Which investors make excess return comovement?

Jie Li, Yongjie Zhang, Xu Feng, Yahui An

PII: S1042-4431(18)30281-6
DOI: https://doi.org/10.1016/j.intfin.2019.01.004
Reference: INTFIN 1102

To appear in: *Journal of International Financial Markets, Institutions & Money*

Received Date: 8 July 2018
Accepted Date: 29 January 2019

Which investors make excess return comovement?

Jie Li[a,b], Yongjie Zhang[a,b], Xu Feng[a,b,1], Yahui An[a,b]

[a] College of Management and Economics, Tianjin University, Tianjin 300072, China

[b] China Center for Social Computing and Analytics, Tianjin 300072, China

Abstract

In this study, we examine whether the trading behavior of different types of investors (individual investors, funds, qualified foreign institutional investors, insurance, and state-owned companies) generate different effects on excess return comovement. The empirical results show that only funds' behavior forecasts more variation in excess return correlation, controlling for the other three types of factors (macroeconomic, stock characteristics, and information diffusion). Moreover, we provide adequate evidence that the comovement arising from funds' trading behavior is due to their specific purchase and redemption mechanism, rather than holding more or trading more. Finally, an additional test is conducted to confirm that purchases cause more comovement than redemptions do, which might be related to larger amount of purchases.

Keywords: Comovement; Investor behavior; Funds; Purchase; Redemption

Acknowledgments: This article is funded by the National Natural Science Foundation of China (71790594, 71771170, 71871157) and the Ministry of Education Fund on Humanities and Social Science (14YJC790029)

[1]Corresponding author at: College of Management and Economics, Tianjin University, Tianjin 300072, China.
E-mail address: fengxu@tju.edu.cn (X. Feng).

Which kind of investor causes comovement?

Abstract

In this study, we examine whether the investment flows of different types of investors (individual investors, funds, qualified foreign institutional investors, insurance, and state-owned enterprises) generate different effects on return comovement. The empirical results show that only the flows of funds forecast the variation in comovements, controlling for the three other types of factors (macroeconomic, stock characteristics, and information diffusion). Moreover, we provide adequate evidence to show that the comovement arising from fund flows is due to the purchase and redemption mechanism. Finally, an additional test is conducted to confirm that fund purchases cause more comovement than redemptions, which might be related to more herding of funds on the buy side.

Keywords: Comovement; Investment flows; Purchase; Redemption; Herding

1. Introduction

Comovement is a universal feature of the price formation process in capital markets and has direct impacts on investment allocation and asset pricing (Barberis, Shleifer and Wurgler, 2005). Despite its importance, comovement remains a relatively unexplored area of financial economics. For example, it is the consensus that comovement stems from the homodromous investment flow pressure on the prices of multiple stocks at the same time. However, which kind of investor buys/sells multiple stocks simultaneously? What are the determinants of the simultaneous trading of investors? These questions have not yet been answered.

One strand of the prior literature directly tests the factors that affect stock price comovement. The investors' trading behavior in these papers is ignored. These studies on the explanations for comovement can be roughly divided into the following three categories: (1) macroeconomic factors, (2) stock characteristic factors, (3) information diffusion factors.

Some prior studies suggest that stock prices move together in response to common changes in macroeconomic factors, such as GDP, business cycles, and interest rates. For example, when the economic situation improves, investors allocate more wealth into the stock market. The influx of cash will make the stock market rise as a whole, leading to the comovement of stock returns. Shiller (1989) and Pindyck and Rotemberg (1990) find that the comovement of commodities is partially explained by macroeconomic fundamentals. Batten et al. (2010) explain that macroeconomic factors jointly influence the volatility processes of the four precious metal price series. Pindyck and Rotemberg (1993) point out that comovements of individual stock prices can be justified by economic fundamentals. Ai et al. (2006) and Lescaroux (2009) state that an obvious reason for prices to move together is that they respond to common macroeconomic shocks.

In addition to macroeconomic factors, stock-level factors, such as those factors based on market capitalization (e.g., small-cap stocks), industry (e.g., oil industry stocks), and liquidity (e.g., low liquidity stocks), also affect comovement. Comovement often occurs in these types of stocks because investors tend to allocate funds at the level of categories rather than at the individual asset level. Greenwood (2008) points out that comovement is driven by correlated shocks to investor demand for a particular set of securities. Investors are always good at grouping assets into categories, such as small-cap stocks and same-industry stocks. Barberis and Shleifer (2003) find that some investors categorize risky assets into different styles. Assets in the same style comove too much, and assets in different styles comove too little.

A third category of comovement is affected by information diffusion. Due to some market friction, information is incorporated more quickly into the price of some stocks than others. For example, some stocks may have a lower trading cost or be held by investors with faster access to information. When macroeconomic news is released, the stocks with fast information dissemination respond immediately, resulting in comovement among these stocks. The remaining stocks also move up together but only after some delay. This body of research usually predicts a rise in the comovement of the stocks included in a major index. The index constituents are expected to have lower trading costs and to be more liquid. Their prices reflect aggregate information more quickly than other stocks excluded from the index. A large body of literature examines whether index inclusion has a contemporaneous effect on price levels. Harris and Gurel (1986), Shleifer (1986), Lynch and Mendenhall (1997), and Wurgler and Zhuravskaya (2002) find strong price effects for S&P 500 inclusion, while Greenwood (2008), Claessens and Yafeh (2012), and Kougoulis and Coakley (2014) find similar effects for the Nikkei 225 indexes and FTSE 100 Index.

The abovementioned factors can be viewed as rational factors. In this study, we show that stock price-return comovements may also be driven by irrational factors. Investors' herding behavior is the most likely irrational factor to explain these comovements. As described by Bikhchandani and Sharma (2001), herding is the excessive and irrational tendency of traders to learn from and imitate one another during a transaction, which leads them to buy and sell the same stock for a certain period of time[2]. If the imitative trading behavior only occurs for one stock, the simultaneous trading caused by the herding will only lead to an increase or fall in one stock price. However, if the irrational imitation transaction occurs for multiple stocks, simultaneous trading will result in the comovement of the stocks. This herding behavior has been extensively researched during the past two decades with evidence from a large array of long-estimated developed and emerging markets, including those of the U.S. (Sias, 2004; Li & Yung, 2004; Liao et al., 2011; Deng et al., 2018), China (Demirer and Kutan, 2006; Chiang et al., 2010) and the global market (Chiang and Zheng, 2010; Gebka and Wohar, 2013; Economou et al., 2015). Additionally, many kinds of investors appear to engage in herd buying or selling behavior in the stock market. Deng et al. (2018) investigate the impact of mutual fund herding behaviors on stock price crashes. Chang et al. (2010) examine herding behavior by qualified foreign institutional investors (QFIIs) in an emerging equity market. Iihara et al. (2010) describe individual investors' herding behavior on the Tokyo stock exchange, but its impact on stock prices is less than that of institutional and foreign investors. Although there has been little literature associating irrational herding with comovement to date, irrational herding has a natural mechanism that leads cash flows into or out of multiple stocks to move consistently in the same direction. Thus, we add irrational herding as another factor that may affect stock price comovement.

There is also another strand of literature that tries to link investors' trading behavior with comovement and ignores the reason why investors trade multiple stocks simultaneously. For example, Anton and Polk (2014) argue that the simultaneous trading of mutual funds induces the comovement of stocks. The studies of Coval and Stafford (2007), Chen et al. (2008), Frazzini and Lamont (2008), Greenwood and Thesmar (2011), Lou (2012), and Vayanos and Woolley (2013) also provide evidence that mutual funds' trading generates comovement. In addition, Jotikasthira et al. (2012) and Bartram et al. (2015) argue that global fund trading behavior affects variations in international stock return comovement.

There is a clear gap between the two abovementioned strands of literature. Studies on the factors that led to comovement did not indicate which kind of investors the factors affected. Similarly, studies on investor trading behavior have not clarified the determinants that affect investor behavior. Therefore, it is still not clear that the behavior of investors simultaneously buying (or selling) multiple stocks is truly affected by the abovementioned determinants. Considering there are various kinds of investors in the financial market, it is still not clear that any

[2] Bikhchandani and Sharma (2001) offer a seminal review of theoretical aspects of herding, including differentiation between rational and irrational herding. A rational herding is an investor who invests based on fundamental factors or investment category; an irrational herding is a random catharsis that is influenced by external factors. The mechanism of rational herding influence on the comovement is similar to that of the above three types of rational factors. Here, we are more focused on the impact of irrational herd behavior on the comovement.

investors except mutual funds trade on multiple stocks simultaneously and cause the comovement.

Our research aims to contribute to these unsolved issues. We use stock samples from the Chinese stock market. We consider all types of investors, including funds, individual investors, insurance companies, Qualified Foreign Institutional Investors (QFII), and state-owned enterprises (SOE), and study the impact of their investment flows on comovement. Our results show that not all investors' investment flows cause comovement. Although the positions of fund holdings and transactions are both in the median of all of the investors' positions, only the fund simultaneous trading behavior has been found to have an impact on comovement.

We undertake an in-depth examination of why only the funds' trading induces comovement. The existing explanations seem to affect all kinds of investors rather than affecting only funds. Thus, we consider the impact of the fund-specific mechanism to explain the fund-comovement relationship, which is the purchase and redemption pressures of the funds. In the face of investors' subscriptions and redemptions, funds have a tendency to trade stocks at the same time. Specifically, on the one hand, in the face of investors' subscriptions, fund managers typically buy portfolios rather than a single stock to diversify investment risks. On the other hand, in the face of investors' redemptions, fund managers often choose to sell multiple shares in a portfolio to reduce the price shock of selling any single stock. Therefore, we parallel the purchase and redemption mechanism with several other types of factors to jointly explain funds' simultaneously buying and selling of multiple stocks. Even with the addition of the other three types of control variables, the subscription and redemption mechanism can significantly explain the funds' simultaneous trading. More importantly, the mediation analysis shows that funds' trading serves as a significant channel variable through which the purchase and redemption mechanism affects comovement.

In addition, the studies of Ivkovich and Weisbenner (2006) and Lu et al. (2007) have shown that, owing to financial pressure, investors usually apply for fund purchase or redemption in the month before earning announcements are published. Therefore, we test whether comovement for the last month of each quarter (near the earnings announcement time) is significantly different from comovement for the first month of each quarter (far from the earnings announcement time). We find that significant changes in comovement occurred only in funds' holdings, but not in the holdings of other types of investors. This evidence fully shows that the mutual fund-specific subscription and redemption mechanism led to their buying and selling multiple stocks, which in turn led to the comovement of the stocks that funds held.

Finally, we further distinguish the impact of the purchase mechanism and the redemption mechanism on funds' simultaneous buying and selling behavior. We find that the influence of the purchase mechanism on comovement is greater, which may be related to funds herding more strongly on the buy side. Herding occurs when a group of investors buy or sell the same security based on the same, or correlated, information signals over some period of time. When fund investors move in the market as a herd, the stocks they trade will show greater relevance. Further, the buying-herd behavior means that fund managers are more likely to generate imitative investment behaviors when building portfolios. In the face of investor purchases, buying-herding of the fund will lead to the rise of multiple stocks, which in turn triggers the comovement of stock return. Such empirical results are consistent with those of Nofsinger and Sias (1999), Zhou and Lai(2009) and Dasqupta et al. (2011) who also discovered the asymmetry of the fund herding effect. The works of Lakshman et al. (2013) and Chang and Lin (2015) also find institutional herd-buying is more active in emerging markets.

This study contributes to the literature on comovement in the following respects. First, from the perspective of investment flows, we distinguish the impact of the investment flows generated by different types of investors on stock comovement. We find that only fund inflow leads to comovement. In contrast, the investment flows of other types of investors (e.g., insurance, QFII, SOE, and individuals) did not affect comovement. Second, from the perspective of the determinants of comovement, we find that the purchase and redemption mechanism is an important factor in explaining comovement when controlling for the three types of traditional factors. To the best of our knowledge, our research is the first to try to investigate the potential influence from the purchase and redemption mechanism on comovement. We find asymmetric impacts of purchase and redemption of funds on comovement. In conclusion, our research broadens the understanding of the determinants of stock comovement.

The rest of the paper is organized as follows. Section 2 presents the data description and methodology. Section 3 predicts the return comovement with investment flows of different types of investors. Section 4 explains why comovement through investment flows of funds matters. Section 5 includes an additional test. Section 6 is the discussion of how this paper is related with previous literature. Section 7 concludes this paper.

2. Data and methodology

2.1 Data and sample

The samples for this paper come from the Chinese stock market. The stock returns, trading volume, and other relevant market and accounting data of this paper come from the Wind database. All companies listed on the Shanghai Stock Exchange and Shenzhen Stock Exchange are used as our stock samples. To avoid the Equity Split Reform potentially influencing this study, we selected a sample from January 1, 2007,[3] to December 31, 2016. Stocks under Special Treatment or delisted from the market were eliminated. When the transaction duration is overly short, the information cannot be fully reflected in the price. To avoid this situation, we deleted listed companies that had a halt in trading of more than 5 days in a quarter.

Data on shareholdings come from the CSMAR database. We defined shareholders as the top 10 investors included in listed companies' quarterly reports. Moreover, we manually unified the names of the shareholders, as a shareholder might have different names (because of abbreviations or misspelling) in the reports released by different companies. Fund purchase and redemption data also comes from the Wind database. Because the number of closed-end funds is very small, and their purchase and redemption mechanisms are different from those of open-end funds, closed-end funds were removed from the sample.

2.2 Measuring comovement

We measure the variation in comovement based on a five-factor model. Fama and French (2015) introduce a five-factor asset pricing model that adds the profitability and investment factors to the three-factor model of Fama and French (1993). The five-factor regression is

$$r_{it} - r_{ft} = a_i + b_i Mkt_t + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it} \qquad (1)$$

where $r_{it} - r_{ft}$ is the portfolio i's return in excess of risk-free rate r_{ft} for month t. Mkt_t is

[3] The Equity Split Reform was implemented on May 9, 2005. By the end of 2006, a total of 1,301 listed companies had completed reforms in Shanghai and Shenzhen, accounting for 97% of the total listed companies, and only 40 listed companies had not entered the reform process.

the value-weight market portfolio return in excess of risk-free rate. SMB_t, HML_t, RMW_t and CMA_t are respectively the size, value, profitability and investment factors. e_{it} is a zero-mean residual. We obtain the five-factor data of the Chinese stock market from the CSMAR database. Guo et al. (2017) validate the applicability of a five-factor model in the Chinese market and find that the profitability factor in particular significantly improves the description of average return.

If the exposures to the five factors capture all variation in expected returns, the residual e_{it} in (1) is the zero-mean for all securities and portfolio i. In contrast, in addition to the five factors, if there are other factors that lead to returns, they should be reflected in the residual e_i. Based on the method used in the study by Kallberg and Pasquariello (2008), we estimated the within-quarter realized correlation $\rho_{ij,t}$ of two stocks' daily five-factor abnormal returns. That is,

$$\rho_{ij,t} = correl(e_{i,t}, e_{j,t}) \tag{2}$$

where $e_{i,t}$ is the residual of the five-factor model on stock i for period t and similarly for stock j.

2.3 Measuring investor simultaneous trading on stocks

At each quarter end, we follow the algorithm described in Anton and Polk (2014) to measure investors' simultaneous trading (investment flows) as the change in the total value of stock held by F common shareholders of the two stocks, scaled by the total market capitalization of the two stocks, and we label this variable $F_{ij,t}$. Thus,

$$F_{ij,t} = \frac{\sum_{f=1}^{F}(\Delta S_{i,t}^{f} P_{i,t} + \Delta S_{j,t}^{f} P_{j,t})}{S_{i,t}P_{i,t} + S_{j,t}P_{j,t}} \tag{3}$$

$$\Delta S_{i,t}^{f} = S_{i,t}^{f} - S_{i,t-1}^{f} \tag{4}$$

where $S_{i,t}^{f}$ is the number of shares held by shareholder f at time t trading at price $P_{i,t}$ with total shares outstanding of $S_{i,t}$ and similarly for stock j. $S_{i,t-1}^{f}$ is the number of shares held by shareholder f at time $t-1$. In particular, based on Anton and Polk (2014), $S_{i,t-1}^{f}$ is assumed to be zero when shareholders exist in the top 10 list at this quarter end but outside of the list at the next quarter end. We define common shareholders as those shareholders who held both stocks i and j in their portfolios at the end of quarter t. For each cross section, we calculate the normalized (to have zero mean and unit standard deviation) rank-transformed $F_{ij,t}$, which we denote as $F_{ij,t}^{*}$.

To examine the impact of different types of investor behavior on comovement, we classify investors into five categories: individual, fund, QFII, insurance, and SOE. The natural person shareholders are included in the individual type. We remove social security funds, index funds, and international funds from the fund sample by applying standard screening criteria used in the literature (Cremers and Petajisto, 2009; Anton and Polk, 2014). Insurance mainly includes property insurance companies, life insurance companies, reinsurance companies, insurance asset management companies, and the National Social Security Fund. QFII refers to Chinese offshore fund management institutions, insurance companies, securities companies, and other asset management institutions approved by the China Securities Regulatory Commission and obtain amount approval from the State Administration of Foreign Exchange. SOEs refer to the enterprises controlled by the central government or local governments. There are other pairs of stocks formed by other undifferentiated types of shareholders. These stocks account for less than 10% of the total and are not considered in this article.

2.4 Modeling variation in comovement

The data used in this paper are panel data containing information on the two dimensions of cross section and time series. The advantages of panel data are the control of individual heterogeneity, increased freedom, and reduced multicollinearity between variables. We estimate panel regression explaining the within-quarter realized correlation ($\rho_{ij,t}$) of each stock pair's daily five-factor abnormal returns with $F^*_{ij,t}$ and a host of pair characteristics that we used as controls:

$$\rho_{ij,t} = a + b_1 * F^*_{ij,t} + \sum_{k=1}^{n} b_k * CONTROL_{ij,k} + \varepsilon_{ij,t} \tag{5}$$

where t = 1,2,…,T, representing T quarters (actually 40 quarters from 2007 to 2016). ij = 1,2,…,N, representing the stock pair formed by stock i and stock j in cross section. The model is estimated with an unbalanced panel data model with fixed time effect, and the standard errors are clustered in firms. We report t-statistics based on the standard errors of the coefficients after adjusting for serial correlation using the Newey and West (1987) procedure.

2.5 Controls

Based on the prior literature, comovement can be attributed to three types of factors: (1) macroeconomic factors, (2) stock characteristic factors, and (3) information diffusion factors. Our goal is to determine whether investor behavior causes stocks to become more correlated. However, in order to make the results credible, we must control the impact of the other three types of factors on comovement.

First, the economic growth rate is a dynamic indicator that reflects the degree of change in the level of economic development in a given period. The economic growth rate can be measured by the GDP growth rate ($RGDP$), which is the GDP of the following year minus the GDP of the previous year divided by the GDP of the previous year. $RGDP$ can be observed as a proxy variable of macroeconomic factors. Consumer price index (CPI) is an important macroeconomic indicator that reflects changes in the level of consumer goods and services related to the lives of residents. The level of CPI directly affects the introduction of the country's macroeconomic control measures and indirectly affects the changes in capital markets (such as stock markets, futures markets, capital markets, and financial markets). CPI can also be observed as a proxy variable of macroeconomic factors.

Second, based on Barberis and Shleifer (2003), comovement is driven by correlated shocks to investor demand for a particular set of securities. Investors might group stocks according to their industry. Therefore, industry may be an important factor affecting stock price comovement. We adopt China Securities Regulatory Commission (CSRC) industry classification standards. A dummy variable named $Industry$ denotes whether the two stocks are in the same industry. To measure the difference in companies' sizes, we sort the stocks in descending order based on total market value and divide them into 5 groups. A dummy variable named $Size$ denotes whether the two stocks are in the same group. Based on Chan and Faff (2003) and Lesmond (2005), turnover is used as the main proxy for liquidity. Similarly, we sort the stocks in descending order based on turnover and divide them into 5 groups, with a dummy variable named $Turnover$ denoting whether the two stocks are in the same group.

Finally, the information diffusion factor posits that information is incorporated more rapidly in some stocks included in a major index than in other stocks not included in the index. A dummy variable named $Index$ denotes whether the two stocks belong to the CSI 300 Index. We include all the variables as controls in many of our specifications.

3 Results

3.1 Summary statistics

This paper studies the effect of investors' simultaneous trading on the comovement of multiple stocks. Table I presents the summary statistics for the number of stocks, pairs of stocks, and number of shareholders. Among all types of shareholders, funds have the most pairs, accounting for approximately 71.02% of all the common holdings. However, only 213 fund shareholders on average contribute those pairs of stocks, comprising 30.17% of all common shareholders. Correspondingly, individuals prefer to concentrate shareholdings in a few stocks, which is indicated by the fact that the highest number of shareholders (282, individual type) form only 394 pairs of stocks.

[Insert Table I here]

We report snapshots of the distribution of investment flows ($F_{ij,t}$) in Panel A of Table II. The proportion of individual trading is the lowest; on the contrary, the trading of SOEs is relatively higher than that of other types of shareholders. The proportion of fund trading is in the middle of all the investor transactions.

[Insert Table II here]

The results presented in Panel B of Table II show that the trading and holding of funds are also located in the median volume. Specifically, the average holdings (S_t) of funds are more than those of individuals, but less than other types of investors; similar results are obtained for the amounts of holdings (P_t*S_t) in the third column and the trading volume (ΔS_t) in the fourth column. These results suggest fund investors do not trade more than other investors. However, our next result shows their investment flows have the most influence on the comovement of stocks.

3.2 Modeling

Our goal is to determine which type of investor behavior contributes to return comovement. We estimate a version of equation (5) with different types of investment flow F and other control variables. As the rational potential reasons for comovements (FF) risk factors in equation 1 and other controls in equation (5) are removed/controlled for, the F largely captures the irrational comovements between stock prices. If any investors' irrational behavior affects comovement, the coefficient of F of that kind of investor should be positively significant in equation (5).

Table III reports results from different types of investor behavior causing cross-sectional variation in five-factor residual correlation. We distinguish different types of investors such as fund, insurance, QFII, SOE and individual. We report the results of regressions incorporating our controls based on the abovementioned three types of factors, including $RGDP$, CPI, $Industry$, $Size$, $Turnover$ and $Index$. Though these controls are economically and statistically significant in describing return comovement, the fund investment flow (F^*) remains a highly significant coefficient of 0.68 (t-statistic of 19.18). However, the investment flows of other types of investors, such as insurance, QFII, SOEs, and individuals, are not significant, meaning that their behavior is a limited predictor of comovement.

[Insert Table III here]

The empirical results in Coval and Stafford (2007), Lou (2012), and Anton and Polk (2014) provide support for the view that concentrated mutual fund sales forced by capital flows exert significant price pressure in equity markets, often resulting in comovement of stock returns. Our test is consistent with the studies above. Moreover, our results extend their conclusions that only the investment flows generated by funds' simultaneous trading behavior, not other types, helps predict variation in comovement.

In Table III, Panel B, we show the same set of regressions as in Panel A but with the addition of two subsamples (2007-2011, 2012-2016)[4]. We show that the fund behavior effect has increased over time: the coefficient on F^{*}moving from a relatively low 0.52 in the first subsample to a relatively high estimate of 0.76 in the second subsample. The results indicate that the irrational behavior of funds has stronger effects on the comovement, which may due to the rapid growth of the fund industry in the Chinese market since 2012. For the sake of brevity, we only report the coefficients of $Constant$ and F^{*}. The results of other control variables have no significant change, which imply that the impacts of other rational factors on the comovement have no significant increase in the period from 2012-2016.

4 Explanation of results

4.1 Purchase and redemption mechanism

The rest of the analysis focuses on exploring why connecting stocks through fund simultaneous trading matters. Although the holdings and trading volume of different types of investors (excluding individual investors) is roughly the same, the influence of their simultaneous trading on comovement varies greatly. Since the existing explanations seem to affect all kinds of investors rather than affecting only funds, we consider the impact of the fund-specific mechanism to explain the fund-comovement relationship. Our results show sufficient support for simultaneous trading resulting from the fund-specific purchase and redemption mechanism.

Studies by Warther (1995), Fortune (1998), Edelen and Warner (2001), Goetzmann and Massa (2003), Cao et al. (2008), and Boyer and Zheng (2009) show that investor flow into (purchase) and out of mutual funds (redemption) leads to actual trades by mutual fund managers. Specifically, mutual funds experiencing extreme inflows tend to increase their existing positions to cover purchases rather than maintain a high cash level for the purpose of diversification of asset allocation or profitability. Similarly, mutual funds experiencing significant outflows have no choice but to sell some of their holdings to cover redemptions, unless they have excess cash. However, few funds maintain significant cash balances. Therefore, the sale of some existing holdings is the only option. Overall, purchases and redemptions tend to initiate funds' changes in positions, which cause stocks to become more correlated.

4.2 Mediating effect test

To test our null hypothesis that the purchase and redemption mechanism forecasts return comovement through the investment flows of funds, one possibility is to examine whether the investment flows of funds is an important channel for this effect using mediation analysis (e.g., Hammersley, 2006; Lang et al., 2012). Consider the influence of the independent variable X on the dependent variable Y. If X influences Y through M, then M is called a mediator variable. The

[4] There are two reasons for using 2012 as a sample interval split point: First, to keep the time equally divided. Second, and more importantly, 2012 was known as the first year of full circulation of the Shanghai and Shenzhen stock markets. According to the statistics in the Wind database, as of the end of 2011, the Shanghai and Shenzhen stock markets basically realized full circulation.

goal of mediation analysis is to assess whether one variable (e.g., investment flows of funds) serves as a mediator variable through which another variable (e.g., purchase and redemption mechanism) affects a particular dependent variable (e.g., comovement).

The mediation analysis model can be expressed as the following regression form:

$$\rho = a + c_1 * pr^* + \sum_{k=1}^{n} b_k * CONTROL_{ij,k} + \varepsilon_{ij,t} \tag{6}$$

$$F_{ij,t}^* = a + c_2 * pr^* + \sum_{k=1}^{n} b_k * CONTROL_{ij,k} + \varepsilon_{ij,t} \tag{7}$$

$$\rho = a + c_3 * pr^* + c_4 * F_{ij,t}^* + \sum_{k=1}^{n} b_k * CONTROL_{ij,k} + \varepsilon_{ij,t} \tag{8}$$

where pr is calculated by the amounts of purchases plus redemptions divided by the total issued amounts. pr^* is rank-transformed and normalized to have unit standard deviation. The control variables are defined in Table III, including $RGDP$, CPI, $Industry$, $Size$, $Turnover$ and $Index$.

The first step, as in equation (6), is to show that the dependent variable (comovement, ρ) is related to the independent variable (purchase and redemption mechanism, pr^*). Next, as in equation (7), the expected mediator variable (investment flows of funds, F^*) is shown to be related to the independent variable (purchase and redemption mechanism, pr^*). Finally, as in equation (8), the mediator variable (investment flows of funds, F^*) and the original independent variable (purchase and redemption mechanism, pr^*) are included in the same regression along with the dependent variable (comovement, ρ). If the mediator variable (investment flows of funds, F^*) mediates the relation between the dependent variable (comovement, ρ) and the original independent variable (purchase and redemption mechanism, pr^*), then the coefficient of the original independent variable will be reduced over the first-stage regression, and the mediator variable will be significant.

The model is estimated using an unbalanced panel data model with fixed time effects. We follow Hammersley (2006) and Lang et al. (2012) in estimating the regression equations via ordinary least squares (OLS). However, because OLS models estimate comovement and investment flows of funds separately and, thus, cannot account for the correlation across the different models[5], we also use the Zellner (1962) seemingly unrelated regression (SUR) method. This method tests the robustness of the OLS results and acknowledges the simultaneity and codetermination of the endogenous variables. The results of the mediation analysis via OLS and SUR regressions are reported in the Table IV.

The first three columns of Table IV present the results of the OLS regressions. The purchase and redemption mechanism, pr^*, can significantly predict comovement with a coefficient of 0.23 and a *t*-statistic of 22.64. Moreover, the results in the third column show that pr^* is a powerful variable for forecasting fund simultaneous trading, with a coefficient of 0.31 and a *t*-statistic of 20.49. Specifically, comparing the results in the first and second columns, the coefficient of pr^* decreases significantly (0.18) when F^* is added into the regression analysis, suggesting that the trading practices of irrational funds are significant channels through which the purchase and redemption mechanism affects the comovement. Based on a Sobel test, the p-value is less than 0.05, indicating a significant mediating effect. The SUR regression results shown in the last three columns of Table IV are very similar to the OLS results. The mediating effect is equal to 0.21, even larger than that in the OLS regression model. The Breusch-Pagan test of independence of the residuals shows that the null hypothesis of the independence of residuals is rejected at the 1%

---

[5] We thank the anonymous referee for pointing out this issue.

significance level. Evidently, the joint estimations (SUR) improve the efficiency of the estimated coefficients.

[Insert Table IV here]

The studies of Shleifer (1986), Greenwood (2008), and Lescaroux (2009) separately test macroeconomic factors, stock characteristic factors, and information diffusion factors that can affect stock comovement. However, our results extend their studies to show that the purchase and redemption mechanism is another factor causing comovement. In addition, our tests are consistent with the empirical results in Fortune (1998), Goetzmann and Massa (2003), and Boyer and Zheng (2009), in that purchase and redemption both lead to actual transactions by fund managers. To summarize, the results of the mediation analysis in Table IV shows the potential influence from the purchase and redemption mechanism on investment flows of funds, which then influence the return of comovement.

4.3 Robustness test

4.3.1 Quarter-end test

The quarter-end test shows that purchase and redemption cause comovement from another perspective. Mutual fund traders usually address financial pressure or financial distress by moving shares into or out of funds during the issue month of financial statements (quarterly report, semiannual report, or annual report). Certainly, such activities will lead to more frequent purchases or redemptions in the last month rather than in the first month of the quarter. Therefore, if the hypothesis that the purchase and redemption mechanism is the determinant of fund simultaneous trading is established, connecting stocks would comove more in the last month than in the first month of a quarter.

We calculate the difference between the return comovement in the last month and the first month of each quarter separately. The Wilcoxon signed rank test is performed to examine whether the difference is significantly greater than zero. Table V presents the statistics for quarterly changes in the comovement of different types of shareholders. Specifically, the comovement difference by active mutual funds is positive (0.003), while those of other types are negative. To summarize, the result from Table V is that the more purchases and redemptions, the more obvious the comovement of stocks. In short, the conclusion from Table V confirms that of Table IV.

[Insert Table V here]

4.3.2 Group test

Mutual funds are extremely reliant on outside capital to fund investment opportunities. Cash reserves increase liquidity while also reducing earnings. Generally, in order to pursue profit maximization, funds often have insufficient cash reserves. When relatively large capital is provided, mutual funds expand their positions rather than holding large amounts of cash. Conversely, when relatively large capital is in demand, mutual funds without significant cash reserves have to sell holdings. Therefore, whether more capital is provided or needed, more cash flow will be generated. Furthermore, the pair stocks will comove more.

To ensure that the purchases or redemptions occur pr is determined at a higher level (the benchmark is 60%, 80%, 100%, or 150%). Our null hypothesis is as follows: the more purchase-induced (or redemption-induced) flow, the more comovement is significantly predicted through investor simultaneous trading. We split the sample data into two groups based on the size of pr: pr above the benchmark and pr below the benchmark.

Table VI presents the results of regressing the return comovement on investor trading behavior based on the two different groups in each benchmark. Regardless of the groups, cash flow generated by investor simultaneous trading, F^*, is statistically significant. Moreover, in every set of different benchmarks, the coefficient of F^* estimated for the high pr group is larger than that of the low pr group. For example, for the set of 60% benchmark, the beta of F^* in the high pr group is 1.064, while the beta of F^* in the low pr group is down to 0.621. Not surprisingly, as the pr benchmark increases from 60% to 150%, the coefficients of F^* in the high pr group increase from 1.064 to 1.439 (the coefficients of F^* in the low pr group increase from 0.621 to 0.673). To summarize, the results indicate that the more purchase-induced (or redemption-induced) the cash flow, the more comovement is significantly predicted through fund irrational behavior, which supports our null hypothesis. The conclusion from Table VI confirm the results that in Table IV.

[Insert Table VI here]

5. Additional test

5.1 Do purchases or redemptions matter?

Since purchases and redemptions give rise to fund flows in different directions, in this section, we want to check whether purchases and redemptions have a symmetric effect on comovement. Mutual funds increase their holdings to cover purchases, while they decrease their holdings to cover redemptions. Since funds' decisions to cover purchases potentially differ from their decisions to cover redemptions, we analyze the effects of inflows and outflows on comovement separately. We calculate inflow, outflow, net inflow (NI) and net outflow (NO) separately in relation to total issued amounts at the end of the current quarter (Wedow, 2013):

$$Inflow_{i,t} = \frac{Purchases_{i,t}}{Total\ issured\ amounts_{i,t}} \tag{9}$$

$$Outflow_{i,t} = \frac{Redemptions_{i,t}}{Total\ issured\ amounts_{i,t}} \tag{10}$$

$$Net\ inflow_{i,t} = Inflow_{i,t} - Outflow_{i,t} \tag{11}$$

$$Net\ outflow_{i,t} = Outflow_{i,t} - Inflow_{i,t} \tag{12}$$

The regression results for the first set (benchmark is 0) are presented in the first two columns of Table VII. In addition, the regression results of other benchmarks (20%, 30%, and 50%) are listed in the remaining other columns. Similar to Table VI, trading behavior statistically forecasts the return comovement in each group. Moreover, a Wilcoxon signed rank test is performed. The results show that the coefficient of F^* estimates for the NI group is larger than that of the NO group at a 1% level of significance, indicating that the purchases play a greater role in incurring comovement than redemptions do. In the face of investors' purchases and redemptions, the funds' irrational behavior has an asymmetrical effect on the comovement. In the following sections, we explore the causes of this asymmetry effect. Not surprisingly, as the net flow benchmark increases from 0 to 50%, the coefficients of F^* on the NI group increase from 1.101 to 1.636 (the coefficients of F^* on the NO group increase from 0.713 to 1.243). These findings also provide additional support for the view that the purchase and redemption mechanism leads stocks to comove more.

[Insert Table VII here]

5.2 Herding behavior

We also give an explanation for why inflows (purchases) of funds have more influence on comovement. Per the above analysis, inflows mattering more means more funds are buying at the same time, while outflows mattering less means fewer funds are selling at the same time. A likely explanation for such behavior could be that funds herd more strongly on the buy side than on the sell side. This may be related to transaction costs and China's unique trading system. Hirshleifer et al. (1994) point out that late-informed investors appear to follow the "leader," as their trades are positively correlated with those of early-informed investors. The leader gets information first and buys at a low point, so they have lower buying costs. However, when the market reverses, leaders can gradually withdraw from the market and gain profits. The follower's high transaction costs will prevent them from selling at the same time as the leaders. As a result, the buying herd is greater than the selling herd. On the other hand, due to the special "T+1" trading system in the Chinese market, there is no restriction on buying, but buying and selling on the same day is not allowed. This also explains to some extent why the buying herd is greater than the selling herd.

To measure funds' herding, we adopt the approach proposed by Sias (2004), which has also been used in the studies of Gavriilidis et al. (2013) and Holmes et al. (2013) for the purpose of examining whether or not institutional herding is intentional. Specifically, Economou et al. (2015) also suggest that the approach by Sias (2004), which identifies herding through the intertemporal dependence of institutional demand, is superior to that of Lakonishok et al. (1992). In this model's context, institutional demand is defined as the raw fraction of funds increasing their position ($Raw\Delta_{k,t}$) in security k during period (in our case, quarter) t:

$$Raw\Delta_{k,t} = \frac{NFB_{k,t}}{NFB_{k,t} + NFS_{k,t}} \tag{13}$$

where $NFB_{k,t}$ designates the number of mutual funds buying stock k in quarter t, and $NFS_{k,t}$ designates the number of mutual funds selling stock k in quarter t. To directly compare the estimated coefficients across different groups of firms, all the variables are cross-sectionally standardized as:

$$\Delta_{k,t} = \frac{Raw\Delta_{k,t} - \overline{Raw\Delta_t}}{\sigma(Raw\Delta_{k,t})} \tag{14}$$

where $\overline{Raw\Delta_t}$ is the cross-sectional average raw fraction of institutions buying in quarter t, and $\sigma(Raw\Delta_{k,t})$ is the cross-sectional standard deviation of the raw fraction of institutions buying in quarter t. To test for the existence of herding, Sias (2004) assumes that $\Delta_{k,t}$ follows an autoregressive process of order one in order to gauge the temporal dependence in the structure of institutional demand:

$$\Delta_{k,t} = \beta_t \Delta_{k,t-1} + \varepsilon_{k,t} \tag{15}$$

The first-order coefficient (β_t) constitutes the cross-sectional correlation between institutional demand in quarter t and quarter $t-1$, since both sides of Eq. (15) are standardized, and there is only one explanatory variable ($\Delta_{k,t}$). Sias (2004) decomposes the coefficient (β_t) into two components, "funds following their own past trades" and "funds following the trades of their peers". The latter component measures funds' herding (Herding Measure, HM) in stock k during quarter *t*:

$$HM_{k,t} = \left[\frac{1}{(K-1)\sigma(Raw\Delta_{k,t})\sigma(Raw\Delta_{k,t-1})}\right] * \left[\sum_{n=1}^{N_{k,t}} \sum_{m=1,m\neq n}^{N_{k,t-1}} \left[\frac{(D_{n,k,t} - \overline{Raw\Delta_t})(D_{m,k,t-1} - \overline{Raw\Delta_{t-1}})}{N_{k,t}N_{k,t-1}}\right]\right] \tag{16}$$

where K is the number of stocks traded by funds during quarter t, $N_{k,t}$ is the number of funds actively trading stock k in quarter t, and $D_{n,k,t}$ is a dummy variable whose value equals one (zero) if fund n is a buyer (seller) of stock k in quarter t. Similarly, $N_{k,t-1}$ is the number of funds trading stock k in quarter $t-1$. $D_{n,k,t-1}$ is a dummy variable whose value equals one (zero) if fund n is a buyer (seller) of stock k in quarter $t-1$. $D_{m,k,t-1}$ is a dummy variable whose value equals one (zero) if fund m $(m \neq n)$ is a buyer (seller) of stock k in quarter $t-1$. $HM_{k,t}$ reflects the extent to which one fund follows the trades of another fund from one quarter to the next. If it is positive (negative), then funds tend to follow (trade against) each other over adjacent quarters.

To distinguish herding on the buy and sell sides, Wermers (1999) further measures herding conditional on whether a stock has a higher or lower proportion of buys than the average stock. The buy-herding ($BHM_{k,t}$) and sell-herding ($SHM_{k,t}$) measures are defined as:

$$BHM_{k,t} = HM_{k,t}|Raw\Delta_{k,t} > \overline{Raw\Delta_t}, \quad (17)$$

$$SHM_{k,t} = HM_{k,t}|Raw\Delta_{k,t} < \overline{Raw\Delta_t}. \quad (18)$$

The summary statistics for herding are presented in Panel A of Table VIII. The mean level of herding (HM) across all stock quarters is 0.046, and the average levels of sell herding (SHM) and buy herding (BHM) are 0.023 and 0.082, respectively. The Wilcoxon signed rank test results show that BHM is significantly higher than SHM, indicating that funds herd more strongly on the buy side. To investigate whether fund herding has changed over time, we report summary statistics of BHM and SHM for two roughly equal-length subperiods separately: 2007-2011 and 2012-2016. The results show that BHM is higher than SHM in every subperiod. Such empirical results are consistent with those of Lakshman et al. (2013) and Chang and Lin (2015), who also found institutional herd-buying is more active in emerging markets. In addition, the mean level of herding (HM) increases from 0.037 in the former subperiod to 0.062 in the latter subperiod, which confirms the result of Panel B in Table III.

Next, we also examine whether institutional herding in our sample interacts with market volatility. We rank the quarterly values of market volatility in ascending order and then split them into two groups ("high volatility" and "low volatility"). The results from Panel B of Table VIII indicate that HM is significantly positive in both groups. The Wilcoxon signed rank test shows that BHM is significantly higher than SHM in all cases, and the difference between periods of high and periods of low volatility exhibits no statistical significance. Moreover, we examine whether institutional herding varies between bull and bear markets. The identification of the bull and bear market is based on the studies by Zhou et al. (2009) and Zhang et al. (2017). The results in Panel C of Table VIII show that herding is only significant in the bull market but not in the bear market. This result is consistent with the findings of Tan et al. (2008) and Chiang and Zheng (2010), who found that herding among A-share investors is more pronounced under conditions of rising markets. From the results in the last three columns, BHM in the bull market is significantly larger than SHM, while in the bear market they do not appear to be significantly different.

To summarize, the results in Table VIII explain why inflows (purchases) of fund have more influence on comovement. The stronger buy herding of funds means that more (less) stocks are simultaneously bought (sold) when faced with inflows (outflows), leading to more significant comovement.

[Inset Table VIII here]

6. Discussion

Our paper primarily focuses on answering what kind of investor trading induces the comovement of stock returns, and what factors drive this kind of investor to trade stocks simultaneously with others. We associate the trading practices of five kinds of investors' (funds, insurances, QFII, SOE and individuals) with comovement measurements and find that only funds' trading generates stock return comovement. After controlling for rational factors, our results imply that the trading practices of irrational funds play an important role in comovement. Further tests in our study show that the funds' herding behavior of following other funds' purchase and redemption trading is the main reason for this irrational comovement. We also show that herding behavior following purchases and redemptions are asymmetric. More herding behavior follows purchases, which causes more significant stock return comovement.

Several studies relate to our work. One group of studies focuses on the factors that may induce comovements. These factors are mostly rational ones, such as macro news (Pindyck and Rotemberg, 1990; Batten et al., 2010), industry level news (Barberis and Shleifer, 2003; Greenwood, 2008) or other information diffusion factors (Wurgler and Zhuravskaya, 2002; Kougoulis and Coakley, 2014). Our study includes the abovementioned factors in its model and finds these rational factors can significantly explain part of comovement, which supports the conclusions of these other studies. However, going beyond these papers, our study adds the funds' herding behavior (Deng et al., 2018) as an irrational factor that contributes to comovement, a factor which is ignored by the previous studies.

Another group of studies examines the relationship between fund trading and comovement. For example, Coval and Stafford (2007), Greenwood and Thesmar (2011), Lou (2012), and Anton and Polk (2014) provide support for the view that concentrated mutual fund sales forced by capital flows exert significant price pressure in equity markets, often resulting in the comovement of stock returns. However, whether other kinds of investors contribute to the comovement is not mentioned by these studies. Going beyond these papers, our study examines the effects of the trading practices of five kinds of investors on the comovement of stock returns. We provide evidence that only funds' trading practices significantly affect comovement, which deepens the understanding of the relationship between investors' trading practices and stock return comovement.

Our work draws a full picture of why stock return comovement occurs in the stock market. Compared with previous studies, our work describes a clear mechanism by which rational and irrational factors drive investor(s) stock trades simultaneously, causing comovement. We suggest that both rational and irrational factors contribute to comovement. The rational factors are largely related to the fundamental information about listed companies and may affect all kinds of investors' trading. However, the irrational factor only operates through funds' trading practices. We find that funds' herd buying and selling behavior following other funds' purchases and redemptions contributes to the irrational part of comovement. Overall, our paper can be seen as a complement to and extension of previous studies.

7. Conclusion

In this study, we connect stocks based on common shareholders using data from the top 10 stockholders of companies listed on the Chinese stock market from 2007 to 2016. Specifically, we divide shareholders into five categories, namely, individual, fund, insurance, QFII, and SOEs, to examine their investment inflow (simultaneous trading) on comovement separately. The empirical results show that only fund simultaneous trading forecasts more variation in return correlation, controlling for the other three types of factors.

Moreover, we provide evidence from multiple perspectives to illustrate that comovement arising from fund simultaneous trading is due to the unique purchase and redemption mechanism. Fund investors' purchases generate cash inflows, while redemptions generate cash outflows. Inflows or outflows lead to asset reallocation and adjustment of the fund portfolio, which in turn generates variation in comovement.

Finally, we carry out an additional test to explore whether purchases and redemptions have the same effect on comovement. The regression results show that purchases play a greater role in forecasting comovement than redemptions. This asymmetric effect might be due to fund herding more strongly on the buy side.

References

[1] Ai, C., Chatrath, A., & Song, F. (2006). On the comovement of commodity prices. American Journal of Agricultural Economics, 88(3), 574-588.

[2] Anton, M., & Polk, C. (2014). Connected stocks. The Journal of Finance, 69(3), 1099-1127.

[3] Barberis, N., & Shleifer, A. (2003). Style investing. Journal of Financial Economics, 68(2), 161-199.

[4] Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. Journal of Financial Economics, 75(2), 283-317.

[5] Bartram, S. M., Griffin, J. M., Lim, T.-H., & Ng, D. T. (2015). How important are foreign ownership linkages for international stock returns? The Review of Financial Studies, 28(11), 3036-3072.

[6] Batten, J. A., Ciner, C., & Lucey, B. M. (2010). The macroeconomic determinants of volatility in precious metals markets. Resources Policy, 35(2), 65-71.

[7] Batten, J. A., Hogan, W. P., & Szilagyi, P. G. (2009). Asia-pacific perspectives on the financial crisis 2007-2010. SSRN Electronic Journal.

[8] Bikhchandani, S., & Sharma, S. (2001). Herd behavior in financial markets. IMF Staff Papers, 47 (3), 279-310.

[9] Boyer, B., & Zheng, L. (2009). Investor flows and stock market returns. Journal of Empirical Finance, 16(1), 87-100.

[10] Brown, N. C., Wei, K. D., & Wermers, R. (2014). Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices. INFORMS.

[11] Cao, C., Chang, E. C., & Wang, Y. (2008). An empirical analysis of the dynamic relationship between mutual fund flow and market return volatility q. Journal of Banking & Finance, 32, 2111-2123.

[12] Chang, C. H., & Lin, S. J. (2015). The effects of national culture and behavioral pitfalls on investors' decision-making: herding behavior in international stock markets ☆. International Review of Economics & Finance, 37, 380-392.

[13] Chen, J., Hanson, S., Hong, H., & Stein, J. C. (2008). Do hedge funds profit from mutual-fund distress? Nber Working Papers.

[14] Chiang, T.C., Li, J., & Tan, L. (2010). Empirical investigation of herding behavior in Chinese stock markets: evidence from quantile regression analysis. Global Finance Journal, 21(1), 111-124.

[15] Chiang, T.C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock

markets. Journal of Banking and Finance, 34(8), 1911-1921.

[16] Claessens, S., & Yafeh, Y. (2012). Comovement of newly added stocks with national market indices: Evidence from around the world. Review of Finance, 203-227.

[17] Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. Journal of Financial Economics, 86(2), 479-512.

[18] Cremers, K. J. M., & Petajisto, A. (2009). How active is your fund manager? a new measure that predicts performance. Review of Financial Studies, 22(9), 3329-3365.

[19] Dasgupta, A., Prat, A., & Verardo, M. (2011). The price impact of institutional herding. Review of Financial Studies, 24(3), 892-925.

[20] Demirer, R., & Kutan, A.M. (2006). Does herding behavior exist in Chinese stock markets? Journal of International Financial Markets, Institutions and Money, 16(2), 123-142.

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

[22] Economou, F., Gavriilidis, K., Kallinterakis, V., & Yordanov, N. (2015). Do fund managers herd in frontier markets - and why? International Review of Financial Analysis, 40, 76-87.

[23] Edelen, R. M., & Warner, J. B. (2001). Aggregate price effects of institutional trading: a study of mutual fund flow and market returns. Journal of Financial Economics, 59(2), 195-220.

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

[25] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1-22.

[26] Fortune, P. (1998). Mutual funds, Part II: Fund flows and security returns. New England Economic Review, 3.

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

[28] Gavriilidis, K., Kallinterakis, V., & Ferreira, M. P. L. (2013). Institutional industry herding: Intentional or spurious? Journal of International Financial Markets Institutions and Money, 26(1), 192-214.

[29] Gebka, B., & Wohar, M. E. (2013). International herding: Does it differ across sectors? Journal of International Financial Markets, Institutions and Money, 23, 55-84.

[30] Goetzmann, W. N., & Massa, M. (2003). Index Funds and Stock Market Growth. The Journal of Business, 76(1), 1-28.

[31] Greenwood, R. (2008). Excess comovement of stock returns: Evidence from cross-sectional variation in Nikkei 225 weights. The Review of Financial Studies, 21(3), 1153-1186.

[32] Greenwood, R., & Thesmar, D. (2011). Stock price fragility. Journal of Financial Economics, 102(3), 471-490.

[33] Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior. Social Science Electronic Publishing, 85(5), 1088-1105.

[34] Guo, B., Zhang, W., Zhang, Y., & Zhang, H. (2017). The five-factor asset pricing model tests for the Chinese stock market. Pacific-Basin Finance Journal, 43, 84-106.

[35] Hammersley, J. S. (2006). Pattern identification and industry-specialist auditors. The Accounting Review, 81(2), 309-336.

[36] Harris, L., & Gurel, E. (1986). Price and volume effects associated with changes in the S&P

500 list: New evidence for the existence of price pressures. The Journal of Finance, 41(4), 815-829.

[37] Hirshleifer, D., Subrahmanyam, A., & Titman, S. (1994). Security analysis and trading patterns when some investors receive information before others. The Journal of Finance, 49(5), 1665-1698.

[38] Holmes, P. R., Kallinterakis, V., & Ferreira, M. P. L. (2013). Herding in a concentrated market: A question of intent. European Financial Management, 19(3), 497-520.

[39] Howard W. Chan, & Robert W. Faff. (2003). An investigation into the role of liquidity in asset pricing: australian evidence. Pacific-Basin Finance Journal, 11(5), 555-572.

[40] Iihara, Y., Kato, H., & Tokunaga, T. (2010). Investors' herding on the Tokyo stock exchange. International Review of Finance, 2(1&2), 71-98.

[41] Ivkovich, Z., & Weisbenner, S. J. (2006). 'Old' money matters: the sensitivity of mutual fund redemption decisions to past performance. Social Science Electronic Publishing.

[42] Jotikasthira, C., Lundblad, C., & Ramadorai, T. (2012). Asset fire sales and purchases and the international transmission of funding shocks. The Journal of Finance, 67(6), 2015-2050.

[43] Kallberg, J. & Pasquariello, P. (2008). Time-series and cross-sectional excess comovement in stock indexes ☆. Journal of Empirical Finance, 15(3), 481-502.

[44] Kougoulis, P., & Coakley, J. (2014). Comovement and FTSE 100 index changes. International Journal of Behavioural Accounting & Finance, 4(2).

[45] Lakonishok, J., Shleifer, A. & Vishny, R.V. (1992). The impact of institutional trading on stock prices. Journal of Financial Economics, 32(1), 23-43.

[46] Lakshman, M. V., Basu, S., & Vaidyanathan, R. (2013). Market-wide herding and the impact of institutional investors in the Indian capital market. Journal of Emerging Market Finance, 12(2), 197-237.

[47] Lang, M., Lins, K. V., & Maffett, M. (2012). Transparency, liquidity, and valuation: international evidence on when transparency matters most. Journal of Accounting Research, 50(3), 729-774.

[48] Lescaroux, F. (2009). On the excess co-movement of commodity prices—a note about the role of fundamental factors in short-run dynamics. Energy Policy, 37(10), 3906-3913.

[49] Lesmond, D. A. (2005). Liquidity of emerging markets. Journal of Financial Economics, 77(2), 411-452.

[50] Li, D. D., & Yung, K. (2004). Institutional herding in the ADR market. Review of Quantitative Finance and Accounting, 23(1), 5-17.

[51] Liao, T. L., Huang, C. J., & Wu, C. W. (2011). Do fund managers herd to counter investor sentiment? Journal of Business Research, 64(2), 207-212.

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

[53] Lu, R., Xu, L., & Xie, X. (2007). Fund performance and investors' choice-analysis on the redemption puzzle of open-end fund market in China. Economic Research Journal.

[54] Lynch, A. W., & Mendenhall, R. R. (1997). New evidence on stock price effects associated with changes in the s&p 500 index. Journal of Business, 70(3), 351-383.

[55] Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity: an autocorrelation consistent covariance matrix. Econometrica, 55(3), 703-708.

[56] Nofsinger, J. R., & Sias, R. W. (1999). Herding and feedback trading by institutional and

individual investors. Journal of Finance, 54(6), 2263-2295.
[57] Pindyck, R. S., & Rotemberg, J. J. (1993). The comovement of stock prices. The Quarterly Journal of Economics, 108(4), 1073-1104.
[58] Pindyck, R. S., & Rotemberg, J. J. (1990). The excess co-movement of commodity prices. Economic Journal, 100(403), 1173-1189.
[59] Schwert, G.W. (1989). Why does stock market volatility change over time? Journal of Finance, 44 (5), 1115-1553.
[60] Shiller, R. J. (1989). Comovements in stock prices and comovements in dividends. The Journal of Finance, 44(3), 719-729.
[61] Shleifer, A. (1986). Do demand curves for stocks slope down? The Journal of Finance, 41(3), 579-590.
[62] Sias, R. W. (2004). Institutional herding. Review of Financial Studies, 17(1), 165-206.
[63] Tan, L., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding behavior in Chinese stock market: An examination of A and B shares. Pacific-Basin Finance Journal, 16(1), 61-77.
[64] Vayanos, D., & Woolley, P. (2013). An institutional theory of momentum and reversal. The Review of Financial Studies, 26(5), 1087-1145.
[65] Warther, V. A. (1995). Aggregate mutual fund flows and security returns. Journal of Financial Economics, 39(2-3), 209-235.
[66] Wedow, M. (2013). Purchase and redemption decisions of mutual fund investors and the role of fund families. European Journal of Finance, 19(2), 127-144.
[67] Wermers, R. (1999). Mutual fund herding and the impact on stock prices. Journal of Finance, 54(2), 581-622.
[68] Wurgler, J., & Zhuravskaya, E. (2002). Does arbitrage flatten demand curves for stocks? The Journal of Business, 75(4), 583-608.
[69] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association, 57(298), 348-368.
[70] Zhang, W., Huang, K., Feng, X., & Zhang, Y. J. (2017). Market maker competition and price efficiency: Evidence from China. Economic Modelling, 66(11), 121-131.
[71] Zhou, R. T., & Lai, R. N. (2009). Herding and information based trading. Journal of Empirical Finance, 16(3), 388-393.
[72] Zhou, W. C., Xu, H. C., Cai, Z. Y., Wei, J. R., Zhu, X. Y., & Wang, W. (2009). Peculiar statistical properties of Chinese stock indices in bull and bear market phases. Physica A: Statistical Mechanics & Its Applications, 388(6), 891-899.

Table I. Summary statistics

This table lists the total number of stocks, pairs of stocks, and different types of shareholders that invested in those stocks from 2007 to 2016. The second column shows the number of stocks held by a particular type of shareholder (not including the stocks that don't have a common shareholder with other stocks). The third column is the pair of stocks formed by the particular common shareholders. The fourth column is the number of specific types of shareholders. Shareholders are divided into five types (individual, fund, insurance, QFII, and SOE), and the other shareholders in undistinguished categories are not listed. All data are quarterly averages.

Numbers of stocks, pairs, and shareholders			
Type	Stocks	Pairs	Shareholders
Fund	551	18723	213
Insurance	325	5226	58
QFII	151	1042	75
SOE	143	978	78
Individual	382	394	282

Table II. Investment flows, shareholding and trading volume of different types of investors

Panel A reports the distribution of investor trading behavior ($F_{ij,t}$), which measures the change in total value of stocks held by all common shareholders of the two stocks, scaled by the total market capitalization of the two stocks, at the quarter end. The distribution is shown for the average of all different types of samples. Panel B lists the average amount of shareholding and trading volume of different types of shareholders at each quarter end. S_t represents the average holdings of each quarter (million shares), and P_t is the stock price at the end of each quarter. ΔS_t is the difference between the preceding and following quarterly holdings.

Panel A: Trading behavior (F) of different types of investors

Type	Mean	SD	0%	25%	50%	75%	95%	100%
Fund	0.009	0.009	-0.005	0.006	0.010	0.017	0.032	0.165
Insurance	0.008	0.008	-0.006	0.008	0.009	0.014	0.027	0.187
QFII	0.012	0.040	-0.004	0.005	0.008	0.013	0.038	0.516
SOE	0.019	0.060	-0.011	0.011	0.016	0.023	0.034	0.632
Individual	0.002	0.004	-0.001	0.002	0.003	0.005	0.013	0.248

Panel B: shareholding and trading volume of different types of investors

Type	S_t	P_t*S_t	ΔS_t
Fund	7354	99853	430
Insurance	8150	103719	612
QFII	9286	88777	908
SOE	9544	95212	747
Individual	209	1916	62

Table III. Investment flows on comovement

This table reports regressions explaining the correlation of daily Fama and French (2015) five-factor residuals in quarter t for the sample stocks defined in Table I. The independent variables include the measure of investment inflow, F^{*}, and a series of controls at time t. All independent variables, excluding dummy variables, are then rank-transformed and normalized to have unit standard deviation, which we denote with *. We report estimates of regressions for the full sample in Panel A. Panel B shows the same set of regressions as in Panel A, in addition to two different subsamples (corresponding to the two equal five-year periods). We also include all of the control variables in the subsample analysis. For the sake of brevity, only the F^{*} coefficient and the intercept estimates are shown in Panel B. The model is estimated with an unbalanced panel data model with a fixed time effect and the standard errors are clustered in firms. We report t-statistics based on the standard errors of the coefficients, after adjusting for serial correlation using the Newey and West (1987) procedure. The associated t-statistics are presented in parentheses.

Panel A: Full sample (2007-2016)

	Dependent variable: correlation of 5F residuals				
	Fund	Insurance	QFII	SOE	Individual
$Constant$	0.04**	0.05**	0.06**	0.04**	0.07**
	(15.12)	(14.01)	(9.95)	(8.77)	(7.31)
F^*	0.68**	-0.42	-0.05	0.02	0.33
	(19.18)	(-1.21)	(-1.51)	(0.84)	(1.28)
$RGDP^*$	0.07**	0.07**	0.14**	0.02**	0.11**
	(18.07)	(5.69)	(3.75)	(3.14)	(3.51)
CPI^*	0.04**	0.04**	0.08**	0.03**	0.09**
	(16.53)	(4.87)	(4.22)	(3.23)	(3.84)
$Industry$	0.01**	0.01**	0.01**	0.01**	0.01**
	(15.78)	(5.23)	(4.32)	(3.18)	(3.20)
$Size$	0.01**	0.01**	0.01**	0.01*	0.01*
	(10.72)	(4.94)	(3.39)	(2.18)	(2.29)
$Turnover$	0.01**	0.01**	0.00**	0.01*	0.01**
	(17.31)	(5.40)	(2.59)	(2.19)	(2.95)
$Index$	0.01**	0.01**	0.01**	0.01**	0.06**
	(19.30)	(5.84)	(2.70)	(3.56)	(4.04)

Panel B: Subsample analysis of Panel A

	Dependent variable: correlation of 5F residuals				
	Fund	Insurance	QFII	SOE	Individual
	First subsample (2007-2011)				
$Constant$	0.03**	0.04**	0.03**	0.02**	0.05**
	(14.20)	(11.78)	(7.89)	(6.77)	(5.23)
F^*	0.52**	-0.54	-0.06	0.01	0.30
	(17.33)	(-1.33)	(-1.24)	(0.71)	(1.23)
	Second subsample (2012-2016)				
$Constant$	0.05**	0.05**	0.07**	0.07**	0.08**
	(16.01)	(15.04)	(10.02)	(9.43)	(8.07)
F^*	0.76**	-0.38	-0.04	0.03	0.35
	(20.06)	(-1.45)	(-1.65)	(0.98)	(1.45)

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Table IV. The test of investment flows as a mediator

This table reports the mediation analysis on the purchase and redemption mechanism forecasting comovement through the investment flows of funds. The mediator variable is the investment flows of funds, F^*. The columns (1) and (4) correspond to equation (6), and the columns (2) and (5) correspond to equation (8). The columns (3) and (6) correspond to equation (7). All independent variables, excluding dummy variables, are then rank-transformed and normalized to have unit standard deviation, which we denote with *. Mediating Effect is equal to the decrease in the coefficient on pr^* that arises from including F^* in the comovement model. The P-value (two-sided) for the significance of the mediating effect, based on the Sobel test, is presented in parentheses. The control variables are defined in Table III. The equations are estimated by two techniques to test the robustness of the model: ordinary least squares (OLS) and seemingly unrelated regression (SUR). We report t-statistics based on the standard errors of the coefficients after adjusting for serial correlation using the Newey and West (1987) procedure. The associated t-statistics are presented in parentheses.

The test of fund behavior as a mediator						
	OLS			SUR		
	Dependent variable: ρ		Dependent variable: F^*	Dependent variable: ρ		Dependent variable: F^*
	(1)	(2)	(3)	(4)	(5)	(6)
$Constant$	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	(30.78)	(33.78)	(34.14)	(30.43)	(32.45)	(30.94)
$RGDP^*$	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	(20.59)	(23.40)	(21.84)	(23.43)	(25.34)	(20.17)
CPI^*	0.01**	0.01**	0.02**	0.01**	0.01**	0.02**
	(9.77)	(10.11)	(10.88)	(13.23)	(11.53)	(10.98)
$Industry$	0.03**	0.02**	0.01**	0.03**	0.03**	0.02**
	(10.32)	(9.12)	(12.67)	(10.14)	(9.07)	(13.23)
$Size$	0.00**	0.01**	0.01**	0.01**	0.01**	0.01**
	(15.15)	(13.12)	(14.29)	(18.34)	(14.82)	(15.32)
$Turnover$	0.01**	0.01**	0.02**	0.01**	0.01**	0.01**
	(12.86)	(10.86)	(13.56)	(13.21)	(9.23)	(14.29)
$Index$	0.01**	0.01**	0.01**	0.02**	0.01**	0.02**
	(14.92)	(16.84)	(15.93)	(15.22)	(16.02)	(16.74)
F^*		0.54**			0.68**	
		(15.43)			(17.11)	
pr^*	0.23**	0.05**	0.31**	0.28**	0.07**	0.38**
	(22.64)	(12.94)	(20.49)	(24.32)	(14.56)	(21.87)
Mediating effect (p-value)	0.18 (0.00)			0.21(0.00)		

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Table V. Quarter-end test

This table lists the comovement difference between the last month and the first month in a quarter for each investor type. The difference data shown in the first column are quarterly averages. The Wilcoxon signed rank test is performed to examine whether the difference is significantly greater than zero. The p-value of the Wilcoxon test is shown in the second column.

The comovement difference between last month and first month in a quarter		
Type	Comovement difference	Wilcoxon_p
Fund	0.003	0.049*
Insurance	-0.005	0.239
QFII	-0.006	0.299
SOE	-0.003	0.256
Individual	-0.006	0.376

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Table VI. Group test

This table reports coefficients from regressions of comovement (ρ) on investment flows (F^*). The regression is the same as equation (5). However, unlike equation (5), pr is included as a grouping variable in this regression. In particular, pr is the proportion of the amounts of purchases plus redemptions accounting for the total issued amounts. All independent variables, excluding dummy variables are then rank-transformed and normalized to have unit standard deviation, which we denote with *. The control variables are defined in Table III. The model is estimated with unbalanced panel data model with fixed time effect and the standard errors are clustered in firms. We report t-statistics based on the standard errors of the coefficients after adjusting for serial correlation using the Newey and West (1987) procedure. The associated t-statistics are presented in parentheses.

	Dependent variable: correlation of 5F residuals							
	pr>60%	pr<60%	pr>80%	pr<80%	pr>100%	pr<100%	pr>150%	pr<150%
$Constant$	0.057**	0.041**	0.055**	0.042**	0.054**	0.042**	0.056**	0.042**
	(21.21)	(20.55)	(22.17)	(21.81)	(22.38)	(23.7)	(23.6)	(27.88)
F^*	1.064**	0.621**	1.213**	0.623**	1.280**	0.643**	1.439**	0.673**
	(22.37)	(23.59)	(23.29)	(24.2)	(22.65)	(25.36)	(20.37)	(27.41)
$RGDP^*$	0.101**	0.054**	0.099**	0.058**	0.099**	0.061**	0.101**	0.065**
	(26.25)	(23.84)	(24.41)	(26.11)	(23.48)	(27.78)	(19.97)	(30.51)
CPI^*	0.071**	0.042**	0.069**	0.045**	0.070**	0.046**	0.065**	0.044**
	(23.11)	(21.76)	(21.67)	(19.98)	(22.87)	(24.66)	(19.87)	(25.81)
$Industry$	0.006**	0.002**	0.006**	0.002**	0.006**	0.002**	0.005**	0.002**
	(12.43)	(6.74)	(12.65)	(6.32)	(12.39)	(6.63)	(9.44)	(7.72)
$Size$	0.003**	0.008**	0.003**	0.008**	0.003**	0.007**	0.001**	0.008**
	(3.26)	(11.76)	(2.79)	(11.98)	(2.94)	(11.76)	(2.91)	(12.51)
$Turnover$	0.010**	0.008**	0.010**	0.008**	0.010**	0.008**	0.010**	0.009**
	(10.65)	(13.06)	(9.92)	(13.68)	(9.53)	(13.97)	(7.98)	(15.04)
$Index$	0.010**	0.005**	0.012**	0.005**	0.012**	0.006**	0.010**	0.006**
	(10.20)	(6.92)	(10.92)	(7.44)	(10.88)	(7.67)	(8.00)	(8.75)

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Table VII. Do purchases or redemptions matter?

This table reports coefficients from regressions of comovement (ρ) on investment flows (F^*). The regression is the same as equation (5). Different from Table VII, net inflow (NI) and net outflow (NO) are included as a grouping variable in this regression. NI (NO) is the proportion of the amounts of purchases (redemptions) less redemptions (purchases) accounting for the total issued amounts. All independent variables, excluding dummy variables are then rank-transformed and normalized to have unit standard deviation, which we denote with *. The control variables are defined in Table III. The model is estimated with unbalanced panel data model with fixed time effect and the standard errors are clustered in firms. We report t-statistics based on the standard errors of the coefficients, after adjusting for serial correlation using the Newey and West (1987) procedure. The associated t-statistics are presented in parentheses.

	Dependent variable: correlation of 5F residuals							
	NI>0	NO>0	NI>20%	NO>20%	NI>30%	NO>30%	NI>50%	NO>50%
$Constant$	0.051**	0.040**	0.064**	0.048**	0.065**	0.044**	0.067**	0.045**
	(21.06)	(24.71)	(28.93)	(27.23)	(27.38)	(28.17)	(22.89)	(20.37)
F^*	1.101**	0.713**	1.420**	0.903**	1.548**	1.103**	1.636**	1.243**
	(24.44)	(24.05)	(20.54)	(14.14)	(20.97)	(13.12)	(19.73)	(8.15)
$RGDP^*$	0.107**	0.052**	0.144**	0.057**	0.174**	0.031**	0.181**	-0.007
	(27.81)	(18.46)	(24.27)	(10.37)	(28.07)	(4.56)	(24.65)	(-0.73)
CPI^*	0.082**	0.045**	0.095**	0.046**	1.021**	0.037**	1.057**	0.012**
	(21.12)	(16.77)	(22.78)	(16.32)	(23.98)	(14.33)	(25.91)	(7.65)
$Industry$	0.004**	0.003**	0.004**	0.003**	0.005**	0.003**	0.006**	0.002**
	(10.32)	(9.21)	(10.65)	(6.93)	(11.54)	(7.89)	(13.98)	(7.68)
$Size$	0.005**	0.007**	0.000	0.004**	0.001	0.004*	0.000	0.007**
	(5.21)	(9.73)	(0.21)	(2.70)	(0.66)	(2.46)	(-0.21)	(2.86)
$Turnover$	0.010**	0.009**	0.009**	0.013**	0.009**	0.015**	0.007**	0.015**
	(10.41)	(12.86)	(6.34)	(9.42)	(6.04)	(8.89)	(4.56)	(6.63)
$Index$	0.015**	0.007**	0.011**	-0.004**	0.012**	-0.003	0.016**	0.001
	(15.11)	(7.83)	(8.41)	(-3.04)	(8.60)	(-1.50)	(9.96)	(0.50)

* Significant at the 0.05 level

** Significant at the 0.01 level

Table VIII. Test for herding

This table reports the herding results from Eq. (16). The selling herding intensity (SHM) and buying herding intensity (BHM) are presented in columns (2) and (3), respectively. A Wilcoxon signed rank test is performed to examine whether BHM is significantly higher than SHM. The result of the Wilcoxon test is shown in the last column. Panel A reports the tests for herding over all the periods and the two subperiods. Panel B reports the tests for herding conditional upon market volatility. Volatility here is calculated every quarter using the standard deviation of daily returns in line with Schwert (1989) on the basis of HS300 index returns. The sample is divided into two distinctive groups, namely, "high volatility" and "low volatility", contingent upon whether the market's volatility during the contemporaneous quarter falls in the top or bottom half of the sample period's quarterly volatility estimates when ranked in ascending order. Panel C reports the results of herding in the bullish or bearish period. The identification of the bull and bear market is based on the study by Batten et al. (2009), Zhou et al. (2009) and Zhang et al. (2017). The associated t-statistics are presented in parentheses.

	(1)	(2)	(3)	(4)
	HM	SHM	BHM	$BHM - SHM$
Panel A Tests for herding conditional upon time periods				
Full sample (2007-2016)	0.046**	0.023**	0.082**	0.059**
	(3.28)	(2.67)	(3.52)	(3.13)
Subsample (2007-2011)	0.037**	0.014	0.046*	0.032*
	(2.34)	(1.45)	(1.98)	(1.96)
Subsample (2012-2016)	0.062**	0.031**	0.094**	0.063**
	(3.45)	(2.67)	(3.67)	(3.56)
Panel B Tests for herding conditional upon market volatility				
High volatility periods	0.055**	0.040**	0.086**	0.046**
	(3.86)	(2.98)	(3.55)	(3.22)
Low volatility periods	0.039**	0.015	0.065**	0.050**
	(2.76)	(1.67)	(2.97)	(2.04)
Panel C Tests for herding conditional upon market returns				
Bull market periods	0.046**	0.012	0.076**	0.064**
	(3.35)	(1.77)	(3.22)	(2.33)
Bear market periods	0.013	0.022	0.025	0.003
	(1.43)	(1.72)	(1.82)	(0.83)

* Significant at the 0.05 level.

** Significant at the 0.01 level.

Highlights

The flow of fund is the main generator of return comovement

Other investors have no impacts on the return comovement

Comovement arising from fund flow is due to the purchase and redemption

Purchases cause more comovement than redemptions do