Pacific-Basin Finance Journal 83 (2024) 102238

Contents lists available at ScienceDirect

Pacific-Basin Finance Journal

journal homepage: www.elsevier.com/locate/pacfin

Retail investor attention and analyst earnings forecasts: Evidence from China

Zhida Zhang, Qi Luo *

Economics and Management School, Wuhan University, Wuhan 430072, China

ARTICLE INFO

JEL classification:
G14
G20

Keywords:
Retail investor attention
Internet searching
Analyst earnings forecast accuracy
Forecast error
Forecast dispersion

ABSTRACT

This article examines the impact of retail investor attention on analyst earnings forecast accuracy. Using a dataset of 21,238 firm-year observations from the Chinese A-share listed firms between 2011 and 2021, we find that future analyst forecast error and dispersion are higher for firms with higher search volumes on the internet, suggesting that retail investor attention has a significant negative impact on analyst earnings forecast accuracy. Further analysis shows that higher retail investor attention could impair the stock price informativeness and induce a greater level of earnings management of the underlying firms, supporting that the informational feedback effect of stock prices and opportunistic behaviors of firm managers are possible mechanisms through which retail investor attention deteriorates analyst earnings forecast accuracy. Moreover, our results in heterogeneity analysis reveal that the negative relationship between retail investor attention and analyst earnings forecast accuracy is more pronounced for non-Big4 audit firms, non-shortable firms, and firms that are not in the Mainland-Hong Kong stock connect program.

1. Introduction

Financial analyst, as an important information intermediary, has been playing a key role in processing and disseminating information for participants in the stock market. Among the research reports made by financial analysts, earnings forecast draws significant attention from both investors and listed firms. From the perspective of investors, based on the professional skill and information advantage of financial analysts, earnings forecast could convey valuable firm-specific information to investors by analyzing the operating performance and future prospects of the underlying firms, making it an important reference for investors' investment decisions (Frankel et al., 2006). As for the perspective of listed firms, analyst earnings forecasts can not only provide firm managers with insightful views from outsiders on the firms' business outlook but also achieve monitoring effects on firm managers, thus influencing the behavior of firm managers as well as the firms' real investment decisions (Yu, 2008; Chen et al., 2015). For example, Chen et al. (2017) demonstrate that higher analyst forecast quality alleviates over-investment and under-investment of the underlying firm, implying that analyst forecast accuracy benefits corporate investment efficiency. Therefore, analyst earnings forecasts with higher accuracy and better quality are crucial for the improvement of the information environment as well as the optimization of resource allocation in the capital market.

Considering the important function of analyst earnings forecast, a growing literature has documented several factors that determine its accuracy, including but not limited to analyst ability (Clement, 1999), information quality (Behn et al., 2008; Wang et al., 2021a,

* Corresponding author at: Economics and Management School, Wuhan University, 299 Bayi Rd, Wuhan, China.
E-mail address: luoqi@whu.edu.cn (Q. Luo).

https://doi.org/10.1016/j.pacfin.2023.102238
Received 12 April 2023; Received in revised form 11 November 2023; Accepted 19 December 2023
Available online 23 December 2023

2021b), decision fatigue (Hirshleifer et al., 2019), and social learning (Kumar et al., 2022). However, as far as we can tell, little is known about how investor attention affects analyst forecasts. Attention is a cognitive resource that determines how investors assimilate, apprehend, and act on information (Hirshleifer et al., 2009). Due to the paucity of data, earlier studies could only measure investor attention through indirect methods, including trading volume (Gervais et al., 2001), stock turnover (Barber and Odean, 2008), and media coverage (Fang and Peress, 2009). Thanks to the rapid development of the internet in recent years, a large number of studies have employed data, which are directly related to investor activities, from a variety of internet platforms to measure retail investor attention, such as Google (Da et al., 2011), Baidu (Wen et al., 2019), and Twitter (Rakowski et al., 2021).

Extant research has not reached a consensus on the role of retail investor attention in the stock market. On the one hand, a stream of prior research argues that retail investors behave as noise traders and higher retail investor attention is associated with more pronounced stock market anomalies and lower price efficiency (Foucault et al., 2011; Jiang et al., 2022). Hence, we conjecture that retail investor attention could deteriorate analyst earnings forecast accuracy by distorting the price efficiency and operating performance of the underlying firm. On the other hand, previous studies also demonstrate that higher retail investor attention indicates that investors are allocating more effort to acquire and analyze firm-specific information, which is beneficial to reducing information asymmetry and increasing information transparency (Wen et al., 2019; Chen and Wu, 2022). Accordingly, we propose another point of view that retail investor attention could improve analyst earnings forecast accuracy by enhancing the information environment of the underlying firms and providing incremental information to analysts.

We focus our research attention on China for the following reasons. First, the Chinese stock market is the largest emerging market in the world, with relatively low market efficiency and a poor information environment compared to mature markets (Piotroski et al., 2015). In addition, the Chinese stock market is dominated by unprofessional retail investors and has the largest investor base on the internet, which supports our measure of retail investor attention using internet search volume. Second, the analyst market in China is notorious for low barriers to entry and fierce competition, leading to low forecast accuracy and poor report quality (Yang et al., 2023). Third, while the market function and mechanism of the stock market in China still fall behind relative to developed markets, the Chinese stock market has initiated a variety of pilot programs, such as securities lending, which allows the short selling of particular stocks, and the Mainland-Hong Kong stock connect program, which is a form of capital market openness. Evidence related to these pilot programs could not only enrich our research context but also reveal meaningful policy implications for other emerging markets.

Based on 21,238 firm-year observations of Chinese listed firms over the period 2011–2021, we find that future analyst forecast error and forecast dispersion are positively associated with retail investor attention, indicating that retail investor attention has a negative impact on analyst earnings forecast accuracy. To address potential endogeneity problems, two-stage least square (2SLS) analysis, additional control variables, and the PSM method are employed. Moreover, we conduct a series of tests to verify the robustness of our results, including using the alternative analyst forecast accuracy measure, using the alternative retail investor attention measures, and screening out extreme sample periods. In further analysis, we find evidence that the distortion of stock price informativeness and opportunistic behaviors of firm managers could be possible channels through which retail investor attention inflicts analyst earnings forecast accuracy. Besides, cross-sectional analysis shows that the negative association between retail investor attention and analyst earnings forecast accuracy differs among firms with different information environments.

Our contributions to the literature are as follows. First, our study brings new insights into the literature on the economic consequences of retail investor attention. To the best of our knowledge, this is the first research that focuses on studying how retail investor attention affects analyst forecasts. Second, our findings supplement the growing literature on the determinants of analyst earnings forecast accuracy. Financial analysts in the stock market are considered as an important information intermediary that produces and disseminates valuable information for investors and firms. However, we find that analysts themselves could be confused and misled by retail investor attention in the market. Finally, we provide evidence that better audit quality, implementation of short selling mechanism, and stock market liberalization could help alleviate the negative influence of retail investor attention on analyst earnings forecast accuracy, revealing important implications not only for stock market participants but also for policymakers and regulators.

The remainder of this paper is organized as follows. Section 2 discusses the literature review and hypothesis development. Section 3 describes the data, main variables, and model. Section 4 presents the empirical results. Section 5 focuses on further analysis. And Section 6 concludes the study.

2. Literature review and hypothesis development

2.1. The determinants of analyst earnings forecast accuracy

Analyst earnings forecast accuracy measures how well financial analysts estimate the expected quarterly or annual earnings per share of the underlying firm. Athanassakos and Kalimipalli (2003) suggest that the information content of analyst earnings forecast plays an important role in asset pricing, portfolio management, and trading strategies. We categorize prior research relevant to analyst earnings forecast accuracy into two strands of studies. The first strand of literature mainly focuses on analyst optimistic bias. O'Brien (1988) demonstrates that analysts tend to exhibit optimistic bias in earnings forecasts. Later studies have identified several reasons for this phenomenon, including pressure from clients (Lin and McNichols, 1998), misjudgment (Easterwood and Nutt, 1999), pressure from employers (Cowen et al., 2006), and career concern (Hong and Kubik, 2003). More recent studies have broadened this line of research from other perspectives. For example, Wu et al. (2018) show that investor sentiment could aggravate the optimism bias in analyst forecasts, leading to higher forecast errors. Moreover, analysts' optimistic behavior is also driven by their self-interest. Hu et al. (2021) show that share pledging intensifies analyst optimism bias in earnings forecasts, which is the result of the collusion between controlling shareholders and analysts. Huang et al. (2022) find that analysts have the motivation to issue optimistic earnings forecasts after stock dividend announcements in the pursuit of personal gain.

The second strand of literature mainly discusses factors related to the analyst information acquisition process, such as information disclosure (Lang and Lundholm, 1996; Dhaliwal et al., 2012), corporate governance (Bhat et al., 2006), product market competition (Haw et al., 2015), analyst market competition (Merkley et al., 2017), corporate site visits (Han et al., 2018), and stock price informativeness (Wang et al., 2021a, 2021b). Apart from the factors listed above, analysts' ability and skill are the key determinants of their forecast accuracy. Clement (1999) shows that while earnings forecast accuracy increases with analyst ability and available resources, it decreases with the task complexity of the analyst. When faced with uncertainty, analyst's judgment and decision-making process could be affected. Zhang (2010) reveals that when faced with higher information uncertainty, analysts tend to issue earnings forecasts with less accuracy. Moreover, external shocks and exposure to risk also influence analyst behavior. Kong et al. (2021) investigate the impact of natural disasters on analyst earnings forecasts and find that earthquakes could distract analysts' attention, inducing analysts to present irrational pessimism on earnings forecasts. Yusoff et al. (2023) suggest that a firm's exposure to foreign exchange risk is positively associated with analyst forecast error and dispersion.

Nevertheless, extant literature pays little attention to the potential influence of retail investor attention on analyst earnings forecast accuracy. And our study aims to fill this gap in the literature by investigating the relationship between the two.

2.2. The effect of retail investor attention

Investor attention has been a widely discussed research topic in the field of behavioral finance. Prior literature related to this topic mainly focuses on investigating the influence of investor attention on asset pricing and the informational role of investor attention. Due to limited attention, the incomplete information set of investors could influence asset prices (Merton, 1987). Many studies have identified the effect of investor attention on several pricing issues in the stock market, including return comovement (Peng and Xiong, 2006), under- and overreactions (Hirshleifer et al., 2011), IPO pricing (Huang and Zhang, 2020), market anomalies (Jiang et al., 2022), etc. The development of internet search engines and online social media has been conducive to measuring investor attention, thus promoting research in the related field. Da et al. (2011) use the search volume index from Google trends to measure retail investor attention and find that increases in retail investor attention could inflate short-term stock prices. Based on data from Seeking Alpha, which is one of the most used internet stock forums by investors around the world, Da and Huang (2019) document that access to more public information may impair group decision-making, resulting in the loss of useful private information and reduction in the accuracy of the consensus forecast in the financial market. Moreover, noisy information on the internet could also drive investor attention. Clarke et al. (2021) find that fake news stories on the internet attract more investor attention than legitimate articles. However, the market reaction to fake news is discounted when compared with legitimate news.

In terms of the informational role of investor attention, previous studies have obtained mixed results. Most studies argue that investor attention can improve the information dissemination process and mitigate information asymmetry in the stock market (Fang and Peress, 2009; Wen et al., 2019; Chen and Wu, 2022). However, investor attention could also reflect noise and irrational behavior. Hervé et al. (2019) categorize investors into noise traders and smart traders by their online search behavior, and they find that noise traders' attention is positively associated with market volatility. Chen et al. (2022) show that retail investor attention is positively associated with post-earnings-announcement drift, which is inconsistent with prior research, and they suggest that the irrational behavior of investors could be a possible reason for this phenomenon. Furthermore, another strand of research reveals that investor attention could affect corporate actions. Hirshleifer and Teoh (2003) analyze the relationship between the limited attention of investors and firms' actions and find that limited attention could affect corporate decisions on information disclosure. Based on the limited attention theory, DellaVigna and Pollet (2009) point out that firms can take advantage of limited investor attention by disclosing negative information on Fridays. Moreover, there are also studies arguing that investor attention has a monitoring effect on corporate actions. For example, Kempf et al. (2017) demonstrate that firms with distracted shareholders are more likely to conduct inefficient actions, which is consistent with views on the monitoring role of investor attention. He et al. (2022) provide evidence that retail investor attention positively affects corporate green innovation by reducing information asymmetry, which could help relax financial constraints and mitigate agency costs.

However, few studies have been conducted on exploring the relationship between retail investor attention and financial intermediaries. We aim to extend this line of research by examining how retail investor attention affects analyst earnings forecasts and obtain new insights into the economic consequences of retail investor attention.

2.3. Hypothesis development

Attention plays a critical role in determining investors' trading activities and stock prices. Due to limited attention, retail investors are more likely to buy stocks that attract their attention, thus bringing upward pressure and noisy information toward the stock prices of the underlying firms (Barber and Odean, 2008; Foucault et al., 2011). Moreover, stock market anomalies are stronger for periods with higher investor attention, suggesting that investor attention matters for stock price efficiency (Nguyen and Pham, 2021; Jiang et al., 2022). Unlike institutional investors, who have access to professional tools to acquire information, retail investors are prone to obtain information through internet searches (Da et al., 2011). However, their attention is more likely to be attracted by fake news and noise on the internet, leading to misreaction (Clarke et al., 2021). In particular, the Chinese stock market is dominated by retail investors who lack expertise in the field of investing, which could intensify the extent of noise trading resulting from constrained attention (Chen et al., 2013). Besides, listed firms in the Chinese stock market are faced with a high level of short-sale constraints, which deters the integration process of negative information to stock prices. All these factors together could undermine the stock price

Table 1
Descriptive statistics.

Variable	N	Mean	Min	Median	Max	Std
FERROR	21,238	2.384	0.004	0.741	32.820	4.984
FDISP	21,238	1.511	0.011	0.523	20.010	3.047
RelAtt	21,238	11.790	6.629	11.730	13.400	0.607
Size	21,238	22.450	17.780	22.280	26.000	1.323
Lev	21,238	0.427	0.007	0.422	0.894	0.203
ROA	21,238	0.043	−1.859	0.0418	0.216	0.071
SOE	21,238	0.355	0.000	0.000	1.000	0.479
BM	21,238	1.104	0.011	0.703	6.249	1.166
INST	21,238	0.409	0.000	0.423	0.867	0.239
FirmAge	21,238	2.857	0.693	2.890	3.497	0.355
Big4	21,238	0.073	0.000	0.000	1.000	0.261
Loss	21,238	0.088	0.000	0.000	1.000	0.284
ANALYST	21,238	1.788	0.000	1.792	3.761	1.077
Growth	21,238	0.177	−0.985	0.116	2.559	0.393

Table 2
Univariate analysis.

Variables	Low retail investor attention observations ($N = 10,619$)		High retail investor attention observations ($N = 10,619$)		Mean test	Median test
	Mean	Median	Mean	Median		
FERROR	2.154	0.688	2.614	0.803	−0.460***	22.940***
FDISP	1.255	0.461	1.766	0.607	−0.511***	161.150***
Size	22.117	21.996	22.777	22.614	−0.660***	730.193***
Lev	0.387	0.375	0.467	0.474	−0.080***	628.671***
ROA	0.046	0.047	0.041	0.036	0.005***	251.252***
SOE	0.266	0	0.445	0	−0.179***	745.834***
BM	1.024	0.682	1.184	0.736	−0.160***	28.208***
INST	0.396	0.407	0.423	0.437	−0.027***	42.853***
FirmAge	2.86	2.944	2.853	2.89	0.007	10.494***
Big4	0.053	0	0.094	0	−0.041***	131.302***
Loss	0.082	0	0.094	0	−0.012***	9.435***
ANALYST	1.693	1.792	1.882	1.946	−0.189***	109.677***
Growth	0.174	0.122	0.181	0.108	−0.007	16.613***

efficiency of firms associated with high retail investor attention.

The feedback effect theory argues that market participants can learn incremental information from efficient stock prices (Bond et al., 2012). On the one hand, the market learning channel suggests that information in stock prices could affect the real decisions of firm managers and operating performance of the underlying firms (Chen et al., 2007; Xiao, 2020; Bennett et al., 2020). On the other hand, in the information acquisition process, firm managers and analysts treat stock prices as an important source of information for earnings forecasts and more informative stock prices could improve their forecast accuracy (Wang et al., 2021). Meanwhile, distractions from low-quality information and cognitive bias could cause interference in analyst forecast behavior (Wang et al., 2022). Consequently, we would expect that with the price pressure and noisy information brought by retail investor attention, firm operations and analysts' information sets could be distorted by the deteriorated information content of stock prices, thus resulting in lower analyst earnings forecast accuracy. In addition, attention partially reflects investors' irrational expectations toward the underlying firm. The literature has documented that to take advantage of the irrational behaviors of retail investors, firm managers tend to engage in opportunistic behaviors such as earnings management (Simpson, 2013; Li et al., 2023), which may lead to poor earnings quality of the firm and make it more challenging for analysts to predict future earnings. Based on the above discussions, we postulate the following testable hypothesis:

Hypothesis 1a. Retail investor attention has a negative effect on analyst earnings forecast accuracy.

From another perspective, the positive informational role of investor attention could benefit analyst earnings forecasts. While access to information is an essential determinant of investment decisions (Nofsinger, 2001), retail investors' information set is constrained by the cost of searching for and processing new information (Ying et al., 2015). However, internet search engines and online media platforms could significantly reduce the cost of obtaining information related to specific firms for retail investors (Fang and Peress, 2009; Hao and Xiong, 2021). In other words, higher retail investor attention measured by internet searching could effectively mitigate information asymmetry between firms and investors, thus ameliorating the information environment of the financial market. Furthermore, since the internet is an interactive platform, retail investor attention in the internet could also facilitate the dissemination of useful information among capital market participants (Rakowski et al., 2021). In addition, when a firm is associated with higher retail investor attention, it would be more difficult for the firm manager to withhold negative news from investors, which would alleviate the agency cost of the underlying firm (Wen et al., 2019). According to the above analysis, retail investor attention could play

Table 3
Baseline results.

	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)
$ReIAtt_t$	1.1056***	0.6959***	0.9273***	0.6080***
	(6.798)	(7.412)	(5.552)	(6.313)
$Size_t$			1.0393***	0.5615***
			(6.568)	(6.236)
Lev_t			−3.0610***	−1.8993***
			(−5.536)	(−5.599)
ROA_t			−10.1690***	−5.7582***
			(−7.067)	(−6.472)
SOE_t			0.9707*	0.3684
			(1.921)	(1.428)
BM_t			0.2199**	0.1505***
			(2.456)	(2.799)
$INST_t$			−0.6002**	−0.2044
			(−2.108)	(−1.145)
$FirmAge_t$			−1.6600**	−0.8962**
			(−2.174)	(−2.039)
$Big4_t$			−0.3371	0.1034
			(−0.682)	(0.363)
$Loss_t$			0.3013	0.5250***
			(1.045)	(2.805)
$ANALYST_t$			0.0436	−0.0008
			(0.677)	(−0.021)
$Growth_t$			−1.1469***	−0.5281***
			(−9.794)	(−7.609)
Constant	−9.8363***	−6.3168***	−24.3955***	−14.1833***
	(−5.143)	(−5.751)	(−6.278)	(−6.176)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Adj R^2	0.0107	0.0147	0.0402	0.0423
N	18,040	18,040	18,040	18,040

This table reports the results of baseline regressions which investigate how analyst earnings forecast accuracy in the next year is related to retail investor attention. The dependent variables are FERROR and FDISP. The variable of interest is retail investor attention ReIAtt. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

an informational and monitoring role in enhancing firm transparency and corporate governance, which exerts analysts to process firm-specific information more efficiently and enriches their information sets, thus improving analyst earnings forecast accuracy. Therefore, we propose the opposing hypothesis as follows:

Hypothesis 1b. Retail investor attention has a positive effect on analyst earnings forecast accuracy.

3. Data and methodology

3.1. Data

Our sample is based on all Chinese A-share listed firms from 2011 to 2021. We collect data on stock returns and financial data of listed firms from the CSMAR (China Stock Market and Accounting Research) database. Data on analyst earnings forecasts are obtained from the WIND database. Data on retail investor attention is from the Web Search Volume Index of Chinese Listed Companies (WSVI) database. We exclude (1) firms in the financial services industry, (2) Special Treatment (ST) firms and Particular Transfer (PT) firms, and (3) firm-year observations with missing variable data. Our final sample includes 21,238 firm-year observations. All variables are winsorized at the 1st and 99th percentiles to mitigate the effect of outliers.

3.2. Main variables

3.2.1. Analyst forecast accuracy

Consistent with prior literature, we use forecast error (FERROR) and forecast dispersion (FDISP) to measure analyst forecast accuracy. Firstly, we follow Merkley et al. (2017) and compute the forecast error of the underlying stock as the absolute value of the difference between the mean forecasted earnings per share by analysts and the actual earnings per share over the absolute value of the actual earnings per share:

Table 4
2SLS analysis.

	1st Stage	2nd Stage		1st Stage	2nd Stage	
	$ReIAtt_t$	$FERROR_{t+1}$	$FDISP_{t+1}$	$ReIAtt_t$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
$ReIAttPeer_{t-1}$	0.0743***					
	(4.619)					
AD_{t-1}				0.0322***		
				(7.875)		
$\widehat{ReIAtt}_t$		8.2620**	7.2197***		7.5353***	5.6484***
		(2.071)	(2.688)		(3.248)	(3.797)
$Size_t$	0.2063***	−0.5357	−0.8002	0.2052***	−0.3844	−0.4708
	(18.178)	(−0.621)	(−1.376)	(18.228)	(−0.716)	(−1.391)
Lev_t	0.1915***	−4.5880***	−3.5502***	0.1982***	−4.4331***	−3.2265***
	(4.735)	(−4.355)	(−4.951)	(4.940)	(−5.479)	(−6.027)
ROA_t	0.0275	−9.0909***	−4.9494***	−0.0008	−9.0709***	−4.9091***
	(0.373)	(−5.394)	(−4.446)	(−0.011)	(−5.462)	(−4.628)
SOE_t	0.0992***	0.2195	−0.3645	0.1042***	0.2867	−0.2083
	(3.882)	(0.335)	(−0.905)	(4.095)	(0.483)	(−0.621)
BM_t	−0.1308***	1.2551**	1.0704***	−0.1313***	1.1585***	0.8609***
	(−21.429)	(2.365)	(2.975)	(−21.742)	(3.566)	(4.088)
$INST_t$	−0.4864***	2.4301	2.6412**	−0.4797***	2.0678*	1.8667**
	(−21.040)	(1.227)	(1.974)	(−20.981)	(1.788)	(2.515)
$FirmAge_t$	0.3685***	−5.1330***	−4.1865***	0.3892***	−4.8489***	−3.5780***
	(6.022)	(−2.752)	(−3.325)	(6.378)	(−3.658)	(−4.170)
$Big4_t$	0.0293	−0.0104	0.1542	0.0251	0.0122	0.2007
	(1.040)	(−0.019)	(0.427)	(0.882)	(0.023)	(0.603)
$Loss_t$	0.0568***	0.1330	0.3640	0.0554***	0.1760	0.4522**
	(4.743)	(0.333)	(1.330)	(4.667)	(0.507)	(1.986)
$ANALYST_t$	−0.0094**	0.1746**	0.1058*	−0.0116***	0.1667**	0.0908*
	(−2.362)	(2.026)	(1.826)	(−2.939)	(2.016)	(1.718)
$Growth_t$	−0.0221***	−1.0498***	−0.4490***	−0.0243***	−1.0569***	−0.4801***
	(−3.464)	(−6.338)	(−4.157)	(−3.845)	(−7.359)	(−5.354)
Constant	5.6223***	−69.2574***	−54.8282***	6.3390***	−64.5442***	−44.787***
	(15.553)	(−2.679)	(−3.158)	(21.031)	(−4.214)	(−4.518)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
1st stage *p*-value	0.0000			0.0000		
Cragg-Donald Wald F	32.174			115.885		
Kleibergen-Paap rk F	21.338			62.017		
R-squared	0.6805			0.6828		
N	14,905	14,905	14,905	14,905	14,905	14,905

This table shows the regression results of the 2SLS analysis. The first instrument ReIAttPeer is calculated as the lagged natural logarithm of the average retail investor attention of the peer firms in the same industry plus one. The second instrument AD is calculated as the lagged natural logarithm of the advertising spending of the firm plus one. The first-stage regression generates the instrumented value of ReIAtt, which is the independent variable used in the second-stage regression. We also report the 1st stage *p*-value, the Cragg-Donald Wald F statistic, and the Kleibergen-Papp rk F statistic in this table. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

$$FERROR_{i,t} = \frac{\left|mean\left(FEPS_{i,t,j}\right) - EPS_{i,t}\right|}{\left|EPS_{i,t}\right|} \tag{1}$$

where $FERROR_{i,t}$ is the forecast error of stock *i* at time *t*, $FEPS_{i,t,j}$ is the forecasted earnings per share of stock *i* at time *t* by analyst *j*, $EPS_{i,t}$ is the actual earnings per share of stock *i* at time *t*. A higher value of *FERROR* indicates a less accurate analyst forecast.

Secondly, we define forecast dispersion as the standard deviation of analysts' earnings forecasts, scaled by the absolute value of the actual earnings per share:

$$FDISP_{i,t} = \frac{Std\left(FEPS_{i,t}\right)}{\left|EPS_{i,t}\right|} = \frac{1}{\left|EPS_{i,t}\right|} \times \sqrt{\frac{\sum_{j=1}^{N_{i,t}}\left(FEPS_{i,t,j} - \overline{FEPS_{i,t,j}}\right)^2}{N_{i,t}-1}} \tag{2}$$

where $FDISP_{i,t}$ is the forecast dispersion of stock *i* at time *t*, $\overline{FEPS_{i,t,j}}$ is the average forecast error of all analysts who cover stock *i* at time *t*. A higher value of *FDISP* means that analysts have more dispersed opinions toward the earnings per share of the underlying stock, implying a less accurate analyst forecast.

Table 5
Regressions with additional control variables.

	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)
$ReIAtt_t$	0.9371***	0.6152***
	(5.600)	(6.385)
$Size_t$	1.0279***	0.5561***
	(6.504)	(6.185)
Lev_t	−3.0526***	−1.8860***
	(−5.521)	(−5.555)
ROA_t	−10.0616***	−5.7028***
	(−7.002)	(−6.411)
SOE_t	0.9615*	0.3663
	(1.909)	(1.427)
BM_t	0.2179**	0.1478***
	(2.433)	(2.750)
$INST_t$	−0.6180**	−0.2123
	(−2.171)	(−1.188)
$FirmAge_t$	−1.6298**	−0.8901**
	(−2.148)	(−2.036)
$Big4_t$	−0.3161	0.1155
	(−0.640)	(0.406)
$Loss_t$	0.3017	0.5234***
	(1.047)	(2.795)
$ANALYST_t$	0.0436	−0.0005
	(0.676)	(−0.011)
$Growth_t$	−1.1468***	−0.5294***
	(−9.772)	(−7.616)
MA_t	−0.0784	−0.0621
	(−0.842)	(−1.116)
CL_t	−0.4826**	−0.2433**
	(−2.472)	(−2.143)
IP_t	0.1230	−0.0475
	(0.605)	(−0.375)
Constant	−24.1475***	−14.0007***
	(−6.230)	(−6.115)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Adj R^2	0.0407	0.0428
N	18,040	18,040

This table presents the regression results with additional control variables. The dependent variables are FERROR and FDISP. The variable of interest is retail investor attention ReIAtt. We include additional control variables, including MA, CL, and IP, in the regressions. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

3.2.2. *Retail investor attention*

Following Da et al. (2011), we use the internet search volume index of listed firms to measure retail investor attention. Our data is collected from the Web Search Volume Index of Chinese Listed Companies (WSVI) database. As for now, WSVI database is the only platform that provides a comprehensive measure of internet search volume index of listed firms for the Chinese stock market. WSVI is constructed based on the search frequencies of individual stocks' tickers, brief names, and full names on large internet search engines, including Baidu and Sina, both of which are heavily used by retail investors in China. Consistent with Wen et al. (2019), we use 1 plus the natural logarithm of the search volume index of a firm's ticker within a year as our main proxy for retail investor attention *ReIAtt* because, while search activities related to a firm's name may reflect things unrelated to investing, ticker searches are more likely to represent investors' interest and attention in the underlying stock. Since WSVI was established in 2011, we set our sample period from 2011 to 2021. A higher value of *ReIAtt* represents a higher level of retail investor attention.

3.3. *Model*

To examine the effect of retail investor attention on analyst earnings forecasts, we consider our baseline empirical model as follows:

$$Forecast_{i,t+1} = \beta_0 + \beta_1 ReIAtt_{i,t} + \gamma Controls + \mu_t + \mu_i + \varepsilon_{i,t} \quad (3)$$

where $Forecast_{i,t+1}$ measures analyst forecast accuracy of firm i in year $t+1$ and is proxied by $FERROR_{i,t+1}$ or $FDISP_{i,t+1}$. $ReIAtt_{i,t}$ measures retail investor attention of firm i in year t. μ_t and μ_i are dummy variables for year-fixed and firm-fixed effects.

We use *Controls* to represent a set of control variables that may influence analyst earnings forecasts. Following prior studies (Kong

Table 6
Regressions with matched sample after PSM.

	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)
Estimation Method	1:1 matching	
$ReIAtt_t$	0.9408***	0.5678***
	(3.604)	(3.914)
$Size_t$	1.0473***	0.6279***
	(4.013)	(4.194)
Lev_t	−3.3934***	−2.5278***
	(−3.840)	(−4.765)
ROA_t	−7.6617***	−5.5549***
	(−3.393)	(−3.745)
SOE_t	1.3318	0.5518
	(1.634)	(1.259)
BM_t	0.2622	0.1657*
	(1.557)	(1.699)
$INST_t$	−1.4969***	−0.5654**
	(−3.202)	(−1.997)
$FirmAge_t$	−2.9894***	−1.9816***
	(−2.797)	(−3.076)
$Big4_t$	−0.6067	0.0080
	(−0.732)	(0.018)
$Loss_t$	1.0884**	0.8165***
	(2.318)	(2.823)
$ANALYST_t$	0.0481	0.0090
	(0.464)	(0.146)
$Growth_t$	−1.2100***	−0.5652***
	(−7.549)	(−5.923)
Constant	−21.0278***	−12.0367***
	(−3.540)	(−3.411)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Adj. R^2	0.0502	0.0515
N	8565	8565

This table shows the regression results using matched sample after PSM. Each observation in the high retail investor attention group is matched with one observation in the low retail investor attention group using the nearest neighbor matching method. The dependent variables are FERROR and FDISP. The variable of interest is retail investor attention ReIAtt. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

et al., 2020; Huang et al., 2022), *Size* is the natural logarithm of total assets. *Lev* is the debt to asset ratio. *ROA* is the net income over total assets. *SOE* is a dummy variable that equals 1 when the firm is controlled by a state entity and 0 otherwise. *BM* is the book to market ratio of equity. *INST* is the institutional investor ownership percentage of the firm. *FirmAge* is the natural logarithm of the age of the firm. *Big4* is a dummy variable that equals 1 when the firm is audited by a Big 4 audit firm and 0 otherwise. *Loss* is a dummy variable that equals 1 when the net income of the firm is negative and 0 otherwise. *ANALYST* is 1 plus the natural logarithm of the number of analysts covering the firm. *Growth* is the revenue growth rate of the firm.

4. Empirical results

4.1. Descriptive statistics

Table 1 shows the descriptive statistics of analyst forecast accuracy measures, retail investor attention, and firm characteristics in the data set. The mean (median) value of *FERROR* is 2.384 (0.741), with a standard deviation of 4.984. As for *FDISP*, the mean (median) value is 1.511 (0.523), with a standard deviation of 3.047. These numbers are similar to prior studies and suggest that both *FERROR* and *FDISP* exhibit left-skewed distributions and vary across different firms and years. The mean (median) value of *ReIAtt* is 11.79 (11.73), with a standard deviation of 0.607, and the minimum (maximum) value of *ReIAtt* is 6.629 (13.4), implying that retail investor attention on individual firms differs across our sample. Moreover, the descriptive statistics for firm characteristic variables of our sample are consistent with prior literature.

4.2. Univariate analysis

We divide our sample into two groups by the median value of retail investor attention and run a univariate analysis to provide preliminary evidence of our analysis. Table 2 denotes that *FERROR* and *FDISP* are higher for firms associated with higher retail investor

Table 7
Robustness tests: change dependent variable, independent variable, and sample period.

	$FOPT_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ReIAtt_t$	0.0218***					0.9902***	0.6713***
	(2.634)					(5.506)	(6.636)
$ReIAtt2_t$		0.8600***	0.5710***				
		(5.445)	(6.182)				
$ReIAtt3_t$				0.3068***	0.1392***		
				(3.829)	(2.943)		
$Size_t$	0.1030***	1.0745***	0.5832***	1.1507***	0.6537***	1.0014***	0.5314***
	(11.719)	(6.825)	(6.480)	(7.313)	(7.195)	(5.943)	(5.559)
Lev_t	−0.2103***	−3.0616***	−1.9019***	−2.8357***	−1.7434***	−3.1211***	−1.9528***
	(−7.357)	(−5.514)	(−5.588)	(−5.143)	(−5.186)	(−5.208)	(−5.401)
ROA_t	−0.0337	−10.4626***	−5.9509***	−10.4403***	−5.9360***	−10.2748***	−5.9161***
	(−0.642)	(−7.320)	(−6.713)	(−7.267)	(−6.663)	(−6.673)	(−6.306)
SOE_t	0.0243	0.9829*	0.3758	0.9893*	0.3931	1.0575*	0.3903
	(1.206)	(1.935)	(1.442)	(1.957)	(1.521)	(1.946)	(1.410)
BM_t	0.0100*	0.2118**	0.1463***	0.1399	0.0859	0.2131**	0.1491***
	(1.748)	(2.358)	(2.685)	(1.577)	(1.603)	(2.251)	(2.608)
$INST_t$	0.0004	−0.6228**	−0.2168	−0.5905**	−0.2648	−0.5887*	−0.1753
	(0.027)	(−2.193)	(−1.224)	(−1.987)	(−1.420)	(−1.938)	(−0.929)
$FirmAge_t$	−0.0592	−1.5114**	−0.8013*	−1.4120*	−0.6915	−1.7045**	−0.9715**
	(−1.430)	(−2.001)	(−1.835)	(−1.863)	(−1.577)	(−2.224)	(−2.206)
$Big4_t$	−0.0759***	−0.3444	0.0986	−0.3657	0.0893	−0.3684	0.0495
	(−2.796)	(−0.693)	(0.344)	(−0.734)	(0.311)	(−0.695)	(0.163)
$Loss_t$	0.0057	0.2821	0.5121***	0.2841	0.5216***	0.4829	0.5098**
	(0.561)	(0.979)	(2.739)	(0.984)	(2.789)	(1.506)	(2.524)
$ANALYST_t$	0.0221***	0.0279	−0.0111	0.0259	−0.0128	0.0316	−0.0110
	(6.700)	(0.432)	(−0.280)	(0.400)	(−0.322)	(0.456)	(−0.267)
$Growth_t$	−0.0642***	−1.1512***	−0.5310***	−1.1395***	−0.5270***	−1.1894***	−0.5324***
	(−10.816)	(−9.818)	(−7.630)	(−9.764)	(−7.591)	(−9.351)	(−7.074)
Constant	−1.4764***	−25.3424***	−14.8576***	−18.9325***	−10.5948***	−24.1948***	−14.0582***
	(−7.096)	(−6.471)	(−6.421)	(−5.106)	(−4.913)	(−5.864)	(−5.793)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj R^2	0.0429	0.0402	0.0423	0.0385	0.0397	0.0424	0.0449
N	18,040	18,040	18,040	18,040	18,040	16,066	16,066

This table reports the results of tests examining the robustness of the relationship between retail investor attention and analyst earnings forecast accuracy. The dependent variable FOPT in column (1) is an alternative measure of analyst forecast accuracy. The independent variable ReIAtt2 in columns (2) and (3) is an alternative measure of retail investor attention. The independent variable ReIAtt3 in columns (4) and (5) is another alternative measure of retail investor attention. Observations from year 2015 is removed in columns (6) and (7) to exclude potential extreme influence. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

attention, which is consistent with Hypothesis 1a.

4.3. Baseline results

Table 3 presents the estimation results of our baseline regression model. Columns (1) and (2) show that the coefficients for *ReIAtt* are 1.1056 and 0.6959, both statistically significant at the 1% level, denoting that firms with higher retail investor attention are faced with higher future analyst forecast error and dispersion. Similarly, after controlling for the potential determinants of analyst earnings forecast accuracy, Columns (3) and (4) of Table 3 show that the coefficients of *ReIAtt* still meet our expectations. In terms of the control variables, while *Size* is positively associated with analyst forecast error and dispersion, the coefficients of *Lev*, *ROA*, and *FirmAge* are negative, which is consistent with previous studies (Chen et al., 2021; Huang et al., 2022).

Overall, Table 3 provides evidence that after controlling for stock and firm characteristics, retail investor attention can positively affect future analyst forecast error and dispersion, indicating a negative relationship between retail investor attention and future analyst earnings forecast accuracy. These results are consistent with our Hypothesis 1a.

4.4. Endogeneity

In our baseline model, we have included a series of firm characteristics control variables, year-fixed and firm-fixed effects to alleviate some endogeneity concerns. Moreover, we have used the analyst forecast accuracy in year $t + 1$ and retail investor attention in year t to address the potential reverse causality problem. However, our findings may still suffer from potential endogeneity problems. To further address this concern, we employ the 2SLS analysis, include additional possibly omitted variables, and use the PSM method.

Table 8
Possible mechanism: the role stock price informativeness.

	SYN_{t+1}	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)
$ReIAtt_t$	0.3004***	1.0040***	0.6588***
	(13.252)	(5.887)	(6.705)
SYN_t		−1.9429*	−1.2837*
		(−1.882)	(−1.959)
$ReIAtt_t \times SYN_t$		0.1839**	0.1216**
		(2.111)	(2.193)
$Size_t$	0.0952***	0.9917***	0.5300***
	(4.324)	(6.244)	(5.861)
Lev_t	−0.2807***	−2.9911***	−1.8528***
	(−3.394)	(−5.424)	(−5.470)
ROA_t	0.3331*	−10.0338***	−5.6684***
	(1.906)	(−6.977)	(−6.386)
SOE_t	−0.0669	0.9588*	0.3605
	(−1.130)	(1.899)	(1.398)
BM_t	0.0404***	0.1984**	0.1362**
	(2.767)	(2.203)	(2.534)
$INST_t$	−0.0893**	−0.5696**	−0.1841
	(−2.065)	(−1.995)	(−1.027)
$FirmAge_t$	−0.2554**	−1.7772**	−0.9739**
	(−2.510)	(−2.325)	(−2.214)
$Big4_t$	0.0356	−0.3200	0.1147
	(0.499)	(−0.649)	(0.405)
$Loss_t$	−0.1538***	0.3014	0.5251***
	(−4.789)	(1.047)	(2.809)
$ANALYST_t$	0.0634***	0.0453	0.0003
	(6.988)	(0.704)	(0.008)
$Growth_t$	−0.0378**	−1.1163***	−0.5079***
	(−2.434)	(−9.531)	(−7.326)
Constant	−5.0034***	−24.0006***	−13.9203***
	(−9.431)	(−6.208)	(−6.096)
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Adj R^2	0.0406	0.0415	0.0438
N	18,040	18,040	18,040

This table shows the results of regressions examining the mechanism behind the relationship between retail investor attention and analyst earnings forecast. In column (1), how retail investor attention influences stock price synchronicity is investigated. In columns (2) and (3), SYN and its interaction with ReIAtt are added to the baseline regression. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

We first follow Chen and Wu (2022) and use *ReIAttPeer*, which is calculated as the lagged natural logarithm of the average retail investor attention of the peer firms in the same industry plus one, as an instrumental variable to perform a 2SLS regression. The logic behind this is that attention to peer firms is closely related to attention to a firm but not directly related to analyst earnings forecast accuracy of the firm. In other words, *ReIAttPeer* meets the relevance and exogeneity requirements to serve as a valid instrument. We present the 2SLS regression results in Table 4, where column (1) reports the first-stage regression results, and columns (2) and (3) show the results of the second-stage regression results. We can see that the coefficient on *ReIAttPeer* in column (1) is positive and significant at the 1% level. Moreover, the Cragg-Donald Wald F statistic and the Kleibergen-Paap rk F statistic, which are both greater than the 10% bias threshold in the Stock-Yogo weak ID test critical values, suggest that *ReIAttPeer* passes the weak instrumental variable test. Finally, the results in columns (2) and (3) indicate that the coefficients of $\widehat{ReIAtt}$ are both significantly positive, which is consistent with our baseline results. In addition, following Lou (2014), we also use advertising spending (*AD*), which is calculated as the natural logarithm of the advertising spending of the firm plus one, as another instrumental variable. Since advertising spending of the firm is related to the internet searching activity of investors but do not directly affect analyst earnings forecast, the relevance and exogeneity conditions to serve as a valid instrument are satisfied. The 2SLS results in columns (4)–(6) of Table 4 still confirm the robustness of our findings.

As Liu and Krystyniak (2021) and Muradoglu et al. (2024) have pointed out, investor attention increases significantly during the preannouncement period of mergers. Besides, M&A will bring a lot of uncertainty about the firm's operations. As a result, it could be much harder for analysts to forecast earnings. Moreover, Hu et al. (2022) demonstrate that disclosure of comment letters could improve analysts' earnings forecast accuracy. Wang et al. (2022) find that industrial policies lead to a decline in the accuracy of analysts' earnings forecasts. Therefore, to mitigate the potential omitted variable bias and better clarify the relationship between retail investor attention and analyst earnings forecast accuracy in our analysis, we augment our baseline regression model by including additional control variables, including *MA*, which is a dummy variable equals to 1 if a firm makes an M&A announcement in year t, and

Table 9
Possible mechanism: the role of earnings management.

	EM_{t+1}	REM_{t+1}
	(1)	(2)
$ReIAtt_t$	0.0033**	0.0066***
	(2.036)	(2.610)
$Size_t$	−0.0100***	−0.0268***
	(−5.409)	(−7.139)
Lev_t	0.0259***	−0.0031
	(4.336)	(−0.284)
ROA_t	0.0407***	−0.0940***
	(2.637)	(−4.506)
SOE_t	0.0011	−0.0034
	(0.294)	(−0.551)
BM_t	−0.0059***	−0.0011
	(−5.909)	(−0.529)
$INST_t$	0.0025	0.0034
	(0.826)	(0.683)
$FirmAge_t$	0.0080	0.0034
	(1.115)	(0.291)
$Big4_t$	−0.0002	0.0001
	(−0.057)	(0.008)
$Loss_t$	0.0080***	−0.0055
	(3.457)	(−1.425)
$ANALYST_t$	0.0004	−0.0015
	(0.597)	(−1.445)
$Growth_t$	0.0029**	0.0049*
	(2.188)	(1.661)
Constant	0.2003***	0.5073***
	(4.788)	(6.282)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Adj R^2	0.0385	0.0397
N	18,040	18,040

This table shows the results of regressions examining how retail investor attention affects earnings management. In column (1), the dependent variable EM is the absolute value of discretionary accruals. In column (2), the dependent variable REM is the absolute value of real earnings management. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

0 otherwise, *CL*, which is a dummy variable equals to 1 if a firm receives a comment letter from stock exchanges in year t, and 0 otherwise, and *IP*, which is a dummy variable equals to 1 if a firm turns out to be in the industry that experienced an industrial policy change made by the Chinese government in year t, and 0 otherwise. The results are presented in Table 5, we can see that the coefficients on *ReIAtt* are both positive and significant, suggesting that our findings still hold.

To exclude the possibility that our results are driven by the inherent differences between firms receiving high retail investor attention and firms with low retail investor attention, we use propensity score matching (PSM) to address the potential sample selection bias in our analysis. Firstly, we divide our sample into two groups based on the median value of retail investor attention. Secondly, we perform logistic regression on the likelihood of a firm having high retail investor attention with all the control variables in our baseline model to obtain the propensity score. Thirdly, we use the nearest neighbor matching method to match each observation in the high retail investor attention group with one observation in the low retail investor attention group. Finally, we re-estimate our model using the matched sample. Table 6 shows the regression results after PSM, which is consistent with our baseline results.

4.5. Robustness tests

In this subsection, we implement a series of robustness tests to guarantee that our findings are robust. Firstly, we follow Huang et al. (2022) and use analyst optimistic bias (*FOPT*), which is the difference between the average value of all analysts' forecasted earnings per share for the specific firm and the actual earnings per share over the absolute value of the actual earnings per share, as an alternative measure of analyst earnings forecast accuracy. From column (1) of Table 7, we can see that the coefficient on *ReIAtt* remains positive and significant at the 1% level, which is consistent with our previous results.

Secondly, we change our independent variable and use *ReIAtt*2, which is the internet search volume index of all related keywords of the underlying firm from WSVI, as another proxy for retail investor attention. The results are shown in columns (2) and (3) of Table 7. The estimated coefficients on *ReIAtt*2 are 0.8600 and 0.5710 respectively, both positive and significant at the 1% level, indicating that our results are robust. Moreover, following the method used by Antweiler and Frank (2004) and Rakowski et al. (2019), we also use *ReIAtt3*, which is the number of posts related to specific firms from the Internet stock message board called Guba in China, to measure

Table 10
The effect of M&A.

			M&A group	Non-M&A group	M&A group	Non-M&A group
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
MA_t	−0.0683	−0.0542				
	(−0.733)	(−0.972)				
$RelAtt_t$			0.9179**	0.9418***	0.6772***	0.5951***
			(2.505)	(4.840)	(3.258)	(5.189)
$Size_t$	1.2452***	0.6965***	1.0863***	1.0645***	0.4819**	0.6240***
	(7.777)	(7.517)	(3.252)	(5.836)	(2.482)	(5.890)
Lev_t	−2.7858***	−1.7176***	−2.4041*	−3.5278***	−1.4327**	−2.2399***
	(−5.045)	(−5.115)	(−1.904)	(−5.179)	(−2.006)	(−5.254)
ROA_t	−10.4296***	−5.9276***	−5.1836	−12.0545***	−3.8413**	−6.6160***
	(−7.242)	(−6.654)	(−1.609)	(−7.127)	(−1.985)	(−6.232)
SOE_t	1.0481**	0.4187	1.1188	0.9540	0.3879	0.4047
	(2.062)	(1.612)	(1.387)	(1.594)	(0.764)	(1.420)
BM_t	0.0788	0.0579	0.4477***	0.0844	0.2668***	0.0802
	(0.893)	(1.092)	(2.715)	(0.783)	(2.655)	(1.236)
$INST_t$	−0.9163***	−0.4111**	−0.6183	−0.5609	−0.3169	−0.1071
	(−3.184)	(−2.289)	(−0.909)	(−1.610)	(−0.854)	(−0.482)
$FirmAge_t$	−1.2131	−0.6044	−2.1667*	−1.5171	−0.5524	−1.0290*
	(−1.609)	(−1.382)	(−1.756)	(−1.539)	(−0.747)	(−1.844)
$Big4_t$	−0.3449	0.0979	−0.2098	−0.1096	0.1384	0.2306
	(−0.692)	(0.342)	(−0.451)	(−0.190)	(0.419)	(0.706)
$Loss_t$	0.3217	0.5381***	1.0690	0.2370	0.9120*	0.5179**
	(1.115)	(2.874)	(1.447)	(0.713)	(1.830)	(2.398)
$ANALYST_t$	0.0256	−0.0124	−0.0686	0.0471	−0.1055	−0.0024
	(0.396)	(−0.312)	(−0.457)	(0.621)	(−1.182)	(−0.052)
$Growth_t$	−1.1555***	−0.5335***	−1.1596***	−1.1407***	−0.5779***	−0.5398***
	(−9.851)	(−7.657)	(−3.414)	(−8.882)	(−3.056)	(−6.708)
Constant	−18.8601***	−10.5483***	−25.1968***	−24.9196***	−14.6191***	−14.7679***
	(−5.076)	(−4.883)	(−3.194)	(−5.469)	(−2.928)	(−5.526)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R^2	0.0375	0.0391	0.0387	0.0441	0.0471	0.0463
N	18,040	18,040	4092	13,948	4092	13,948

This table shows the results of the analysis examining how M&A affects analyst earnings forecast and whether the relationship between retail investor attention and analyst earnings forecast is driven by M&A. MA is a dummy variable equals to 1 if a firm makes an M&A announcement in year t, and 0 otherwise. We split our sample into two groups by the value of MA. Columns (1) and (2) report how analyst earnings forecast accuracy is related to M&A, columns (3) and (5) report the estimation results from the M&A group, and columns (4) and (6) report the estimation results from the non-M&A group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

retail investor attention. The results in columns (4) and (5) of Table 7 also prove that our findings are robust.

Finally, because the Chinese stock market went through an extraordinary boom and bust period in year 2015, we exclude 2015 from our sample period to rule out the possibility that extreme market condition is the main driver of our results. Columns (6) and (7) of Table 7 show that after re-estimating the regression model without observations in 2015, retail investor attention still has a significantly positive influence on analyst forecast error and dispersion.

In summary, by changing measures of analyst earnings forecast accuracy and retail investor attention, as well as the sample period in our analysis, the negative and significant relationship between retail investor attention and analyst earnings forecast accuracy still holds. These findings provide compelling evidence that our results are robust.

5. Further analysis

In this section, we provide evidence on the possible channels through which retail investor attention affects future analyst earnings forecast accuracy. Specifically, we analyze and examine the role of stock price informativeness and earnings management. Furthermore, considering the indistinct role of retail investor attention in the financial market, we posit that the impact of retail investor attention on analyst earnings forecast accuracy should vary across different circumstances. Hence, we investigate the cross-sectional variation in the relationship between retail investor attention and analyst earnings forecast accuracy, conditional on audit quality, short-sale constraints, and stock market liberalization.

5.1. Possible mechanisms

Firstly, as we have discussed in hypothesis development, retail investor attention could distort stock price informativeness, which

Table 11
The effect of comment letter.

			Comment letter group	Non-Comment letter group	Comment letter group	Non-Comment letter group
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
CL_t	−0.4206**	−0.2451**				
	(−2.293)	(−2.317)				
$RelAtt_t$			0.7578***	1.0079***	0.5840***	0.6201***
			(3.039)	(4.152)	(4.142)	(4.293)
$Size_t$	1.2389***	0.6929***	1.2733***	1.1333***	0.6939***	0.6096***
	(7.743)	(7.486)	(4.137)	(6.063)	(3.880)	(5.395)
Lev_t	−2.7882***	−1.7209***	−2.7141***	−4.1053***	−1.9772***	−2.2610***
	(−5.053)	(−5.126)	(−3.087)	(−5.370)	(−3.839)	(−4.535)
ROA_t	−10.3631***	−5.8911***	−9.6097***	−9.0566***	−7.6260***	−4.4675***
	(−7.193)	(−6.605)	(−3.967)	(−5.297)	(−5.236)	(−4.062)
SOE_t	1.0503**	0.4207	0.7255	1.5644**	0.0961	0.7052*
	(2.072)	(1.629)	(0.980)	(2.187)	(0.285)	(1.873)
BM_t	0.0759	0.0564	0.2542	0.2266**	0.1322	0.1593***
	(0.860)	(1.065)	(1.211)	(2.463)	(1.031)	(2.749)
$INST_t$	−0.9454***	−0.4290**	−1.4119***	−0.2077	−0.7942***	0.1332
	(−3.283)	(−2.386)	(−3.335)	(−0.506)	(−2.940)	(0.511)
$FirmAge_t$	−1.1681	−0.5762	−0.8738	−1.1332	−0.2998	−0.7732
	(−1.554)	(−1.321)	(−0.826)	(−1.046)	(−0.459)	(−1.319)
$Big4_t$	−0.3195	0.1133	−1.2376	0.2444	−0.6458	0.7155*
	(−0.641)	(0.395)	(−1.427)	(0.367)	(−1.291)	(1.850)
$Loss_t$	0.3249	0.5403***	−0.4725	0.8742**	0.0511	0.9262***
	(1.126)	(2.889)	(−0.943)	(2.502)	(0.163)	(3.997)
$ANALYST_t$	0.0221	−0.0148	0.0498	0.1170	0.0212	0.0167
	(0.341)	(−0.371)	(0.518)	(1.278)	(0.347)	(0.307)
$Growth_t$	−1.1616***	−0.5375***	−1.2219***	−0.9613***	−0.5082***	−0.4619***
	(−9.891)	(−7.705)	(−6.236)	(−6.454)	(−4.354)	(−5.170)
Constant	−18.6172***	−10.4153***	−29.0162***	−29.2635***	−17.7041***	−16.0965***
	(−5.011)	(−4.830)	(−4.226)	(−5.412)	(−4.263)	(−5.111)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R^2	0.0380	0.0396	0.0338	0.0457	0.0408	0.0460
N	18,040	18,040	8145	9895	8145	9895

This table shows the results of the analysis examining how comment letter affects analyst earnings forecast and whether the relationship between retail investor attention and analyst earnings forecast is driven by comment letter. CL is a dummy variable equals to 1 if a firm receives a comment letter from stock exchanges in year t, and 0 otherwise. We split our sample into two groups by the value of CL. Columns (1) and (2) report how analyst earnings forecast accuracy is related to comment letter, columns (3) and (5) report the estimation results from the comment letter group, and columns (4) and (6) report the estimation results from the non-comment letter group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

in turn affects decision makers' information set and firm managers' operating decisions, thus making it harder for analysts to forecast earnings. Moreover, Zuo (2016) provides evidence that firm managers do learn from information in the stock prices of their own firms, and higher stock price informativeness helps managers improve their earnings forecast accuracy, which at least partially supports our analysis of the role of stock price informativeness on analyst earnings forecast accuracy. To prove our hypothesis, we follow Wang et al. (2021) and use stock price synchronicity (*SYN*), which measures how much firm-specific information is reflected in stock price movement, as the proxy for stock price informativeness and conduct more tests. Higher *SYN* represents lower stock price informativeness. We expect that higher retail investor attention will increase stock price synchronicity, which implies a decrease in stock price informativeness. And consistent with evidence provided by Wang et al. (2021) that better stock price informativeness could improve analyst forecast quality, we expect that analyst earnings forecast accuracy should be less sensitive to retail investor attention for firms associated with higher stock price informativeness. In column (1) of Table 8, we can see that retail investor attention has a significant positive association with future stock price synchronicity, which supports our hypothesis. In columns (2) and (3) of Table 8, the coefficients on $RelAtt \times SYN$ are both positive and significant, which is consistent with our argument that the negative effect of retail investor attention on analyst forecast accuracy is less pronounced for firms with higher stock price informativeness.

Secondly, given that investor attention partially reflects investors' irrational expectations ofs stock prices, firm managers with opportunistic purposes are prone to cater to investors through earnings management (Simpson, 2013; Li et al., 2023). Lower earnings quality will thus make it harder for analysts to forecast earnings (Peterson et al., 2015). Therefore, following Dechow et al. (1995) and Roychowdhury (2006), we use the absolute value of discretionary accrual to measure accrued earnings management (*EM*) and the absolute value of real activities manipulation to measure real earnings management (*REM*) and investigate the relationship between retail investor attention and earnings management. The results are presented in Table 9. In column (1), the coefficient of *EM* is positive and significant at the 5% level. Meanwhile, the coefficient of *REM* in column (2) is positive and significant at the 1% level. Overall, the results in Table 9 demonstrate that more retail investor attention will be followed by a higher level of accrued and real earnings

Table 12
The effect of industrial policy.

			Industrial policy group	Non-Industrial policy group	Industrial policy group	Non-Industrial policy group
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
IP_t	−0.0703	−0.1409				
	(−0.365)	(−1.187)				
$RelAtt_t$			0.8346***	0.9801***	0.5982***	0.6035***
			(3.769)	(3.740)	(4.680)	(3.979)
$Size_t$	1.2462***	0.6980***	1.2983***	0.9063***	0.7009***	0.4474***
	(7.781)	(7.527)	(5.110)	(4.387)	(4.813)	(3.602)
Lev_t	−2.7909***	−1.7164***	−3.3409***	−2.7206***	−2.3000***	−1.3067**
	(−5.051)	(−5.105)	(−4.122)	(−3.336)	(−4.891)	(−2.341)
ROA_t	−10.4424***	−5.9400***	−10.1995***	−9.7535***	−6.0282***	−5.3444***
	(−7.244)	(−6.662)	(−4.793)	(−5.187)	(−4.902)	(−4.227)
SOE_t	1.0534**	0.4258	0.3941	1.6183**	0.0692	0.6615*
	(2.074)	(1.640)	(0.595)	(1.973)	(0.205)	(1.686)
BM_t	0.0786	0.0564	0.1347	0.2477***	0.0910	0.1574***
	(0.892)	(1.065)	(0.792)	(2.603)	(0.872)	(2.716)
$INST_t$	−0.9218***	−0.4163**	−0.8239**	−0.6139	−0.4214*	−0.0647
	(−3.203)	(−2.317)	(−2.167)	(−1.324)	(−1.789)	(−0.218)
$FirmAge_t$	−1.2066	−0.6031	−2.4815**	−0.2708	−1.4489**	−0.2653
	(−1.599)	(−1.379)	(−2.337)	(−0.255)	(−2.394)	(−0.467)
$Big4_t$	−0.3410	0.1029	−0.6501	−0.1670	−0.2798	0.5734
	(−0.684)	(0.360)	(−0.847)	(−0.232)	(−0.647)	(1.371)
$Loss_t$	0.3229	0.5381***	0.0753	0.5838	0.4278	0.6579**
	(1.119)	(2.875)	(0.187)	(1.435)	(1.623)	(2.511)
$ANALYST_t$	0.0239	−0.0139	0.0492	0.1184	−0.0263	0.0681
	(0.370)	(−0.348)	(0.555)	(1.178)	(−0.483)	(1.134)
$Growth_t$	−1.1588***	−0.5377***	−1.3175***	−0.9098***	−0.5790***	−0.4413***
	(−9.869)	(−7.705)	(−7.826)	(−5.174)	(−5.941)	(−4.172)
Constant	−18.8633***	−10.5062***	−26.0430***	−26.9231***	−15.0837***	−14.1415***
	(−5.072)	(−4.862)	(−4.421)	(−4.751)	(−4.319)	(−4.233)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj R^2	0.0375	0.0392	0.0373	0.0420	0.0405	0.0412
N	18,040	18,040	10,579	7461	10,579	7461

This table shows the results of the analysis examining how industrial policy affects analyst earnings forecast and whether the relationship between retail investor attention and analyst earnings forecast is driven by industrial policy. IP is a dummy variable equals to 1 if a firm turns out to be in the industry that experience an industrial policy change made by the Chinese government in year t, and 0 otherwise. We split our sample into two groups by the value of IP. Columns (1) and (2) report how analyst earnings forecast accuracy is related to industrial policy, columns (3) and (5) report the estimation results from the industrial policy group, and columns (4) and (6) report the estimation results from the non-industrial policy group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

management, which supports our proposed mechanism.

To sum up, consistent with our conjecture, the results above show that retail investor attention could not only distort the stock price informativeness of underlying firms, but also drive firm managers to engage in higher levels of accrued and real earnings management, which in turn deteriorates the accuracy of analyst earnings forecast.

5.2. Alternative explanations

So far, our evidence shows that retail investor attention could affect analyst earnings forecast by influencing stock price informativeness and earnings management. Since firm-related events such as M&A, comment letter, and industrial policy tend to attract investor attention and bring uncertainty to the operation of underlying firms, analyst earnings forecasts could thus be influenced. In this section, we discuss and examine whether the events mentioned above affect our findings.

Firstly, by using *MA* as the independent variable in model (3), we examine the effect of M&A on analyst earnings forecast. To explore whether M&A determines the relationship between retail investor attention and analyst earnings forecast, we then split our sample into two groups by the value of *MA*. The results are presented in Table 10. Columns (1) and (2) show that *MA* does not have a significant effect on analyst earnings forecast. From columns (3)–(6) we can see that the coefficients of *RelAtt* are all significantly positive, which implies that retail investor attention negatively influences analyst earnings forecast accuracy even without M&A.

Secondly, we use *CL* as the independent variable to investigate how comment letters affect analyst earnings forecast. Moreover, to explore whether comment letter is the key driver of the relationship between retail investor attention and analyst earnings forecast, we split our sample into two groups by the value of *CL*. The results are presented in Table 11. Columns (1) and (2) show that *CL* could

Table 13
Heterogeneity: audit quality.

	Big 4 auditor group		Non-Big 4 auditor group	
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)
$ReIAtt_t$	−0.5185	−0.2773	0.9952***	0.6635***
	(−0.773)	(−0.643)	(5.737)	(6.697)
$Size_t$	1.5349***	0.8063**	1.0947***	0.6020***
	(2.705)	(2.122)	(6.386)	(6.202)
Lev_t	−7.0974**	−4.1327**	−2.9661***	−1.8772***
	(−2.548)	(−2.252)	(−5.242)	(−5.458)
ROA_t	−12.7051*	−11.0452**	−10.2800***	−5.6201***
	(−1.886)	(−2.399)	(−6.996)	(−6.204)
SOE_t	0.2871	0.2280	1.0881**	0.4055
	(0.368)	(0.290)	(1.994)	(1.453)
BM_t	0.2194	0.1287	0.1772*	0.1148*
	(1.367)	(1.344)	(1.708)	(1.865)
$INST_t$	−0.2722	−0.1568	−0.6058**	−0.2285
	(−0.314)	(−0.237)	(−2.026)	(−1.231)
$FirmAge_t$	−1.3575	0.3796	−1.4524*	−0.8356*
	(−0.920)	(0.380)	(−1.745)	(−1.754)
$Loss_t$	4.0906***	2.9141***	0.1218	0.4150**
	(2.955)	(3.184)	(0.416)	(2.178)
$ANALYST_t$	−0.3441*	−0.1520	0.0575	−0.0041
	(−1.669)	(−1.153)	(0.853)	(−0.098)
$Growth_t$	−0.6609	−0.5245	−1.1695***	−0.5318***
	(−1.152)	(−1.581)	(−9.732)	(−7.431)
Constant	−19.3188	−12.0221	−26.7143***	−15.7034***
	(−1.199)	(−1.090)	(−6.450)	(−6.468)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Adj. R^2	0.0859	0.1076	0.0406	0.0417
N	1336	1336	16,704	16,704

This table presents the results of the heterogeneity analysis examining how the relationship between retail investor attention and analyst earnings forecast varies with audit quality. We split our sample into two groups by audit quality. Columns (1) and (2) report the estimation results from the Big 4 auditor group, and columns (3) and (4) report the estimation results from the non-Big 4 auditor group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

reduce analyst earnings forecast error and dispersion, which is consistent with Hu et al. (2022). The coefficients of *ReIAtt* from columns (3)–(6) are still significantly positive, suggesting that comment letters do not drive the relationship between retail investor attention and analyst earnings forecast accuracy.

Thirdly, we use *IP* as the independent variable to analyze how industrial policy affect analyst earnings forecast. Likewise, to explore the impact of industrial policy on the relationship between retail investor attention and analyst earnings forecast, we divide our sample into two groups by the value of *IP*. The results are presented in Table 12. The coefficients of *IP* are negative, but not statistically significant in columns (1) and (2). The results in columns (3)–(6) demonstrate that the negative impact of retail investor attention on analyst earnings forecast accuracy still exists even if a firm does not experience industrial policy change.

Overall, the evidence presented in this section not only confirms that firm related events such as M&A, comment letter, and industrial policy are not the main drivers of the relationship between retail investor attention and analyst earnings forecast accuracy, but also provides additional support to our proposed analysis.

5.3. Heterogeneity

As documented by Kong et al. (2020), Big 4 audit firms, including E&Y, KPMG, Deloitte, and PWC, could improve their clients' earnings quality and information transparency, thus providing analysts with high-quality information and enhancing earnings forecast accuracy. Therefore, we follow Kong et al. (2020) and define *Big4* as a dummy variable that equals 1 if a firm is audited by Big 4 auditors and 0 otherwise. We then investigate the effect of audit quality by estimating the retail investor attention-analyst earnings forecast accuracy relation for the Big 4 auditor group and non-Big 4 auditor group, respectively. As denoted in Table 13, the coefficients of *ReIAtt* for firms in the Big 4 auditor group are −0.5185 and − 0.2773, both insignificant as reported in columns (1) and (2), while the coefficients of *ReIAtt* for firms in the non-Big 4 auditor group are 0.9952 and 0.6635, both significant at the 1% level as reported in columns (3) and (4). The results in Table 13 reveal that retail investor attention mainly inflicts analyst earnings forecast accuracy for firms with low audit quality.

Saffi and Sigurdsson (2011) argue that short selling could enhance stock price efficiency by facilitating the integration of negative information into stock prices. Moreover, Hou et al. (2021) find that short selling could mitigate the optimistic bias of analyst earnings

Table 14
Heterogeneity: short-sale constraints.

	Shortable group		Non-shortable group	
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)
$RelAtt_t$	0.1825	0.1880	1.1260***	0.7432***
	(0.661)	(1.086)	(5.348)	(6.097)
$Size_t$	1.9387***	1.0711***	0.8441***	0.5071***
	(6.367)	(5.956)	(3.931)	(4.194)
Lev_t	−5.3337***	−3.7523***	−1.9907***	−1.0067**
	(−4.674)	(−5.333)	(−2.966)	(−2.435)
ROA_t	−12.7110***	−8.9926***	−8.9139***	−4.4730***
	(−5.411)	(−5.825)	(−4.849)	(−3.925)
SOE_t	1.5931**	0.4681	0.2956	0.2270
	(2.168)	(1.196)	(0.387)	(0.591)
BM_t	0.1868	0.1243*	0.0972	0.0212
	(1.628)	(1.785)	(0.666)	(0.236)
$INST_t$	−0.2609	0.0220	−0.5971	−0.2596
	(−0.383)	(0.059)	(−1.638)	(−1.152)
$FirmAge_t$	−0.0410	0.1832	−2.5157**	−1.3254*
	(−0.031)	(0.280)	(−2.091)	(−1.854)
$Big4_t$	0.9262	0.6954	−1.3886**	−0.1925
	(1.049)	(1.417)	(−2.365)	(−0.557)
$Loss_t$	0.2283	0.6220**	0.3167	0.5993**
	(0.487)	(2.123)	(0.879)	(2.473)
$ANALYST_t$	0.1605	0.0533	0.0171	−0.0470
	(1.364)	(0.722)	(0.211)	(−0.953)
$Growth_t$	−0.9874***	−0.4796***	−1.1723***	−0.5565***
	(−3.977)	(−3.210)	(−8.308)	(−7.004)
Constant	−41.3361***	−23.4067***	−19.5829***	−13.2121***
	(−4.955)	(−4.906)	(−3.865)	(−4.479)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Adj R^2	0.0426	0.0594	0.0389	0.0368
N	6505	6505	11,535	11,535

This table shows the results of the heterogeneity analysis examining how the relationship between retail investor attention and analyst earnings forecast varies with short-sale constraints. We split our sample into two groups by shortability. Columns (1) and (2) report the estimation results from the shortable group, and columns (3) and (4) report the estimation results from the non-shortable group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

forecasts by improving analysts' information sets. Therefore, to explore how short selling affects the relationship between retail investor attention and analyst earnings forecast accuracy, we split our sample into two groups by the shortability of the underlying stocks. We expect retail investor attention to have a more severe impact on analyst earnings forecast accuracy for non-shortable firms. Columns (1) and (2) of Table 14 show that the coefficients of *RelAtt* for firms in the shortable group are 0.1825 and 0.1880, both insignificant, while the results in columns (3) and (4) of Table 14 report that the coefficients of *RelAtt* for firms in the non-shortable group are 1.1260 and 0.7432, both significant at the 1% level. Our findings in Table 14 suggest that retail investor attention mainly distorts analyst earnings forecast accuracy for firms faced with severe short-sale constraints.

Based on the Shanghai-Hong Kong stock connect program, Chen et al. (2021) find that stock market liberalization could facilitate the entry of foreign institutional investors, which are conducive to improving information environment and governance of firms in the program, thus correcting analyst earnings forecast bias to a certain extent. To examine how stock market liberalization affects the relationship between retail investor attention and analyst earnings forecast accuracy, we split our sample into two groups based on whether the underlying firm is in the Mainland-Hong Kong stock connect program (MHKC), which is the combination of Shanghai-Hong Kong stock connect program and Shenzhen-Hong Kong stock connect program. We believe the negative impact of retail investor attention on analyst earnings forecast accuracy should be weakened for firms in the MHKC group. The results are presented in Table 15. It is worth noting that the coefficients of *RelAtt* in columns (1) and (2) for firms in the MHKC group are −1.5038 and −0.9435, both negative and significant. This may be due to the fact that improved information environments and corporate governance brought about by stock market liberalization could facilitate the informational and motoring effects of retail investor attention on the underlying firms, which stimulates the flow of firm-specific information and enhances analyst forecasts. The results in columns (3) and (4) of Table 15 report that the coefficients of *RelAtt* for firms in the non-MHKC group are 1.0497 and 0.6924, both significant at the 1% level. The evidence in Table 15 implies that stock market liberalization could be an effective mechanism in rectifying the negative effect of retail investor attention on analyst earnings forecast accuracy in the Chinese stock market.

To sum up, the results in further analysis confirm our proposed mechanism. Retail investor attention could distort analyst earnings forecast accuracy by deteriorating stock price informativeness of the underlying stocks. Besides, the evidence from heterogeneity analysis suggests that the forecast accuracy of firms audited by Big 4 accounting firms, shortable firms, and firms in the Mainland-Hong

Table 15
Heterogeneity: stock market liberalization.

	Mainland-Hong Kong stock connect group		Non-Mainland-Hong Kong stock connect group	
	$FERROR_{t+1}$	$FDISP_{t+1}$	$FERROR_{t+1}$	$FDISP_{t+1}$
	(1)	(2)	(3)	(4)
$RelAtt_t$	−1.5038**	−0.9435**	1.0497***	0.6924***
	(−2.165)	(−2.345)	(5.943)	(6.872)
$Size_t$	1.6961***	1.0126***	1.0894***	0.5935***
	(2.974)	(2.891)	(6.337)	(6.082)
Lev_t	−3.7293	−1.8466	−3.1498***	−1.9104***
	(−1.427)	(−1.086)	(−5.429)	(−5.412)
ROA_t	−2.2494	−2.4820	−10.2848***	−5.7279***
	(−0.367)	(−0.576)	(−6.884)	(−6.246)
SOE_t	−0.3586	−0.4637**	0.8944*	0.3914
	(−1.492)	(−2.322)	(1.717)	(1.454)
BM_t	0.1321	0.0893	0.2658***	0.1717***
	(0.621)	(0.676)	(2.593)	(2.801)
$INST_t$	−1.8844*	−0.4836	−0.6336**	−0.2696
	(−1.930)	(−1.120)	(−2.096)	(−1.425)
$FirmAge_t$	−2.0212	−2.1174	−1.9006**	−1.0553**
	(−0.876)	(−1.292)	(−2.277)	(−2.221)
$Big4_t$	0.1634	0.1850	−0.3795	0.1399
	(0.455)	(0.526)	(−0.673)	(0.432)
$Loss_t$	1.7322	1.2832	0.2622	0.5235***
	(1.138)	(1.552)	(0.885)	(2.711)
$ANALYST_t$	0.2938	0.1054	0.0488	−0.0085
	(1.439)	(0.852)	(0.719)	(−0.204)
$Growth_t$	−0.6150	−0.5068*	−1.1561***	−0.5227***
	(−1.474)	(−1.943)	(−9.487)	(−7.252)
Constant	−11.3752	−3.5273	−26.0801***	−15.3408***
	(−0.907)	(−0.463)	(−6.249)	(−6.243)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Adj R^2	0.0445	0.0571	0.0412	0.0431
N	1174	1174	16,866	16,866

This table presents the results of the heterogeneity analysis examining how the relationship between retail investor attention and analyst earnings forecast varies with stock market liberalization. We split our sample into two groups based on whether the underlying firm is in the Mainland-Hong Kong stock connect program (MHKC). Columns (1) and (2) report the estimation results from the Mainland-Hong Kong stock connect group, and columns (3) and (4) report the estimation results from the non-Mainland-Hong Kong stock connect group. Firm and year fixed effects are controlled. Robust standard errors are clustered at the firm level. t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Kong stock connect program is less negatively influenced by retail investor attention.

6. Conclusions

In this paper, we investigate how retail investor attention influences analyst earnings forecast accuracy by analyzing a sample of listed firms in the Chinese stock market. We discover that firms associated with a higher level of retail investor attention have higher future forecast error and forecast dispersion, supporting our hypothesis that retail investor attention negatively affects analyst earnings forecast accuracy. Moreover, our results remain robust throughout a series of additional tests.

Further study on the possible influencing channels behind this relationship reveals that the information content of stock prices and opportunistic behaviors of firm managers play an important role in explaining the mechanism of how retail investor attention affects analyst forecasts. This result is consistent with our conjecture that retail investor attention to an individual firm can instigate noise trading and undermine the firm's stock price efficiency, as well as drive firm managers to engage in earnings management, which in turn worsens analyst earnings forecast accuracy. Furthermore, heterogeneity analysis shows that the negative influence of retail investor attention on analyst earnings forecast accuracy is weakened in firms with higher audit quality, firms with lower short-sale constraints, and firms with broader access to foreign institutional investors.

In conclusion, our findings not only shed new light on the extant literature related to investor attention and analyst forecast but also have several valuable implications for stock market participants and policymakers. Firstly, as far as we can tell, our study is the first to document the negative impact of retail investor attention on analyst earnings forecast accuracy. Future work on the real effect of retail investor attention on other agents in the financial market could be a meaningful and promising research direction. Secondly, while using the internet to gather firm-specific information for investing has become more prevalent than before, retail investors and financial analysts in emerging markets should be aware of the noisy information on the internet and remain independent during their decision-making process. Finally, to diminish the negative impact of retail investor attention on analyst earnings forecast accuracy, policymakers in emerging markets could take further steps to relax short-sale constraints for qualified stocks and make more effort to expand the scope of capital market liberalization policy.

Funding information

This research is supported by the National Natural Science Foundation of China (No. 72273099).

CRediT authorship contribution statement

Zhida Zhang: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Qi Luo:** Supervision, Funding acquisition, Resources, Writing – review & editing.

References

Antweiler, W., Frank, M.Z., 2004. Is all that talk just noise? The information content of internet stock message boards. J. Financ. 59 (3), 1259–1294.
Athanassakos, G., Kalimipalli, M., 2003. Analyst forecast dispersion and future stock return volatility. Q. J. Bus. Econ. 42 (1–2), 57–78.
Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818.
Behn, B.K., Choi, J., Kang, T., 2008. Audit quality and properties of analyst earnings forecasts. Account. Rev. 83 (2), 327–349.
Bennett, B., Stulz, R., Wang, Z., 2020. Does the stock market make firms more productive? J. Financ. Econ. 136 (2), 281–306.
Bhat, G., Hope, O., Kang, T., 2006. Does corporate governance transparency affect the accuracy of analyst forecasts? Account. Finance 46 (5), 715–732.
Bond, P., Edmans, A., Goldstein, I., 2012. The real effects of financial markets. Annu. Rev. Financ. Econ. 4, 339–360.
Chen, X., Wu, C., 2022. Retail investor attention and information asymmetry: evidence from China. Pac. Basin Financ. J. 75, 101847.
Chen, Q., Goldstein, I., Jiang, W., 2007. Price informativeness and investment sensitivity to stock price. Rev. Financ. Stud. 20 (3), 619–650.
Chen, Z., Lin, W.T., Ma, C., Zheng, Z., 2013. The impact of individual investor trading on stock returns. Emerg. Mark. Financ. Trade 49 (3), 62–69.
Chen, T., Harford, J., Lin, C., 2015. Do analysts matter for governance? Evidence from natural experiments. J. Financ. Econ. 115 (2), 383–410.
Chen, T., Xie, L., Zhang, Y., 2017. How does analysts' forecast quality relate to corporate investment efficiency? Finance 43, 217–240.
Chen, S., Lin, B., Lu, R., Ma, H., 2021. Stock market openness and analyst forecast bias. J. Account. Public Policy 40 (6), 106874.
Chen, X., Diao, X., Wu, C., 2022. Heterogeneous investor attention and post earnings announcement drift: evidence from China. Econ. Model. 110, 105796.
Clarke, J., Chen, H., Du, D., Hu, Y.J., 2021. Fake news, investor attention, and market reaction. Inf. Syst. Res. 32 (1), 35–52.
Clement, M.B., 1999. Analyst forecast accuracy: do ability, resources, and portfolio complexity matter? J. Account. Econ. 27 (3), 285–303.
Cowen, A., Groysberg, B., Healy, P., 2006. Which types of analyst firms are more optimistic? J. Account. Econ. 41 (1–2), 119–146.
Da, Z., Huang, X., 2019. Harnessing the wisdom of crowds. Manag. Sci. 66 (5), 1783–2290.
Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. J. Financ. 66 (5), 1461–1499.
Dechow, P.M., Sloan, R.G., Sweeney, A.P., 1995. Detecting earnings management. Account. Rev. 70 (2), 193–225.
DellaVigna, S., Pollet, J.M., 2009. Investor inattention and Friday earnings announcements. J. Financ. 64 (2), 709–749.
Dhaliwal, D.S., Radhakrishnan, S., Tsang, A., Yang, Y.G., 2012. Nonfinancial disclosure and analyst forecast accuracy: international evidence on corporate social responsibility disclosure. Account. Rev. 87 (3), 723–759.
Easterwood, J.C., Nutt, S.R., 1999. Inefficiency in analysts' earnings forecasts: systematic misreaction or systematic optimism. J. Financ. 54 (5), 1777–1797.
Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. J. Financ. 64 (5), 2023–2052.
Foucault, T., Sraer, D., Thesmar, D.J., 2011. Individual investors and volatility. J. Financ. 66 (4), 1369–1406.
Frankel, R., Kothari, S.P., Weber, J., 2006. Determinants of the informativeness of analyst research. J. Account. Econ. 41 (1–2), 29–54.
Gervais, S., Kaniel, R., Mingelgrin, D.H., 2001. The high-volume return premium. J. Financ. 56 (3), 877–919.
Han, B., Kong, D., Liu, S., 2018. Do analysts gain an information advantage by visiting listed companies? Contemp. Account. Res. 35 (4), 1843–1867.
Hao, J., Xiong, X., 2021. Retail investor attention and firms' idiosyncratic risk: evidence from China. Int. Rev. Financ. Anal. 74, 101675.
Haw, I., Hu, B., Lee, J.J., 2015. Product market competition and analyst forecasting activity: international evidence. J. Bank. Financ. 56, 48–60.
He, F., Yan, Y., Hao, J., Wu, J., 2022. Retail investor attention and corporate green innovation: evidence from China. Energy Econ. 115, 106308.
Hervé, F., Zouaoui, M., Belvaus, B., 2019. Noise traders and smart money: evidence from online searches. Econ. Model. 83, 141–149.
Hirshleifer, D., Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. J. Account. Econ. 36 (1–3), 337–386.
Hirshleifer, D., Lim, S.S., Teoh, S.H., 2009. Driven to distraction: extraneous events and underreaction to earnings news. J. Financ. 64 (5), 2289–2325.
Hirshleifer, D., Lim, S.S., Teoh, S.H., 2011. Limited investor attention and stock market misreactions to accounting information. Rev Asset Pric Stud 1 (1), 35–73.
Hirshleifer, D., Levi, Y., Lourie, B., Teoh, S.H., 2019. Decision fatigue and heuristic analyst forecasts. J. Financ. Econ. 133 (1), 83–98.
Hong, H., Kubik, J.D., 2003. Analyzing the analysts: career concerns and biased earnings forecasts. J. Financ. 58 (1), 313–351.
Hou, D., Meng, Q., Chan, K.C., 2021. Does short selling reduce analysts' optimism bias in earnings forecasts? Res. Int. Bus. Financ. 56, 101356.
Hu, J., Long, W., Luo, L., Peng, Y., 2021. Share pledging and optimism in analyst earnings forecasts: evidence from China. J. Bank. Financ. 132, 106245.
Hu, N., Xu, J., Xue, S., 2022. Mandatory disclosure of comment letters and analysts' forecasts. Int. Rev. Financ. Anal. 84, 102426.
Huang, Y., Zhang, H., 2020. Does individual investors' attention influence underwriters' IPO pricing? Appl. Econ. 52 (34), 3680–3687.
Huang, L., Li, W., Wang, H., Wu, L., 2022. Stock dividend and analyst optimistic bias in earnings forecast. Int. Rev. Econ. Financ. 78, 643–659.
Jiang, L., Liu, J., Peng, L., Wang, B., 2022. Investor attention and asset pricing anomalies. Rev. Fin. 26 (3), 563–593.
Kempf, E., Manconi, A., Spalt, O., 2017. Distracted shareholders and corporate actions. Rev. Financ. Stud. 41, 1660–1695.
Kong, D., Liu, L., Liu, S., 2020. Market information traveling on high-speed rails: the case of analyst forecasts. Pac. Basin Financ. J. 61, 101320.
Kong, D., Lin, Z., Wang, Y., Xiang, J., 2021. Natural disasters and analysts' earnings forecasts. Finance 66, 101860.
Kumar, A., Rantala, V., Xu, R., 2022. Social learning and analyst behavior. J. Financ. Econ. 143 (1), 434–461.
Lang, M.H., Lundholm, R.J., 1996. Corporate disclosure policy and analyst behavior. Account. Rev. 71 (4), 467–492.
Li, C., Liu, X., Hou, Z., Li, Y., 2023. Retail investor attention and equity mispricing: the mediating role of earnings management. Financ. Res. Lett. 53, 103621.
Lin, H., McNichols, M.F., 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. J. Account. Econ. 25 (1), 101–127.
Liu, H., Krystyniak, K., 2021. Investor attention and merger announcements. J. Behav. Financ. 22 (1), 97–112.
Lou, D., 2014. Attracting investor attention through advertising. Rev. Financ. Stud. 27 (6), 1798–1829.
Merkley, K., Michaely, R., Pacelli, J., 2017. Does the scope of the sell-side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports. J. Financ. 72 (3), 1285–1334.
Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. J. Financ. 42 (3), 483–510.
Muradoglu, G., Peng, N., Qin, H., Xia, C., 2024. Investor attention and market reactions to early announcements in mergers and acquisitions. Int. Rev. Financ. Anal. 91, 102993.
Nguyen, H.T., Pham, M.H., 2021. Does investor attention matter for market anomalies? J. Behav. Exp. Financ. 29, 100451.
Nofsinger, J.R., 2001. The impact of public information on investors. J. Bank. Financ. 25 (7), 1339–1366.
O'Brien, P.C., 1988. Analysts' forecasts as earnings expectations. J. Account. Econ. 10 (1), 53–83.
Peng, L., Xiong, W., 2006. Investor attention, overconfidence, and category learning. J. Financ. Econ. 80 (3), 563–602.

Peterson, K., Schmardebeck, R., Wilks, T.J., 2015. The earnings quality and information processing effects of accounting consistency. Account. Rev. 90 (6), 2483–2514.
Piotroski, J.D., Wong, T.J., Zhang, T., 2015. Political incentives to suppress negative information: evidence from Chinese listed firms. J. Account. Res. 53 (2), 405–459.
Rakowski, D., Shirley, S.E., Stark, J.R., 2021. Twitter activity, investor attention, and the diffusion of information. Financ. Manag. 50 (1), 3–46.
Roychowdhury, S., 2006. Earnings management through real activities manipulation. J. Account. Econ. 42 (3), 335–370.
Saffi, P.A.C., Sigurdsson, K., 2011. Price efficiency and short selling. Rev. Financ. Stud. 24 (3), 821–852.
Simpson, A., 2013. Does investor sentiment affect earnings management? J. Bus. Financ. Acc. 40 (7–8), 869–900.
Wang, J., Gan, J., Li, Z., 2021a. Industrial policy, uncertainty, and analysts' earnings forecast: evidence from China. Econ. Anal. Policy 70, 249–258.
Wang, J., Huo, D., Shi, Y., Zhang, K., 2021b. Can information content of stock prices improve analyst forecast quality? Syst. Engi. The. Pract. 41 (8), 1974–1989 (*in Chinese*). 10.12011/SETP2020-3100.
Wang, M., Zhao, W., Tao, Y., 2022. Does voluntary disclosure on an internet platform affect analyst forecast accuracy? Appl. Econ. Lett. 29 (17), 1547–1555.
Wen, F., Xu, L., Ouyang, G., Kou, G., 2019. Retail investor attention and stock price crash risk: evidence from China. Int. Rev. Financ. Anal. 65, 101376.
Wu, Y., Liu, T., Han, L., Yin, L., 2018. Optimistic bias of analysts' earnings forecasts: does investor sentiment matter in China? Pac. Basin Financ. J. 49, 147–163.
Xiao, S., 2020. Do noisy stock prices impede real efficiency? Manag. Sci. 66 (12), 5485–6064.
Yang, M., Chen, S., Wang, J., Wang, Y., 2023. Corporate social responsibility and stock price crash risk: does analyst coverage matter? Appl. Econ. 55 (10), 1045–1061.
Ying, Q., Kong, D., Luo, D., 2015. Investor attention, institutional ownership, and stock return: empirical evidence from China. Emerg. Mark. Financ. Trade 51 (3), 672–685.
Yu, F., 2008. Analyst coverage and earnings management. J. Financ. Econ. 88 (2), 245–271.
Yusoff, I., Chen, C., Lai, K., Naiker, V., Wang, J., 2023. Foreign exchange exposure and analysts' earnings forecasts. J. Bank. Financ. 146, 106715.
Zhang, X.F., 2010. Information uncertainty and analyst forecast behavior. Contemp. Account. Res. 23 (2), 565–590.
Zuo, L., 2016. The informational feedback effect of stock prices on management forecasts. J. Account. Econ. 61 (2–3), 391–413.