



## OPEN ACCESS

## EDITED BY

Raul Matsushita,  
University of Brasilia, Brazil

## REVIEWED BY

Regina C. B. Da Fonseca,  
Instituto Federal de Goiás (IFG), Brazil  
Joao Plinio Juchem Neto,  
Federal University of Rio Grande do Sul,  
Brazil

## \*CORRESPONDENCE

Shan Lu,  
✉ shan.lu@cufe.edu.cn

## SPECIALTY SECTION

This article was submitted  
to Social Physics,  
a section of the journal  
Frontiers in Physics

RECEIVED 12 January 2023

ACCEPTED 02 February 2023

PUBLISHED 15 February 2023

## CITATION

Ma Y, Zhao J and Lu S (2023), Mutual fund  
net flows in China: A co-holding  
network perspective.  
*Front. Phys.* 11:1142905.  
doi: 10.3389/fphy.2023.1142905

## COPYRIGHT

© 2023 Ma, Zhao and Lu. This is an open-  
access article distributed under the terms  
of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original author(s)  
and the copyright owner(s) are credited  
and that the original publication in this  
journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Mutual fund net flows in China: A co-holding network perspective

Yue Ma<sup>1</sup>, Jichang Zhao<sup>1</sup> and Shan Lu<sup>2\*</sup>

<sup>1</sup>School of Economics and Management, Beihang University, Beijing, China, <sup>2</sup>School of Statistics and Mathematics, Central University of Finance and Economics, Beijing, China

As the largest part of the mutual fund, open-end funds play a critical role in the Chinese mutual fund market. In this work, we construct an undirected weighted fund network of Chinese open-end funds based on the asset co-holding and explore whether funds' co-holdings influence their net flows. The network structure traits are introduced into a panel data model to examine the impact of the asset co-holding network on fund flow changes. It is revealed that when a fund has more connections with other funds, indicated by a higher degree in the co-holding network, it receives more net flows. However, when the average co-holding value of a fund increases, its outflows outnumber the inflows. Furthermore, heterogeneity analysis suggests that this effect particularly exists for low-performance funds. The implications of these findings for fund investor behavior and the fund market are also discussed.

## KEYWORDS

fund networks, co-holding network, net flows, mutual funds, financial networks

## 1 Introduction

For nearly two decades, the mutual fund market has enjoyed a tremendous growth as capital market reform progressed in China. Mutual funds also have contributed significantly to the maturation and professionalization of China's capital market. In particular, investors prefer open-end funds for their high liquidity, transparent and reliable information disclosure, and investment diversification. According to the China Securities Regulatory Commission, open-end funds accounted for about 90% of the total mutual fund at the end of 2019. The changes in fund flows reflect the preferences of funds' investors and are also closely associated with funds' sizes that directly impact the fund managers' investment decisions and future performance [1]. Additionally, since open-end funds are substantial stock market investors, the changes in the fund market are inextricably linked to stock market fluctuations and returns [2]. Therefore, it is essential to pay attention to the changes in the fund flow for the benefit of both fund investors and managers and to maintain financial market stability.

Owing to the fact that open-end funds can be purchased and redeemed freely, their sizes can vary due to different strategies taken by investors. The main influential factor in fund investors' behavior is the historical performance of funds, and lots of studies have shown that open-end fund investors tend to purchase funds with excellent historical performance [3, 4]. However, some studies based on the Chinese fund market have obtained the opposite result, i.e., investors sell funds when the performance of their holdings improves, a phenomenon known as the redemption puzzle [5]. Some scholars have used risk-adjusted returns to measure the performance of China's open-end funds and found no redemption puzzle [6]. The current empirical studies on the flow-performance relationship of China's fund market have not yet reached a consistent conclusion, which is probably because additional factors that influence funds' net flows have not yet been thoroughly considered in such an analysis.

By the end of 2019, the number of China's open-end funds had reached more than 5,600 with nearly 4,000 listed companies in Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE). Despite the large number of underlying assets, the phenomenon of funds holding common assets is prevalent in China, which is usually recognized as the herding effect. The reasons behind the herding effect are multifold. Deliberately imitating others' portfolio may arise from asymmetric information among fund managers [7]. When faced with severe asymmetric information, managers who get inside information ahead adjust their portfolios accordingly [8]. Although anyone is able to access the information of funds' asset holdings since mutual funds must periodically reveal them according to the information disclosure requirements, managers outside the information scope may tend to follow the foregoers. A herding strategy may bring copycat funds comparable to the performances of their target funds [9]. In addition, fund managers are committed to make asset allocation decisions in the guidance of incentive provisions [10]. These provisions might be very similar to each other and may let the managers pinpoint a certain group of stocks that share common characteristics, which further result in the prevalence of portfolio overlaps and concentrated holdings among mutual funds.

Herding behavior has been criticized as an irrational investment strategy [11, 12]. Although either proactive or unintentional herding behavior can be described by portfolio overlaps [13], whether portfolio overlaps of funds could affect fund net flows conditioning their historical performance is still unclear. To answer this, the present paper first depicts portfolio overlaps among funds by constructing fund co-holding networks and then interprets the impact of fund portfolio overlaps on fund net flows from the co-holding perspective. To begin with, we construct fund co-holding networks on a quarterly basis, using the asset-holding data of open-end equity funds that have always been in duration from 2011 to 2019. For reasons of data availability and to highlight major assets, this study focuses on funds' connections originating from overlapping portfolios of their top stock holdings. Each edge represents a portfolio-overlapping relationship between two funds. The overlapping portfolios among funds create links between funds which emerge as a co-holding network. Previous works have focused on the stock network of sharing common mutual funds to unveil panic contagion which leads to an abrupt crash [14] and have focused on how the risk tolerance of individuals affects this procedure [15]. In this study, we conceive the different perspectives of [15] examining how a fund co-holding network influence the fund itself, that is, the net flow of a fund.

The node degree and average co-holding value for each fund in the co-holding network, which portray the extent of portfolio overlapping from the perspectives of the number of connections and the average value of each co-holding connection, respectively, are considered as the potential factors that have an effect on fund net flows. Specifically, a higher node degree indicates that a fund has broader assets that are overlapped with other funds, and the average value of each co-holding connection indicates how much a fund's portfolio is overlapped with other funds in each overlapping equity. Although these might be contributed by various reasons associated with funds' managers' decisions as discussed previously, they could also influence fund performance and in turn affect the investors' net flows of this fund. Alternatively, investors who are familiar with

funds' portfolios through public announcement would either choose funds with higher node degrees for the sake of following the general trend or decline funds that are too similar to others that are easily affected by market fluctuations. To examine these, we construct a fixed-effect regression model to examine the impact of fund network effects on fund net flows with fund network structure variables and control variables, and furthermore, heterogeneity analysis is conducted to explore the effect with respect to fund size and performance. Finally, comparing the results of the fixed-effect model and the ordinary least squares estimation method proves the robustness of this paper's model.

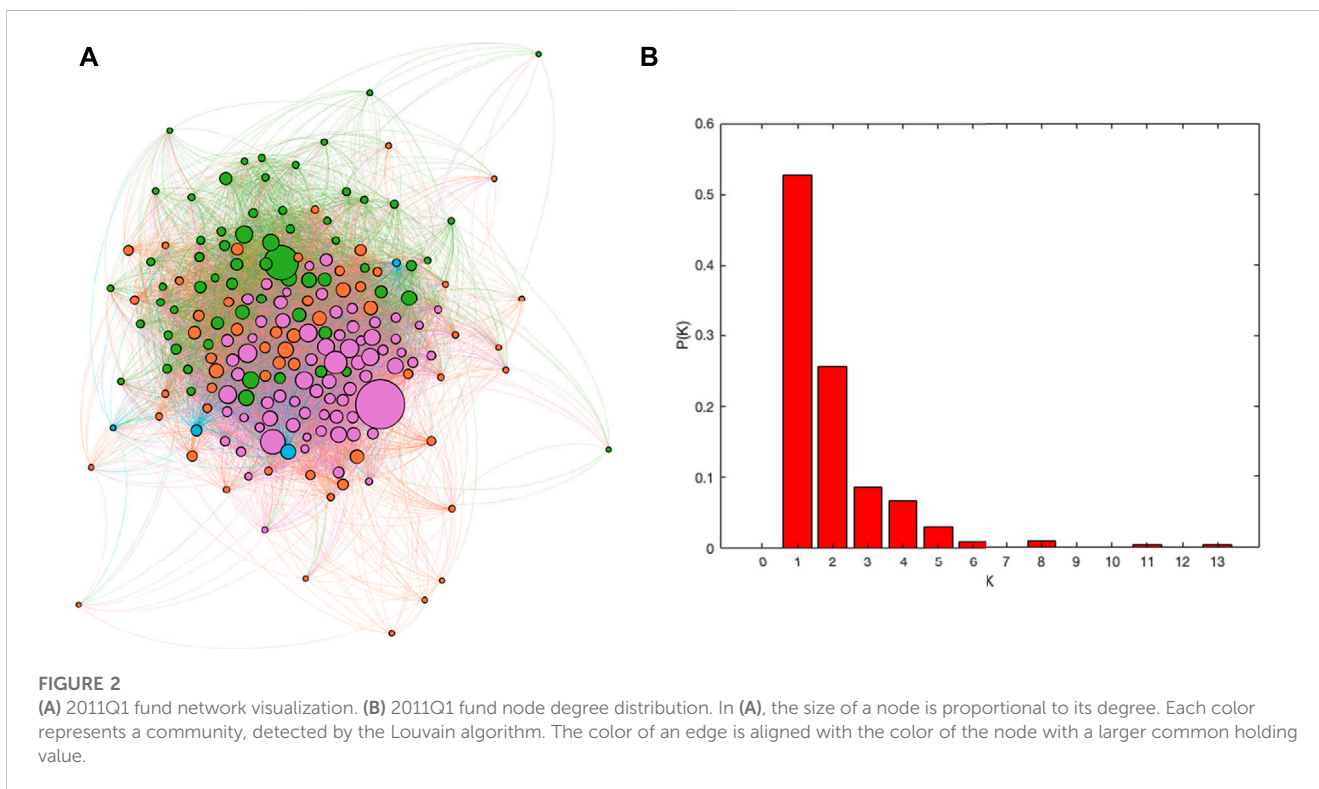
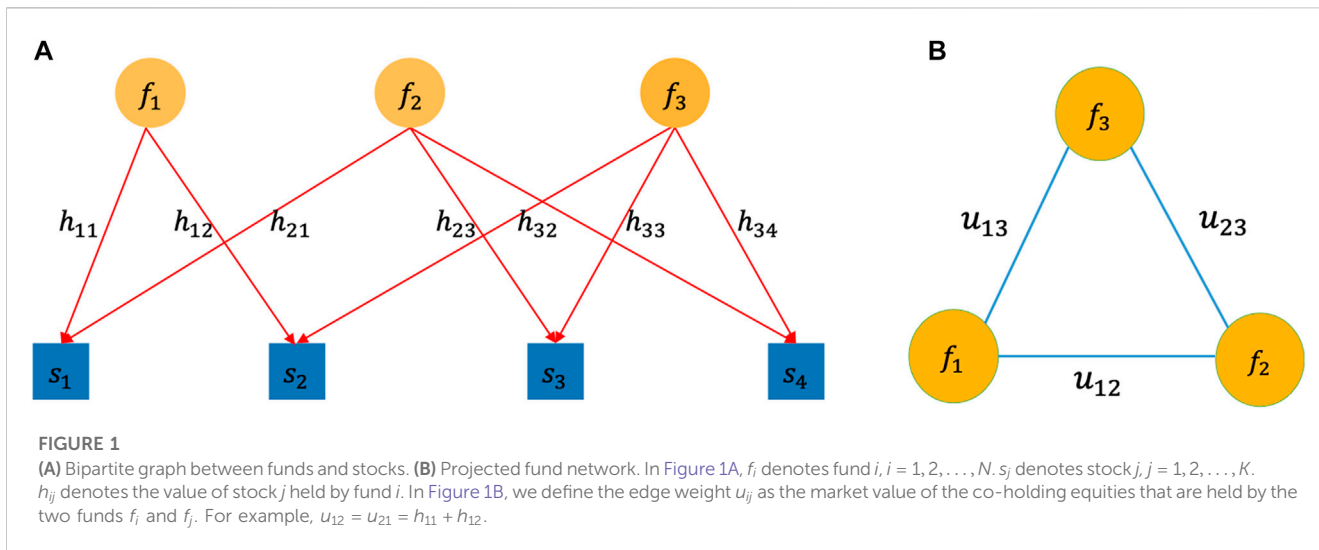
This study has shown that the funds' co-holding network has an effect on the fund net flows. The most prominent findings are that the central funds of the co-holding network, implied by high node degrees, attract more net flows, while funds with higher average values of the common assets receive less net flows. We further set the heterogeneity analysis which confirms that the node degree shows significance at a level of 5% only in the large-size funds. This implies that as the large fund's node degree rises, it would attract more inflows or fewer outflows. In either the high-performance group or the low-performance group, the greater the fund node degree is, the more the fund inflow will be, i.e., investors' purchases outnumber redemptions. However, the fund average co-holding value embodies significance at a level of 1% only in the low-performance group, indicating that the average influence of fund's connections in the co-holding network does not significantly affect the net flows for high-performance funds. More generally, we offer a new view of connecting funds in terms of asset co-holdings and supplement the extant understanding towards fund net flows with network factors such as the degree and average co-holding value. Overall, this study sheds light on the importance of co-holding relationships in determining fund net flows in China, and we believe this novel perspective would inspire follow-up studies that discuss related topics.

## 2 Methods and data

### 2.1 Fund co-holding network

To capture the correlations among funds due to common stock holdings, this paper first builds the fund common-holding network [16]. The co-holding network is projected from a pre-constructed network between stocks and funds, as shown in Figure 1A, by assuming that  $N$  open-end funds hold a total of  $K$  stocks using the market value of the bulk-holding stocks of each fund. The projected network, i.e., the co-holding network, is shown in Figure 1B. An edge between fund nodes represents the existence of co-holding between two funds. The edge weight measures the degree of impact between the two funds, which is defined as the market value of the co-holding equities that are held by the two funds.

Figure 2 shows the network visualization of 195 funds and the fund network degree distribution based on data of 2011Q1 in China. As can be seen, small communities of nearly even sizes can be detected in the fund network derived from common assets. In the meantime, as Figure 2B shows, the node degrees also have a relatively homogeneous distribution rather than a heterogeneous one (e.g., scale free), implying that the co-holding network of funds



is different from previous financial networks, for instance, those among stocks [15]. The observations from other snapshots are similar, indicating stable structures between funds with portfolio overlaps. In addition, the possible effect from these structures on fund net flows accordingly deserves further explorations through econometric models.

Greenwood (2011) analyzed the relationship between the stock ownership structure and stock vulnerability based on United States fund holdings data [17]. The study testified that fragile stock caused by concentrated ownership or its owners faces correlated or volatile liquidity shocks, and fragility strongly predicts price volatility.

Similarly, liquidation behavior of one or more funds on common assets can cause interactions between funds that are denoted by nodes in a fund network, affecting the market price of common assets [18]. To be more specific, price fluctuations in the sold stock have spillover effects that include, but are not limited to, changes in the net asset value of other co-holding funds owning the stock and a drop in the total net value [13, 19]. “Herding effects” within a group of investors can also cause inter-nodal effects. Mutual assets based on common assets may cause greater losses than balance sheet exposures and exhibit risk contagion characteristics [20]. Nevertheless, these extant advances focus on the phenomenon of

liquidity risk that spreads in common asset networks and the antecedents and effects of risk propagation. The impact of the structure of common asset networks on fund flows, however, is rarely assessed and profoundly motivates the present study. The issues we intend to address are whether the structural qualities of funds in the network promote fund inflows or outflows and whether the explanatory fund structure varies under different heterogeneity situations.

Specifically, the present study introduces two network-structure indicators, namely, the node degree ( $degree_{it}$ ) and average co-holding value ( $ach_{it}$ ), to further investigate the impact of co-holding on open-end fund flows. The node degree ( $degree_{it}$ ) denotes the number of the edges that are connected to the node  $i$  at time  $t$ , and the larger the degree is, the more funds that the node has co-holding relationships with, which indicate that the node is in a position of higher interconnectedness in the fund network. The definition is as follows:

$$degree_{it} = \sum_j^N n_{ij} = \sum_j^N n_{ji} \tag{1}$$

where if node  $i$  connected with node  $j$ ,  $n_{ji} = 1$ , otherwise  $n_{ji} = 0$  (we deliberately ignore subscript  $t$  in  $n_{ji}$  for clarity). The fund node degree directly reflects how funds in the market hold common assets with other funds. Central node funds are considered more important in the network, and such nodes usually have greater influence and professional advantage, and their investment behavior is easily referenced or even imitated by other funds [21, 22]. More importantly, a node degree implies how broadly a fund's portfolio is overlapped with others.

In addition, nodes with the same degree value may correspond to different edge weights. Therefore, we believe that edge weights should be combined with a node degree, which can describe the extent of the portfolio overlapping of a fund in a more accurate and comprehensive manner. We construct a new indicator called the average co-holding value ( $ach_{it}$ ), which indicates the average per-connection's influence of each fund. It is calculated by the node strength and degree, where node strength denotes the sum of the edge weights of a node. The definition is as follows; in particular,  $ach_{it}$  denotes the average co-holding of fund  $i$  in quarter  $t$ .

$$ach_{i,t} = \frac{node\ strength_{it}}{degree_{it}} = \frac{\sum_j^N u_{ij}}{degree_{it}} \tag{2}$$

Intuitively, the greater the average co-holding value is, the greater is the extent to which the fund has portfolio overlapping with others.

## 2.2 Panel data model

To demonstrate the interaction between individual funds in the aforementioned network, two fund network indicators, the node degree and average co-holding value, are included as explanatory variables. Based on the theory of the performance-flow relationship, fund performance is included as a control variable in this paper. We only consider equity funds and hybrid funds in this study because their assets mainly allocate stocks.

In addition, funds' fundamental characteristics are also influential factors that cannot be ignored, such as the fund size,

fund age, and fund return volatility. In summary, this paper examines the influence of portfolio overlapping, measured from the perspective of a co-holding network, on fund net flows and sufficiently controls fund performance and funds' fundamental characteristics. A panel data model is therefore established as follows:

$$flow_{it} = \beta_0 + \beta_1 degree_{it} + \beta_2 ach_{it} + \beta_3 \alpha_{it} + \beta_4 size_{it} + \beta_5 age_{it} + \beta_6 sigma_{it} + \epsilon_{it} \tag{3}$$

where  $flow_{it}$  refers to the net flow of fund  $i$  in quarter  $t$  as the dependent variable,  $degree_{it}$  refers to the node degree of fund  $i$  in quarter  $t$ , and  $ach_{it}$  refers to the average co-holding value of fund  $i$  in quarter  $t$ . The fund node degree and average co-holding value are used as explanatory variables for the fund network structure and they have been defined and inferred before. The other variables are regarded as control variables.  $\alpha_{it}$  refers to the risk-adjusted quarterly excess returns benchmarked against the Fama–French five-factor model as the fund performance variables. In addition,  $size_{it}$  refers to the size of fund  $i$  in quarter  $t$ ,  $age_{it}$  refers to the age of fund  $i$  in quarter  $t$ , and  $sigma_{it}$  refers to the return volatility of fund  $i$  in quarter  $t$ .  $\epsilon_{it}$  is the residual term. The details are given in the following sections.

### 2.2.1 Dependent variable

The dependent variable  $flow_{it}$  is defined and measured according to the current dominant methods [4]. Here, we further use the change of proxy in fund size in terms of the growth rate of the total fund net asset value. We define  $flow_{it}$  in the view of the growth rate to remove the effect of fund size magnitude on fund flows to precisely reflect the extent of change in fund flows. Specifically, defining  $TNA_{it}$  as the total net asset value of fund  $i$  at the end of quarter  $t$  and the fund return over the prior year as  $R_{i,t}$ , the net flow of fund  $i$  in quarter  $t$ ,  $flow_{it}$ , is as follows:

$$flow_{it} = \frac{TNA_{it} - TNA_{it-1}(1 + R_{it})}{TNA_{it-1}} = \frac{TNA_{it}}{TNA_{it-1}} - (1 + R_{it}) \tag{4}$$

### 2.2.2 Control variables

The control variables for the model in this paper include fund performance  $\alpha_{it}$ , fund size  $size_{it}$ , fund age  $age_{it}$ , and fund return volatility  $sigma_{it}$ .

#### 1) Fund performance ( $\alpha_{it}$ )

Previous research studies on fund flows of open-end funds mainly focus on performance-flow relationship, which generally explains that fund performance has a significant positive effect on fund flows. For example, historical performance has a positive effect on fund flows [3]. This finding was validated, and then it proved that funds with better historical performance attract more fund flows [4]. Study on German equity funds found that investors choose to redeem funds with both excellent and poor performance and that large fund internal rankings also influence investors' redemption decisions [23].

We select the Fama–French five-factor model-adjusted returns to calculate the fund's current performance [24]. The Fama–French five-factor model can explain the portfolio's excess returns more effectively. The regression model is as follows:

$$R_{it} - R_{ft} = \alpha'_{it} + b_i(R_{mt} - R_{ft}) + s_iSMB_t + h_iHML_t + r_iRMW_t + c_iCMA_t + \epsilon_{it} \tag{5}$$

where  $R_{it}$  is the monthly return of fund  $i$  for month  $t$  and  $R_{ft}$  is the risk-free return for month  $t$ . We use the 1-year fixed deposit rate as the risk-free rate.  $R_{mt}$  is the monthly market return on the value-weighted market portfolio. In addition,  $b_i, s_i, h_i, r_i,$  and  $c_i$  are the coefficients of the excess market return, the size factor  $SMB_t$ , the orthogonal version of the BM factor  $HML_t$ , the profitability factor  $RMW_t$ , and the investment factor  $CMA_t$ . The intercept term  $\alpha'_{it}$  obtained from the regression is the monthly excess return independent of the risk exposure.

The performance  $\alpha'_{it}$  obtained from the model (5) is calculated in a monthly manner, and to maintain consistency with the data frequency of fund's net flow, its quarter value is aggregated here by calculating cumulative excess returns. Here,  $\alpha_{it}$  denotes the risk factor-adjusted excess return for quarter  $t$  and  $\alpha_{u,3v}$  denotes the risk factor-adjusted excess return for month  $3v$ . The quarterly excess return is as follows:

$$\alpha_{it} = (1 + \alpha_{u,3v-2})(1 + \alpha_{u,3v-1})(1 + \alpha_{u,3v}) - 1 \tag{6}$$

In summary, we obtain risk-adjusted quarterly excess returns  $\alpha_{it}$  with the Fama–French five-factor model as a benchmark and take it as the fund performance for the panel model.

2) Fund size ( $size_{it}$ )

In addition to performance factors, it has been shown that the fund flow is negatively correlated with prior fund size, with large-scale funds growing more slowly than smaller funds [25, 26]. For active funds, size is the enemy of performance. The larger the fund size is, the more difficult it is to capture excess returns. Fund flows are negatively correlated with fund size in a prior period; in particular, larger funds grew more slowly than smaller funds [26, 27]. Therefore, we take fund size as a control variable. The indicator is defined as  $size_{it} = \ln(TNA_{i,t-1})$ , where  $TNA_{i,t-1}$  is the total net asset of fund  $i$  in quarter  $t - 1$ .

3) Fund age ( $age_{it}$ )

Funds with a long history have sufficient historical data for reference and attract some conservative investors who prefer funds with relatively rich market experience. Changes in net flows of young funds are more sensitive to historical performance [28]. Young funds have higher net flows than senior funds [27]. Therefore, we take fund age as a control variable. The indicator is defined as  $age_{it} = \ln(date_{i,t-1})$ , where  $date_{i,t-1}$  is the number of days since the fund has been operational originally.

4) Fund return volatility ( $sigma_{it}$ )

Fund net flows are also affected by return volatility. According to portfolio theory proposed by Markowitz, risk-averse investors tend to choose funds with smoother returns when faced with the same level of returns, and funds with higher asset risk have lower net flows [4]. In this study, we take the standard deviation of fund net growth

rate data to calculate this indicator.  $\sigma_{i,t}$  denotes the standard deviation of daily returns of fund  $i$  in quarter  $t$ , and  $n$  is the number of trading days in quarter  $t - 1$ . The indicator is defined as  $sigma_{it} = \sigma_{i,t-1} \sqrt{n}$ .

In summary, we use panel regression, where we take the fund flow as the dependent variable, the metrics derived from the fund networks that demonstrate the interaction between individual funds in the constructed network as the explanatory variables, and many other influential factors related to fund flow as control variables. The primary metrics are  $degree_{it}$  and  $ach_{i,t}$  that are derived from the fund networks. The networks are constructed according to the funds' holding data, inspired by the valuable insights that network science has brought to finance and economics [29]. The control variables include fund performance, age, size, and return volatility. In particular, fund performance is based on performance-flow relationship theory and calculated using the Fama–French five-factor model. These variables have been proven to be useful in predicting the fund flow as discussed. We conducted essential tests such as the Spearman correlation test and the Hausman test before panel regression modeling. Figure 3 shows an illustration of the model in this work, revealing the main variables and corresponding methodologies.

### 2.3 Data

The dataset is provided by the CSMAR (<https://www.gtarsc.com>) and RESSET databases (<http://www.resset.cn>), ranging from 2011 to 2019 for 36 quarters. The characteristics of different types of open-end funds vary. Although the overall size of money funds is significant, they have lower risk and higher liquidity characteristics and are usually regarded as risk-free or low-risk investment instruments. In contrast, equity-biased funds usually have positions of up to 50%–95% in equities and have the characteristics of high risk and a high return. This paper selects equity funds and hybrid funds as research subjects based on this characteristic. We select funds that have always survived during the sample period, resulting in the balanced panel dataset for 195 funds over 36 quarters.

As for the programming language, in this study, we use Python to construct the fund network and calculate the fund network variables ( $degree_{it}$  and  $ach_{i,t}$ ). We also use MATLAB and Gephi to draw the fund network graphs. In addition, Stata is used to obtain the results of regression models.

Table 1 first shows that the absolute value of fund returns is smaller than that of the market index returns in “Mean” column, but the standard deviation is much higher than the market index, indicating a higher dispersion degree of fund returns. The high standard deviation indicates that the fund returns are highly volatile because this paper aims at stock funds, with which the stock market itself is pretty volatile. In addition, the remarkable dispersion in fund returns also suggests that it is necessary to examine the effect of co-holding networks on funds of high and low performance, respectively.

Table 2 gives the descriptive statistics of the data. From 2011 to 2019, the average fund flow for full sample is –3.6%, which indicates that these funds are overall in a state of outflow during the observation period. The minimum fund node degree is zero, indicating that there are isolated nodes in the fund network; in summary, the co-holding network might not be connected. The last column in Table 2 “Obs” shows the sample size.

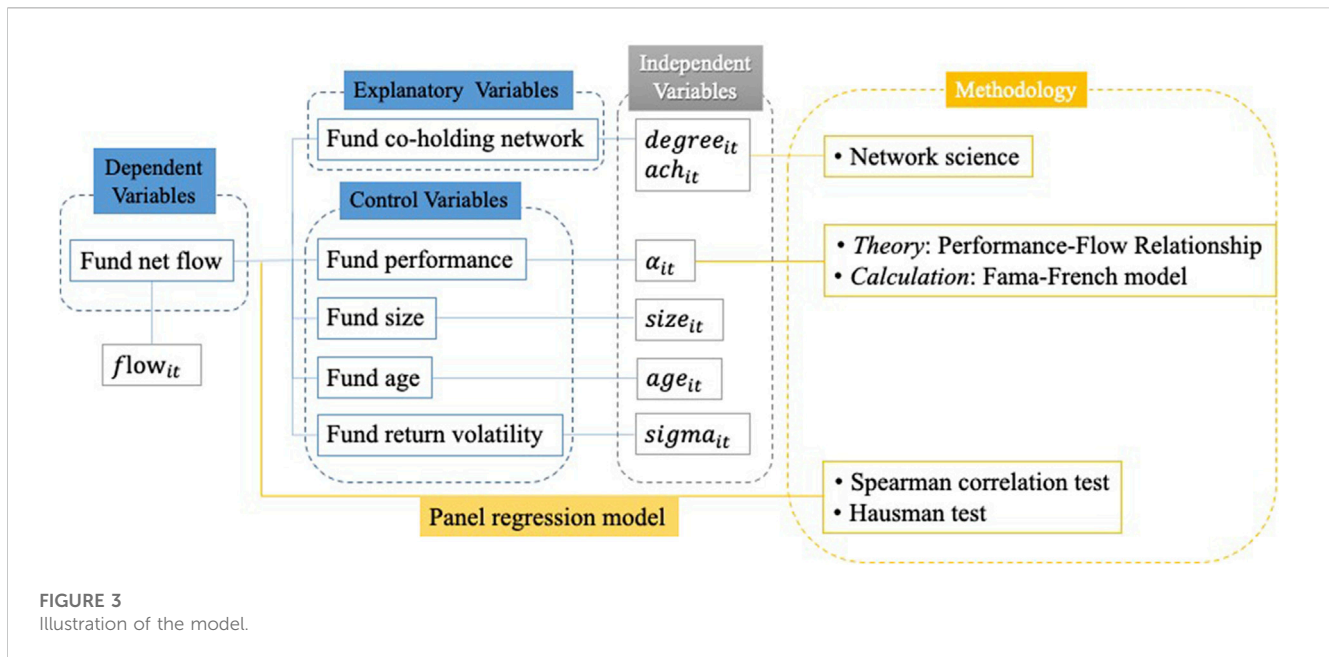


FIGURE 3 Illustration of the model.

TABLE 1 Summary statistics of the Fama–French five-factor model.

Variable	Obs	Mean	Median	Std (%)	25% quantile	75% )quantile
$R_{it} - R_{ft}$	28080	0.12%	3.71%	7.81	-0.88%	1.78
$\alpha_{it}$	21060	0.02%	-0.03%	0.79	-0.54%	0.56
$R_{mt} - R_{ft}$	28080	0.35%	0.01%	4.53	-1.17%	0.94
$SMB_t$	28080	1.02%	0.85%	4.91	-2.02%	4.12
$HML_t$	28080	-0.21%	-0.12%	3.54	-2.08%	1.56
$RMW_t$	28080	-0.33%	-0.49%	2.83	-1.96%	1.18

TABLE 2 Descriptive statistics of the total data.

Variable	Mean	Std	Min	Max	Obs
$flow_{it}$	-0.036	0.188	-1.227	7.434	7020
$degree_{it}$	4.218	0.732	0	5.187	7020
$ach_{it}$	1.296	0.149	-0.191	4.281	7020
$\alpha_{it}$	0.001	0.235	-0.073	0.098	7020
$size_{it}$	3.111	0.984	-0.517	5.554	7020
$age_{it}$	8.146	0.331	7.089	8.804	7020
$sigma_{it}$	0.013	0.006	0.0004	0.049	7020

### 3 Results

#### 3.1 Panel data model

To ensure that the variables in model (3) meet the prerequisites for regression analysis, we first verify whether

the variables are correlated by conducting the Spearman correlation test on the panel data. The test results show that fund flows significantly correlate with explanatory and control variables at the level of 5%. The Hausman test indicates that a fixed-effect model fits more for the main model, while a random effect model fits more for the comparison model with explanatory variables only. Table 3 shows the empirical results of model (3).

Column (a) in Table 3 examines only the explanatory variables on fund flows, whereas column (b) adds control variables. The models pass the effectiveness test in terms of both the F-statistics and the fitness degree. The two network variables in model (a) are both significantly at the level of 1%. The greater the fund node degree is, the greater the fund’s net flows are, indicating that funds with a higher connectedness in the network also have relatively higher net flows. It implies that the central nodes of the network also have relatively higher inflows or lower outflows. The greater the average co-holding value is, the lower its net flow is, which indicates that a fund with a comparatively higher average co-holding value gets more outflows or less inflows.

TABLE 3 Regression results of model (3).

Independent variable	(a) RE	(b) FE
$degree_{it}$	0.024*** (.000)	-0.011*** (.006)
$ach_{it}$	-0.071*** (.000)	-0.081*** (.001)
$\alpha_{it}$	/	0.570*** (.000)
$size_{it}$	/	-0.005 (.365)
$age_{it}$	/	-0.021* (.063)
$sigma_{it}$	/	-3.402*** (.000)
R-square	.001	0.023
F-value	24.89	26.98

*p*-values for the *t*-test of regression coefficients are in parentheses; \*\*\*, \*\*, and \* indicate rejection of the original hypothesis at the 1%, 5%, and 10% significance levels, respectively.

The inclusion of control factors in model (b) reveals that including control variables does not change the explanatory ability of  $degree_{it}$  and  $ach_{it}$  for fund flows in general. Additionally, fund performance statistically and economically explains fund flows, consistent with previous studies. When fund performance is outstanding in the present period, the fund has more net flows, which indicates that investors tend to purchase funds. Current period performance is considerably and positively connected with fund flows. In summary, with or without control variables in the model, both the fund network variables we focus on are statistically significant for fund net flows. In practice, funds with more “neighbors” receive larger net flows. In contrast, funds with a higher average co-holding value than their neighbors experience more withdrawals or a less additional investment from investors.

### 3.2 Heterogeneity analysis

The previous regression results show that a greater node degree and higher fund performance have a significant and beneficial impact on fund net flows, while a higher average co-holding value has a significant inverse effect on fund net flows. To see if this result holds true for different types of funds, this study divides the funds into groups according to fund size and performance and then investigates the impact of performance and the fund network structure on fund flows within each group.

To avoid sample selection bias, the top 1/3 funds in sizes are classified as large funds, while the bottom 1/3 funds are labeled as small funds. A similar approach is used for fund performance grouping. The number of funds in the large-sized group is 144, and there are 140 funds in the small-sized group, while there are 176 funds in the high-performance group and 171 funds in the low-performance group. Table 4 shows descriptive statistics for the grouped data.

Table 4 shows that the mean values of both the fund’s degree and average co-holding value in the large-funds group are slightly higher than the mean value of the small-funds group. Fund flows for large funds on average are lower than those for small funds, showing that large funds have lower net flows. The mean of fund performance ( $\alpha_{it}$ ) in the high-performance group is double that of the low-performance group, demonstrating the grouping’s necessity and

rationale. Simultaneously, fund flows in the high-performance group are much larger than those in the low-performance group. Furthermore, additional variables such as fund age, return volatility, and the two network-related indicators do not differ much between the high- and low-performance groups.

Based on the Hausman tests, we choose the random-effects model (RE) for both large-size and small-size groups. For the high-performance fund group, the fixed-effect model (FE) should be selected, while for the low-performance fund group, the RE model should be chosen. Table 5 shows the regression results. The last row in Table 5 “Obs” shows the sample size.

Table 5 shows that the different fund network indicators have different significances in different size groups. In the large-funds group, a degree is significant at a level of 5% while the average co-holding value is not significant. However, a degree shows no significance in the small-size group, while  $ach_{it}$  shows significance at a 10% significance level. This suggests that when the node degree of a large-size fund rises, implying the fund stays in the network’s core, it will attract more net flows. However, for small-size funds, as the average co-holding value of small funds’ common holdings increases, fund net flows decrease.

In addition, the impact of funds’ performance on net fund flows is also statistically significant in either large-size group or small-size group, with different significance levels. The fund return’s volatility is significant in both large and small fund models at a level of 1%, and regardless of fund size, the fund flow is better explained in the large-funds group when the fund return’s volatility increases. Fund size and age are not significant for small funds’ net flows.

For heterogeneity analysis on the fund’s performance, the fund node degree is significant at a level of 1% of significance in the high-performance group. The average co-holding value factor has no significance, implying that the average market value of a fund’s holding in a co-holding network has no relevance on the high-performance funds’ flows. The fund return’s volatility is also statistically significant and economically significant, specifically showing a significant negative movement, that is, the higher the return volatility is, the more the investors will tend to redeem the fund to hedge their risk when making investment strategies.

In the low-performance group, a fund node degree shows less significance (at 10% significance level) than the average co-holding value (at 1% significance level) on fund net flows. The effect of a

**TABLE 4** Descriptive statistics of variables of heterogeneity analysis.

Variable	Grouping by size		Grouping by performance	
	Large fund	Small fund	High-performance fund	Low-performance fund
$flow_{it}$	−.038	−.031	−.021	−.044
$degree_{it}$	4.293	4.105	4.271	4.148
$ach_{it}$	1.385	1.214	1.334	1.268
$\alpha_{it}$	.004	−.003	.017	−.016
$size_{it}$	4.053	2.093	3.230	2.966
$age_{it}$	8.124	8.174	8.137	8.144
$sigma_{it}$	.013	.013	.013	.013

**TABLE 5** Regression results of model (3) in different fund groups.

Variable	Grouping by size		Grouping by performance	
	Large fund	Small fund	High-performance fund	Low-performance fund
	RE	RE	FE	RE
$degree_{it}$	.009** (0.028)	.013 (.125)	.041** (.000)	.007* (.064)
$ach_{it}$	−.004 (0.854)	−.100* (.065)	.058 (.236)	−0.164*** (.000)
$\alpha_{it}$	.417*** (.001)	.447* (.099)	.455 (.211)	−.063 (.674)
$size_{it}$	−.017** (.013)	−.014 (.119)	−.0002 (.985)	.005 (.239)
$age_{it}$	−.015* (.097)	−.020 (.404)	.027 (.360)	−.031*** (.005)
$sigma_{it}$	−3.821*** (.000)	−2.454*** (.008)	−2.239*** (.016)	−3.397*** (.000)
Obs	2340	2340	2340	2340

*p*-values for the *t*-test of regression coefficients are in parentheses; \*\*\*, \*\*, and \* indicate rejection of the original hypothesis at the 1%, 5%, and 10% significance levels, respectively.

node degree on fund net flows is significantly positive, while the average co-holding value is significantly negative. In other words, an increase in the node degree will bring more inflows or less outflows to the fund for low-performance funds. On the contrary, when the average co-holding value increases, i.e., an increase in the market value of the fund’s common holdings in each connection, it will result in less fund net flows for low-performance funds. Furthermore, the fund degree is significant, but the average co-holding value is not in the high-performance fund group. In addition, it is worth noting that the return volatility is statistically significant in both groups, and fund age only reaches the 1% significance level in the low-performance fund group. Specifically, if a fund’s performance is at a low stage for an extended period of time, investors would prefer to redeem to reduce losses rather than adding additional investment to the fund.

### 3.3 Robustness test

We test the robustness of our main findings further through regression on the ordinary least squares (OLS) model. In

addition, a similar effect of both the degree and average co-holding value on the fund net flow is obtained by using OLS. In particular, comparing the regression results in terms of both F-statistics and R-square, as revealed in Table 6, the fixed-effect model is superior to the OLS estimation method on significance. For example, in the OLS model, the degree and average co-holding value, which are the primary explanatory variables, are only significant at a level of 5%, while its significance level in the FE model is at a level of 1%. Overall, the regression results support the main findings of this paper and further illustrate the robustness of our panel data model. The last column in Table 6, “Obs,” shows the sample size.

## 4 Discussion

This study explores the influencing factors on fund net flows in the Chinese open-end fund market through constructing a co-holding fund network to leverage the phenomenon of common stock holdings and incorporating network indicators at the node level in explaining the fund net flows. We find that, first and



TABLE 6 Regression results of the FE model and OLS.

Variable	FE	OLS	Obs
$degree_{it}$	−.010*** (.006)	.008** (.023)	7020
$ach_{it}$	−.081*** (.001)	−.045** (.028)	7020
$\alpha_{it}$	.570*** (.000)	.536*** (.000)	7020
$size_{it}$	−.005 (.365)	−.006* (.051)	7020
$age_{it}$	−.021* (.063)	−.011 (.173)	7020
$sigma_{it}$	−3.402*** (.000)	−3.207*** (.000)	7020
Constant	.256	.140	7020
F-value	26.98	26.53	/
R-square	.023	.022	/

*p*-values for the *t*-test of regression coefficients are in parentheses; \*\*\*, \*\*, and \* indicate rejection of the original hypothesis at the 1%, 5%, and 10% significance levels, respectively.

foremost, the fund degree and average co-holding value do significantly influence the fund flow. A fund with a greater degree gets more net flows, and meanwhile, a fund with a greater average co-holding value gets less net flows. Since a higher node degree indicates that a fund has broader assets that are overlapped with other funds and a higher average value of a co-holding connection indicates more that a fund's portfolio is overlapped with other funds in each overlapping equity, our findings verified that the herd behavior among funds are indeed associated with investment decisions of funds' investors, which closely link to fund net flows. More precisely, the results suggest that investors prefer funds with higher node degrees but decline funds that are too similar to others. This might arise from the fact that an investor would like to follow the general trend to choose high-degree funds but avoid funds with high co-holding values as they are more easily affected by market fluctuations. Second, we find that there is a performance chasing tendency among Chinese open-end funds. Investors prefer to purchase funds with excellent performance, which proves the non-existence of the "redemption anomaly" conditioning on the co-holding factors. This finding is in line with the existing research [6].

Furthermore, the results from heterogeneity analysis show that a degree and average co-holding value have different significances with respect to different sizes and performances. A node degree is significant in large-size funds but not in the small-size ones. On the contrary, an average co-holding value is significant in small-size funds but not in the large-size ones. When we focus on fund performance groups, a node degree plays a significant impact on high-performance fund flows, while an average co-holding value does not, indicating that the average value of a fund's connections in a co-holding network does not significantly affect the fund's net flows. The low-performance group did the opposite; only the average co-holding value is strongly significant for fund flows. These findings suggest that the co-holding network prompts varied impacts on different funds in

terms of how many co-holding relationships they have (i.e., degrees) and how much co-holding values they have in these relationships (average co-holding values).

Based on the foregoing findings, we suggest that by creating a co-holding network, we may intuitively depict the potential relationships across funds. These relationships vividly reflect either proactive or unintentional herding behavior featured by portfolio overlaps. The fund network structure approach discussed in this paper offers new perspectives on the potential factors that influence fund flows. Aside from the impact of the fund network structure on fund flows, the network structure may also play a significant role in liquidity risk propagation and prevention, which is of great interest for further study. In addition, identifying major fund nodes in the fund market using the proposed co-holding fund network to accurately aid regulators in improving the precision of the financial risk control in the fund market is also worth attention in the future.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by YM, JZ, and SL. The first draft of the manuscript was written by YM, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## Funding

This research was financially supported by the National Natural Science Foundation of China (Grant No. 71871006).

## Acknowledgments

SL thanks the support from the Emerging Interdisciplinary Project of CUFU, the Program for Innovation Research, and the disciplinary funding from the Central University of Finance and Economics. SL is a member of the Financial Sustainable Development Research Team in CUFU.

## References

- Pollet JM, Wilson M. How does size affect mutual fund behavior? *J Finance* (2008) 63(6):2941–69. doi:10.1111/j.1540-6261.2008.01417.x
- Alexander GJ, Cici G, Gibson S. Does motivation matter when assessing trade performance? An analysis of mutual funds. *Rev Financ Stud* (2007) 20(1):125–50. doi:10.1093/rfs/hhl008
- Ippolito RA. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *J L Econ* (1992) 35(1):45–70. doi:10.1086/467244
- Sirri ER, Tufano P. Costly search and mutual fund flows. *J Finance* (1998) 53(5):1589–622. doi:10.1111/0022-1082.00066
- Lu R, Chen BZ, Xu LB, Xie XH. Fund performance and investors' choice analysis on the redemption puzzle of open-end fund market in China. *Econ Res J* (2007) 6:39–50.
- Xiao J, Shi J. Historical performance and fund flows: Does “redemption anomaly” exist in China's open-end fund market? *Econ Res J* (2011) 46(1).
- Cai F, Han S, Li D, Li Y. Institutional herding and its price impact: Evidence from the corporate bond market. *J Financ Econ* (2019) 131(1):139–67. doi:10.17016/FEDS.2016.091
- Ke B, Deloitte KP, Licata M, Broad AE. How informed are actively trading institutional investors? Evidence from their trading behavior before a break in a string of consecutive earnings increases. *J Account Res* (2004) 42(5):895–927. doi:10.1111/j.1475-679X.2004.00160.x
- Verbeek M, Wang Y. Better than the original? The relative success of copycat funds. *SSRN Electron J* (2013) 37:3454–71. doi:10.2139/ssrn.1566794
- Maug E, Naik N. Herding and delegated portfolio management: The impact of relative performance evaluation on asset allocation. *Q J Finance* (2011) 01(02):265–92. doi:10.1142/S2010139211000092
- Sias RW. Institutional herding. *Rev Financ Stud* (2004) 17(1):165–206. doi:10.1093/rfs/hhg035
- D'Arcangelis AM, Rotundo G. Herding in mutual funds: A complex network approach. *J Bus Res* (2021) 129:679–86. doi:10.1016/j.jbusres.2019.11.016
- Coval J, Stafford E. Asset fire sales (and purchases) in equity markets. *J Financ Econ* (2007) 86(2):479–512. doi:10.1016/j.jfineco.2006.09.007
- Lu S, Zhao J, Wang H, Ren R. Herding boosts too-connected-to-fail risk in stock market of China. *Physica A* (2018) 505:945–64. doi:10.1016/j.physa.2018.04.020
- Lu S, Zhao J, Wang H. The emergence of critical stocks in market crash. *Front Phys* (2020) 8:8. doi:10.3389/fphy.2020.00049

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Braverman A, Minca A. Networks of common asset holdings: Aggregation and measures of vulnerability. *J Netw Theor Financ* (2018) 4(3):53–78. doi:10.21314/JNTE.2018.045
- Greenwood R, Thesmar D. Stock price fragility. *J Financ Econ* (2011) 102(3):471–90. doi:10.1016/j.jfineco.2011.06.003
- Guo W, Minca A, Wang L. The topology of overlapping portfolio networks. *Statist Risk Model* (2016) 33(3):139–55. doi:10.1515/strm-2015-0020
- Duarte F, Eisenbach TM. Fire-sale spillovers and systemic risk. *J Finance* (2021) 76(3):1251–94. doi:10.1111/jofi.13010
- Blocher J. Network externalities in mutual funds. *J Financial Mark* (2016) 30:1–26. doi:10.1016/j.finmar.2016.04.001
- Ozsoylev H, Walden J, Deniz Yavuz M, Bildik R, Beshears J, Cohn J, et al. Investor networks in the stock market. *Rev Financ Stud* (2014) 27(5):1323–66. doi:10.1093/rfs/hht065
- Chung SL, Liu W, Liu WR, Tseng K. Investor network: Implications for information diffusion and asset prices. *Pac Basin Finance J* (2018) 48:186–209. doi:10.1016/j.pacfin.2018.02.004
- Stephan J, Wedow M. Purchase and redemption decisions of mutual fund investors and the role of fund families. *Eur J Finance* (2012) 19(2):127–44. doi:10.1080/1351847X.2012.662908
- Fama EF, French KR. A five-factor asset pricing model. *J Financ Econ* (2015) 116(1):1–22. doi:10.1016/j.jfineco.2014.10.010
- Fant LF, O'neal ES. Temporal changes in the determinants of mutual fund flows. *J Financial Serv Res* (2000) 23(3):353–71. doi:10.1111/j.1475-6803.2000.tb00747.x
- Kempf A, Ruenzi S. Family matters: Rankings within fund families and fund inflows. *J Bus Finance Account* (2008) 35(1–2):177–99. doi:10.1111/j.1468-5957.2007.02062.x
- Ferreira MA, Keswani A, Miguel AF, Ramos SB. The flow-performance relationship around the world. *J Bank Financ* (2012) 36(6):1759–80. doi:10.1016/j.jbankfin.2012.01.019
- Chevalier J, Ellison G. Risk taking by mutual funds as a response to incentives. *J Polit Econ* (1997) 105(6):1167–200. doi:10.1086/516389
- Robert MM, Simon AL, George S. Complex systems: Ecology for bankers. *N* (2008) 451:893–5. doi:10.1038/451893a