strategy $(1,1)$.

[PLEASE INSERT FIGURE 2 HERE]

In case 3 and case 4, we will demonstrate the impact of investors' sentiment on the evolutionary process. Panel C and Panel D of Figure 2 present the status of positive sentiment ($\tau = 0.5$) and negative sentiment ($\tau = 1.5$). In case 3, $X^{**} = 0.1000$ and $Y^{**} = 0.0667$. In case 4, $X^{**} = 0.6000$ and $Y^{**} = 0.4000$. Obvious, we can tell from these two panels that, positive sentiment will increase the possibility that strategy $(1,1)$ is the final stable state, while negative sentiment will have the opposite effect. The findings above have proved our hypothesis in Section 2.2 with the simulation results.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

3 Empirical design

3.1 Data sources

The Shanghai–Hong Kong Stock Connect and Shenzhen–Hong Kong Stock Connect officially began trading in November 2014 and December 2016, respectively. Under this mechanism, mainland and Hong Kong investors can purchase the listed shares of the other exchange through local brokers. To empirically identify the herding behavior around the smart money, i.e., NCF, we incorporate all Shanghai and Shenzhen A-share listed companies listed before 2017 as the research samples, spanning from Q1 2017 to Q1 2022.[2]

We obtain data from several sources. Buyer-initiated trade and seller-initiated trade data, stock-level Northbound Capital Flow data, market data, and corporate financial data are all collected from the Wind database.[3] We construct our proxy variables of the investor sentiment from post data of financial post bar and news data of financial newspapers and periodicals. The data are all from the Chinese Research Data Services Platform (hereafter CNRDS). Finally, in further analysis, we need to obtain NCF generated by different types of institutions, sourced from HKEX.[4]

Following the prior research, we exclude the stocks that had been specially treated during the sample period because their financial status and trading status were in an abnormal period. After matching data from multiple sources, these filtering criteria yield a final sample of over 48,000 observations with 2,330 listed companies over 21-quarters period.

3.2 Variable definitions

3.2.1 Stock-level herding measurement

Our dependent variable is stock-level herding. Lakonishok et al. (1992) proposed the traditional herding measurement (hereafter LSV), which has been commonly used in prior studies

[2] Since we introduce financial data in the empirical analysis, the sample frequency is determined to be quarterly.

[3] Our paper also selects the iFind database for cross-validation of original data. iFind database and Wind database are the leading financial data service providers in China, and the latter has a stronger influence.

[4] See details from https://www.hkexnews.hk/index.htm. The website provides the number of specific stocks held by different institutions on a specific date.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

(Wermers, 1999; Grinblatt et al., 1995). Following Jiao and Ye (2014) and Brown et al. (2014), the following equation gives our first stock-level herding measurement $AdjHM$. Different from the LSV measurement, a high (low) $AdjHM$ measure indicates that the stock is heavily bought (sold) by herds of investors.

$$AdjHM_{i,t} = \begin{cases} |P_{i,t} - E[P_{i,t}]| - E|P_{i,t} - E[P_{i,t}]|, & P_{i,t} > E[P_{i,t}] \\ -(|P_{i,t} - E[P_{i,t}]| - E|P_{i,t} - E[P_{i,t}]|), & P_{i,t} \leq E[P_{i,t}] \end{cases} \tag{19}$$

In the classical LSV measurement, $P_{i,t}$ denotes the proportion of money managers buying stock i in quarter t relative to the total number of money managers trading the stock. However, we cannot obtain detailed investor information to construct the stock-level herding measurement. Thus, We follow Christoffersen and Tang (2010) and Cai et al. (2019), adjusting the herding measures by using trading values rather than the number of investors. Specifically, we use buyer-initiated trade and seller-initiated trade data. Buyer-initiated trades are executed at an ask price while seller-initiated trades are executed at a bid price.

The buyer-initiated and seller-initiated trade data obtained from the Wind database are categorized into four types based on the traded value of the specific order. Small orders are trades with a value smaller than 40 thousand yuan; Medium orders are trades with a value greater than 40 thousand yuan but smaller than 200 thousand yuan; Large orders are trades with a value greater than 200 thousand yuan but smaller than a million yuan; Super-large orders are trades with a value exceeding a million yuan.

Thus for the $P_{i,t}$ in Eq.(19), it is the ratio of stock i's buyer-initiated trade value in quarter t divided by the sum of buyer-initiated trade value and seller-initiated trade value. $E[P_{i,t}]$ is the expected fraction of buy-initiated trade value, proxied by the cross-sectional average of $P_{i,t}$. $E|P_{i,t} - E[P_{i,t}]|$ is an adjustment factor for random variation around $E[P_{i,t}]$, assuming investors trade independently. We calculate the numerical value of $E|P_{i,t} - E[P_{i,t}]|$ using the normal approximation method given by Venezia et al. (2011).

To empirically examine the heterogeneous behavior between retail investors and institutional investors, we need to distinguish trades conducted by the two types of investors. Since institutional investors have more pooled money than retail investors, an order executed by institutional investors is likely to have a larger traded value. Thus, we calculate the value of retail investor trades as the sum of medium orders and small orders, while

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

the value of institutional investor trades is calculated as the sum of super-large orders and larger orders.

To sum up, we can calculate the $P_{i,t}$ by the following equation:

$$P_{i,t} = \frac{\sum_{j=n}^{N} buy_{i,j,t}}{\sum_{j=n}^{N} (buy_{i,j,t} + sell_{i,j,t})} \tag{20}$$

where $buy_{i,j,t}$ and $sell_{i,j,t}$ denote the buyer-initiated and sell-initiated trading value. i, j and t represent the firm, order size and quarter. Combined with Eq.(19), we can set $n = 1$ and $N = 4$ when calculating the stock-level aggregate investors' herding, $n = 1$ and $N = 2$ when calculating the stock-level retail investors' herding, $n = 3$ and $N = 4$ when calculating the stock-level institutional investors' herding.

Following Cai et al. (2019), the second measurement we construct is the amount-based herding measures (hereafter AHM) with the same data set. All variables and notations are the same as those in Eq.(20).

$$AHM_{i,t} = \frac{\sum_{j=n}^{N} buy_{i,j,t} - sell_{i,j,t}}{\sum_{j=n}^{N} buy_{i,j,t} + sell_{i,j,t}} \tag{21}$$

Lastly, as discussed in many previous literature (Wermers, 1999; Celiker et al., 2015), the traditional LSV measurement does not distinguish whether the imbalance is on the buy or the sell side. Thus, our third herding measure is used when exploring the heterogeneous herding behavior of investors. Buy_HM and $sell\_HM$ are constructed based on the LSV model to distinguish buy-side herding from sell-side herding.

$$Buy\_HM_{i,t} = |P_{i,t} - E[P_{i,t}]| - E|P_{i,t} - E[P_{i,t}]|, P_{i,t} > E[P_{i,t}] \tag{22}$$

$$Sell\_HM_{i,t} = |P_{i,t} - E[P_{i,t}]| - E|P_{i,t} - E[P_{i,t}]|, P_{i,t} < E[P_{i,t}] \tag{23}$$

3.2.2 Northbound capital flow

Using northbound capital flow as a proxy for smart money, the paper's central aim is to investigate whether the NCF can trigger herding. Thus, we construct our independent variable northbound capital flow as the change in the proportion of stock value held by

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

northbound capital. Specifically, NCF is given by:

$$NCF_{i,t} = \frac{MVN_{i,t}}{TVN_{i,t}} - \frac{MVN_{i,t-1}}{TVN_{i,t-1}} \tag{24}$$

where $MVN_{i,t}$ is the market value of stock i held by the northbound capital at quarter t, and $TVN_{i,t}$ is the total market value of stock i at quarter t.

3.2.3 Control variables

Since we are investigating the influence of northbound capital flow on investors' herding, we list any factors that may affect the magnitude as well as the direction of herding in our control variables. Previous studies identify stock characteristics related to the herding behavior of institutional and retail investors. Wermers (1999) showed that herding is stronger in stock with high past returns. Sias (2004) found a positive correlation between herding intensity and the information uncertainty about a stock. Thus, following previous studies (Brown et al., 2014; Barber et al., 2022), we set the following as our control variables: $Return$, $Return\_lag$, $|Return|$, $|Return\_lag|$, $Size$, Std, $Turnover$, BM, Cap, and Vol. A detailed definition of all control variables can be found in Table 4.

[PLEASE INSERT TABLE 4 HERE]

To mitigate the effect of potential outliers, we winsorize all variables (except for dummy variables) at both the first and 99th percentiles. Table 5 reports the descriptive statistics for our sample. The mean and median level of stock-level aggregate herding given by $AdjHM$ are both -0.053. Our second herding measure AHM has a mean and median of 0.001 and 0.000, respectively.

[PLEASE INSERT TABLE 5 HERE]

3.3 Empirical model

We proceed by designing our regression model to best match our theoretical counterparts. Recall that our main theoretical predictions about the herding behavior induced by the smart money flow in **Proposition 1** and **Corollary 1**. As stated above, we proxy for the herding level and smart money flow by constructing $AdjHM$, AHM and NCF. Thus,

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

we begin our assessment of the relation between smart money and investors' herding by employing the following general-form panel regression model:

$$Y_{i,t} = \beta_0 + \beta_1 \times NCF_{i,t} + X'\gamma + \eta_i + \mu_t + \varepsilon_{i,t} \tag{25}$$

where $Y_{i,t}$ is the proxy variables of investors' herding, mainly including $AdjHM_{i,t}$ and $AHM_{i,t}$. The subscript i and t denote the specific stock and the quarter respectively. We incorporate all control variables mentioned in Section 3.2.3 into the vector X. η_i and μ_t accommodate firm fixed effects and quarter fixed effects, respectively, and $\varepsilon_{i,t}$ reflects the model's residual term. We estimate Eq.(25) using ordinary least squares (OLS), with standard errors adjusted for heteroskedasticity and clustered at the firm level.

Since we are testing whether the smart money flow could induce investors' herding behavior, the coefficient of interest here is β_1. To verify our Proposition 1 and Corollary 1, that is, under certain conditions, investors choose to herd around smart money. We likewise anticipate the $\hat{\beta}_1$ to be significantly positive. When examining the buy-side or sell-side herding and heterogeneous investors' herding behavior, we replace $Y_{i,t}$ with the corresponding variable stated in Section 3.2.1, and other settings remain unchanged.

Further, our **Proposition 2** predicts that more positive sentiment would stimulate the herding behavior among investors. To test the marginal effects of investors' sentiment on herding behavior, we incorporate the interaction term of investors' sentiment proxy and NCF. In the specific demonstration, we introduce two kinds of opposite investor sentiment, namely, mania and panic, to comprehensively reflect the effects of different sentiments. The regression models are presented as follows.

$$Y_{i,t} = \beta_0 + \beta_1 \times NCF_{i,t} \times Panic_t + \beta_2 \times NCF_{i,t} + X'\gamma + \eta_i + \mu_t + \varepsilon_{i,t} \tag{26}$$

$$Y_{i,t} = \beta_0 + \beta_1 \times NCF_{i,t} \times Mania_{i,t} + \beta_2 \times NCF_{i,t} + \beta_3 \times Mania_{i,t} + X'\gamma + \eta_i + \mu_t + \varepsilon_{i,t} \tag{27}$$

where $Panic_t$ and $Mania_{i,t}$ are the proxy variables for the two opposite sentiments, and other settings are consistent with Eq.(25). In these specifications, we mainly focus on the coefficient estimates of the interaction term, i.e., $\hat{\beta}_1$, which is expected to be negative and positive, respectively.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

4 Empirical results

4.1 Main results

To test whether investors herd on smart money, we relate herding measures with the northbound capital flow. Specifically, we estimate the regression given in our empirical model Eq.(25). A large positive (negative) value for our herding measures $AdjHM$ and AHM indicates buy-side (sell-side) herding.

Table 6 presents the result of the regression. Columns (1) and (4) present regression results with no control variables and fixed effects. Results in columns (2) and (5) include only control variables, while results in columns (3) and (6) include both. The significance of coefficients is consistent with or without control variables and fixed effects, verifying the robustness of our findings. Consistent with our conjecture, the coefficient of NCF is significantly positive at the 1% level for both herding measurements. This demonstrates investors' tendency to herd around northbound capital. Thus, these findings provide empirical evidence in line with our theoretical anticipation, especially **Proposition 1** and **Corollary 1**. As for the coefficients of control variables, both $|Return|$ and $|Return\_lag|$ are positive at a 1% significance level which is intuitive: investors also tend to chase past returns. We also find that volatility (measured by Std) is negative at a 1% significance level, in accordance with the finding of Kremer and Nautz (2013).

[PLEASE INSERT TABLE 6 HERE]

Next, we aim to determine the heterogeneous herding behavior of institutional and retail investors. Using the same proxy for herding, we first construct the herding measurements separately for retail and institutional investors using their buyer-initiated and seller-initiated data as stated in Section 3.2.1. We then regress the measurements on NCF and various controls and fixed effects. The results are presented in Panel A of Table 7. All explanatory variables have the expected signs and are highly statistically significant. The results for both types of investors are significantly positive at the 1% level. To expand on our study, we examine buy-side herding and sell-side herding separately with Buy_HM and $Sell\_HM$. The result is displayed in Panel B of Table 7. All coefficients have the expected signs — positive for buy-side herding and negative for sell-side herding. When examining buy-side herding, NCF is significantly positive at a 1% level for retail investors and

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

a 10% level for institutional investors; when examining sell-side herding, NCF is significantly negative at 1% for both institutional and retail investors. The absolute values of retail investors' coefficients are consistently greater than that of institutional investors, in accordance with the finding of Li et al. (2017). This can be explained by the finding that institutional investors are typically more rational and possess more information than retail investors, which is integral to **Corollary 2** in our theoretical framework.

[PLEASE INSERT TABLE 7 HERE]

4.2 The role of investor sentiment

4.2.1 Panic suppress herding

We now explore the effect of panic on investors' herding level. The Chicago Board Options Exchange (CBOE), the largest options marketplace in the United States, has published a number of implied volatility indices, including the Chinese stock implied volatility index, namely CBOE China ETF Volatility Index (VXFXI) (Xiao et al., 2019). The implied volatility indices' changes are considered as a better measure of uncertainty in the financial markets and of investors' panic (Li et al., 2019).

Using VXFXI as a proxy of investors' panic, we perform a regression of herding measurements on the VXFXI index multiplied by NCF as listed in Eq.(26). In line with our expectation and as shown in Panel A of Table 8, the coefficients are negative at a 1% significance level for retail investors and for the aggregate market herding behavior.

In addition, we investigate the influence of panic-inducing events. Specifically, we examine the effect of the U.S.-China Trade War on herding using the dummy variables *Dispute*. The results in Panel B of Table 8 indicate that, the panic induced by the trade war suppresses retail investors' herding: the coefficient is significantly negative at a 5% level for U.S.-China Trade War. This is in accordance with the finding of Wu et al. (2020) which showed that herding behavior is significantly lower than usual in Chinese stock markets during the panic status. As for the aggregate herding, the results remain negative but insignificant.

Summarizing the results above, the herding behavior, especially retail investors' herding, would decrease with the panic. Chen et al. (2022) analyzed the spillover effects of the

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

extreme risk from the oil and USD/CNY exchange rate market to the Chinese stock market and found that herding behavior would trigger market panic and fear. Our results complement the reverse perspective, that is, the impact of panic on herding. The underlying economic logic is that, though investors are keen on chasing signals to guide their investment, the panic would push them to reduce their risk exposure. However, the marginal effects of panic are not significant for institutional investors, implying that they are more experienced and trained to cope with the panic.

[PLEASE INSERT TABLE 8 HERE]

4.2.2 Mania ignite herding

Next, we explore the effect of mania on investors' herding levels. We utilize Guba posts and financial News as two proxies of investor mania. Guba posts and News are widely used as sources to reflect public opinion and investor sentiment (Sun et al., 2018; Gao et al., 2019). Both proxies are constructed with data collected on the Chinese Research Data Service Platform. Guba posts data are found in Stock Comments Database, and financial news data are found in the Financial News Database of Chinese Listed Companies. The two proxies are then constructed as the following formula.

$$Guba_{i,t} = \frac{PP_{it}}{PP_{it} + NP_{it}} \tag{28}$$

$$News_{i,t} = \frac{PN_{it}}{PN_{it} + NN_{it}} \tag{29}$$

where $PP_{i,t}$ and $NP_{i,t}$ are the numbers of positive and negative posts on Guba, and $PN_{i,t}$ and $NN_{i,t}$ are the numbers of positive and negative financial news from over 500 important newspaper media.

We estimate Eq.(27), and the regression results are displayed in Table 9. Panel A uses Guba posts to measure investors' sentiment. The results are significantly positive for retail investors at a 1% level, while it is significantly negative for institutional investors at a 1% level. Intuitively, Guba is a platform developed for retail investors to share their thoughts. As a result, it strongly reflects retail investors' sentiment. However, this does not apply to institutional investors, which may even overturn their views on the market. Panel B uses

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

financial news data to measure investors' mania. The result is significantly positive at a 1% level for institutional investors. The findings are accordant with Nofsinger (2001). Using data from Wall Street Journal and Macro-economics announcements, Nofsinger (2001) discovered that institutional investors react strongly to news releases. The result is also positive at a 5% significant level for the aggregate market herding. It is worth noting from the two panels that, the mania could stimulate the aggregate market herding behavior. Nevertheless, different sentiment construction methods will reflect the mania of different types of investors.

[PLEASE INSERT TABLE 9 HERE]

Thus far, combining the results in Section 4.2.1 and Section 4.2.2, we have found strong supportive evidence of our **Proposition 2**. That is, we have demonstrated that more positive sentiment would amplify the herding effect of investors on smart money.

4.3 Robustness tests

Our results thus far indicate that investors would chase and herd around NCF, and are heterogeneous among different investors. Nevertheless, to the extent that our herding measurement is originated from LSV method, which involves relatively complex design and calculation, one could argue whether the original LSV method has the same results. In addition, the significant results may be related to the construction of $P_{i,t}$.

To test the robustness of our main findings, we draw on a battery of herding measures and repeat our estimation from Eq.(25), and the results are presented in Table 10.

First, Panel A displays the result of incorporating the unadjusted LSV herding measurement as the dependent variable. We regress the measurement on the absolute value of NCF and other control variables and fixed effects. The LSV model cannot distinguish between buy-side and sell-side herding, making it a measure of magnitude rather than direction. As both an increase and decrease in NCF can trigger herding, it is necessary to regress on the absolute value of NCF. Column (1) shows the aggregate herding of a specific stock, while columns (2) and column (3) separately examine institutional investors and individual investors' herding respectively. The coefficient of $|NCF|$ is positive at a 1% significance level for the overall market behavior. The coefficients are significantly positive

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

at a 5% level and a 1% level for institutional investors and retail investors, respectively. In accordance with the previous result, the coefficient of retail investors is greater than that of institutional investors.

Second, Panel B adjusts the heterogeneous investors' herding measurement caliber. Any trade with a value exceeding 200 thousand will now be categorized as institutional buy or sell trading. The result is still significantly positive at a 1% level for both institutional and retail investors across the two different herding measurements. Thus, the result is not affected by $P_{i,t}$ construction.

[PLEASE INSERT TABLE 10 HERE]

Our next set of robustness tests aims to perform regression with different independent variables, and the results are shown in Table 11. Following Venezia et al. (2011), we incorporate lagged terms of NCF and dependent variables into the regression. Columns (1) and (3) in Panel A include NCF and lagged herding measurements in the regression. Columns (2) and (4) in Panel A include the lagged NCF term. The results are all significantly positive at a 1% level. Correspondingly, Panel B displays the results after incorporating the absolute value of northbound capital flow in the regression. We first include the absolute value of NCF ($AVNCF$) in columns (1) and (3). To prevent the impact of extreme value, we also include $log(1 + AVNCF)$ and present the results in columns (2) and (4). The results are also positive at a 1% significant level.

[PLEASE INSERT TABLE 11 HERE]

As another robustness test, Petersen (2009) argued that in the presence of a time effect, Fama-MacBeth produces unbiased standard errors and correctly sized confidence intervals. Thus, we additionally perform a Fama and MacBeth (1973) regression of investor herding on northbound capital flow. As shown in Table 12, columns (1) and (3) display the results with no control variables, while columns (2) and (4) include control variables. The coefficients are still positive and statistically significant at a 1% level.

[PLEASE INSERT TABLE 12 HERE]

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

5 Further analysis

5.1 The perception of smart money

Akre et al. (2011) argued that decisions are based on the perception of stimuli. This relationship is critical to understanding decision-making. Thus, to further our research, we explore whether and how the investors' perception of smart money impacts herding behavior from mainly three perspectives.

First, we explore how the sensitivity to northbound capital flow varies across different stocks. We examine the relationship between the northbound capital holding ratio Holding (the proportion of the stock value held by northbound capital) and our herding measurements. Panel A of Table 13 reports the results. The coefficient estimates of all columns are significantly negative at a 1% level, indicating that a high proportion of NCF holding may decrease investors' herding impulsion. In this more precise setting, the findings align with our pure investment experience and psychological knowledge, that is, constant stimulation will desensitize investors (Tryon, 2005).

Second, Venezia et al. (2011) and Hsieh et al. (2020) all suggested that small capitalization firms generate intense herding. To demonstrate whether such effect holds in our setting, we incorporate the interaction term of Cap and NCF to reflect the marginal effects. Panel B of Table 13 presents the findings, where we examine the relationship between herding measurement and stock market capitalization. Consistent with the existing literature (Zhou and Lai, 2009; Venezia et al., 2011), a monotonic reverse relationship is found between our herding measurements and market capitalization. The results are mostly negative at a 1% significance level, with some at a 5% significance level. Investors are typically insensitive to stocks with high market capitalization, since investors' search costs of small capitalization stocks may be higher, and they will be more likely to chase smart money in these stocks.

[PLEASE INSERT TABLE 13 HERE]

Third, we also explore investors' sensitivity to the two different types of northbound capital: allocation-oriented and trade-oriented NCF. Long-term investors, such as overseas pension funds, who trade A shares through foreign banks, are called allocation-oriented northbound capital. On the contrary, short-term investors, such as overseas hedge funds,

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

who trade A-shares through foreign securities companies, are known as trade-oriented northbound capital. Allocation-oriented northbound capital prefers to hold high-quality A-share stocks for a long time. Thus, the turnover rate and trading frequency of trade-oriented investors are significantly higher than those of allocation-oriented investors.

As a result, trade-oriented capital induces more trading fluctuations, consequently receiving more public attention, and investors are more likely to herd around this type of NCF. The results presented in Table 14 are in line with our expectations. We find that herding is only significantly related to trade-oriented capital but not to allocation-oriented capital. The results are significantly positive at a 1% level for trade-oriented capital. This result proves that investors only pay attention to the fluctuations of NCF and do not pay attention to the more detailed fund types within NCF.

[PLEASE INSERT TABLE 14 HERE]

5.2 Regression discontinuity design

In this section, following Barber et al. (2022), we perform a sharp regression discontinuity design to further establish the causal impact of the NCF on investors' herding. Ideally, if the exchanges announce the stock list of the largest NCF on a quarterly basis, one could consider a regression discontinuity design that explores the threshold between stocks at the end of the list and those almost on the list. However, from financial institutions' reports and media news, we can see many daily, weekly or monthly stock lists of northbound capital inflow, including the top 10, top 30, etc. We are unsure what investors have observed on the list, and the stock lists are likely to change throughout the quarter. Thus, even if we get the daily list for quarterly analysis, the noise potentially introduced by using the approximated lists may be too large to warrant a clean identification.

Instead, we calculate the net inflow of the northbound capital for all stocks on a quarterly basis and then sort them. We choose the ranking of 100 as the threshold for the following two reasons: On the one hand, many news reports chose to report a stock list of the top 100 NCF net inflows in the current quarter; On the other hand, if the threshold ranking is too small, there will exist a large bias. Due to the discrete nature of the rank variable, the exact cutoff threshold in the [100,101] interval is an empirical design choice. Following Kaniel and Parham (2017), we choose the average 100.5 as our breaking point. If

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

there's no tendency to follow northbound capital, there should be no discernable differences in investors' herding intensity around the threshold. Thus intuitively, the estimation should exploit the discontinuity at the threshold of ranking 100.5 and test for discontinuities in investor herding behavior around this threshold. Following Barber et al. (2022), We estimate the following sharp RD specification:

$$
\begin{aligned}
Y_{i,t} = \beta_0 + \beta_1 I_{rank<100.5} + \sum_{n=1}^{N} \beta_2^n \left(rank_{i,t} - 100.5\right)^n \\
+ \sum_{n=1}^{N} \beta_3^n \left(rank_{i,t} - 100.5\right)^n \times I_{rank<100.5} + X'\delta + \eta_i + \mu_t + \varepsilon_{i,t}
\end{aligned}
\tag{30}
$$

where $Y_{i,t}$ is our herding measures, and $I_{rank<100.5}$ is a dummy variable that equals one if the stock rank is lower than 100.5. As the controls, we incorporate different polynomial functions of rank ($N = 2, 3$) so that the point estimate on the above-cutoff indicator variable (β_1) is identified under the assumption that the way of investors' herding behavior related to the rank is not discontinuous exactly at the rank threshold of 100.5.

[PLEASE INSERT FIGURE 3 HERE]

Fitting our data with quadratic and cubic functions, we find a sharp discontinuity around the breaking point. Figure 3 presents a visual illustration of our results. The diagrams are in accordance with the results of our regression. Two regressions are performed with different sample bandwidths. Panel A shows the result with a sample bandwidth of 15. Panel B of Table 15 displays the results with a sample bandwidth of 25. With a bandwidth of 15, the regressions in Panel A are all significantly positive at a 5% level. The results for regression with a sample bandwidth of 25 are positive at a 5% significant level for a quadratic function and at a 10% significant level for a cubic function. Overall, these results are consistent with the idea that the stocks ranked high in NCF can attract investors' attention and affect investors' herding behavior.

[PLEASE INSERT TABLE 15 HERE]

仅用于2022年成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

5.3 Real smart money or just nominal title

To test how "smart" northbound capital is, we construct an equally weighted stocks portfolio with stocks that have northbound capital holding. We take long (short) positions for stocks that experienced net inflow (outflow) of northbound capital. The return of the portfolio is shown in Figure 4. The diagram also includes the return of the Shanghai Composite Index as a reference. With an overall return of 38.51%, which is 20% higher than the return of SSEC, northbound capital is indeed "smart" as it can obtain high returns.

[PLEASE INSERT FIGURE 4 HERE]

Table 16 examines the stock selection ability of northbound capital. Columns (1) and (2) regress the return of assets on northbound capital holding. The coefficients are significantly positive at a 5% level and 10% level for the one-quarter forward ROA and two-quarter forward ROA respectively. Northbound capital is also positively related to the one-quarter forward stock return. The coefficient is significantly positive at a 1% level. In line with our expectation, northbound capital is smart and exhibits strong stock selection ability.

[PLEASE INSERT TABLE 16 HERE]

6 Conclusions

Our paper answers some important outstanding questions in the literature on the opposition between boundedly rational investors and redundant information. Hirshleifer and Teoh (2003) argued that, owing to limits to investor attention, the information presented in salient, easily processed form is assumed to be absorbed more easily. In this paper, we take advantage of the smart money concept, which is a widely accepted investment guidance signal, to demonstrate the consequences of investors chasing simplified signals, i.e., herding behavior, which is new to the literature. We show the investors' herding around smart money, theoretically and empirically, and also verify the role of sentiment and perception.

We construct a theoretical model to examine heterogeneous investors' strategy in response to exogenous smart money signals, that is, to follow the signal or to conduct self-analysis, within the framework of the evolutionary game theory. Our model shows that

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

institutional and retail investors choose to herd around smart money under certain conditions, and investors' sentiment would impact the herding level by affecting investors' mental accounts. Specifically, investors' positive sentiment would exacerbate the herding behavior, vice versa.

The Stock Connect Program provides an important cornerstone for our empirical research. By consensus, financial news and social media habitually refer to northbound capital as smart money, which is also supported by previous literature (Huang et al., 2016; Lu et al., 2021). Thus, we conduct an empirical analysis using the stock-level northward capital flow (NCF) and herding measurements with the Shanghai and Shenzhen A-share listed companies to analyze the causality effect. Our paper shows that investors will have significant herding behavior around the NCF, and such effect is more pronounced in the retail investors' group. Our results are robust with a battery of robustness tests, including altering herding and NCF measurement caliber, incorporating lagged terms, and conducting RDD analysis.

We further expand our research on the impact of investors' sentiment and perception of smart money flow. Evidence shows that investors' panic would suppress the herding behavior around smart money, while the mania would have the opposite effect. We also confirm that investors' perception of smart money plays a role on their herding, that is, investors may be insensitive to the capital flow towards stocks with higher northbound capital holding ratios or with large capitalization. When we divide the NCF into allocation-oriented type and trading-oriented type, investors would herd around the fund type inducing more fluctuations in stock price.

Collectively, this study contributes to the literature on the consequences of boundedly rational investors in the information explosion era, while our theoretical model and empirical analysis bring a unique insight to the literature on the nature of herding behavior. With the Stock Connect Program, we also supplement the evidence of the impact of investor sentiment and perception on the degree of herding.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

References

Akbas, F., W. J. Armstrong, S. Sorescu, and A. Subrahmanyam (2015). Smart money, dumb money, and capital market anomalies. *Journal of Financial Economics 118*(2), 355–382.

Akre, K. L., H. E. Farris, A. M. Lea, R. A. Page, and M. J. Ryan (2011). Signal perception in frogs and bats and the evolution of mating signals. *Science 333*(6043), 751–752.

Barber, B. M., X. Huang, T. Odean, and C. Schwarz (2022). Attention induced trading and returns: Evidence from robinhood users. *Journal of Finance, forthcoming*.

Ben-David, I., J. Li, A. Rossi, and Y. Song (2022). What do mutual fund investors really care about? *The Review of Financial Studies 35*(4), 1723–1774.

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

Cai, F., S. Han, D. Li, and Y. Li (2019). Institutional herding and its price impact: Evidence from the corporate bond market. *Journal of Financial Economics 131*(1), 139–167.

Cardinaels, E., S. Hollander, and B. J. White (2019). Automatic summarization of earnings releases: attributes and effects on investors' judgments. *Review of Accounting Studies 24*(3), 860–890.

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

Chen, L., F. Wen, W. Li, H. Yin, and L. Zhao (2022). Extreme risk spillover of the oil, exchange rate to chinese stock market: Evidence from implied volatility indexes. *Energy Economics 107*, 105857.

Christoffersen, S. K. and Y. Tang (2010). Institutional herding and information cascades: Evidence from daily trades. *Available at SSRN 1572726*.

Cole, S., A. Paulson, and G. K. Shastry (2014). Smart money? the effect of education on financial outcomes. *The Review of Financial Studies 27*(7), 2022–2051.

Daniel, K., D. Hirshleifer, and S. H. Teoh (2002). Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics 49*(1), 139–209.

Demirer, R., A. M. Kutan, and C.-D. Chen (2010). Do investors herd in emerging stock markets?: Evidence from the taiwanese market. *Journal of Economic Behavior & Organization 76*(2), 283–295.

Devenow, A. and I. Welch (1996). Rational herding in financial economics. *European Economic Review 40*(3-5), 603–615.

Dietrich, J. R., S. J. Kachelmeier, D. N. Kleinmuntz, and T. J. Linsmeier (2001). Market efficiency, bounded rationality, and supplemental business reporting disclosures. *Journal of Accounting Research 39*(2), 243–268.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Doellman, T. W., J. Itzkowitz, J. Itzkowitz, and S. H. Sardarli (2019). Alphabeticity bias in 401 (k) investing. *Financial Review 54*(4), 643–677.

Elliott, W. B., J. L. Hobson, and B. J. White (2015). Earnings metrics, information processing, and price efficiency in laboratory markets. *Journal of Accounting Research 53*(3), 555–592.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Frazzini, A. and O. A. Lamont (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics 88*(2), 299–322.

Galariotis, E. C., W. Rong, and S. I. Spyrou (2015). Herding on fundamental information: A comparative study. *Journal of Banking & Finance 50*, 589–598.

Gao, Y., X. Xiong, X. Feng, Y. Li, and S. A. Vigne (2019). A new attention proxy and order imbalance: Evidence from china. *Finance Research Letters 29*, 411–417.

Gibson, S., A. Safieddine, and R. Sonti (2004). Smart investments by smart money: Evidence from seasoned equity offerings. *Journal of Financial Economics 72*(3), 581–604.

Giglio, S. and K. Shue (2014). No news is news: do markets underreact to nothing? *The Review of Financial Studies 27*(12), 3389–3440.

Grinblatt, M. and B. Han (2005). Prospect theory, mental accounting, and momentum. *Journal of Financial Economics 78*(2), 311–339.

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

Hervé, F., M. Zouaoui, and B. Belvaux (2019). Noise traders and smart money: Evidence from online searches. *Economic Modelling 83*, 141–149.

Hirshleifer, D., S. S. Lim, and S. H. Teoh (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance 64*(5), 2289–2325.

Hirshleifer, D. and S. H. Teoh (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics 36*(1-3), 337–386.

Hsieh, S.-F., C.-Y. Chan, and M.-C. Wang (2020). Retail investor attention and herding behavior. *Journal of Empirical Finance 59*, 109–132.

Huang, Y., G. Jacoby, and C. X. Jiang (2016). The bonding hypothesis and the home market liquidity of chinese cross-listed stocks. *Journal of International Financial Markets, Institutions and Money 43*, 146–157.

Jegadeesh, N. and W. Kim (2010). Do analysts herd? an analysis of recommendations and market reactions. *The Review of Financial Studies 23*(2), 901–937.

Jiao, Y. and P. Ye (2014). Mutual fund herding in response to hedge fund herding and the impacts on stock prices. *Journal of Banking & Finance 49*, 131–148.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Kaniel, R. and R. Parham (2017). Wsj category kings–the impact of media attention on consumer and mutual fund investment decisions. *Journal of Financial Economics 123*(2), 337–356.

Kizys, R., P. Tzouvanas, and M. Donadelli (2021). From covid-19 herd immunity to investor herding in international stock markets: The role of government and regulatory restrictions. *International Review of Financial Analysis 74*, 101663.

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

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

Li, J., X. Yu, and X. Luo (2019). Volatility index and the return–volatility relation: Intraday evidence from chinese options market. *Journal of Futures Markets 39*(11), 1348–1359.

Li, W., G. Rhee, and S. S. Wang (2017). Differences in herding: Individual vs. institutional investors. *Pacific-Basin Finance Journal 45*, 174–185.

Lin, W. (2017). Modeling volatility linkages between shanghai and hong kong stock markets before and after the connect program. *Economic Modelling 67*, 346–354.

Liu, C., S. Wang, and K. J. Wei (2021). Demand shock, speculative beta, and asset prices: Evidence from the shanghai-hong kong stock connect program. *Journal of Banking & Finance 126*, 106102.

Lou, D. (2012). A flow-based explanation for return predictability. *The Review of Financial Studies 25*(12), 3457–3489.

Lu, J., X. Bu, and J. Chen (2021). Do inflowing sophisticated investors induce classification shifting? new evidence from market liberalisation in china. *Accounting & Finance 61*(5), 6193–6223.

Milkman, K. L. and J. Beshears (2009). Mental accounting and small windfalls: Evidence from an online grocer. *Journal of Economic Behavior & Organization 71*(2), 384–394.

Nofsinger, J. R. (2001). The impact of public information on investors. *Journal of Banking & Finance 25*(7), 1339–1366.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis 40*(2), 307–329.

Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics 80*(3), 563–602.

Peng, L., L. Zhang, and W. Chen (2021). Capital market liberalization and investment efficiency: evidence from china. *Financial Analysts Journal 77*(4), 23–44.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies 22*(1), 435–480.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

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

Sørensen, M. (2007). How smart is smart money? a two-sided matching model of venture capital. *The Journal of Finance 62*(6), 2725–2762.

Sun, Y., M. Fang, and X. Wang (2018). A novel stock recommendation system using guba sentiment analysis. *Personal and Ubiquitous Computing 22*(3), 575–587.

Thaler, R. H. (2008). Mental accounting and consumer choice. *Marketing Science 27*(1), 15–25.

Thaler, R. H., S. Benartzi, et al. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review 91*(1), 79–98.

Tryon, W. W. (2005). Possible mechanisms for why desensitization and exposure therapy work. *Clinical Psychology Review 25*(1), 67–95.

Venezia, I., A. Nashikkar, and Z. Shapira (2011). Firm specific and macro herding by professional and amateur investors and their effects on market volatility. *Journal of Banking & Finance 35*(7), 1599–1609.

Wang, J., Y. Qin, and J. Zhou (2021). Incentive policies for prefabrication implementation of real estate enterprises: An evolutionary game theory-based analysis. *Energy Policy 156*, 112434.

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

Wu, G., B. Yang, and N. Zhao (2020). Herding behavior in chinese stock markets during covid-19. *Emerging Markets Finance and Trade 56*(15), 3578–3587.

Xiao, J., C. Hu, G. Ouyang, and F. Wen (2019). Impacts of oil implied volatility shocks on stock implied volatility in china: Empirical evidence from a quantile regression approach. *Energy Economics 80*, 297–309.

Xu, K., X. Zheng, D. Pan, L. Xing, and X. Zhang (2020). Stock market openness and market quality: evidence from the shanghai–hong kong stock connect program. *Journal of Financial Research 43*(2), 373–406.

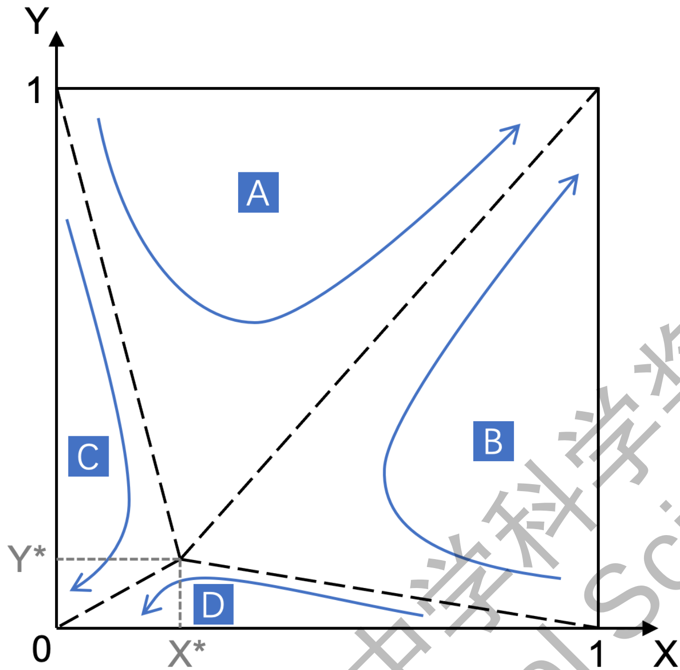
Ying, Q., Y. Zhu, M. Yao, and Z. Li (2021). Does stock market liberalisation restrain corporate financialisation? *Accounting & Finance 61*(5), 6263–6294.

Zhao, Y., C. Xiang, and W. Cai (2021). Stock market liberalization and institutional herding: Evidence from the shanghai-hong kong and shenzhen-hong kong stock connects. *Pacific-Basin Finance Journal 69*, 101643.

Zhou, R. T. and R. N. Lai (2009). Herding and information based trading. *Journal of Empirical Finance 16*(3), 388–393.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

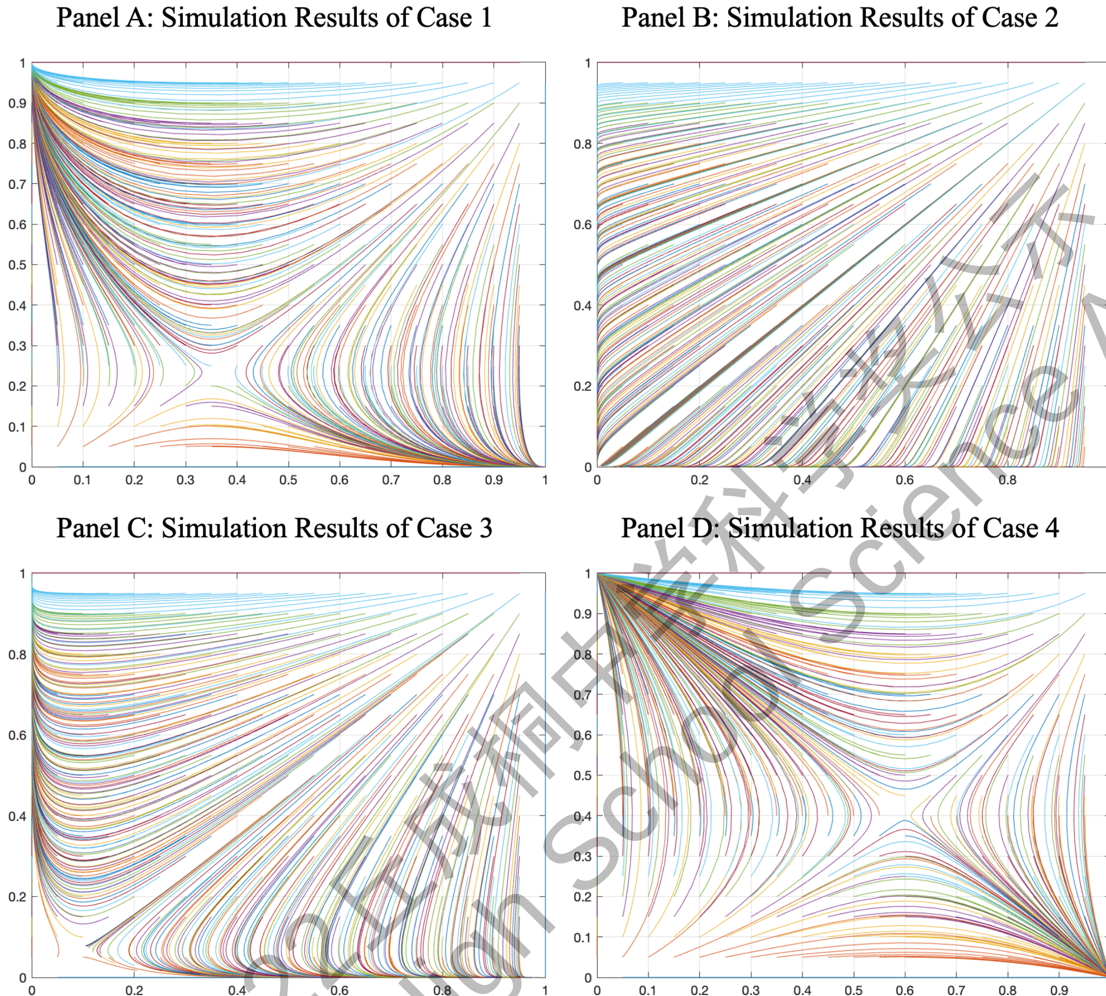
Figure 1: Depict of Evolutionary Game Model



Note: This figure illustrates the vector fields for the five possible stable equilibria. As demonstrated above, the local equilibrium points $(0,1)$ and $(1,0)$ are unstable, (X^*,Y^*) is a saddle point, and only $(0,0)$ or $(1,1)$ is the stable point. From this figure, we can find the dynamic evolution process of all initial states.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Figure 2: The Simulation Results of EGM

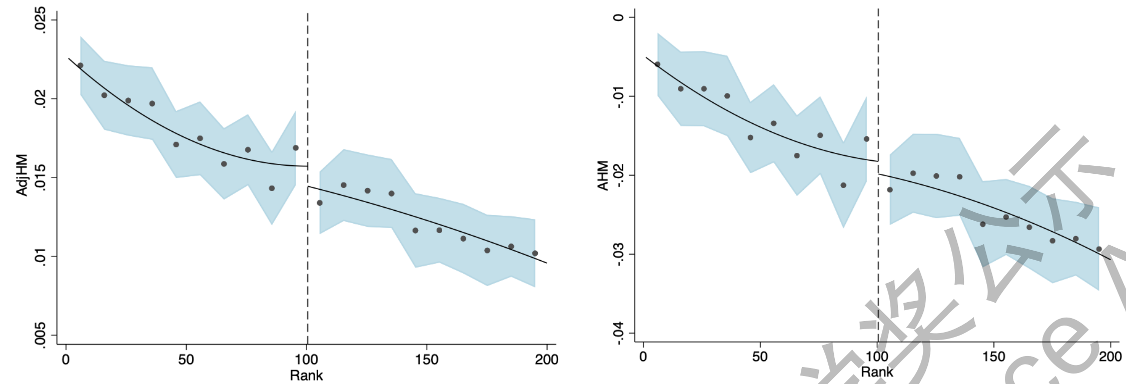


Note: This figure illustrates our simulation results of four typical cases. We utilize the MATLAB system to simulate the dynamic evolution, and then the evolution process of the investors' strategies would be observed more intuitively.

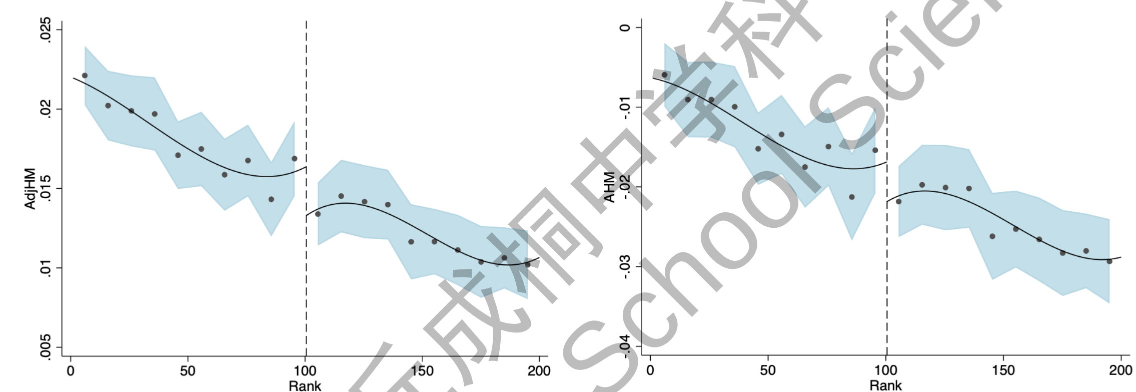
仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Figure 3: Herding Behavior Change around the Threshold of Ranking 100.5

Panel A: Quadratic Fitted Line



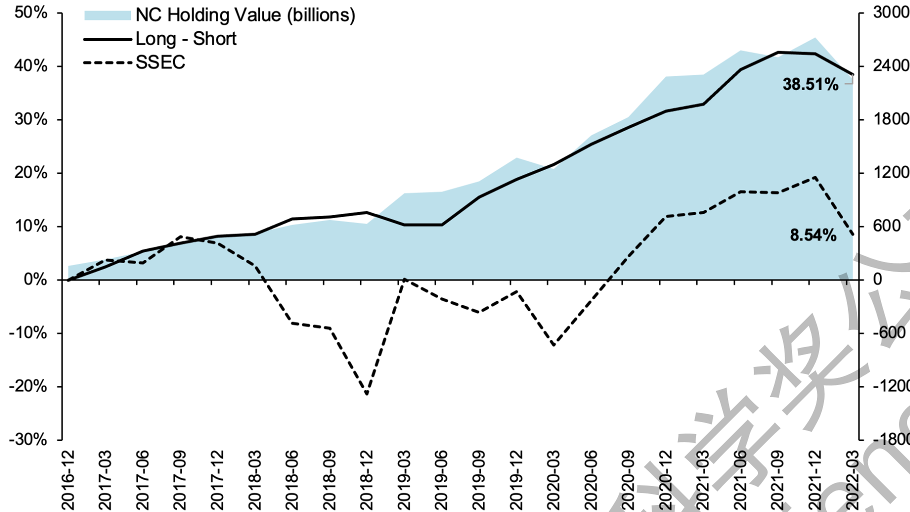
Panel B: Cubic Fitted Line



Note: This figure presents a visual illustration of our results. In Panel A, we incorporate the quadratic functions of the rank variable. In Panel B, we incorporate the cubit functions of the rank variable. Each panel contains two results for $AdjHM$ and AHM, respectively.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Figure 4: Cumulative Returns of Portfolios Sorted by NCF



Note: In this figure, we construct an equally weighted stocks portfolio with stocks that have northbound capital holding. We take long (short) positions for stocks that experienced net inflow (outflow) of northbound capital. The return of the portfolio is shown in Figure 4.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 1: Payment Matrix of EGM

Exogeneous Smart Money Behavior		Retail Investor	
		***F*ollow**	***S*elf-Analysis**
Institutional Investor	***F*ollow**	$\alpha_1 r(\theta_1 + \theta_2) - C_1$ $\alpha_2 r(\theta_1 + \theta_2) - C_2$	$\alpha_1 r(\theta_1 + \theta_2') - C_1$ $\alpha_2' r(\theta_1 + \theta_2') - C_2'$
	***S*elf-Analysis**	$\alpha_1' r(\theta_1' + \theta_2) - C_1'$ $\alpha_2 r(\theta_1' + \theta_2) - C_2$	$\alpha_1' r(\theta_1' + \theta_2') - C_1'$ $\alpha_2' r(\theta_1' + \theta_2') - C_2'$

Note: According to our hypothesis, after seeing the exogenous smart money behavior, in the game system of institutional investors and retail investors, there are four circumstances that occur, that is $\{$*Follow*, *Self-Analysis*$\}\times\{$*Follow*, *Self-Analysis*$\}$. This table presents the payment matrix.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 2: Payment Matrix of EGM Incorporating Investor Sentiment

Exogeneous Smart Money Behavior		Retail Investor	
		***F**ollow*	***S**elf-Analysis*
Institutional Investor	***F**ollow*	$\alpha_1 r(\theta_1+\theta_2)-\tau C_1$ $\alpha_2 r(\theta_1+\theta_2)-\tau C_2$	$\alpha_1 r(\theta_1+\theta_2')-\tau C_1$ $\alpha_2' r(\theta_1+\theta_2')-\tau C_2'$
	***S**elf-Analysis*	$\alpha_1' r(\theta_1'+\theta_2)-\tau C_1'$ $\alpha_2 r(\theta_1'+\theta_2)-\tau C_2$	$\alpha_1' r(\theta_1'+\theta_2')-\tau C_1'$ $\alpha_2' r(\theta_1'+\theta_2')-\tau C_2'$

Note: To clarify the impact of investor sentiment on herding behavior around smart money, we can incorporate an additional parameter τ to modify the costs of investors' decisions. Investors' sentiment is divided into positive type and negative type, which will reduce or increase the costs of investors' mental accounts respectively, i.e., $0 < \tau < 1$ or $1 < \tau$. Then, the payment matrix will be fine-tuned as shown in the table.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 3: Parameter Setting for Simulation

Parameter	Case 1	Case 2	Case 3	Case 4
θ_1	0.11500	0.11500	0.11500	0.11500
θ_2	0.11000	0.11000	0.11000	0.11000
θ_1'	0.01500	0.01500	0.01500	0.01500
θ_2'	0.01000	0.01000	0.01000	0.01000
α_1	0.01500	0.01500	0.01500	0.01500
α_2	0.01000	0.01000	0.01000	0.01000
α_1'	0.11500	0.11500	0.11500	0.11500
α_2'	0.11000	0.11000	0.11000	0.11000
r	0.03000	0.11000	0.03000	0.03000
C_1	0.00010	0.00010	0.00010	0.00010
C_2	0.00015	0.00015	0.00015	0.00015
C_1'	0.00020	0.00020	0.00020	0.00020
C_2'	0.00030	0.00030	0.00030	0.00030
τ	1.00000	1.00000	0.50000	1.50000

Note: This table illustrates four typical cases in our model simulation. In case 1, the generalized and benchmark case is simulated. In case 2, we will show the evolutionary process when the condition of Proposition 1 is satisfied. In case 3 and case 4, we will demonstrate the impact of investors' sentiment on the evolutionary process.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 4: Variable Definitions

Variable	Definition
AHM	Amount-based herding measurement which is constructed by the market trading data as shown in Eq.(21).
AdjHM	Adjusted herding measurement which is constructed by the market trading data as shown in Eq.(19).
NCF	The proportion of the stock value held by northbound capital at quarter t minus the proportion of the stock value hold by northbound capital at quarter $t-1$.
Size	Natural logarithm of the total assets.
Std	Standard deviation of the market return.
Turnover	Position turnover rate, equal to the trading volume divided by the total number of outstanding shares.
BM	The ratio of book value of equity divided by market value of equity.
Cap	Total market value, equal to the number of shares multiplied by the price per share.
Vol	Trading volume of stocks in the current quarter.
Return	Stock return in the current quarter.
Return_lag	Stock return in the last quarter.
$\|Return\|$	Absolute value of stock return in the current quarter.
$\|Return\_lag\|$	Absolute value of stock return in the last quarter.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 5: Descriptive Statistics

Variable	Obs	Mean	S.D.	P10	Median	P90
AHM	48627	-0.053	0.048	-0.108	-0.054	0.002
AdjHM	48627	0.001	0.019	-0.019	0.000	0.022
NCF	48567	0.000	0.004	-0.001	0.000	0.003
Size	48622	8.923	1.440	7.337	8.688	10.770
Std	48627	0.025	0.009	0.014	0.024	0.038
Turnover	48627	1.160	1.095	0.259	0.794	2.530
BM	48627	0.506	0.312	0.174	0.437	0.939
Cap	48627	9.039	1.069	7.868	8.818	10.470
Vol	48627	6.349	1.101	4.939	6.325	7.810
Return	48627	0.011	0.200	-0.198	-0.024	0.268
Return_lag	48627	0.015	0.196	-0.189	-0.019	0.267
$\|Return\|$	48627	0.146	0.138	0.020	0.109	0.307
$\|Return\_lag\|$	48627	0.142	0.138	0.018	0.106	0.302

Note: This table summarizes the summary statistics and our research sample spans a 21-quarter period from Q1 2017 to Q1 2022. After the essential data processing, we have a final sample of over 48,000 firm-quarter observations. All variables are defined in Table 4.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 6: Herding Behavior around Smart Money

	AdjHM			*AHM*		
	(1)	(2)	(3)	(4)	(5)	(6)
NCF	0.691***	0.345***	0.245***	1.691***	0.751***	0.590***
	(0.022)	(0.020)	(0.019)	(0.055)	(0.047)	(0.044)
Return		0.024***	0.035***		0.084***	0.091***
		(0.001)	(0.001)		(0.002)	(0.002)
Return_lag		0.008***	0.010***		0.023***	0.027***
		(0.000)	(0.001)		(0.001)	(0.001)
$\|Return\|$		-0.006***	-0.010***		-0.030***	-0.032***
		(0.001)	(0.001)		(0.002)	(0.002)
$\|Return\_lag\|$		0.001*	-0.000		0.003**	-0.002
		(0.001)	(0.001)		(0.002)	(0.002)
Size		-0.001***	-0.003***		-0.003***	-0.007***
		(0.000)	(0.001)		(0.001)	(0.002)
Std		0.031	-0.298***		0.247***	-0.732***
		(0.019)	(0.030)		(0.046)	(0.067)
Turnover		-0.000	-0.001***		-0.001	-0.002***
		(0.000)	(0.000)		(0.000)	(0.001)
BM		-0.003***	-0.005***		-0.004***	-0.017***
		(0.001)	(0.001)		(0.001)	(0.003)
Cap		0.007***	0.011***		0.016***	0.025***
		(0.000)	(0.001)		(0.001)	(0.002)
Vol		-0.002***	0.001**		-0.002***	0.004***
		(0.000)	(0.000)		(0.000)	(0.001)
Constant	0.001***	-0.038***	-0.061***	-0.053***	-0.159***	-0.211***
	(0.000)	(0.001)	(0.006)	(0.000)	(0.003)	(0.014)
Firm FE	No	No	Yes	No	No	Yes
Quarter FE	No	No	Yes	No	No	Yes
Observations	48567	48562	48562	48567	48562	48562
Adj R^2	0.020	0.185	0.304	0.019	0.241	0.368

Note: This table presents the results of the regression in Eq.(25).The dependent variables are *AdjHM* and *AHM* defined in Eq.(19) and Eq.(21), respectively. *NCF* indicates the level of inflow of northbound capital to a specific stock. Control variables are all defined in Table 4. Robust standard errors (clustered at the firm level) are provided in parentheses. *, **, ***, denotes statistical significance at the 10%, 5%, and 1% level, respectively.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 7: Heterogeneous Investor's Herding Level around Smart Money

Panel A: Heterogeneous investor's herding results

	AdjHM		*AHM*	
	Institution	Retail	Institution	Retail
NCF	0.218^{***}	0.276^{***}	0.447^{***}	0.635^{***}
	(0.044)	(0.043)	(0.094)	(0.089)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	48562	48562	48562	48562
Adj R^2	0.467	0.531	0.493	0.544

Panel B: Heterogeneous investor's herding results across buy- and sell- side

	Buy_HM		*Sell_HM*	
	Institution	Retail	Institution	Retail
NCF	0.102^{*}	0.276^{***}	-0.160^{***}	-0.395^{***}
	(0.055)	(0.021)	(0.041)	(0.096)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	20973	29954	27451	18328
Adj R^2	0.472	0.341	0.138	0.742

Note: Panel A of this table shows the results with the herding measurements separately for retail and institutional investors as independent variables. Panel B of this table shows the results about the buy-side herding and sell-side herding separately with Buy_HM and $Sell\_HM$ as defined in Eq.(22) and Eq.(23). All control variables are consistent with those in Table 6.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 8: Panic Suppress Herding

	AdjHM			*AHM*		
	Full	Institution	Retail	Full	Institution	Retail
Panel A: Volatility index						
VXFXI × *NCF*	-0.013***	-0.004	-0.033***	-0.034***	-0.018	-0.075***
	(0.003)	(0.006)	(0.007)	(0.006)	(0.014)	(0.015)
NCF	0.587***	0.330*	1.148***	1.470***	0.927**	2.599***
	(0.080)	(0.171)	(0.194)	(0.183)	(0.373)	(0.395)
Observations	48562	48562	48562	48562	48562	48562
Adj R^2	0.304	0.467	0.532	0.368	0.493	0.544
Panel B: U.S. – China Trade War						
Dispute × *NCF*	-0.028	0.288	-0.473**	-0.059	0.452	-0.952**
	(0.077)	(0.203)	(0.204)	(0.178)	(0.437)	(0.420)
NCF	0.271***	-0.053	0.721***	0.646***	0.021	1.531***
	(0.077)	(0.207)	(0.209)	(0.178)	(0.442)	(0.430)
Observations	48562	48562	48562	48562	48562	48562
Adj R^2	0.304	0.467	0.531	0.368	0.493	0.544
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table explores the effect of panic on investors' herding level. In Panel A, Using VXFXI as a proxy of investors' panic, we perform a regression of herding mea- surements on the VXFXI index multiplied by NCF as listed in Eq.(26). In Panel B, we examine the effect of the U.S.-China Trade War on herding using the dummy variables Dispute.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 9: Mania Ignite Herding

	AdjHM			*AHM*		
	Full	Institution	Retail	Full	Institution	Retail
Panel A: Investors' sentiment measured Guba post						
Guba × *NCF*	0.383*	-1.619***	2.152***	0.656	-3.553***	4.356***
	(0.219)	(0.544)	(0.568)	(0.501)	(1.164)	(1.158)
Guba	0.029***	0.058***	0.023***	0.079***	0.139***	0.059***
	(0.002)	(0.005)	(0.004)	(0.005)	(0.011)	(0.008)
NCF	0.016	1.110***	-0.959***	0.184	2.396***	-1.874***
	(0.121)	(0.311)	(0.338)	(0.279)	(0.663)	(0.686)
Observations	46218	46218	46218	46218	46218	46218
Adj R^2	0.314	0.469	0.526	0.378	0.494	0.537
Panel B: Investors' sentiment measured financial news						
News × *NCF*	0.226**	0.642***	-0.143	0.527**	1.425***	-0.370
	(0.092)	(0.229)	(0.213)	(0.217)	(0.496)	(0.437)
News	-0.003***	-0.005***	-0.000	-0.007***	-0.013***	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
NCF	0.099	-0.192	0.366***	0.250*	-0.469	0.865***
	(0.061)	(0.148)	(0.141)	(0.144)	(0.322)	(0.289)
Observations	43942	43942	43942	43942	43942	43942
Adj R^2	0.310	0.471	0.528	0.372	0.495	0.538
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table explores the effect of mania on investors' herding levels. Panel A uses Guba posts to measure investors' sentiment. Panel B uses financial news data to measure investors' mania.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 10: Robustness Tests – Adjusting the Herding Measurement Caliber

Panel A: Incorporating unadjusted LSV herding measurement (HM)			
	Full	Institution	Retail
$\|NCF\|$	0.075^{***}	0.186^{**}	0.341^{***}
	(0.024)	(0.074)	(0.064)
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Observations	48562	48562	48562
Adj R^2	0.155	0.286	0.563

Panel B: Adjusting the heterogenous investors' herding measurement caliber				
	AdjHM		*AHM*	
	Institution	Retail	Institution	Retail
NCF	0.232^{***}	0.301^{***}	0.502^{***}	0.743^{***}
	(0.028)	(0.064)	(0.062)	(0.130)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	48562	48562	48562	48562
Adj R^2	0.426	0.590	0.426	0.590

Note: In this table, we draw on a battery of herding measures and repeat our estimation from Eq.(25). Panel A displays the result of incorporating the unadjusted LSV herding measurement as the dependent variable. Panel B adjusts the heterogeneous investors' herding measurement caliber.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 11: Robustness Tests – Altering the independent variables

	AdjHM		*AHM*	
	(1)	(2)	(3)	(4)
Panel A: Incorporating lagged terms				
NCF	0.445***		1.101***	
	(0.022)		(0.053)	
L.AdjHM / L.AHM	0.155***		0.158***	
	(0.016)		(0.013)	
L.NCF		0.057***		0.147***
		(0.020)		(0.047)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	48307	48291	48307	48291
Adj R^2	0.221	0.197	0.275	0.252
Panel B: Incorporating the absolute value of the NCF (AVNCF)				
AVNCF	0.004***		0.009***	
	(0.001)		(0.001)	
log(1+AVNCF)		0.005***		0.011***
		(0.001)		(0.001)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	48622	48622	48622	48622
Adj R^2	0.302	0.302	0.365	0.365

Note: Robustness tests in this table aim to perform regression with different independent variables. Columns (1) and (3) in Panel A include NCF and lagged herding measurements in the regression. Columns (2) and (4) in Panel A include the lagged NCF term. Panel B displays the results after incorporating the absolute value of northbound capital flow in the regression.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 12: Robustness Tests – Fama-Macbeth Regressions

	AdjHM		*AHM*	
	(1)	(2)	(3)	(4)
NCF	0.747***	0.235***	1.875***	0.558***
	(0.067)	(0.049)	(0.170)	(0.120)
Return		0.036***		0.092***
		(0.004)		(0.010)
Return_lag		0.007***		0.018***
		(0.002)		(0.004)
$\|Return\|$		-0.003		-0.014
		(0.005)		(0.011)
$\|Return\_lag\|$		0.008**		0.018**
		(0.003)		(0.008)
Size		-0.001***		-0.003***
		(0.000)		(0.001)
Std		-0.244**		-0.530**
		(0.095)		(0.239)
Turnover		0.001		0.002*
		(0.000)		(0.001)
BM		-0.005***		-0.014***
		(0.001)		(0.003)
Cap		0.006***		0.015***
		(0.000)		(0.001)
Vol		-0.001***		-0.002***
		(0.000)		(0.001)
Constant	0.001***	-0.028***	-0.054***	-0.132***
	(0.000)	(0.003)	(0.003)	(0.006)
Firm FE	No	No	No	No
Quarter FE	No	No	No	No
Observations	21	21	21	21
Avg R^2	0.025	0.261	0.026	0.284

Note: In this table, we additionally perform a Fama and MacBeth (1973) regression of investor herding on northbound capital flow. Columns (1) and (3) display the results with no control variables, while columns (2) and (4) include control variables.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 13: The Perception of Smart Money: Shareholding Ratio and Market Capitalization

	AdjHM			*AHM*		
	Full	Institution	Retail	Full	Institution	Retail
Panel A: The Northbound Capital holding ratio						
Holding $\times$*NCF*	-3.089***	-3.272***	-2.203***	-7.983***	-7.181***	-5.853***
	(0.350)	(1.128)	(0.802)	(0.844)	(2.363)	(1.645)
NCF	0.376***	0.266***	0.461***	0.941***	0.588***	1.103***
	(0.024)	(0.060)	(0.058)	(0.056)	(0.130)	(0.119)
Holding	0.002	0.136***	-0.133***	-0.013	0.245***	-0.319***
	(0.009)	(0.041)	(0.040)	(0.021)	(0.086)	(0.081)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48560	48560	48560	48560	48560	48560
Adj R^2	0.305	0.468	0.533	0.368	0.493	0.545
Panel B: The stock's market value						
Cap $\times$*NCF*	-0.077***	-0.094**	-0.140**	-0.283***	-0.276***	-0.375***
	(0.019)	(0.042)	(0.062)	(0.042)	(0.090)	(0.125)
NCF	1.016***	1.155***	1.677***	3.419***	3.207***	4.379***
	(0.186)	(0.420)	(0.598)	(0.423)	(0.911)	(1.209)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	48562	48562	48562	48562	48562	48562
Adj R^2	0.304	0.467	0.531	0.368	0.493	0.544

Note: This table reports the impact of investors' perception on herding behavior. In Panel A, we explore how the sensitivity to northbound capital flow varies across different stocks. In Panel B, we incorporate the interaction term of *Cap* and *NCF* to reflect the marginal effects.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 14: The Perception of Smart Money: Allocation- or Trading-oriented NCF

	AdjHM			*AHM*		
	(1)	(2)	(3)	(4)	(5)	(6)
NCF_trade	0.014***		0.014***	0.038***		0.038***
	(0.003)		(0.003)	(0.008)		(0.008)
NCF_allocation		-0.000	-0.001		-0.002	-0.003
		(0.002)	(0.002)		(0.005)	(0.005)
Return	0.034***	0.034***	0.034***	0.088***	0.088***	0.088***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Return_lag	0.010***	0.010***	0.010***	0.025***	0.025***	0.025***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
|*Return*|	-0.011***	-0.011***	-0.011***	-0.034***	-0.034***	-0.034***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
|*Return_lag*|	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Size	-0.003***	-0.003***	-0.003***	-0.005**	-0.005**	-0.005**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Std	-0.322***	-0.321***	-0.322***	-0.837***	-0.834***	-0.837***
	(0.031)	(0.031)	(0.031)	(0.070)	(0.070)	(0.070)
Turnover	-0.001***	-0.001***	-0.001***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
BM	-0.010***	-0.010***	-0.010***	-0.030***	-0.029***	-0.030***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Cap	0.010***	0.010***	0.010***	0.023***	0.023***	0.023***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Vol	0.000	0.000	0.000	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Constant	-0.056***	-0.056***	-0.056***	-0.195***	-0.195***	-0.195***
	(0.007)	(0.007)	(0.007)	(0.017)	(0.017)	(0.017)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37945	37945	37945	37945	37945	37945
Adj R^2	0.341	0.341	0.341	0.392	0.392	0.392

Note: In this table, we explore investors' sensitivity to the two different types of northbound capital: allocation-oriented and trade-oriented NCF.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 15: Regression Discontinuity Analysis

	AdjHM		*AHM*	
	(1)	(2)	(3)	(4)
Panel A: Sample Bandwidth = 15				
Rank less than 100.5	0.018**	0.015**	0.038**	0.033**
	(0.008)	(0.007)	(0.017)	(0.016)
Polynomial order	2	3	2	3
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	384	384	384	384
Adj R^2	0.305	0.299	0.317	0.310
Panel B: Sample Bandwidth = 25				
Rank less than 100.5	0. 010**	0.010*	0.021**	0.021*
	(0.005)	(0.006)	(0.010)	(0.012)
Polynomial order	2	3	2	3
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	812	812	812	812
Adj R^2	0.347	0.345	0.355	0.354

Note: In this table, we perform a sharp regression discontinuity design to further establish the causal impact of the NCF on investors' herding. Two regressions are performed with different sample bandwidths. Panel A shows the result with a sample bandwidth of 15. Panel B displays the results with a sample bandwidth of 25.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Table 16: Real Smart Money or Nominal Title

	(1)	(2)	(3)	(4)
	ROA_{t+1}	ROA_{t+2}	$Return_{t+1}$	$Return_{t+2}$
NCF	0.133^{**}	0.101^{*}	1.078^{***}	0.068
	(0.058)	(0.052)	(0.240)	(0.260)
Return	0.021^{***}	0.026^{***}	0.041^{***}	0.090^{***}
	(0.002)	(0.003)	(0.007)	(0.007)
Return_lag	0.021^{***}	0.019^{***}	0.080^{***}	0.051^{***}
	(0.003)	(0.002)	(0.006)	(0.007)
$\|Return\|$	-0.012^{***}	-0.010^{***}	0.016^{*}	0.003
	(0.004)	(0.003)	(0.009)	(0.009)
$\|Return\_lag\|$	-0.008^{***}	0.002	0.016^{**}	0.002
	(0.003)	(0.003)	(0.008)	(0.009)
Size	-0.018^{***}	-0.019^{***}	0.051^{***}	0.031^{***}
	(0.003)	(0.003)	(0.006)	(0.006)
Std	-0.297^{***}	-0.205^{**}	0.438^{**}	-0.407^{*}
	(0.075)	(0.083)	(0.220)	(0.218)
Turnover	0.003^{**}	0.002^{**}	-0.011^{***}	-0.007^{***}
	(0.001)	(0.001)	(0.002)	(0.002)
BM	-0.015^{***}	-0.016^{***}	-0.017^{*}	0.019^{**}
	(0.005)	(0.005)	(0.010)	(0.009)
Cap	0.036^{***}	0.023^{***}	-0.172^{***}	-0.147^{***}
	(0.003)	(0.003)	(0.006)	(0.005)
Vol	-0.004^{***}	-0.004^{***}	-0.016^{***}	0.002
	(0.001)	(0.001)	(0.003)	(0.003)
Constant	-0.080^{**}	0.038	1.221^{***}	1.066^{***}
	(0.031)	(0.031)	(0.061)	(0.058)
Firm FE	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes
Observations	46055	43735	46055	43735
Adj R^2	0.440	0.423	0.316	0.298

Note: This table examines the stock selection ability of northbound capital. Columns (1) and (2) incorporate return on assets as the dependent variable, while columns (3) and (4) incorporate stock return.

仅用于2022年成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

Acknowledgement

1. Topic Origin

With my father being an experienced retail investor and my brother studying quantitative finance in Hong Kong, I have long been surrounded by financial news discussions. With them sharing their investment strategies, I soon realized that they, as well as other investors, sometimes choose to follow simple market signals as they lack the necessary skills to perform in-depth evaluation and analysis. I was drawn to the idea of what specific market signals can trigger herding and the inherent mechanism behind such irrational behavior.

2. Research Background

China, with an emerging capital market, provides a unique environment for our analysis. The establishment of Shanghai-Hong Kong Stock Connect in 2014 has increased the role of overseas capital in China's equity market. It is a common consensus that the foreign capital, which enters China's capital market through this cross-boundary investment channel, possesses private information about China and is therefore reported as "smart money". With massive media attention, this overseas capital can be seen as a market signal that causes herding. Thus, this ideal situation offers a great opportunity to study investors' herding behavior around smart money.

3. Research Process

I started the investigative process by doing extensive reading on literature relating to investors' herding behavior and smart money. With my profound interest in game theory, combined with Professor Chen's expert advice in the field of model construction, we realized that evolutionary game theory might provide an ideal basis for our theoretical model. To gain a better understanding of how evolutionary game theory can be applied in this scenario, I reviewed literatures that also applied evolutionary game theory as their core theoretical model in order to gain a better understanding of its application. Following the standard approach, I managed to finish the model construction in the early stage of the research.

Next, to empirically show the effect of smart money on investors' herding behavior, I extracted related data from WIND and iFind, eliminating firms that experienced special

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards

treatment. To match our theoretical model, we further our empirical result by analyzing the effect of sentiment, proxy by East Money Guba post and financial news data, on herding intensity. The data are collected from the Chinese Research Data Service Platform (CNRDS). Throughout the process, Professor Chen taught me the necessary skills in data extraction and data analysis using python and deepen my ability to use Stata to perform regressions.

The robustness test is a result of collective brainstorming. Both Professor Chen and I came up with a few concerns that might undermine the credibility of our results. I proposed incorporating the traditional LSV herding measurement and adjusting the heterogenous investors' herding caliber. On the other hand, Professor Chen prompted me to learn the regression discontinuity design and see whether it is suitable for our research. Looking through statistics textbooks to acquire the concept, I find it a perfect match for our study. Since attention to conspicuous market signals is the main focus of our investigation, a regression discontinuity design that examines the threshold between stocks at the end of the stock list (ranked by northbound capital flow on a quarterly basis) and those almost on the list is a suitable robustness test.

All in all, I would like to again express my gratitude for Professor Chen's support throughout the long and challenging research process.

仅用于2022丘成桐中学科学奖公示
2022 S.-T. Yau High School Science Awards