International Journal of Finance and Investment
ISSN: 3078-6002 (Print), ISSN: 3078-2295 (Online) | Vol. 6, No. 1, 2026

Mutual Fund Crowding in Disclosed Top Holdings and Stock Price Efficiency: Evidence from Chinese A-Shares

Xihao Zhang

Bange College, Central South University of Forestry and Technology, Electronic Information Engineering, Changsha, Hunan 410000, China
yangyng611@163.com

Abstract: Using disclosed top-ten holdings data of mutual funds in China's A-share market, this paper constructs a disclosure-based crowding measure and examines its relationship with stock price efficiency. Combining CSMAR data, the paper builds a Q1 disclosure-based panel covering 38,536 stock-quarter observations for 4,871 A-share stocks from 2011-03-31 to 2023-03-31. Using PriceDelay as the main proxy, the baseline results show that the coefficient on Top10Crowding_Ln1p_w remains significantly negative and reaches -0.0046 in the main specification, indicating that stronger fund crowding is associated with faster incorporation of information into prices. This conclusion remains stable across alternative variable definitions and sample treatments. Mechanism tests provide clearer support for the turnover channel, whereas illiquidity behaves more like a complex channel or suppressor effect. The efficiency-improving effect is mainly concentrated in large-cap, low-illiquidity, and high-volatility stocks.

Keywords: Mutual funds, disclosed top-holdings crowding, stock price efficiency, liquidity, China A-share market.

1. Introduction

In recent years, the institutionalization of China's A-share market has continued to deepen. Mutual funds have become an important pricing force, and concentrated allocation in core assets has become increasingly prominent [1-5].

This raises the central question of the paper: does concentrated fund ownership improve stock price efficiency, or does crowded trading weaken the price discovery function [6-11]?

Existing studies discuss the market consequences of institutional investors, herding, and crowded positions, but direct evidence on the relationship between concentrated fund allocation and stock price efficiency remains limited [3, 4, 8-15].

The institutional setting of China's A-share market makes this question especially relevant [1, 2, 12-14].

This paper defines its research object as mutual fund crowding in disclosed top holdings, that is, the concentration pattern formed when multiple funds simultaneously place the same stock among their heavy holdings within the same disclosure period, and it uses PriceDelay to measure stock price efficiency [16-18].

Based on the Q1 disclosure-based sample from 2011-03-31 to 2023-03-31, this paper finds that stronger fund crowding is significantly associated with lower price delay. The turnover channel receives support, and the effect is mainly concentrated in large-cap, low-illiquidity, and high-volatility stocks.

The contribution of this paper is that it characterizes institutional concentration through the identifiable concept of mutual fund crowding in disclosed top holdings and directly takes stock price efficiency as the core outcome variable.

The disclosed fund-holdings source data currently cover only the March 31 quarter of each year, so the formal sample is essentially a Q1 disclosure-based panel.

The remainder of the paper is organized as follows. Section 2 reviews the literature and develops the hypotheses. Section 3 presents the research design. Section 4 reports the empirical results. Section 5 concludes and discusses implications.

2. Literature Review and Research Hypotheses

2.1. Institutional Investors and Stock Price Efficiency

The literature offers two broad views on the relationship between institutional investors and stock price efficiency: one emphasizes that institutions can accelerate the incorporation of information into prices, whereas the other argues that highly concentrated holdings may bring crowded trading and liquidity fragility [3-11, 15].

Therefore, whether mutual fund crowding in disclosed top holdings improves or weakens stock price efficiency still needs to be tested in a concrete market setting.

2.2. Literature on Herding, Common Ownership, and Crowding

More directly related to this paper is the literature on herding, common ownership, and crowded positions. The concept of "mutual fund crowding in disclosed top holdings" used here is closer to concentration intensity under a disclosure-based definition [3-5, 10, 11, 14, 15].

Existing studies focus more on returns, volatility, crash risk, and liquidity consequences, while direct tests of stock price efficiency are relatively scarce [3-5, 15].

Accordingly, this paper narrows the research object to heavy-holding crowding that matches the fund disclosure frequency, thereby improving the consistency between variable definition and empirical identification.

2.3. Studies on Fund Heavy Holdings in the Chinese Market

In China's A-share market, mutual fund heavy holdings have long attracted considerable attention, and concentrated allocation in core assets together with style co-movement has become increasingly prominent [1-5].

Existing China-focused studies have discussed the return, risk, and stability consequences of fund holdings, but direct and systematic evidence on how mutual fund crowding in disclosed top holdings affects stock price efficiency remains limited [3-5, 12-15].

Against this background, this paper directly links heavy-holding crowding to stock price efficiency in the institutional setting of China's A-share market.

2.4. Literature Assessment and Research Entry Point

Therefore, the identifiable concept of mutual fund crowding in disclosed top holdings provides a useful entry point for re-examining the relationship between institutional investors and price efficiency. The entry point of this paper is not a broad and weakly observable notion of institutional concentration, but a disclosure-based measure that records whether multiple funds simultaneously place the same stock among their top-ten holdings. Compared with broader concepts such as herding or common ownership, this measure is directly observable from public disclosures, matches the frequency of the available sample, and can be linked more cleanly to subsequent stock-level price-efficiency outcomes. This is why disclosed top-holdings crowding serves as the core identification entry point of the paper.

2.5. Research Hypotheses

Based on the preceding literature, fund crowding may either improve price efficiency or weaken it through crowded trading and liquidity frictions [3, 4, 8-13, 15].

2.5.1. Mutual Fund Crowding in Disclosed Top Holdings Significantly Affects Stock Price Efficiency

More specifically, fund crowding may affect the speed with which prices absorb information by changing trading structure and liquidity conditions.

2.5.2. Mutual Fund Crowding in Disclosed Top Holdings May Affect Stock Price Efficiency Through Liquidity-Related Channels

At the current stage, this paper mainly uses Turnover and Illiquidity to test whether this channel exists.

At the same time, the effect of fund crowding is likely to be constrained by firm size, baseline liquidity conditions, and the volatility environment.

2.5.3. The Effect of Mutual Fund Crowding in Disclosed Top Holdings on Stock Price Efficiency Is Significantly Heterogeneous

This proposition is tested in the subsequent subgroup regressions.

2.6. Summary of This Chapter

This chapter reviews the literature on institutional investors, crowded positions, and fund heavy holdings in the Chinese market, and accordingly develops three sets of hypotheses on the main effect, mechanism, and heterogeneity.

3. Research Design and Variable Definitions

3.1. Research Logic and Empirical Framework

This paper focuses on whether fund crowding affects stock price efficiency and whether this effect operates through liquidity-related channels.

Empirically, the paper constructs a stock-quarter fund crowding indicator and develops the analysis through baseline regressions, mechanism tests, heterogeneity analysis, and robustness checks.

3.2. Sample Selection and Data Sources

This paper uses Chinese A-share listed firms as the sample. Fund disclosure data, daily trading data, market returns, and financial data are all drawn from the CSMAR database.

In sample processing, the paper keeps non-financial A-share firms, excludes ST, *ST, PT, and abnormal-trading observations, drops observations with missing key variables, and aligns financial definitions as much as possible.

Because the disclosure data cover only the March 31 quarter of each year, the formal regression sample is a Q1 disclosure-based panel from 2011-03-31 to 2023-03-31, containing 38,536 stock-quarter observations.

3.3. Variable Definitions

The following subsection briefly explains the core variable definitions, and the corresponding formulas are shown in the text.

3.3.1. Dependent Variable: Stock Price Efficiency

Stock price efficiency reflects the speed and adequacy with which prices absorb information. This paper mainly focuses on the speed with which new information is incorporated into prices [6, 7].

This paper adopts PriceDelay proposed by Hou and Moskowitz (2005) as the main measure [16]. If stock returns are mainly explained by contemporaneous market and industry returns, information enters prices relatively quickly; otherwise, the response is more delayed.

Specifically, at each quarter-end t, this paper estimates the following two models using daily return data over the previous 365 calendar days:

Model 1: Restricted model

$$R_{i,d} = \alpha_i + \beta_1 R_d^m + \beta_2 R_d^{ind} + \epsilon_{i,d} \quad (3\text{-}1)$$

Model 2: Full model

$$R_{i,d} = \alpha_i + \beta_1 R_d^m + \beta_2 R_d^{ind} + \sum_{k=1}^{4} \gamma_k R_{d-k}^m + \sum_{k=1}^{4} \delta_k R_{d-k}^{ind} + \epsilon_{i,d} \quad (3\text{-}2)$$

Definition of PriceDelay

$$\text{PriceDelay}_{i,t} = 1 - \frac{R^2_{\text{restricted},i,t}}{R^2_{\text{full},i,t}} \quad (3\text{-}3)$$

$$\text{PriceDelay}_w = \text{Winsorize}_{1\%,99\%}(\text{PriceDelay}) \quad (3\text{-}4)$$

The larger the difference in goodness of fit between the restricted and full models, the stronger the explanatory power of lagged information for current returns, the higher the PriceDelay, and the lower the stock price efficiency.

Accordingly, the subsequent empirical analysis mainly relies on the regression results for PriceDelay.

For robustness, this paper further uses SYNCH as a supplementary measure [17, 18]. However, the core judgment throughout the paper remains based on PriceDelay.

Based on the goodness of fit of the full model, this paper further constructs SYNCH as a supplementary indicator as follows:

$$\text{SYNCH}_{i,t} = \ln\left(\frac{R^2_{\text{full},i,t}}{1 - R^2_{\text{full},i,t}}\right) \quad (3\text{-}5)$$

$$\text{SYNCH}_w = \text{Winsorize}_{1\%,99\%}(\text{SYNCH}) \quad (3\text{-}6)$$

SYNCH is essentially a log-odds transformation of the goodness of fit of the full model.

3.3.2. Main Explanatory Variable: Mutual Fund Crowding in Disclosed Top Holdings

The main explanatory variable in this paper is the intensity of mutual fund crowding in disclosed top holdings, namely the concentrated allocation phenomenon formed when multiple funds simultaneously place the same stock among their top-ten holdings within the same disclosure period [3-5, 10, 11, 14].

In terms of construction, the paper first counts how many times each stock appears in funds' top-ten holdings within the same disclosure period to obtain Top10Crowding_Count, and then constructs the log-transformed indicator Top10Crowding_Ln1p_w as the core explanatory variable.

Accordingly, this paper formalizes the intensity of mutual fund crowding in disclosed top holdings as follows:

$$\text{Top10Crowding_Count}_{i,t} = \sum_f I\left(Holding_{f,i,t} = 1\right) \quad (3\text{-}7)$$

$$\text{Top10Crowding_Ln1p}_{i,t} = \ln\left(1 + \text{Top10Crowding_Count}_{i,t}\right) \quad (3\text{-}8)$$

$$\text{Top10Crowding_Ln1p}_{w} = Winsorize_{(1\%,99\%)(\text{Top10Crowding_Ln1p})} \quad (3\text{-}9)$$

A higher value of this indicator means that the same stock is jointly listed as a heavy holding by more funds, implying stronger disclosure-based crowding.

3.3.3. Control Variables

The baseline regressions control for firm size, leverage, return on assets, turnover, volatility, past return, and illiquidity.

Among them, Size_MV_w, Leverage_w, and ROA_w capture firm fundamentals, whereas Turnover_w, Volatility_w, PastReturn_w, and Illiquidity_w capture trading activity, risk, and liquidity conditions.

Accordingly, the formulas for the main control variables can be written as follows:

$$Size_MV_{i,t} = \ln\left(TotalMV_Yuan_qe_{i,t}\right) = \ln\left(Dsmvtll_last_{i,t} \times 1000\right) \quad (3\text{-}10)$$

$$Leverage_{i,t} = \frac{TotalLiabilities_{i,t}}{TotalAssets_{i,t}} \quad (3\text{-}11)$$

$$ROA_{i,t} = \frac{NetProfit_ForROA_{i,t}}{TotalAssets_{i,t}} \quad (3\text{-}12)$$

$$Turnover_{i,t} = \frac{1}{N_{i,t}} \sum_{d \in t} \left(\frac{Dnvaltrd_{i,d}}{Dsmvosd_{i,d} \times 1000} \right) \quad (3\text{-}13)$$

$$Volatility_{i,t} = sd\left(Dretwd_{i,d}\right), d \in t \quad (3\text{-}14)$$

$$PastReturn_{i,t} = exp\left(\sum_{d \in t} \ln\left(1 + Dretwd_{i,d}\right)\right) - 1 \quad (3\text{-}15)$$

$$Illiquidity_{i,t} = \frac{1}{N_{i,t}} \sum_{d \in t} \left(\frac{10^6 \times |Dretwd_{i,d}|}{Dnvaltrd_{i,d}} \right) \quad (3\text{-}16)$$

$$X_w = Winsorize_{1\%,99\%}(X) \quad (3\text{-}17)$$

Unless otherwise stated, all continuous variables used in the formal regressions are winsorized at the 1% and 99% levels.

3.3.4. Mechanism Variables and Heterogeneity Grouping Variables

In the mechanism analysis, this paper mainly examines the two liquidity-related variables Turnover_w and Illiquidity_w. The heterogeneity analysis groups stocks by the medians of Size_MV_w, Illiquidity_w, and Volatility_w.

3.4. Baseline Regression Model

To test the relationship between mutual fund crowding in disclosed top holdings and stock price efficiency, this paper specifies the following baseline regression model:

$$PriceDelay_{i,t} = \alpha + \beta\, Crowding_{i,t} + \boldsymbol{\gamma}' Controls_{i,t} + YearFE_t + IndustryFE_j + \varepsilon_{i,t} \quad (3\text{-}18)$$

Here, Crowding denotes the intensity of fund crowding and Controls denotes the set of control variables. The model includes year fixed effects and industry fixed effects, and standard errors are clustered at the stock level.

Specifically, M1 includes only fixed effects, M2 adds firm characteristic variables, and M3 further adds trading-behavior and market-microstructure variables and is treated as the main specification.

3.5. Mechanism Test Design

The mechanism analysis adopts a two-step approach: it first tests whether fund crowding significantly affects the candidate mechanism variables, and then includes the mechanism variables in the baseline regression to examine the change in the coefficient on the main explanatory variable.

Step 1: test whether fund crowding significantly affects the candidate mechanism variable:

$$Mechanism_{i,t} = \alpha + \beta\, Crowding_{i,t} + \boldsymbol{\gamma}' Controls_{i,t} + YearFE_t + IndustryFE_j + \varepsilon_{i,t} \quad (3\text{-}19)$$

Step 2: include the corresponding mechanism variable in the baseline regression and examine both the change in the fund crowding coefficient and the significance of the mechanism variable itself:

$$PriceDelay_{i,t} = \alpha + \beta\, Crowding_{i,t} + \delta\, Mechanism_{i,t} + \boldsymbol{\gamma}' Controls_{i,t} + YearFE_t + IndustryFE_j + \varepsilon_{i,t} \quad (3\text{-}20)$$

If fund crowding significantly affects the mechanism variable and the coefficient on the main explanatory variable shrinks noticeably after the mechanism variable is added, the variable is treated as providing partial mechanism evidence.

3.6. Heterogeneity Test Design

The heterogeneity analysis groups the sample by the medians of firm size, illiquidity, and return volatility, and then estimates the baseline model separately in the high and low subsamples to identify conditional differences in the effect of fund crowding.

3.7. Robustness Test Design

The robustness analysis has two layers: replacing variable definitions, and deleting extremely crowded observations while adjusting the fixed-effects specification.

3.8. Summary of This Chapter

This chapter establishes the basic framework for the subsequent empirical analysis, taking mutual fund crowding in disclosed top holdings as the core explanatory variable and PriceDelay as the main dependent variable.

4. Empirical Results

4.1. Sample Description

After data cleaning and variable construction, this paper

obtains a final sample of 38,536 stock-quarter observations covering 4,871 A-share stocks. The sample is essentially an annual Q1 disclosure-based panel from 2011-03-31 to 2023-03-31.

Before the formal regressions, all continuous variables are winsorized at the 1% and 99% levels. The main explanatory variable is Top10Crowding_Ln1p_w, the main dependent variable is PriceDelay_w, and the supplementary indicator is SYNCH_w.

4.2. Descriptive Statistics

Table 1 shows substantial cross-sectional variation in the intensity of mutual fund crowding, price delay, and stock price synchronicity, providing the necessary data basis for the subsequent regressions.

Table 1. Descriptive Statistics of Main Variables

Variable	N	Mean	SD	Min	P25	Median	P75	Max
Crowding	38,536	0.9161	1.2224	0.0000	0.0000	0.0000	1.6094	4.7875
Price Delay	38,536	0.0916	0.0983	0.0053	0.0317	0.0595	0.1119	0.5822
SYNCH	38,536	-0.4412	0.8151	-2.5667	-0.9618	-0.4282	0.0835	1.6223
Size	38,536	22.6000	0.9498	20.9536	21.9107	22.4478	23.1374	25.5580
Leverage	38,536	0.4137	0.2104	0.0441	0.2413	0.4042	0.5701	0.8983
ROA	38,536	0.0083	0.0134	-0.0266	0.0013	0.0066	0.0144	0.0575
Turnover	38,536	0.0271	0.0244	0.0023	0.0109	0.0194	0.0346	0.1343
Volatility	38,516	0.0287	0.0103	0.0101	0.0213	0.0272	0.0345	0.0586
Past Return	38,536	0.0408	0.2301	-0.3666	-0.1132	-0.0000	0.1535	0.8923
Illiquidity	38,536	0.0004	0.0004	0.0000	0.0001	0.0003	0.0006	0.0022

Note: This table reports descriptive statistics for the main variables in the formal regression sample. Continuous variables are winsorized at the 1% and 99% levels.

4.3. Correlation Analysis

Table 2 shows that Top10Crowding_Ln1p_w is negatively correlated with PriceDelay_w and positively correlated with SYNCH_w, preliminarily suggesting that stronger fund crowding may be associated with more timely stock price responses.

Table 2. Correlation Matrix of Main Variables

Variable	Crowding	Delay	SYNCH	Size	Lev	ROA	Turn	Vol	Ret	ILLIQ
Crowding	1.000	-0.103	0.161	0.707	0.028	0.306	-0.144	0.050	0.104	-0.387
Delay	-0.103	1.000	-0.686	-0.106	-0.087	-0.004	0.252	0.122	-0.015	0.044
SYNCH	0.161	-0.686	1.000	0.219	0.109	0.003	-0.246	-0.061	-0.035	-0.069
Size	0.707	-0.106	0.219	1.000	0.124	0.277	-0.185	-0.013	0.058	-0.602
Lev	0.028	-0.087	0.109	0.124	1.000	-0.308	-0.149	-0.041	0.006	-0.111
ROA	0.306	-0.004	0.003	0.277	-0.308	1.000	-0.054	-0.061	0.020	-0.129
Turn	-0.144	0.252	-0.246	-0.185	-0.149	-0.054	1.000	0.588	0.190	-0.099
Vol	0.050	0.122	-0.061	-0.013	-0.041	-0.061	0.588	1.000	0.091	0.004
Ret	0.104	-0.015	-0.035	0.058	0.006	0.020	0.190	0.091	1.000	-0.052
ILLIQ	-0.387	0.044	-0.069	-0.602	-0.111	-0.129	-0.099	0.004	-0.052	1.000

Note: This table reports Pearson correlations among the main variables in the formal regression sample. Continuous

variables are winsorized at the 1% and 99% levels.

4.4. Baseline Regression Results

Table 3 reports the baseline regression results. In M1, M2, and M3, the coefficient on Top10Crowding_Ln1p_w is significantly negative, indicating that stronger fund crowding is significantly associated with lower price delay.

After controlling for firm characteristics, trading behavior, and market-microstructure variables, the negative relationship remains robust.

M3 is treated as the main specification and supports the core conclusion that fund crowding is associated with higher stock price efficiency [16].

Among the control variables, turnover, volatility, and illiquidity all have significantly positive coefficients, suggesting that trading frictions and noise may weaken the speed at which information enters prices [16].

Table 3. Baseline Regression Results

Variable	M1	M2	M3
Crowding	-0.0099***	-0.0036***	-0.0046***
	(0.0004)	(0.0006)	(0.0006)
Size		-0.0045***	0.0041***
		(0.0009)	(0.0010)
Leverage		-0.0230***	-0.0178***
		(0.0032)	(0.0031)
ROA		-0.1209**	-0.0787*
		(0.0486)	(0.0475)
Turnover		1.0241***	1.0141***
		(0.0301)	(0.0407)
Volatility			0.7100***
			(0.0932)
Past Return			-0.0208***
			(0.0033)
Illiquidity			31.7142***
			(1.8293)
N	38,536	38,536	38,516
Clusters	4,871	4,871	4,871
R2	0.1740	0.2367	0.2483
Adj. R2	0.1720	0.2348	0.2463
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

Note: This table reports the baseline regression results. The dependent variable is PriceDelay_w. All models include year fixed effects and industry fixed effects.

4.5. Robustness Test Results

Table 4 shows that the core conclusion does not change materially after replacing the dependent variable, changing the definition of the main explanatory variable, and returning to the unwinsorized sample [17, 18].

This indicates that the baseline conclusion is not driven by a particular variable form or by the treatment of extreme values.

Table 4. Robustness Test Results

Variable	R1	R2	R3
Crowding	0.0223***		
	(0.0052)		
Crowding Count		-0.0001***	
		(0.0000)	
Crowding (Raw)			-0.0039***
			(0.0007)
Size	0.1090***	0.0022**	
	(0.0097)	(0.0009)	
Leverage	0.1199***	-0.0176***	
	(0.0255)	(0.0031)	
ROA	-0.5383	-0.1107**	
	(0.3879)	(0.0473)	
Turnover	-6.2239***	1.0420***	
	(0.2225)	(0.0404)	
Volatility	-9.7702***	0.6457***	
	(0.6066)	(0.0928)	
Past Return	0.0783***	-0.0227***	
	(0.0204)	(0.0033)	
Illiquidity	-144.6582***	32.5842***	
	(13.3515)	(1.8488)	
Size (Raw)			-0.0044***
			(0.0010)
Leverage (Raw)			-0.0243***
			(0.0034)
ROA (Raw)			-0.0819
			(0.0507)
Turnover (Raw)			0.9292***
			(0.0441)
Volatility (Raw)			0.6189***
			(0.1036)
Past Return (Raw)			-0.0154***
			(0.0036)
Illiquidity (Raw)			1.7568***
			(0.6795)
N	38,516	38,516	38,516
Clusters	4,871	4,871	4,871
R2	0.4879	0.2487	0.2254
Adj. R2	0.4865	0.2467	0.2234
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

Note: This table reports the first layer of robustness test results. R1 replaces the dependent variable with SYNCH_w, R2 replaces the main explanatory variable with Top10Crowding_Count, and R3 uses the unwinsorized sample and unwinsorized variables.

4.6. Mechanism Test Results

Table 5 reports the mechanism test results. Fund crowding significantly reduces turnover. After Turnover_w is included in the main regression, the absolute value of the fund crowding coefficient shrinks noticeably, suggesting that the turnover channel receives support.

This result is more naturally interpreted as follows: concentrated fund holdings stabilize trading structure to some extent and reduce short-term noise [9, 11].

When Illiquidity_w is used as the mechanism variable, fund crowding significantly increases illiquidity, but after it is added to the main regression, the absolute value of the main coefficient instead becomes larger, indicating that this channel is closer to a complex channel or suppressor effect.

Therefore, the current evidence provides stronger support for a "trading-structure stabilization" mechanism than for an "illiquidity improvement" mechanism [15].

Table 5. Mechanism Test Results

Variable	ME1	ME2	ME3	ME4	ME5
Crowding	-0.0069***	-0.0031***	-0.0042***	0.0000***	-0.0074***
	(0.0006)	(0.0001)	(0.0006)	(0.0000)	(0.0006)
Turnover			0.8651***		
			(0.0394)		
Illiquidity					16.0813***
					(1.7493)
Size	-0.0061***	-0.0019***	-0.0044***	-0.0003***	-0.0019*
	(0.0009)	(0.0002)	(0.0009)	(0.0000)	(0.0010)
Leverage	-0.0313***	-0.0085***	-0.0239***	-0.0002***	-0.0288***
	(0.0032)	(0.0007)	(0.0032)	(0.0000)	(0.0032)
ROA	-0.0075	0.0903***	-0.0857*	-0.0006***	0.0028
	(0.0482)	(0.0104)	(0.0482)	(0.0002)	(0.0478)
Volatility	2.4101***	1.8251***	0.8312***	-0.0048***	2.4865***
	(0.0761)	(0.0196)	(0.0940)	(0.0003)	(0.0766)
Past Return	-0.0155***	0.0018**	-0.0171***	0.0001***	-0.0173***
	(0.0033)	(0.0007)	(0.0033)	(0.0000)	(0.0033)
N	38,516	38,516	38,516	38,516	38,516
Clusters	4,871	4,871	4,871	4,871	4,871
R2	0.2167	0.5028	0.2396	0.4699	0.2191
Adj. R2	0.2147	0.5016	0.2376	0.4686	0.2171
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes

Note: This table reports the mechanism test results. ME2 and ME4 use the candidate mechanism variables as dependent variables, whereas ME1, ME3, and ME5 use PriceDelay_w as the dependent variable.

4.7. Heterogeneity Test Results

Table 6 shows that the price-efficiency-improving effect of fund crowding is mainly concentrated in large-cap, low-illiquidity, and high-volatility stocks; in the high-illiquidity group, the effect does not hold.

Table 6. Heterogeneity Test Results

Variable	H1	H2	H3	H4	H5	H6
Crowding	0.0017	-0.0047***	-0.0039***	0.0022**	-0.0017**	-0.0068***
	(0.0013)	(0.0007)	(0.0007)	(0.0011)	(0.0008)	(0.0008)
N	19,253	19,263	19,260	19,256	19,258	19,258
Clusters	3,680	3,479	3,981	4,192	4,478	4,405
R2	0.2797	0.2386	0.2476	0.3028	0.1999	0.2810
Adj. R2	0.2760	0.2346	0.2437	0.2992	0.1957	0.2774
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the heterogeneity test results. The sample is grouped by the medians of firm size, illiquidity, and return volatility. All models include year fixed effects and industry fixed effects.

4.8. Second-Layer Robustness Test Results

Table 7 further shows that after deleting extremely crowded observations and adjusting the fixed-effects specification, the

negative relationship between fund crowding and lower PriceDelay remains stable.

Table 7. Additional Robustness Test Results

Variable	SR1	SR2	SR3
Crowding	-0.0041***	-0.0043***	-0.0043***
	(0.0006)	(0.0006)	(0.0006)
N	38,129	38,516	38,516
Clusters	4,868	4,871	4,871
R2	0.2491	0.2187	0.2336
Adj. R2	0.2471	0.2183	0.2328
FE Setting	Year + Industry FE	Year FE Only	Year + Broad Industry FE
Sample Treatment	Exclude obs. with Top10Crowding_Count >= 119.000000	Full sample	Full sample

Note: This table reports the second layer of robustness test results. SR1 shows the results after deleting extremely crowded observations, SR2 reports the results with only year fixed effects, and SR3 adopts a broad-industry fixed-effects definition.

5. Conclusion and Implications

Taken together, the results in Sections 4.4 to 4.8 provide relatively stable evidence within the current Q1 disclosure-based sample: fund crowding is significantly associated with lower price delay, and the robustness, mechanism, and heterogeneity analyses broadly support this conclusion. This evidence should nevertheless be interpreted within a clear sample boundary, because the relationship identified here is between Q1 disclosure-based crowding in mutual fund heavy holdings and stock price efficiency. Against this background, this section summarizes the main findings and implications of the paper.

First, the main-effect results show that mutual fund crowding in disclosed top holdings is significantly associated with higher stock price efficiency, specifically in the sense that stronger crowding corresponds to lower PriceDelay [8, 9, 12, 13, 16].

Second, the robustness tests show that the basic relationship between fund crowding and price efficiency remains stable after changing variable definitions, sample treatment, and fixed-effects settings [16-18].

Third, the mechanism tests indicate that the turnover channel receives relatively clear support, whereas illiquidity is closer to a complex channel or suppressor effect [9, 11, 15].

Fourth, the heterogeneity results show that the price-efficiency-improving effect of fund crowding is mainly concentrated in large-cap, low-illiquidity, and high-volatility stocks [8, 9].

Based on the above findings, this paper argues that concentrated institutional allocation does not necessarily imply a deterioration in efficiency, and "crowding" should not be mechanically equated with "pricing failure."

This study also has clear limitations: the sample covers only a Q1 disclosure-based panel, and mechanism identification is still mainly concentrated in the liquidity dimension.

Future research may supplement fund disclosure data for other quarters and strengthen causal identification by combining institutional changes or exogenous shocks.

Overall, the evidence in this paper shows that within the current sample range, mutual fund crowding in disclosed top holdings does not weaken stock price efficiency; instead, it is more likely to be associated with faster price adjustment [6-9].

References

[1] China Securities Regulatory Commission. (2024, September 24). State Council Information Office press conference: Financial support for high-quality economic development. https://www.csrc.gov.cn/csrc/c106311/c7508374/content.shtml

[2] China Securities Regulatory Commission. (2025, January 23). State Council Information Office press conference: Promoting medium- and long-term capital into the market and advancing the high-quality development of the capital market. https://www.csrc.gov.cn/csrc/c106311/c7535350/content.shtml

[3] Su, Z., He, X., & Zhang, Y. J. (2024). Fund crowding: Net asset value growth and crash risk. Management Review, 36(3), 30–44.

[4] Hu, X. F., Xu, R., & Xu, Y. Z. (2022). Crowded investment and stock market stability: A mechanism study from the perspective of concentrated asset sales. International Finance Research, (11), 86–94.

[5] Wang, C. F., Li, B., Yao, S. Y., & Fang, Z. M. (2024). Crowding for warmth: Fund groups and quarter-end stock price run-ups. Journal of Systems & Management, 33(4), 1025–1042.

[6] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417.

[7] Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. American Economic Review, 70(3), 393–408.

[8] Boehmer, E., & Kelley, E. K. (2009). Institutional investors and the informational efficiency of prices. The Review of Financial Studies, 22(9), 3563–3594.

[9] Breugem, M., & Buss, A. (2019). Institutional investors and information acquisition: Implications for asset prices and informational efficiency. The Review of Financial Studies, 32(6), 2260–2301.

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

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

[12] Lin, X., He, W., Yu, J. F., & Xiong, X. (2023). Do mutual funds improve market pricing efficiency? Fund holding quality and stock returns. Journal of Financial Research, 514(4), 149–167.

[13] Zhu, P. F., & Li, W. J. (2025). Institutional investor behavior and stock market pricing efficiency. Shanghai Economic Research, (8).

[14] Guo, B. M., Yi, M., & Liu, P. C. (2022). Common ownership within fund families: Market structure and pricing efficiency. Journal of Financial Economics Research, (5).

[15] Jia, L. N., Bian, W. X., & Zhang, W. G. (2015). The impact of fund crowded trading on A-share stock price bubbles. Operations Research and Management Science, 24(5), 237–244.

[16] Hou, K., & Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. The Review of Financial Studies, 18(3), 981–1020.

[17] Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? Journal of Financial Economics, 58(1-2), 215–260.

[18] Durnev, A., Morck, R., Yeung, B., & Zarowin, P. (2003). Does greater firm-specific return variation mean more or less informed stock pricing? Journal of Accounting Research, 41(5), 797–836.

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