Information cherry-picking:
When confirmation bias met anchoring bias*

Shiyang Huang†, Tse-Chun Lin‡, Yan Luo§, and Ningyu Zhou**

Abstract

We hypothesize that a surge in information availability coupled with investors' confirmation bias aggravates retail investors' anchoring bias. We build a stylized model and find supportive evidence from a controlled laboratory experiment. We then use the staggered EDGAR implementation to provide the causal inference that investors' anchoring tendency is amplified for treated stocks after the EDGAR implementation due to information cherry-picking. The effect is more evident (i) among stocks with lower reporting readability and higher information uncertainty, (ii) shortly after releases of corporate 10K/Q reports, and (iii) for Internet users and sophisticated investors. The exacerbated anchoring bias also worsens investment performance.

Keywords: Confirmation bias, Anchoring bias, 52-week highs, Information cherry-picking, Controlled laboratory experiment, EDGAR implementation

JEL: D14; D83; D91; G12; G14; G38; G41

* We are grateful for helpful comments from Svetlana Bryzgalova, Yen-Cheng Chang, Stefano Cassella, J. Anthony Cookson, Ed deHaan, Francisco Gomes, Pingyang Gao, Claire Yurong Hong, Jiekun Huang, Travis L. Johnson, OSub Kwon, Kevin Tseng, and seminar and conference participants at the Copenhagen Business School Nankai University, National Chengchi University, National Taiwan University, PBCSF Tsinghua University, 2023 China International Conference in Finance (CICF), 2023 Taiwan Economic Research, 2024 Future of Financial Information Conference (poster session), and 2024 Research in Behavioral Finance Conference (RBFC). We thank Terrance Odean for sharing the large discount brokerage data. All remaining errors are ours.
† HKU Business School, The University of Hong Kong, Email: huangsy@hku.hk.
‡ HKU Business School, The University of Hong Kong, Email: tsechunlin@hku.hk.
§ School of Management, Fudan University, Email: luoyan@fudan.edu.cn
** School of Management, Fudan University, Email: 20110690045@fudan.edu.cn

Information cherry-picking:
When confirmation bias met anchoring bias

Abstract

We hypothesize that a surge in information availability coupled with investors' confirmation bias aggravates retail investors' anchoring bias. We build a stylized model and find supportive evidence from a controlled laboratory experiment. We then use the staggered EDGAR implementation to provide the causal inference that investors' anchoring tendency is amplified for treated stocks after the EDGAR implementation due to information cherry-picking. The effect is more evident (i) among stocks with lower reporting readability and higher information uncertainty, (ii) shortly after releases of corporate 10K/Q reports, and (iii) for Internet users and sophisticated investors. The exacerbated anchoring bias also worsens investment performance.

Keywords: Confirmation bias, Anchoring bias, 52-week highs, Information cherry-picking, Controlled laboratory experiment, EDGAR implementation

JEL: D14; D83; D91; G12; G14; G38; G41

"Do you remember, before the Internet, that people thought the cause of stupidity was the lack of access to information?

Well…it wasn't that."

----Unknown

1. Introduction

Confirmation bias, which describes people's tendency to give undue weight to evidence supporting prior beliefs while neglecting or discounting evidence that would be against them, suggests that people often interpret information in a biased way. It is driven by the fundamental nature of human cognition (Kahneman, 2011) and could be the bias that deserves attention above all other problematic aspects of human reasoning (Evans, 1989). Anecdotal evidence of confirmation bias is observed frequently in daily lives, from the selective consumption of ideologically congruent news to the search for belief-consistent information in evaluating vaccine effectiveness and safety. In this study, we shed light on an important yet unanswered question: in the presence of confirmation bias, how does an exogenous increase in the accessibility of corporate information influence investors' inherent behavioral biases?

In conventional wisdom, the enhancement of information dissemination from corporates to stock markets should be beneficial to investors in improving their investment decisions as information asymmetry is mitigated, and valuation uncertainty is reduced. For instance, Hirshleifer (2001) states that "…people are likely to be more prone to bias in valuing securities for which information is sparse." Baker and Wurgler (2006, 2007) conjecture that when valuation uncertainty is higher, biases are more insidious, and valuation mistakes are more likely to happen. Kumar (2009) provides empirical evidence in line with these arguments.

However, we argue that the surge in information also makes it easier for individuals with confirmation bias to cherry-pick information that only confirms their biased priors, which has been widely documented in studies on media industries. For example, Mullainathan and Shleifer (2005) find that although the Internet reduces information acquisition costs, individuals may cherry-pick the information to self-segregate ideologically. Sunstein (2001) contends that "Our communications market is rapidly moving" to a situation where "people restrict themselves to their own points of view—liberals watching and reading mostly or only liberals; moderates, moderates; conservatives, conservatives; Neo-Nazis, Neo-Nazis."

Ample follow-up studies demonstrate that individuals tend to tailor information exposure on the Internet or social media in ways that only reinforce their prior beliefs (e.g., Gentzkow and Shapiro, 2011; Bessi, 2016; Moravec, Minas, and Denni, 2019; Slechten, Courtois, Coenen, and Zaman, 2022). These studies suggest that although the Internet offers (i) pathways for enhancing democratic participation, (ii) increased interconnectedness, and (iii) access to knowledge, the uncritical use of Internet platforms and indiscriminating consumption of the information provided can impede achieving the very goals it promises. Based on this stream of literature, retail investors might exhibit the same confirmation bias in their investment decision process. We build a suggestive and stylized model in Section 2.2 to develop our hypotheses: *more information coupled with retail investors' confirmation bias could amplify rather than alleviate these investors' pre-existing behavioral biases*.

The behavioral bias we focus on is the anchoring tendency of salient reference points, which offers a comparison basis to disentangle the prior investors' bias level and its changes after the information shock. Moreover, it is also one of the most well-documented and acknowledged human behavioral biases. For example, Heath, Huddart, and Lang (1999) show that exercise activities of employee stock options nearly double when the stock price exceeds its maximum level over the previous year. It demonstrates that individual option-holders set a psychological reference point based on the stock price's past 52-week high. A large volume of subsequent studies also shows that stock market participants are affected by a stock's 52-week high in making various financial decisions even though it contains no new information and should be irrelevant.[1] Our experimental analysis in Section 2.3 confirms the existence of the 52-week high anchoring bias in simulated investment decisions. The analysis also provides a preliminary result on the influence of confirmation bias on anchoring bias.

Although confirmation bias has been claimed to be pervasive in our daily lives, detecting its influence on investors in stock markets is challenging, and the current empirical evidence is limited. Providing the causal inference in a real-world high-stakes setting faces several challenges, including (i) identifying exogenous shocks on the accessibility of information to investors and (ii) obtaining investors' detailed trading records and detecting their well-

[1] Practitioners that adopt technical analysis on the stock price pattern term the 52-week high as the resistance level as if the price is facing strong resistance to breaking through such a price barrier. When the stock price is close to this resistance level, technical traders tend to believe that the stock has relatively small upside potential and thus are likely to either close their long positions by selling the stocks or take short positions by short selling. For instance, using retail trading data from Euroclear Finland, Vedova, Grant, and Westerholm (2023) document that retail investors sharply increase their selling of stocks at the 52-week high.

identified prior beliefs or pre-existing biases. Addressing the two challenges, the novelty of our empirical setup is that we exploit the staggered implementation of the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system between 1993 and 1996 to test how increased information accessibility affects the degree of retail investors' anchoring bias. Specifically, when a firm's financial statements become publicly available for free through the EDGAR system, we view it as a positive shock to information dissemination and expect enhanced investor information cherry-picking due to the confirmation bias, which in turn strengthens the 52-week high anchoring bias. The staggered implementation of the EDGAR system allows us to use a quasi-natural experiment setting and provides causal evidence.

To capture the anchoring bias, we utilize the large discount brokerage database of Barber and Odean (2000) and trace the retail investors' trading. We first demonstrate that retail investors are significantly more inclined to sell (buy) stocks when prices are near (away from) the 52-week high using a stacked difference-in-differences (DiD) approach with propensity score matching based on firm size, after controlling for various known factors affecting trading decisions.[2] More importantly, as corporate fundamental information becomes more easily accessible through the Internet after stocks are subject to the mandatory filing on EDGAR, retail investors' inclination to anchor their trading decisions on stocks' 52-week high is intensified. Such evidence contradicts the conventional wisdom that more information, which reduces valuation uncertainty, helps alleviate investors' behavioral biases. The results, however, align with our hypothesis and the findings from research on social media that confirmation bias coupled with more easily available information from the Internet may exacerbate prejudice.

We then examine the EDGAR implementation effect conditional on the information uncertainty of financial disclosure at the stock level. When information uncertainty is high, retail investors have more discretion in cherry-picking information that confirms their biased priors, as the information can be subject to diverse interpretations. In contrast, if the disclosed information is more precise, it is less likely that retail investors will engage in information cherry-picking when interpreting the information.[3] Our results show that the increase in

[2] To address potential estimation biases arising from time-varying treatments and treatment effect heterogeneity in the staggered DiD tests, we use the stacked DiD method in all regressions following Baker, Larcker, and Wang (2022) and Chang, Hsiao, Ljungqvist, and Tseng (2022). Chang et al. (2022) note that the SEC assigned firms to EDGAR phase-in waves randomly conditional on size, and we follow them to select control firms using propensity score matching based on firm size.

[3] We measure information uncertainty mainly using three sets of measures: (i) the textual readability of corporate 10-K reports measured by the Fog index, the Flesch Reading Ease Index, and the Kincaid Index constructed following Li (2008) (ii) market reaction surrounding most recent earnings announcements following Landsman

investors' anchoring tendency following EDGAR implementation is greater when information uncertainty of a treated stock is higher, e.g., when the readability of financial reports is lower or when the divergence of opinions upon information release is greater. Moreover, we find that the effect is especially evident shortly after the releases of corporate 10K/Q reports, confirming that our findings result from retail investors' cherry-picking of new information which becomes easily accessible after the EDGAR implementation.

Next, we perform tests conditional on retail investors' access to the Internet and their sophistication. We find that before the EDGAR implementation, both Internet and sophisticated investors are less likely to be subject to anchoring bias than their counterparts.[4] However, since the EDGAR system increases the accessibility of corporate filings through the Internet, it should exert a stronger effect on Internet users than non-Internet users. Similarly, sophisticated investors are more likely to go through corporate filings when making investment decisions. Unsophisticated investors who lack investment experience and knowledge may pay less attention to corporate disclosures on EDGAR. We document supporting evidence. It renders further support for our hypothesis: the increase in anchoring tendency results from investors' confirmation bias due to better information dissemination after the EDGAR implementation.

We further show that amplified anchoring bias in the post-EDGAR period leads to worse mark-to-market investment performance. Our analysis indicates that buying a stock close to its 52-week high is more profitable in the post-EDGAR period. However, as shown in our main results, these retail investors are more likely to sell stocks when the price is close to the 52-week high in the post-EDGAR period, which worsens their overall investment performance.

We perform a series of tests to address potential influences of confounding factors. Chang et al. (2023) suggest that retail investors pay more attention to a firm when its mandatory disclosures are more easily accessible. This heightened attention may trigger more behavioral biases. To exclude such an alternative hypothesis, we control for changes in investors' attention at the stock level after the EDGAR implementation using abnormal trading volume and returns following Barber and Odean (2008). The scale of our focal estimators remains unchanged. Additionally, we account for changing investor characteristics after the EDGAR

et al. (2012), and (iii) basic firm characteristics including size and age.

[4] Sophisticated investors' lower anchoring propensity than unsophisticated investors is intuitive. Internet users' lower anchoring tendency relative to non-internet users is also compatible with the findings of Barber and Odean (2002) that internet users outperform before online trading, which is indicative of their lower tendency to make investment mistakes, including being anchored by stock price 52-week high.

implementation by including investor-EDGAR joint fixed effects. We also control for potential changing return patterns after EDGAR by incorporating interaction terms between EDGAR and various past returns. Our findings withstand all these tests. Moreover, given the high *t*-statistic values of our estimated coefficients, the concern of multiple hypotheses testing pointed out by Heath, Ringgenberg, Samadi, and Werner (2023) is not a major concern of our study. We also conduct a set of sensitivity checks where we use alternative EDGAR inclusion dates, exclude transitional files, employ alternative regression specifications, or adopt alternative information uncertainty measures. The results remain qualitatively similar.

Our study contributes to the literature in several ways. First, although it is generally believed that easier access to stock information reduces valuation uncertainty and helps mitigate investors' behavioral biases (e.g., Hirshleifer, 2001), we show that confirmation bias plausibly induces an opposite effect on the degree of anchoring bias. The surge in corporate information facilitates retail investors' cherry-picking of information that only confirms their pre-existing anchoring bias. Our results provide the first causal evidence that more information available to retail investors does not necessarily bring better welfare when confirmation bias is at play.[5] Our findings bring novel insight to policymaking that aims to protect retail investors' welfare via better accessibility of corporate disclosure such as the EDGAR implementation. The results suggest to policymakers that educating retail investors to process information objectively may be equally or more important than simply providing information access.

Second, our study sheds light on individuals' confirmation bias. The empirical evidence of confirmation bias in stock markets is sparse due to the challenges of identifying a clean testing environment. For example, Chang and Cheng (2015) use market capitalization to measure a stock's information environment and show that firms with rich information exhibit a drift in intangible returns (Daniel and Titman, 2006). They attribute their findings to investors' confirmation bias in interpreting firm intangible information when there is a rich information environment. Park, Konana, Gu, Kumar, and Raghunathan (2013) and Cookson, Engelberg, and Mullins (2023) propose to identify investors' confirmation bias based on how they process information from message boards or social networks of investors. We extend this limited line of research by providing the first causal evidence on how investors' confirmation bias, coupled

[5] In this sense, our paper is also related to the review article by Blankespoor, deHaan, and Marinovic (2020) who show that disclosure processing costs affect investors, stock returns, and other market outcomes. In particular, they categorize the costs into awareness costs, acquisition costs, and integration costs. Our paper shows that the reduction of acquisition costs via EDGAR implementation affects the integration process of disclosure when the information cherry-picking effect kicks in.

with exogenously increased corporate information accessibility, may exacerbate their inherent behavioral biases. Consistent with the confirmation bias interpretation, this effect is more prominent when retail investors have greater discretion in interpreting the disclosed information. We also add to the confirmation bias literature by providing evidence in a real high-stakes decision-making scenario.

Third, our study is related to the emerging literature that examines the real effects brought by the implementation of the EDGAR system but with a different focus. The existing studies along this line mainly investigate how modern information technologies, represented by the implementation of the EDGAR system, affect information production (Gao and Huang, 2020; Chang, Ljungqvist, and Tseng, 2023), corporate investments and financing (Goldstein, Yang, and Zuo, 2023; Bird, Karolyi, Ruchti, and Truong, 2021), external information acquisition (Goldstein et al., 2023; Bird et al., 2021), and investor disagreement in asset pricing (Chang, Hsiao, Ljungqvist, and Tseng, 2022). We differ from these studies by investigating the net effect of increased information accessibility on retail investors' behavioral biases. Our findings suggest that information dissemination after the EDGAR implementation may interact with retail investors' confirmation bias, thereby amplifying their inherent behavioral biases.

Finally, we add to the literature on the anchoring bias by providing account-level evidence on how the 52-week high reference point influences retail investors' trading decisions. The 52-week high has been shown to serve as a robust anchor for stock market participants.[6] Surprisingly, there is little account-level evidence that directly shows the influence of the 52-week high stock price on retail investors' stock trading behavior. An exception is a contemporaneous study by Vedova, Grant, and Westerholm (2023) who use data from Euroclear Finland and show that retail investors' limit order selling drastically increases leading up to and on the 52-week high day. Our study not only documents the 52-week anchoring bias in retail trading in the U.S. but, more importantly, shows that anchoring bias becomes stronger after better access to corporate information. We further show that the worsened bias plausibly leads to more mark-to-market trading losses when stock prices are close to the 52-week high.[7]

[6] The literature has shown that the 52-week high affects the exercise of options (Heath et al., 1999; Poteshman et al., 2003), cross-sections of stock returns (George and Hwang, 2004; George, Hwang, and Li, 2018; Huang, Lin, and Xiang, 2021), trading decisions (Huddart, Lang, and Yetman, 2009), implied volatility (Driessen, Lin, and Van Hemert, 2013), offer prices in mergers and acquisitions (Baker et al., 2012; Ma et al., 2019), and analyst recommendation revisions (Li, Lin, and Lin, 2021).

[7] Gao and Huang (2020) show that after the EDGAR implementation, retail trades become more informative about future stock returns. Our results may not be at odds with their findings. Although we show that retail

2. Literature review and hypothesis development

2.1 Literature review

2.1.1 Confirmation bias and information cherry-picking

Confirmation bias refers to individuals' tendency to give undue weight to evidence supporting their prior beliefs (Kahneman, 2011). This bias has attracted great attention in the literature on how people consume information on the Internet and social media. At its advent, the Internet and social media seem to be ideal solutions to behavioral biases as they make access to various kinds of information more convenient. However, vast studies find just the opposite: the rise of the Internet and social media facilitate individuals' information cherry-picking, as it becomes easier for people to find messages supporting what they want to believe.

Slechten et al. (2022) examine participants' search for social and political information on Google, and find that they tailor their information exposure by formulating search queries to maintain prior beliefs. Mullainathan and Shleifer (2005) illustrate that the increase in the number of news outlets leads consumers to become more segregated ideologically. Gentzkow et al. (2011) find that ideological segregation of online news consumption is higher than the segregation of most offline news consumption. Bessi (2016) shows that social media users consume information according to their preferences and join virtual polarized communities wherein they reinforce their prior beliefs. Moravec et al. (2019) reveal that Facebook "fake news" flags are not effective in altering users' prior beliefs.

The investigation of confirmation bias in stock markets, however, is rather limited. Pouget et al. (2017) propose a model and show that confirmation bias provides a unified rationale for excess volatility, excess volume, and momentum in stock markets. Park et al. (2013) find that investors exhibit confirmation bias when processing information from message boards. They further show that investors with stronger confirmation bias engage in excessive trading and experience lower investment returns. Chang and Cheng (2015) argue that investors exhibit confirmation bias in interpreting firm intangible information when there is a rich information environment. Cookson et al. (2023) examine posts and follower connections on StockTwits, and show that users build a personalized newsfeed that supports their original views.

investors suffer more from their anchoring bias after the EDGAR implementation via information cherry-picking, such bias only becomes relevant and prominent when stock prices are close to the 52-week high. In other situations, retail investors could still produce more information with the help of EDGAR as shown in Gao and Huang (2020).

However, to draw a causal inference, it is ideal to have exogenous shocks on the information environment to test whether more information results in information consumers' greater tendency to stick to their prior beliefs. We utilize the exogenous shock from the EDGAR implementation on information supply to show that confirmation bias increases information cherry-picking, thereby strengthening the anchoring bias, as we will illustrate below.

2.1.2 Anchoring propensity and the 52-week high effect

The seminal work of Tversky and Kahneman (1974) shows that when forming estimations, individuals tend to use some salient, easily available number as a reference point or an anchor, even when the number contains no relevant information. Such an anchoring propensity is perhaps one of the most well-documented and acknowledged human behavioral biases.

In the financial decisions, Heath et al. (1999) show that employee option exercise activity nearly doubles when the stock price exceeds its maximum level over the previous year, suggesting that the 52-week high influences individuals' option exercise decisions. George and Hwang (2004) show that investors tend to use the 52-week high as a reference point when trading, which explains a significant proportion of the profits from momentum investing. Huddart et al. (2009) find that trading volume spikes when stock prices cross either the upper or lower limit of the 52-week trading range, implying that past price extremes influence investors' trading decisions. Baker et al. (2012) document that in mergers and acquisitions, offer prices are biased towards the 52-week high of the target firm's price, which is a salient but irrelevant anchor. Huang et al. (2021) find that investors' tendency to use the 52-week high as an anchor when making investment decisions generates underreaction to the news of a focal firm's economically linked peers. Utilizing individual investor data from Finland, Vedova et al. (2023) show that households do sharply increase their selling at the 52-week high.

Given the ample evidence on the 52-week anchoring bias, we treat it as the prior bias retail investors have and focus on how the confirmation bias could amplify it via information cherry-picking. Such a well-documented behavioral bias addresses one of the empirical challenges for studying confirmation bias in the stock market.

2.1.3 EDGAR implementation, information dissemination, and the real effects

Until the early 1990s, corporate files can only be accessed physically in a few SEC's public reference rooms. In 1993, the SEC starts its staggered implementation of the EDGAR system, through which listed firms are required to submit their electronic corporate files online. This

substantially reduces investors' information acquisition costs and greatly increases the dissemination of corporate information. Such exogenous information shocks address the other empirical challenges for studying confirmation bias in the stock market.

Existing studies mainly utilize the staggered implementation of the EDGAR system to explore its influences on the information environment and the information production and consumption of various market participants. For example, Gao and Huang (2020) document that greater dissemination of corporate information increases information production by corporate outsiders, including both individual investors and analysts. Chang et al. (2022) show that the EDGAR implementation reduces investor disagreement, leading to stock price corrections. On the other hand, Bird et al. (2021) show that the implementation of EDGAR has an unintended consequence of crowding out external information gathering, resulting in market prices being significantly less informative for corporate managers and thus reducing their incentives to learn from the stock price for investment decisions. Goldstein et al. (2023) find that EDGAR implementation increases corporate investment via improved equity financing and enhanced managerial incentives. In addition, Chang et al. (2023) show that mandatory disclosure such as the EDGAR implementation can reduce a strategic component of sell-side equity analyst behavior as investors can verify their forecasts using corporate filings.

2.2 A suggestive model and hypothesis development

Based on these studies and combining the confirmation bias with anchoring bias, we propose another unintended consequence of the EDGAR implementation: it may worsen some investors' welfare when the stock prices are close to salient reference points. To clarify this argument, we propose a suggestive model of information cherry-picking to develop our hypotheses. This model is built upon the work of Rabin and Schrag (1999) and Pouget, Sauvagnat, and Villeneuve (2016) and presumes investors have both anchoring and confirmation biases. The novel insight from the model is that when investors have confirmation bias, public disclosure will amplify their anchoring bias.

2.2.1 Model setup

There are two dates in the economy, and each date is indexed by $t \in \{0, 1\}$. There is one risky asset that has a return R at $t = 1$. At $t = 1$, R takes high or low values with equal probability. We normalize the high value of R to be 1 and normalize the low value of R to be -1.

There is one investor with behavioral biases (biased investor). This investor can buy or sell one unit of the risky asset but incur a random cost c, which will be realized when the investor makes her trading decision. c follows a uniform distribution from $[0, 1]$. We can interpret this cost as an outside opportunity. Noted that the cost, c, is to avoid the corner solution.

Anchoring bias. Before trading, the biased investor observes a signal that serves as an anchor to evaluate the risky asset's cash flow. We denote this signal as S_A. It can represent the 52-week values of the asset price. S_A takes a value of $52W_High$ or $52W_Low$. When S_A takes a value of $52W_High$, it represents that the current price is close to the 52-week high, and the biased investor thinks that the risky asset has been overvalued such that R is more likely to be -1. Thus, the investor tends to sell the risky asset. When S_A takes a value of $52W_Low$, it represents that the current price is close to the 52-week low (far away from the 52-week high), and the biased investor thinks that the risky asset has been undervalued such that R is more likely to be 1. As a result, the investor tends to buy the risky asset.

Noted that while the investor uses S_A as the anchor, she misinterprets the signal. We set S_A to be informative about the return of the risky asset:

$$P(R = 1|S_A = 52W_High) = P(R = -1|S_A = 52W_Low) = p.$$

But the biased investor has the belief as:

$$P^B(R = -1|S_A = 52W_High) = P^B(v = 1|S_A = 52W_Low) = p_B, \text{ with } p_B > \frac{1}{2}.$$

As we can see, whether the investor correctly interprets the signal depends on p. For example, when $p < \frac{1}{2}$, $P(R = -1|S_A = 52W_High) = P(R = 1|S_A = 52W_Low) = 1 - p > \frac{1}{2}$, which means that the investor correctly interprets the signal. In contrast, when $p > \frac{1}{2}$, the investor misinterprets the signal. The context of the 52-week high is corresponding to the case of $p > \frac{1}{2}$ (George and Hwang, 2004).

Public disclosure. Before trading, the biased investor receives a public signal S_P, whose probability distribution depends on the state of the risky asset. We assume:

$$P(S_P = 0) = 1 - \pi, P(S_P = 1|R = 1) = P(S_P = -1|R = -1) = \pi\theta, \text{ with } \theta > \frac{1}{2}.$$

Parameters π and θ represent the likelihood that the signal is informative and the level of informativeness, respectively. Specifically, $\pi = 0$ represents the case in which there is no

public disclosure. In this sense, a higher π means more public information is available for the investor. Noted that this model setup is based on Pouget, Sauvagnat, and Villeneuve (2016), as introducing an uninformative signal, $S_P = 0$, so that the biased investor might neglect the informational content of news and does not realize that she is biased.

Confirmation bias. In addition to the anchoring bias, the biased investor is subject to confirmation bias. Following Rabin and Schrag (1999) and Pouget, Sauvagnat, and Villeneuve (2016), we assume that the biased investor misinterprets public information when it conflicts with her prior belief. That is, we consider that a confirmation-biased trader treats some conflicting signals as uninformative and is otherwise completely Bayesian.

Let $\mu^B = P^B(R = 1|S_A)$. The biased investor believes that:

$$\sigma^B = S_p I_{\left(\mu^B - \frac{1}{2}\right)S_p \geq 0} + zS_p I_{\left(\mu^B - \frac{1}{2}\right)S_p < 0},$$

where z takes a value of 0. It suggests that when S_P conflicts S_A, the biased investor ignores S_P and considers it as an uninformative signal.[8]

Biased investor's trading Our goal is to study how an investor's anchoring bias is affected when exposed to more public disclosure with confirmation bias. To guide our empirical exercises, we focus on the case of $S_A = 52W_High$. This case corresponds to the well-documented 52-week high anchoring bias. That is, when the biased investor observes that the current price is close to the 52-week high, she thinks that the risky asset has been overvalued and thus tends to sell it. However, the increasing price can be driven by fundamental reasons, but the biased investor overlooks the information content in the 52-week high and treats it as the anchor to make trading decisions. Such an anchoring bias leads to a positive return drift and explains part of the momentum anomaly as shown by George and Hwang (2004).

To measure how the biased investor responds to $S_A = 52W_High$, we calculate the likelihood that this biased investor sells the risky asset. For this investor, she trades-offs the cost, c, and the expected trading profit $E(R|F)$, where F represents the investor's information set.

[8] Instead, we can assume that z is a random variable taking the value between $\{0, 1\}$ with probability distribution $P(z = 0) = q$, while q measures the degree of confirmation bias. Our main analysis is corresponding to the case with $q = 1$.

2.2.2 A Benchmark without public disclosure

We start our analysis with the benchmark case without public disclosure. This case is corresponding to the case with $\pi = 0$. In this case, the investor believes that $P^B(R = -1|S_A = 52W_High) = p_B$, suggesting that $\mu^B = P^B(R = 1|S_A = 52W_High) = 1 - p_B < \frac{1}{2}$. Thus, upon receiving $S_A = 52W_High$, if the investor submits a sell order, and the expected trading profit is:

$$E(R|S_A = 52W_High) = 2p_B - 1.$$

Considering randomness of c, the likelihood of selling the risky asset is:

$$L = prob(2p_B - 1 > c) = 2p_B - 1.$$

2.2.3 Public disclosure and confirmation bias

We move to the analysis with public disclosure and confirmation bias. Conditional on $S_A = 52W_High$, the distribution of S_p is as follows:

$$S_p = \begin{cases} 0, \text{with prob } 1 - \pi \\ 1, \text{with prob } \pi[p\theta + (1-p)(1-\theta)] \\ -1, \text{with prob } \pi[p(1-\theta) + (1-p)\theta] \end{cases}.$$

Considering the confirmation bias, the biased investor observes the following signal:

$$\sigma^B = \begin{cases} 0, \text{with prob } 1 - \pi + \pi[p\theta + (1-p)(1-\theta)] \\ 1, \text{with prob } 0 \\ -1, \text{with prob } \pi[p(1-\theta) + (1-p)\theta] \end{cases}.$$

While the biased investor has confirmation bias, she updates her belief using Bayes' rule as if the perceived signal were the one truly generated. That is, upon receiving the signal σ^B, the investor believes that:

$$P^B(R = -1|S_A = 52W_High, \sigma^B = 0) = p_B,$$

$$P^B(R = -1|S_A = 52W_High, \sigma^B = 1) = \frac{p_B(1-\theta)}{p_B(1-\theta)+(1-p_B)\theta},$$

which is larger than $\frac{1}{2}$ only if $\frac{p_B}{1-p_B} > \frac{\theta}{1-\theta}$.

$$P^B(R=-1|S_A=52W_High,\sigma^B=-1)=\frac{p_B\theta}{p_B\theta+(1-p_B)(1-\theta)}>p_B>\frac{1}{2}.$$

Considering randomness of c, the likelihood of selling the risky asset is:

$$L^C=\{1-\pi+\pi[p\theta+(1-p)(1-\theta)]\}(2p_B-1)$$
$$+\pi[p(1-\theta)+(1-p)\theta]\left(\frac{2p_B\theta}{p_B\theta+(1-p_B)(1-\theta)}-1\right)$$

$$=2p_B-1+\pi[p(1-\theta)+(1-p)\theta]\frac{2p_B(1-p_B)(2\theta-1)}{p_B\theta+(1-p_B)(1-\theta)}$$

$$=\underbrace{2p_B-1}_{L}+\underbrace{\pi[p+(1-2p)\theta]\frac{2p_B(1-p_B)(2\theta-1)}{p_B\theta+(1-p_B)(1-\theta)}}_{Addtional\ Selling}.$$

Based on the above equation of the selling, we first characterize the effect of public disclosure on the biased investor's selling behavior in the following proposition.

Proposition 1: *The additional selling driven by anchoring bias is* $\pi[p+(1-2p)\theta]\frac{2p_B(1-p_B)(2\theta-1)}{p_B\theta+(1-p_B)(1-\theta)}$*; it is positive when* $\pi>0$.

As shown in Proposition 1, the presence of public disclosure amplifies the selling behavior driven by the anchoring bias. The direct implication from **Proposition 1** is that the additional selling driven by anchoring bias is increasing with π.

We also study the effect of confirmation bias on the selling. That is, if the investor only has anchoring bias but no confirmation bias. The signal she interprets is as follows:

$$\sigma^B=\begin{cases}0,\text{with prob } 1-\pi\\ 1,\text{with prob } \pi[p\theta+(1-p)(1-\theta)].\\ -1,\text{with prob } \pi[p(1-\theta)+(1-p)\theta]\end{cases}$$

Considering randomness of c, the likelihood of selling the risky asset is:

$$L^{NC}=(1-\pi)(2p_B-1)+\pi[p\theta+(1-p)(1-\theta)]\}max\left(\frac{2p_B(1-\theta)}{p_B(1-\theta)+(1-p_B)\theta},0\right)+$$
$$\pi[p(1-\theta)+(1-p)\theta]\left(\frac{2p_B\theta}{p_B\theta+(1-p_B)(1-\theta)}-1\right).$$

The comparison between L^C and L^{NC} shows the effect of confirmation bias on her

selling behavior. Thus, we have the following lemma regarding the effect of confirmation bias.

Lemma 1: *The likelihood of selling is higher when the investor has confirmation bias. That is,* $L^C > L^{NC}$.

We push further and conduct comparative statics regarding the uncertainty in public disclosure. Specifically, we can measure the informativeness of public disclosure by θ. Taking derivatives of L^C with θ, we have

$$\frac{\partial L^C}{\partial \theta} = \pi 2 p_B (1 - p_B) \left((1-2p) \frac{1}{p_B + \frac{1-\theta}{2\theta - 1}} + [p + (1-2p)\theta] \frac{1}{\left(\frac{1-\theta}{2\theta-1} + p_B\right)^2} \frac{1}{(2\theta-1)^2} \right).$$

Interestingly, we find that when $p < \frac{1}{2}$ (e. g., investor correctly interprets S_A), $\frac{\partial L^C}{\partial \theta} > 0$. This result suggests that when the investor has no anchoring bias and correctly interprets S_A, selling driven by the negative signal of S_A is always increasing with the informativeness of public disclosure. The intuition lies in the correlation between S_A and S_P. When S_A is informative, S_P is more likely to have the same direction of S_A when S_P is more informative. Thus, more informative S_P amplifies more anchoring on S_A.

However, this is not the case in the context of the 52-week high anchoring bias. In the context of the 52-week high anchoring bias, the investor misinterprets S_A. That is, upon receiving $S_A = 52W_High$, the biased investor believes that risky asset tends to have a low return. But the truth is that $S_A = 52W_High$ indicates that the risky asset tends to have a high return. In this case, when S_P is more informative, it is less likely to generate the signal in the same direction as the "signal" perceived by the investor, damping the anchoring on S_A. We formally show the results as follows.

Proposition 2: The informativeness of public disclosure has impacts on the likelihood of selling as follows:

1) When the investor correctly interprets his original signal (that is, $p < \frac{1}{2}$), the likelihood of selling is increasing with the public disclosure informativeness. $\frac{\partial L^C}{\partial \theta} > 0$.

2) When the investor has anchoring bias and incorrectly interprets her original signal (that is, $p > \frac{1}{2}$), there are two cases.

   - When $\theta < \theta^*$(here, $\theta \geq \frac{1}{2}$), the likelihood of selling is increasing with the public disclosure informativeness;
   - When $\theta > \theta^*$, the likelihood of selling is decreasing with the public disclosure informativeness; Here, θ^* is defined in the Appendix.

See Proof in Internet Appendix B.

2.2.4 Hypothesis development

Based on the above propositions, we propose three hypotheses. First, **Proposition 1** shows public disclosure can amplify investors' anchoring tendency when the investor also has confirmation bias. Focusing on the EDGAR implementation and anchoring on the 52-week high, we propose our first hypothesis as follows.

Hypothesis 1*: More information coupled with investors' confirmation bias could amplify rather than alleviate investors' pre-existing anchoring biases*.

Proposition 2 highlights the role played by the anchoring bias. When the investor with anchoring bias incorrectly interprets the signal, the less informative public disclosures can amplify the anchoring bias even more. Thus, we have the following hypothesis.

Hypothesis 2*: The effect of EDGAR implementation on retail investors' anchoring tendency towards the 52-week high is more pronounced among stocks with less informative public disclosure.*

The above hypotheses are centered around public disclosure. Intuitively, when investors pay more attention to public disclosure (measured by increasing π), the anchoring bias has a stronger effect. Thus, we have the following third hypothesis.

Hypothesis 3*: The effect of EDGAR implementation on retail investors' anchoring tendency towards the 52-week high is more pronounced among investors who pay close attention to public disclosure.*

2.3 The experimental analysis

This section provides preliminary evidence for our main hypothesis by simulating individuals' information acquisition process and investment decision-making within a controlled experimental environment. The setup allows us to detect the presence of anchoring bias and investigate how the interplay of increased information dissemination and individuals' selective information consumption amplifies their inherent anchoring bias. The experiment sets the stage for the subsequent empirical investigation in the high-stakes stock market.

2.3.1 Methods

We conduct the experiment on Credamo, a China-based online platform with a participant pool of three million, comparable to Amazon Mechanical Turk or Prolific in the U.S.[9] We randomly recruit two groups of participants through Credamo, each initially comprising 125 subjects, culminating in a final sample of 246 participants. Group 1, the treatment group, consists of 125 participants (average age: 28.46; female: 58.40%), while Group 2, the control group, has 121 participants (average age: 28.66; female: 57.85%).

The experiment employs a 2-stage process for both groups. In the first stage, after viewing a simulated historical price chart of a stock, subjects are required to make initial predictions about the stock's future price movement (increase, decrease, or remain stable) and rate their confidence on a scale from 1 to 7. Subjects in Group 1 (treatment group) are shown a price chart where the current stock price is close to its 52-week high, whereas subjects in Group 2 (control group) are presented with a chart displaying a very similar price trend but with the current price still far from its 52-week high. The critical difference between the groups is thus the proximity of the current stock price to its 52-week high.

After submitting their initial assessments, subjects in both groups proceed to the second stage, where they receive an identical set of fundamental financial information that includes two positive and two negative pieces of information. The four pieces of information are balanced in sentiment to offer an overall neutral perspective on the stock's business condition.

[9] Our experimental results also show that this amplification effect of confirmation bias on prior biases is not limited to the United States.

Subjects are asked to select the piece of information they consider most important for their decision-making. They are then required to make another price direction prediction and reassess their confidence level.

Our hypothesis predicts that subjects in Group 1, compared to those in Group 2, are more likely to anticipate a stock price decline due to the current price's proximity to the 52-week high. Furthermore, in the second stage, they are expected to reinforce this downward prediction with increased confidence by selectively focusing on negative information due to their confirmation bias.

2.3.2 Experimental results

In stage 1, 52.8% of Group 1 subjects predict a price decline. In contrast, only 10.2% of subjects in Group 2 anticipate a price decline. The difference between the two groups is significant at the 1% level. The evidence confirms that subjects are influenced by a stock's proximity to its 52-week high when predicting its future price movements, demonstrating the role of anchoring bias in shaping their initial beliefs.

We then examine whether subjects cherry-pick information provided in stage 2 conditional on their initial predictions made in stage 1. The results indicate that subjects tend to prioritize information that aligns with their initial predictions. Specifically, subjects who initially anticipate a price increase show an emphasis on positive information (71.19% in Group 1, 66.04% in Group 2), while those expecting a price decline focus on negative information (77.27% in Group 1, 75.00% in Group 2). This pattern is consistent with individuals' cherry-picking of information that confirms their priors.

The highlight of our experiment is the exploration of the interplay between anchoring bias and confirmation bias. For this analysis, we focus only on subjects who have exhibited the anchoring bias in stage 1, e.g., subjects in Group 1 who predict a price decline. We find that among subjects who have displayed anchoring bias in stage 1 and selectively focus on negative information in stage 2, their self-reported confidence in their second predictions increases from 4.57 in stage 1 to 5.0 in stage 2. This increase in self-reported confidence in the prediction accuracy is statistically significant at the 1% level.

The results provide initial evidence that more information can amplify inherent anchoring bias through confirmation bias. Individuals cherry-pick information and become more

confident with their biased priors. Next, we test hypotheses in a real-world high-stakes setting.

3. Institutional background, data, and sample

3.1 Institutional background

Before the implementation of the EDGAR system in 1993, public firms in the U.S. are required to submit paper copies of corporate filings to the SEC, which are stored physically in the SEC's public reference rooms. Investors who want to check these corporate filings need to either visit one of the three SEC's public reference rooms in person or pay a nontrivial subscription fee to commercial data vendors. Given that corporate filings can only be viewed by one reader at a time, the dissemination of corporate information is further restricted by the limited number of corporate filing copies. Further, the large volumes of corporate files in the reference room make it difficult to identify and analyze data of a specific firm. It is also difficult to maintain the orders of corporate filings, and files are often misplaced or stolen (Gao and Huang, 2020).

To facilitate the dissemination of corporate filings, the SEC develops an automated system through which firms can submit electronic corporate filings through the Internet, or EDGAR. This system substantially increases investors' access to corporate filings at much lower costs. Corporate filings can be easily accessed online with few restrictions, and it becomes much easier for investors to locate specific pieces of corporate information, e.g., by using the search function when reading electronic documents.

The SEC announces the phase-in schedule for the mandatory implementation of the EDGAR system on February 23, 1993. All public firms are divided into 10 groups, and each group is phased at different times. The firms in the first group have to start filing their electronic corporate files (e.g., 10-K, 8-K) through EDGAR in April 1993, while the firms in the last group have to do so in May 1996.

As Chang et al. (2022) note, although the staggered implementation of the EDGAR system starts in 1993, access to these online files is not publicly free for the first four implementation waves which take place before 1994. In October 1993, the National Science Foundation (NSF) decides to acquire Mead Data Central's historic EDGAR filings and provides funding to a project to make all past and current EDGAR filings available online for free. Such a project is hosted by New York University's Stern School of Business. The free online access to EDGAR goes live on January 17, 1994, when the historic and current filings of firms in the first four

implementation waves become available via the online-access system of NYU. Given that retail investors are less likely to pay for corporate filings, we follow Chang et al. (2022) to consider the actual shocks to information access to take place only after the NSF-funded NYU online-access system became available.

3.2 Data and sample

We obtain the list of firms on the EDGAR phase-in schedule between 1993 and 1996 from Appendix B of SEC Release No.33-6977 (released on February 23, 1993) following Gao and Huang. (2020). The file contains each firm's Central Index Key (CIK) and phase-in group number (CF 01 to 10). Detailed information is shown in Table IA1 of the Internet Appendix.[10]

We acquire retail trading data from the large discount brokerage database of Barber and Odean (2000), the LDB data set hereafter, which covers trades made by 78,000 households between 1991 and 1996. Gao and Huang (2020) point out that about a quarter of the investors in the LDB data set reside in California, which is the most active state of participation in the early years of the Internet. With more tech-savvy individual investors, the sample is particularly appropriate for assessing the influence of information dissemination via the Internet on retail investors' information cherry-picking in their trading decisions.

In our sample, we include only trades that are made for common stocks traded on NYSE, NASDAQ, or AMEX and for which the stock phase-in schedule can be obtained from the SEC release mentioned above. To mitigate the concern that IPO timing might be endogenous to the EDGAR implementation, we exclude firms that go public after the release of the rules by requiring firms' financial information to be available in Compustat as of January 31, 1993, following Gao and Huang (2020). We obtain financial data from Compustat, stock price, and return data from the CRSP. In total, we obtain 983,932 individual-stock-date observations over the period from 1991 to 1996 in our final sample.

Table 1 shows the summary statistics for the main variables. $EDGAR_{i,t-1}$ is a dummy indicating whether the stock has become subject to the mandatory filings to EDGAR with free access online. For firms involved in the first four implementation waves, given that their electronic EDGAR filings do not become publicly available online for free until January 17,

[10] Note that for the first four implementation waves, firms' electronic fillings become freely accessible online only after January 17, 1994. For the rest six implementation waves that take place after 1994, firms' corporate files are free online once they have been uploaded to the EDGAR system according to the schedule.

1994, $EDGAR_{i,t-1}$ equals one for these firms only after that date. For firms involved in the following six implementation waves, $EDGAR_{i,t-1}$ equals one once included in the EDGAR system. Such an identification strategy is employed by Chang et al. (2022). The mean value of $EDGAR_{i,t-1}$ is 0.402. That is, 40.2% of the retail trading in our sample occurred after EDGAR implementation. We examine retail investors' tendency to be anchored by a stock's 52-week high in trading based on $P_{i,t-1}^{52h}$, which is the price of stock i on day $t-1$ scaled by its highest level over the past 52 weeks. The mean value of $P_{i,t-1}^{52h}$ is 0.795, indicating that the average stock price before the trading day is about 79.5% of the highest price during the previous year.

[Insert Table 1 about here]

A set of stock characteristics are constructed and included in our investigation following the literature (e.g., Li et al., 2021), including stock i's firm size measured by its natural logarithm of market value on day $t-22$ ($Size_{i,t-1}$), book value over market value at the end of the previous year ($B/M_{i,t-1}$), cumulative returns over different horizons before the trading date ($Runup_{i,t-5\sim t-1}$, $Runup_{i,t-21\sim t-6}$, $Runup_{i,m-6\sim m-2}$, $Runup_{i,m-12\sim m-7}$), and its average daily turnover ratio, bid-ask spread, idiosyncratic volatility, and idiosyncratic skewness in the previous month, which are denoted as $Turnover_{i,t-1}$, $Bid\text{-}ask\ spread_{i,t-1}$, $Idiosyncratic\ volatility_{i,t-1}$, and $Idiosyncratic\ skewness_{i,t-1}$, respectively.

4. EDGAR implementation and retail investors' anchoring propensity

4.1 Baseline tests

EDGAR phase-in schedule provides a natural staggered DiD design as our identification strategy (e.g., Gao and Huang, 2020). However, recent studies (e.g., Baker et al., 2022) contend that such regressions are susceptible to estimation biases introduced by treatment effect heterogeneity and propose three alternative estimators for addressing the bias: the Callaway and Sant'Anna (2021) estimator, the Sun and Abraham (2021) estimator, and the stacked regression estimator (e.g., Cengiz, Dube, Lindner, and Zipperer, 2019). The first two types of estimators are closely related and are estimated based on changes in the average treatment effect (ATT). Given that our investigation is conducted based on the changes in retail trading's sensitivity to a stock's 52-week high, which is captured by regression coefficients, remedies utilizing ATT are not feasible. Therefore, we employ the stacked regression estimator, which is also used by Chang et al. (2022) in exploiting the staggered implementation of EDGAR, to perform the main tests for alleviating concerns about potential bias.

We follow Baker et al. (2022) to create event-specific "clean 2 × 2" datasets to make sure that the control firms within treatment windows have not yet been treated. In other words, in examining treated firms in each EDGAR phase-in group, only firms that have not yet become subject to the mandatory filing on EDGAR are included as controls. Given that there are 10 EDGAR phase-in groups, and the last group lacks "clean" controls (untreated firms), only firms in the first 9 EDAGR phase-in groups are used as treated firms in the staggered regressions.[11]

Further, Chang et al. (2022) argue that SEC assigned firms to EDGAR phase-in waves randomly conditional on firm size. Hence, the control firms selected should be similar in size to the treated firms. We follow their study and use the propensity score matching methodology to select the control firms that not only are from the "clean 2 × 2" datasets but also have the nearest-neighbor propensity score to treated firms on market capitalization. These event-specific datasets are then stacked together for next-step regressions.

We analyze retail buys and sales separately. We first run the following regression to examine retail buys:

$$Buy_{i,j,t} = c_{i,w} + c_{j,w} + c_{t,w} + \beta_1 \times P_{i,t-1}^{52h} + \beta_2 \times P_{i,t-1}^{52h} \times EDGAR_{i,t-1} + Controls + \varepsilon_{i,j,t,} \quad (1)$$

where $Buy_{i,j,t}$ is proxied by three alternative retail-buy measures: (i) a dummy that equals one if investor *j* buys stock *i* on day *t* and zero otherwise ($Buy_{i,j,t}$), (ii) the logarithm of the number of shares of stock *i* bought by investor *j* on day *t* ($BuyShr_{i,j,t}$), and (iii) the logarithm of the value of stock *i* bought ($BuyValue_{i,j,t}$). The variable $P_{i,t-1}^{52h}$ is the price of stock *i* on day *t–1* scaled by its highest level over the past 52 weeks, and $EDGAR_{i,t-1}$ is a dummy that equals one if stock *i*'s filing are available online for free via EDGAR on date *t–1* and zero otherwise.

The control variables are selected following Li et al. (2021), which include the momentum effect, firm size, book-to-market ratio, turnover ratio, bid-ask spread, idiosyncratic volatility, and idiosyncratic skewness. The momentum effect is captured by cumulative return over the periods of [*t–5*, *t–1*] [*t–21*, *t–6*], [*m–6*, *m–2*], and [*m–12*, *t–7*], where *t* represents the day *t*, and *m* represents the month *m*. Firm size is measured by the natural logarithm of market value on day *t–22*; book-to-market ratio is book value over market value at the end of the previous year; other controlled firm characteristics are all calculated in the previous month. Wave-specific

[11] Chang et al. (2022) also exclude waves 8 and 9 based on a lack of clean controls. However, our focus on retail investors' trading behaviors allows us to maintain enough observations compared with a firm-level analysis. Thus, retaining all 9 waves is one point of departure from Chang et al. (2022) in terms of regression setup.

Account, *stock*, and *date* fixed effects are all controlled for, and standard errors are double clustered by wave-specific account and date.[12]

If the stock price 52-week high acts as an anchor that affects investors' trading decisions, the coefficient on $P_{i,t-1}^{52h}$ in Eq. (1) should be negative. If the increased information dissemination after EDGAR implementation alleviates investors' anchoring tendency due to reduced valuation uncertainty, the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ should be positive. If, however, the increased information dissemination coupled with confirmation bias exacerbates investors' anchoring bias, the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ will be negative.

We investigate retail sales using the following regressions:

$$Sell_{i,j,t} = c_{i,w} + c_{j,w} + c_{t,w} + \beta_1 \times P_{i,t-1}^{52h} + \beta_2 \times P_{i,t-1}^{52h} \times EDGAR_{i,t-1} + Controls + \varepsilon_{i,j,t,} \quad (2)$$

where $Sell_{i,j,t}$ is proxied by three alternative retail-sale measures: (i) a dummy that equals one if investor *j* sells stock *i* on day *t* and zero otherwise ($Sell_{i,j,t}$), (ii) the logarithm of the number of shares of stock *i* sold by investor *j* on day *t* ($SellShr_{i,j,t}$), and (iii) the logarithm of the value of stock *i* sold ($SellValue_{i,j,t}$.). The variables $P_{i,t-1}^{52h}$, $EDGAR_{i,t-1}$, and control variables are as defined in Eq. (1). Again, we control for wave-specific *account*, *stock*, and *date* fixed effects, and double cluster standard errors by wave-specific account and date.

The coefficient on $P_{i,t-1}^{52h}$ in Eq. (2) is expected to be significantly positive, indicating that retail investors are more likely to sell a stock when its stock price is close to the 52-week high, and vice versa. A positive coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ in Eq. (2) would imply that investors' anchoring tendency is amplified after EDGAR implementation, while a negative coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ would suggest the opposite.

The results are reported in Table 2. First, the coefficient on $P_{i,t-1}^{52h}$ is significantly negative (positive) in Panel A (B) when retail buys (retail sales) are examined. This result confirms that retail investors are anchored by stock price 52-week high in both buying and selling stocks. More importantly, the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ is significantly negative for retail buys but significantly positive for retail sales. It suggests that for both buys

[12] We have tried various standard error clustering methods, including account and date, account and stock, stock and date, and the results are qualitatively similar. Note that a stacked regression alternative defines the main variables within each wave-specific dataset, so that all the fixed effects are saturated with the indicators for phase-in waves.

and sales, retail investors' tendency to anchor on stock price 52-week high in trading is aggravated after EDGAR implementation. The finding supports our **Hypothesis 1** that when more information can be easily accessed, investors' confirmation bias is enhanced, which outweighs the effect of reduced valuation uncertainty and leads to amplified investors' anchoring bias.[13]

[Insert Table 2 about here]

4.2 Influence of information uncertainty

The EDGAR system facilitates investors' information acquisition, but does not eliminate the uncertainty of corporate information. Some corporations disclose information precisely while others do it vaguely. When information uncertainty is higher, it is more flexible and plausible for investors with confirmation bias to cherry-pick information. We thus expect the results to be stronger when information uncertainty is higher (**Hypothesis 2**).

We use three sets of measures to capture a firm's information uncertainty: (i) readability indices of corporate 10-K reports, constructed following Li (2008), (ii) abnormal trading volume and abnormal return volatility surrounding the most recent earnings announcements, constructed following Landsman et al. (2012), and (iii) basic firm characteristics including size and age.

The information uncertainty measures constructed based on textual features of corporate 10-K reports cannot be measured before the EDGAR implementation when the electronic files are not available. We thus measure the readability of 10-K reports for each firm during our entire sample period, assuming that a firm's disclosure behavior and thus corporate files' textual feature is stable over our sample period. We then use the time-series average of the corporate file features to capture the firm's information uncertainty. We acknowledge that constructing the measures in such a way will bring estimation biases. We thus employ our second and third set of information uncertainty measures, which could be measured before the implementation of the EDGAR system for each stock to provide complementary evidence.

[13] We conduct a supplementary experiment, detailed in Internet Appendix C for Experimental Evidence, to examine the effect of additional information on individuals' anchoring biases. The results indicate that providing more information tends to amplify rather than mitigate pre-existing behavioral biases, further supporting our baseline findings.

4.2.1 Textual features of EDGAR files

For each stock, we download all its corporate 10-K reports during the sample period and calculate their readability. We expect retail investors' confirmation bias to be weaker when corporate files have greater readability, which reduces potential ambiguity in interpreting the information conveyed. Hence, the increase in retail investors' anchoring bias following the EDGAR implementation is expected to be lower for the 10-K files with better readability.

We construct three commonly used readability indices, the Fog, Flesch Reading Ease, and Kincaid indices, following Li (2008). The indices are further normalized. A higher Flesch Reading Ease Index indicates greater readability, while the Fog and the Kincaid Index are inversed indicators. To mitigate the influence of measurement errors caused by a single index, we construct two overall readability measures based on these three readability indices. The first measure $Readability^{PCA}_{i,t}$ is constructed based on the first principal component of the three readability indices. It is an inversed indicator which positively associated with both the Fog and Kincaid indices but negatively associated with the Flesch Reading Ease Index. A higher value of $Readability^{PCA}_{i,t}$ indicates lower readability and thus greater information uncertainty.

The second measure $Readability^{Rank}_{i,t}$ is constructed based on decile rankings, from low to high, of the three readability indices. For both the Fog and Kincaid indices, stocks in the bottom decile are assigned a score of one while stocks in the top decile are assigned a score of 10. For the Flesch Reading Ease Index, we assign a score of 10 to stocks in the bottom decile and one to stocks in the top decile. For each stock, we add up its rankings of the three indices and denote the sum as $Readability^{Rank}_{i,t}$, which is also an inversed indicator for readability.

In Panels A and B of Table 3, we partition each trade into the high (low) information uncertainty subgroup if the corporate 10-K files of the stock being traded have an average $Readability^{PCA}_{i,t}$ above the sample 70th percentile (below the 30th percentile). We repeat the classification but replace $Readability^{PCA}_{i,t}$ with $Readability^{Rank}_{i,t}$ in Panels C and D.

[Insert Table 3 about here]

Subsample tests conducted based on $Readability^{PCA}_{i,t}$ and $Readability^{Rank}_{i,t}$ yield consistent results. The coefficient on $P_{i,t-1}{}^{52h}$ is significantly negative for retail buys and significantly positive for retail sales. The results hold for both low and high readability groups, confirming that the 52-week high effect is pervasive before the EDGAR implementation. Moreover, the

sign of the coefficients on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ across columns and panels are supportive of our hypothesis that investors' anchoring propensity increases after the EDGAR implementation. The last row of each panel reports the difference in the increase in investors' anchoring propensity between subgroups. The result confirm that the increase in investors' anchoring propensity after EDGAR implementation is greater when information uncertainty is higher.

4.2.2 Alternative measures of information uncertainty

One drawback of capturing information uncertainty based on textual features of the EDGAR files is that the measures cannot be constructed before EDGAR implementation when the electronic files are not available. Although using time-series averages of the measures may be sensible as firms' information disclosure behavior should be relatively stable over time, it cannot be refuted that these measures are subject to potential estimation biases.

To address this concern, we construct two time-varying measures of information uncertainty based on the market reaction around the most recent 10-K/Q report releases, following Landsman et al. (2012), to provide additional evidence. The market-reaction-based measures include abnormal trading volume ($AVOL_{i,t}$) and abnormal return volatility ($AVAR_{i,t}$) surrounding the earnings announcements. We partition sample trades into subgroups conditional on whether firms' abnormal trading volume or abnormal return volatility surrounding the most recent earnings announcements is above the sample 70th percentile or below the 30th percentile. Information uncertainty is expected to be higher for both higher $AVOL_{i,t}$ and $AVAR_{i,t}$ groups. Results in Table IA2 in the Internet Appendix confirm that the results are significantly stronger among stocks with higher uncertainty.

Further, in Table IA3, we adopt alternative information uncertain measures constructed using firm characteristics including size and age. Information uncertainty is likely to be higher in smaller or younger firms. Again, the results are consistent with previous findings. Altogether, we show that retail investors' reliance on the 52-week in trading increases when more corporate information is easily available in the post-EDGAR period, especially when information uncertainty is higher. The findings support our **Hypothesis 2** that higher information uncertainty enables retail investors to interpret the disclosed information in ways confirming their priors. Hence, their confirmation bias plays a greater role in incorporating the cherry-picked information, resulting in a stronger anchoring tendency after EDGAR implementation.

4.3 Results around information events

If the findings in Table 2 are driven by the fact that increased information accessibility adds to retail investors' anchoring bias as they cherry-pick information contained in corporate filings, we would expect such an effect to be most evident shortly after the corporate filings are released. We follow Chang et al. (2022) to select earnings announcements as the information events. We divide our sample trades into two subgroups based on whether the retail trading is made within 10 days after the most recent release of stock *i*'s 10K/Q reports. We then investigate whether retail investors' increased anchoring tendency after the EDGAR implementation is more prominent right after 10K/Q reports are released. The results are reported in Table 4.

[Insert Table 4 about here]

Table 4 shows that the coefficients on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ are more significant for retail buys and sales transacted shortly after the information events. The results confirm that investors' increased tendency to be anchored by stock price 52-week high in trading is particularly apparent when new corporate information, such as earnings news, has just been released on the EDGAR system. The result also provides a validity check to our argument that retail investors' biased consumption of information adds to their irrational behavior.

4.4 Internet vs. non-Internet users

The LDB data set includes an indicator for online trading, which allows us to examine the change in investors' anchoring tendency after the EDGAR implementation conditional on their access to the Internet. We identify an investor as an Internet user if she has access to the Internet before the trading date and as a non-Internet user otherwise. If the EDGAR implementation adds to investors' confirmation bias as it is easier for them to obtain information confirming their priors, the effect should be more prominent for Internet users than their counterparts, given what EDGAR provides is free access to corporate filings via the Internet (**Hypothesis 3**).

We examine retail trading using Eqs. (1) and (2) for Internet and non-Internet users separately, and the results are reported in Table 5. Panels A and B report results for retail buys and sales, respectively. Across all the six columns in both panels, the coefficients on $P_{i,t-1}^{52h}$ are significantly negative (positive) for retail buys (retail sales), confirming that both Internet users and non-Internet users are subject to the anchoring bias before EDGAR implementation. Moreover, the magnitude of the coefficient on $P_{i,t-1}^{52h}$ for non-Internet users is about or more

than two times that of the coefficient on $P_{i,t-1}^{52h}$ for Internet users. It is not surprising that Internet users have a lower anchoring tendency prior to EDGAR implementation as Barber and Odean (2002) document that Internet users have stronger performance than counterparts before they switch to trading online. Thus, they are likely to have greater investment sophistication than non-Internet users, which is also manifested by their lower anchoring propensity.

[Insert Table 5 about here]

Further, the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ is consistently and significantly negative (positive) in Panel A (B) when retail buys (retail sales) are examined. It suggests that the increased information dissemination after the EDGAR implementation adds to the anchoring bias of both Internet and non-Internet users. Although non-Internet users are less likely to view corporate information through the EDGAR system online, the increased information dissemination following EDGAR might also enrich the information distributed offline. For instance, newspapers are likely to report more corporate news carried over from EDGAR.

Confirming our **Hypothesis 3**, the magnitude of the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ for Internet users is about 50% greater than that on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ for non-Internet users. Results reported at the bottoms of Panels A and B indicate that the magnitude of the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ is significantly larger for Internet users than for non-Internet users. For intuitive visualization, we also plot the results in Panel A of Figure 1.

4.5 Sophisticated vs. non-sophisticated investors

Sophisticated investors are more likely than non-sophisticated investors to view corporate files before making investment decisions. Hence, we anticipate that the implementation of the EDGAR system should exert a greater influence on sophisticated investors, which is also an empirical test of **Hypothesis 3**. We define retail investors' investment sophistication according to the number of trades they have executed following past studies (e.g., Kuo, Lin, and Zhao, 2015; Chang, Lin, Luo, and Ren, 2019). The gist is that investors with more trades executed are expected to have more trading experience and are thus considered to have greater investment sophistication than their counterparts Therefore, we classify sample investors as sophisticated investors if their transaction executed numbers rank in the top 30% among our entire sample before February 1993, and as non-sophisticated investors for the bottom 30%.

We repeat tests in Eqs. (1) and (2) for sophisticated and non-sophisticated retail investors

separately and report the results in Table 6. The coefficients on $P_{i,t-1}^{52h}$ are significantly negative (positive) in Panel A (B) with retail buys (retail sales), confirming that the anchoring bias is influential and pervasive in the pre-EDGAR period. Furthermore, the magnitude of the coefficient on $P_{i,t-1}^{52h}$ is greater for unsophisticated investors than for sophisticated investors. The finding is as expected since investors with greater investment sophistication should be less affected by behavioral biases, including the anchoring bias.

[Insert Table 6 about here]

When retail buys are examined in Panel A of Table 6, the magnitudes of the negative coefficients on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ among sophisticated investors are much greater than those of unsophisticated investors. In Panel B where retail sales are examined, the coefficients on the interaction terms are again of greater magnitudes for sophisticated investors than for unsophisticated investors. The differences in the coefficients on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ for the two groups are highly significant at the 1% level. The evidence supports our conjecture that the increase in retail investors' anchoring propensity after the EDGAR implementation is greater for sophisticated investors. We also plot the empirical results in Panel B of Figure 1.

5. Assessment of trading performance

In this section, we examine whether and how investors' amplified anchoring bias in the post-EDGAR period affects their trading performance. We calculate the mark-to-market cumulative market-adjusted returns 10, 20, 30, 40, 50, and 60 days after a sample stock *i* is traded by investor *j* on day *t* in month *m*. For a retail buy, the cumulative abnormal return measures the excess profits earned by investor *j* by buying stock *i*, and is used as the dependent variable of Eq. (3). For a retail sale, we multiply cumulative abnormal return by a negative one as the dependent variable to reflect investor *j*'s potential profits forgone (if the stock *i*'s post-sale cumulative abnormal return is positive) or losses avoided (if the stock *i*'s post-sale cumulative abnormal return is negative). The regression model is specified as follows:

$$CAR_{i,j,t+x} = c_{i,w} + c_{j,w} + c_{t,w} + \beta_1 \times P_{i,t-1}^{52h} \times EDGAR_{i,t-1} + \beta_2 \times P_{i,t-1}^{52h} + Controls + \varepsilon_{i,j,t}, \quad (3)$$

where the control variables are the same as in Eq.(1), and β_1 measures the change in the mark-to-market profit made by trading a stock close to its 52-week high after EDGAR implementation. Buying (selling) a stock close to its 52-week high is more profitable in the post-EDGAR period if β_1 is positive (negative), and vice versa. Results are shown in Table 7.

[Insert Table 7 about here]

Panel A of Table 7 shows that the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ is significantly positive for buys when cumulative market-adjusted return is calculated 10, 20, 30, 40, 50, and 60 days after a trade is made. The magnitude of the coefficient increases with the length of post-trading windows. It suggests that buying a stock close to its 52-week high is more profitable in the post- than in the pre-EDGAR period. It is consistent with our conjecture that investors' anchoring bias is amplified in the post-EDGAR period, resulting in higher selling pressures and thus deeper undervaluation of stocks with prices close to their 52-week high. Therefore, such stocks experience greater upward price correction afterward, which continues for more than two months. By the same token, investors' increased tendency to purchase stocks far away from their 52-week high brings greater losses for them in the post-EDGAR period.

In comparison, Panel B of Table 7 shows that the coefficient on $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ is significantly negative for sell orders. Again, the results hold when different windows are used to calculate post-trading performance, and the coefficient on the interaction term increases monotonically with the length of the post-trading windows. The results indicate that compared with trades made in the pre-EDGAR period, selling a stock close to its 52-week high is more detrimental to the investor's trading performance in the post-EDGAR period as the price of the stock sold is more likely to increase in the following periods. The findings are compatible with our previous discussion that investors' greater anchoring tendency in the post-EDGAR period increases their tendency to sell stocks when the stock prices are approaching their 52-week high, and such behavioral biases make them lose more money.

6. Addressing confounding factors and sensitivity checks

6.1 Confounding factors

In previous tests, the staggered nature of the implementation of the EDGAR system provides a unique setting for establishing causal inference. For any confounding factor to explain our findings, it would have to keep in line with the specified phase-in schedule (Gao and Huang, 2020). In the meantime, the stacked DiD approach (Chang et al. 2022) helps further alleviate the concern introduced by treatment heterogeneity (e.g., Baker et al., 2022).

However, there are still possible variations accompanied by EDGAR implementation that may act as confounding factors. For instance, as Chang et al. (2023) suggest, retail investors

pay more attention to a firm when its mandatory disclosures become easily accessible after EDGAR implementation. More attention allocated a stock with online electronic filings may also remind investors of its closeness to the 52-week high, thus resulting in enhanced anchoring propensity. Further, investors might go through unobservable feature variations due to the information event, or stock price patterns might change after the event. In this section, we perform additional tests to filter out the influence of these potential confounding factors.

6.1.1. Investor attention

To address the alternative explanation of changing investor attention after EDGAR, we add stock-level attention measures into our baseline regression. Barber and Odean (2008) identify attention-grabbing stocks conditional on whether they have just experienced abnormally high trading volume or extreme one-day return. We follow their study to capture stock-level investor attention using two measures: $Attention_vol_{i,t}$ and $Attention_ret_{i,t}$. We set $Attention_vol_{i,t}$ to be one for stock i on day t if its abnormal trading volume is in the top 10^{th} percentile, and set $Attention_ret_{i,t}$ to be one for stock i on day t if its return is in the top or bottom 5^{th} percentile. We add these attention measures as well as their interaction terms with stocks' 52-week highs into our baseline regression to filter out any confounding effect of investor attention. The results are robust and reported in Table IA4.

6.1.2. Other post-EDGAR heterogeneity: investor characteristics or return behaviors

To control for the influence arising from changing investor characteristics after EDGAR implementation, we include an $Account \times EDGAR_{i,t-1}$ fixed effect in Eqs. (1) and (2), and report the results in Table IA5. To control for the possibility that stock return patterns, such as momentum, change systematically in the post-EDGAR period, we further add interactions between $EDGAR_{i,t}$ and stocks' past returns in our model specifications to control for the joint effect of EDGAR inclusion and past return patterns. The results are reported in Table IA6. The results in both tables are robust, confirming that our findings are not driven by changing account-level or return behavior after EDGAR implementation.

6.1.3 Multiple hypotheses testing

The staggered implementation of EDAGR has been exploited to test several different hypotheses. Heath et al. (2023) argue that reusing a setting without accounting for the other outcomes that have already been examined will lead to *t*-statistics that cannot be interpreted in the usual manner. They provide the simulated adjusted *t*-statistic critical values conditional on

the number of prior findings. For instance, they suggest that a new hypothesis should have a *t*-statistic of at least 2.5 if there are five prior findings and 3.0 if 20 prior findings use the same setting. When there are 293 prior findings, the adjusted critical *t*-statistic critical value increases to 3.66-3.69. Given that the *t*-statistics of our variable of interest in most cases are much higher than 3.69, multiple hypotheses testing should not be an important issue for our study.

6.2 Sensitivity checks

6.2.1. Alternative EDGAR inclusion dates

We follow Chang et al. (2022) to identify sample stocks' EDGAR inclusion dates based on the availability of free online access to their EDGAR filings. Such an identification strategy is suitable in our setting as retail investors are less likely to pay for access to the EDGAR system when it is not free in the early stage, e.g., before January 17, 1994.

In robustness checks, we also follow earlier studies on EDGAR implementation (e.g., Gao and Huang, 2020) to set the event date to be the one that a sample firm is subject to mandatory EDGAR filings as specified in the phase-in schedule of SEC Release 33-6977. Moreover, for two of the ten EDGAR batches (CF-05 and CF06), the actual phase-in dates are revised later as specified in SEC Release 33-7122 and shown in Table IA1. Thus, we follow Goldstein et al. (2023) to use the revised dates as alternative EDGAR inclusion dates to re-perform our main tests. The results are robust to these alternative dates and reported in Tables IA7 and IA8.

6.2.2. Exclusion of the "transitional" filers

Prior to the mandatory phase-in of the EDGAR system, the SEC allowed some firms to use the system voluntarily for testing purposes. These voluntary filers are then assigned to the first EDGAR batch, or CF-01, and are referred to as the "transitional" filers. Our findings are not affected qualitatively if these transitional filers are excluded as shown in Table IA9.

6.2.3. Alternative regression specification

We have performed tests using stacked regressions with propensity-score-matched sample stocks in previous tests. The results are generally unchanged if we apply the stacked regression with all clean controls (without PSM) or use the standard DiD regression with the PSM sample, as shown in Tables IA10 and IA11.

7. Conclusions

In this study, we investigate the influence of increased information dissemination on retail investors' trading behavior in stock markets by exploiting the staggered implementation of the EDGAR system from 1993 to 1996. Utilizing retail trading data from a large discount brokerage data set between 1991 and 1996, we examine the changes in investors' anchoring tendency after the implementation of the EDGAR system in a quasi-natural experiment setting.

We find that retail investors' investment decisions are indeed influenced by the stock price's 52-week high. More intriguingly, investors' anchoring tendency is amplified following the implementation of EDGAR: they become less likely to buy a stock and more likely to sell a stock when its price is close to the 52-week high. Such evidence is especially strong when information uncertainty is higher, as retail investors have greater flexibility in interpreting the information from EDGAR filings, plausibly inducing a greater confirmation bias. It is also stronger shortly after the release of corporate 10K/Q reports, confirming that it is driven by retail investors' biased interpretation of information in corporate filings.

Furthermore, retail investors who are more likely to go through corporate filings on the EDGAR system, such as Internet users and sophisticated investors, experience a greater increase in the anchoring tendency than their counterparts after EDGAR implementation. The finding is consistent with our hypothesis that greater information dissemination through the Internet facilitates investors' information cherry-picking, and thus could aggravate rather than alleviate their inherent behavioral biases. Lastly, we show that retail investors' amplified anchoring bias in the post-EDGAR period leads to worse investment performance.

Overall, from the perspective of retail investors' behavioral biases, we reveal an unexpected and unintended consequence of greater information dissemination brought about by the advent of modern technologies. When more can be less, it all depends on whether the disseminated information can be exploited objectively.

Reference

Baker, A.C., Larcker, D.F. and Wang, C.C.Y., 2022. How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics* (144), 370-395.

Baker, M. and Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* (4), 1645-1680.

Baker, M. and Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* (21), 129-151.

Baker, M., Pan, X. and Wurgler, J., 2012. The effect of reference point prices on mergers and acquisitions. *Journal of Financial Economics* (106), 49-71.

Barber, B.M. and Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* (55), 773-806.

Barber, B.M. and Odean, T., 2001. The Internet and the investor. *Journal of Economic Perspectives* (15), 41-54.

Barber, B.M. and Odean, T., 2002. Online investors: Do the slow die first? *The Review of Financial Studies* 15(2), 455-487.

Bessi, A., 2016. Personality traits and echo chambers on Facebook. *Computers in Human Behavior* (65), 319-324.

Blankespoor, E., deHaan, E., and Marinovic, I., 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics* 70(2-3), 101344.

Bird, A., Karolyi, S.A., Ruchti, T. and Truong, P., 2021. More is less: Publicizing information and market feedback. *Review of Finance* 25(3), 745-775.

Callaway, B. and Sant'Anna, P.H.C., 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics* 225(2), 200-230.

Chang, Y.C., Hsiao, P.J., Ljungqvist, A., and Tseng, K., 2022. Testing disagreement models. *Journal of Finance*, forthcoming.

Chang, E.C., Lin, T.C., Luo, Y., and Ren, J., 2019. Ex-day returns of stock distributions: An anchoring explanation. *Management Science* 65(3), 1076-1095.

Chang, Y.C., Ljungqvist, A., and Tseng, K., 2023. Do corporate disclosures constrain strategic analyst behavior? *Review of Financial Studies* 36(8), 3163-3212.

Cengiz, D., Dube A., Lindner A., and Zipperer B., 2019, The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics* 134(3), 1405-1454.

Cookson, J.A., Joseph, E.E., and Mullins, W., 2023. Echo chambers. *Review of Financial Studies* 36(2), 450-500.

Driessen, J., Lin, T.C., Van Hemert, O., 2013. How the 52-week high and low affect option-implied volatilities and stock return moments. *Review of Finance* 17, 369-401.

Evans, J. St. B. T., 1989. Bias in human reasoning: Causes and consequences. NJ: Erlbaum.

Gao, M., Huang, J., 2020. Informing the market: The effect of modern information technologies on information production. *Review of Financial Studies* (4), 1368-1411

Gentzkow, M., and Shapiro, J.M., 2011. Ideological segregation online and offline. *Quarterly Journal of Economics* (126), 1799-1839.

George, T.J. and Hwang, C.-Y., 2004.The 52-week high and momentum investing. *Journal of Finance*(59), 2145-2176.

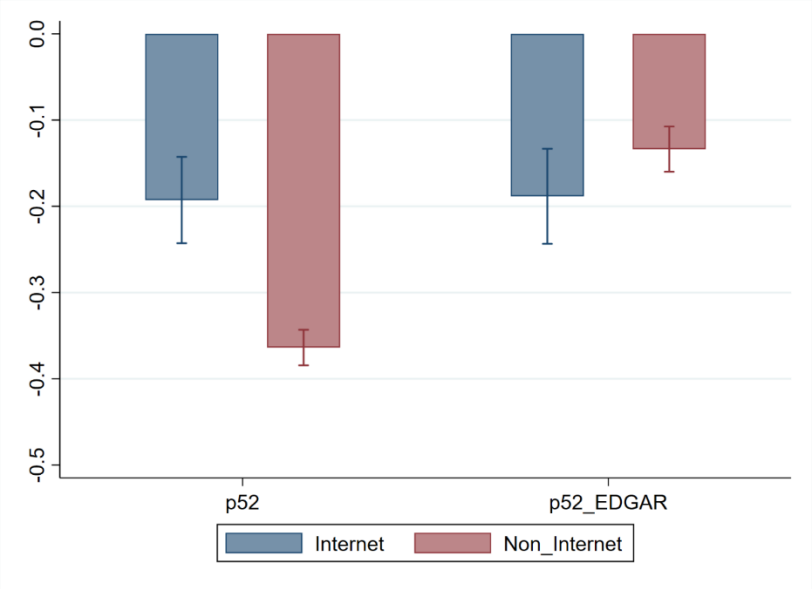
George, T.J. and Hwang, C.-Y. and Li, Y., 2018. The 52-week high, q-theory, and the cross section of stock returns. *Journal of Financial Economics* 128(1), 148-163.

Goldstein, I., Yang, S. and Zuo, L., 2023. The real effects of modern information technologies: Evidence from the EDGAR implementation. *Journal of Accounting Research*, 61(5), 1699-1733.

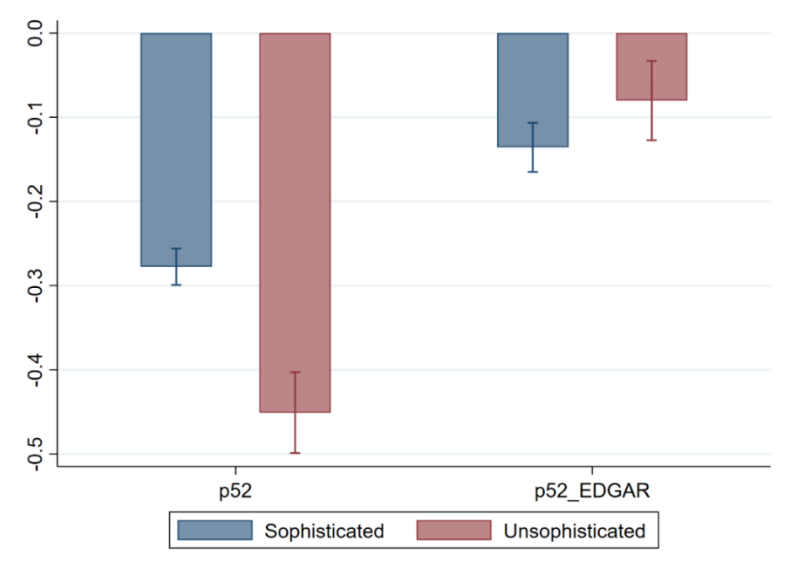
Heath, C., Huddart, S. and Lang, M., 1999. Psychological factor and stock option exercise.

Quarterly Journal of Economics 114(2), 601-627.
Heath, D., Ringgenberg, M., Samadi, M. and Werner, I.M., 2023. Reusing natural experiments. *Journal of Finance* 78(4)*,* 2329-2364.
Hirshleifer, D.A., 2001. Investor psychology and asset prices. *Journal of Finance* 56(4), 1533-1597.
Huang, S., Lin, T.C. and Xiang, H., 2021. Psychological barrier and cross-firm return predictability. *Journal of Financial Economics* 142, 228-356.
Huddart, S., Lang, M., Yetman, M. H., 2008. Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science* 55(1), 16-31.
Kahneman, D., 2011. Thinking, fast and slow. Macmillan.
Kumar, A., 2009. Hard-to-value stocks, behavioral biases, and informed trading. *Journal of Financial and Quantitative Analysis* 44(6), 1375-1401.
Kuo, W.Y., Lin, T.C. and Zhao, J., 2015. Cognitive limitation and investment performance: Evidence from limit order clustering. *Review of Financial Studies* 28(3), 838-875.
Landsman, W.R., Maydew, E. L., and Thornock, J.R., 2012. The information content of annual earnings announcements and mandatory adoption of IFRS. *Journal of Accounting and Economics* 53, 34-54.
Li, F., 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45(2-3), 221-247.
Li, F., Lin, C. and Lin, T. C., 2021. Salient anchor and analyst recommendation downgrade. *Journal of Corporate Finance* 69, Article:102033.
Li, J. and Yu, J.,2012. Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics* 104(2), 401-419.
Loughran, T. and Mcdonald, B. 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-K. *Journal of Finance* 66(1), 35-65.
Ma, Q., Whidbee, D.A., and Zhang, W. A., 2019. Acquirer reference prices and acquisition performance. *Journal of Financial Economics* 132(1), 175-199.
Moravec, P., Minas, R. and Dennis, A.R., 2019, Fake news on social media: People believe what they want to believe when it makes no sense at all. *MIS Quarterly*, 43(4), 1343-1360.
Mullainathan, S. and Shleifer, A., 2005. The market for news. *American Economic Review* 95(4), 1031-1053.
Park, J., Konana, P., Gu, B., Kumar, A. and Raghunathan, R.,2013. Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research*, 24(4), 883-1167.
Poteshman, A.M. and Serbin, V., 2003. Clearly irrational financial market behavior: Evidence from the early exercise of exchange traded stock options. *Journal of Finance* 58, 37-70.
Pouget, S., Sauvagnat, J., and Villeneuve, S., 2017. A mind is a terrible thing to change: Confirmation bias in financial markets. *Review of Financial Studies* 30(6), 2066-2109.
Rabin, M., and Schrag, J. 1999. First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics* 114(1), 37-82.
Slechten L., Courtois, C., Coenen, L. and Zaman, B., 2022. Adapting the selective exposure perspective to algorithmically governed platforms: The case of Google search. *Communication Research* 49(8), 1039-1065.
Sun, L., Abraham, S., 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2), 175-199.
Sunstein, C.R. 2001. Republic.com. Princeton, NJ: Princeton University Press.
Tversky, A. and Kahneman D., 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185(4157), 1124-1131.
Vedova, J.D., Grant, A., and Westerholm, P.J., 2023. Investor behavior at the 52-week high. *Journal of Financial Quantitative Analysis* 58(7), 2852-2889.

Figure 1. Subsample analysis of the 52-week high effect before and after EDGAR inclusion



Panel A. Internet vs. non-Internet traders



Panel B. Sophisticated vs. unsophisticated investors.

Note: This figure shows the retail investors' anchoring propensity and its variations after EDGAR implementation conditional on investors' access to the Internet and sophistication. Panels A and B plot the regression coefficients on $P_{i,t-1}^{52h}$ and $P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$ for buy orders in each subgroup. (Columns (1) and (2) in Panel A of Tables 5 and 6)

Table 1. Summary statistics.

This table reports the summary statistics of the main variables used in the empirical investigation. $P_{i,t-1}^{52h}$ is the price of stock *i* on day *t–1* scaled by its highest level over the past 52 weeks. Following Chang et al. (2022), we set $EDGAR_{i,t-1}$ as a dummy variable that equals one if stock *i*'s filings are available online for free via EDGAR on date *t–1* and zero otherwise. $Buy_{i,j,t}$ ($Sell_{i,j,t}$) is a dummy variable that equals one if investor *j* buys (sells) stock *i* on day *t* and zero otherwise. $BuyShr_{i,j,t}$ ($SellShr_{i,j,t}$) is the natural logarithm of the number of shares of stock *i* bought (sold) by investor *j* on day *t*. $BuyValue_{i,j,t}$ ($SellValue_{i,j,t}$) is the natural logarithm of the value of stock *i* bought (sold) by investor *j* on day *t*. $Size_{i,t-1}$ and $B/M_{i,t-1}$ are stock *i*'s market capitalization and book-to-market ratio, respectively. $Runup_{i,t-5\sim t-1}$, $Runup_{i,t-21\sim t-6}$, $Runup_{i,m-6\sim m-2}$, $Runup_{i,m-12\sim m-7}$ are cumulative returns over different horizons before the trading date. $Turnover_{i,t-1}$, $Bid\text{-}ask\ spread_{i,t-1}$, $Idiosyncratic\ volatility_{i,t-1}$, and $Idiosyncratic\ skewness_{i,t-1}$ are stock *i*'s average daily turnover ratio, bid-ask spread, idiosyncratic volatility, and idiosyncratic skewness in the previous month.

	# of obs	*Mean*	*SD*	*P25*	*Median*	*P75*
$P_{i,t-1}^{52h}$	983,932	0.795	0.196	0.687	0.852	0.955
$EDGAR_{i,t-1}$	983,932	0.402	0.490	0.000	0.000	1.000
$Buy_{i,j,t}$	983,932	0.535	0.499	0.000	1.000	1.000
$Sell_{i,j,t}$	983,932	0.465	0.499	0.000	0.000	1.000
$BuyShr_{i,j,t}$	983,932	2.960	2.899	0.000	3.910	5.300
$SellShr_{i,j,t}$	983,932	2.609	2.936	0.000	0.000	5.300
$BuyValue_{i,j,t}$	983,932	4.628	4.396	0.000	7.030	8.670
$SellValue_{i,j,t}$	983,932	4.067	4.451	0.000	0.000	8.620
$Size_{i,t-1}$	983,741	7.267	2.209	5.590	7.330	8.910
$B/M_{i,t-1}$	961,597	0.538	0.665	0.203	0.379	0.693
$Runup_{i,t-5\sim t-1}$	983,923	0.020	0.326	–0.033	0.004	0.048
$Runup_{i,t-21\sim t-6}$	983,913	0.024	0.165	–0.049	0.011	0.081
$Runup_{i,m-6\sim m-2}$	983,786	0.145	0.452	–0.095	0.076	0.288
$Runup_{i,m-12\sim m-7}$	983,643	0.177	0.487	–0.092	0.090	0.332
$Turnover_{i,t-1}$	983,737	0.007	0.007	0.002	0.004	0.008
$Bid\text{-}ask\ spread_{i,t-1}$	983,737	0.023	0.029	0.009	0.016	0.029
$Idiosyncratic\ volatility_{i,t-1}$	983,719	0.029	0.032	0.015	0.023	0.034
$Idiosyncratic\ skewness_{i,t-1}$	983,711	0.211	0.935	–0.249	0.218	0.700

Table 2. EDGAR implementation and the 52-week high effect.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation. Panels A and B report results for buy and sell orders, respectively. $Buy_{i,j,t}$ ($Sell_{i,j,t}$) is an indicator that equals one if investor j buys (sells) stock i on day t and zero otherwise. $BuyShr_{i,j,t}$ ($SellShr_{i,j,t}$) is the logarithm of the number of shares of stock i bought (sold), and $BuyValue_{i,j,t}$ ($SellValue_{i,j,t}$) is the logarithm of the value of stock bought (sold). $P_{i,t-1}^{52h}$ is stock i's price on day $t-1$ scalded by its highest level over the past 52 weeks. Following Chang et al. (2022), we employ a stacked DiD regression with propensity score matching sample on market capitalization and set $EDGAR_{i,t}$ as a dummy variable that equals one if stock i's filings are available online for free via EDGAR on date t and zero otherwise. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$		–0.139***		–0.721***		–1.275***
		(–10.72)		(–9.67)		(–11.17)
$EDGAR_{i,t}$		0.172***		0.936***		1.526***
		(14.88)		(13.83)		(14.99)
$P_{i,t-1}^{52h}$	–0.376***	–0.334***	–2.191***	–1.974***	–3.124***	–2.738***
	(–38.86)	(–32.93)	(–39.68)	(–34.25)	(–36.82)	(–31.18)
$Runup_{i,t-5\sim t-1}$	0.060***	0.055***	0.084	0.060	0.600***	0.561***
	(4.82)	(4.58)	(1.16)	(0.85)	(5.52)	(5.30)
$Runup_{i,t-21\sim t-6}$	–0.019***	–0.020***	–0.284***	–0.292***	–0.030	–0.042
	(–2.69)	(–2.91)	(–6.46)	(–6.69)	(–0.50)	(–0.70)
$Runup_{i,m-6\sim m-2}$	0.033***	0.032***	0.153***	0.145***	0.336***	0.325***
	(11.59)	(11.12)	(9.39)	(8.91)	(13.54)	(13.08)
$Runup_{i,m-12\sim m-7}$	0.034***	0.033***	0.166***	0.158***	0.316***	0.304***
	(16.16)	(15.51)	(13.66)	(13.03)	(17.10)	(16.47)
$Size_{i,t-1}$	0.085***	0.088***	0.350***	0.365***	0.844***	0.869***
	(31.81)	(32.70)	(23.50)	(24.42)	(36.21)	(37.05)
$B/M_{i,t-1}$	0.005***	0.007***	0.064***	0.071***	0.016	0.027*
	(2.80)	(3.44)	(5.48)	(6.07)	(1.01)	(1.67)
$Turnover_{i,t-1}$	2.286***	2.095***	10.417***	9.405***	20.245***	18.511***
	(10.57)	(9.65)	(8.11)	(7.28)	(10.84)	(9.88)
$Bid\text{-}ask\ spread_{i,t-1}$	0.245***	0.301***	1.685***	1.992***	1.339**	1.830***
	(3.37)	(4.12)	(3.68)	(4.33)	(2.22)	(3.02)
$Idiosyncratic\ volatility_{i,t-1}$	0.636***	0.678***	4.655***	4.874***	4.844***	5.231***
	(9.48)	(10.01)	(11.58)	(12.00)	(8.71)	(9.35)
$Idiosyncratic\ skewness_{i,t-1}$	–0.020***	–0.020***	–0.112***	–0.111***	–0.176***	–0.176***
	(–21.29)	(–21.40)	(–20.55)	(–20.65)	(–20.76)	(–20.89)
Wave-specific Account, stock, and date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337	1,086,337	1,086,337	1,086,337
R-Squared	0.208	0.209	0.218	0.219	0.198	0.199

Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$		0.139***		0.648***		1.325***
		(10.72)		(8.65)		(11.60)
$EDGAR_{i,t}$		–0.172***		–0.913***		–1.538***
		(–14.88)		(–13.44)		(–15.12)
$P_{i,t-1}^{52h}$	0.376***	0.334***	2.052***	1.859***	3.531***	3.129***
	(38.86)	(32.93)	(36.38)	(31.50)	(41.34)	(35.38)
$Runup_{i,t-5\sim t-1}$	–0.060***	–0.055***	–0.206***	–0.184**	–0.135	–0.095
	(–4.82)	(–4.58)	(–2.80)	(–2.55)	(–1.25)	(–0.91)
$Runup_{i,t-21\sim t-6}$	0.019***	0.020***	0.143***	0.151***	0.374***	0.385***
	(2.69)	(2.91)	(3.26)	(3.48)	(5.96)	(6.20)
$Runup_{i,m-6\sim m-2}$	–0.033***	–0.032***	–0.175***	–0.167***	–0.285***	–0.273***
	(–11.59)	(–11.12)	(–10.36)	(–9.88)	(–11.34)	(–10.90)
$Runup_{i,m-12\sim m-7}$	–0.034***	–0.033***	–0.193***	–0.185***	–0.285***	–0.272***
	(–16.16)	(–15.51)	(–15.53)	(–14.91)	(–15.34)	(–14.71)
$Size_{i,t-1}$	–0.085***	–0.088***	–0.539***	–0.554***	–0.563***	–0.589***
	(–31.81)	(–32.70)	(–35.43)	(–36.30)	(–23.93)	(–24.86)
$B/M_{i,t-1}$	–0.005***	–0.007***	–0.028**	–0.036***	–0.060***	–0.070***
	(–2.80)	(–3.44)	(–2.43)	(–3.09)	(–3.80)	(–4.43)
$Turnover_{i,t-1}$	–2.286***	–2.095***	–12.495***	–11.555***	–14.569***	–12.788***
	(–10.57)	(–9.65)	(–9.61)	(–8.88)	(–7.71)	(–6.74)
$Bid\text{-}ask\ spread_{i,t-1}$	–0.245***	–0.301***	–0.576	–0.881*	–3.943***	–4.433***
	(–3.37)	(–4.12)	(–1.28)	(–1.95)	(–6.55)	(–7.34)
$Idiosyncratic\ volatility_{i,t-1}$	–0.636***	–0.678***	–3.486***	–3.684***	–5.864***	–6.265***
	(–9.48)	(–10.01)	(–8.97)	(–9.49)	(–10.42)	(–11.05)
$Idiosyncratic\ skewness_{i,t-1}$	0.020***	0.020***	0.113***	0.112***	0.183***	0.182***
	(21.29)	(21.40)	(20.47)	(20.54)	(21.28)	(21.43)
Wave-specific Account, stock, and date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337	1,086,337	1,086,337	1,086,337
R-Squared	0.208	0.209	0.212	0.212	0.202	0.203

Table 3. Influence of information uncertainty measured by 10-K report readability.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation conditional on the readability of 10-K reports. Two overall readability measures are constructed based on the Fog, Flesch Reading Ease, and Kincaid indices of Li (2008). The variable $Readability^{PCA}_{i,t}$ is the first principal component of the three indices. The variable $Readability^{Rank}_{i,t}$ is the sum of the decile rankings of the three indices, with a higher ranking indicating greater readability. In Panels A and B, a 10-K report with a $Readability^{PCA}_{i,t}$ above (below) 70th percentile (30th percentile) is considered to be low (high) readability. In Panels C and D, a 10-K report with a $Readability^{Rank}_{i,t}$ above (below) 70th percentile (30th percentile) is considered to be low (high) readability. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,i,t}$		$BuyShr_{i,i,t}$		$BuyValue_{i,i,t}$	
Readability measure $Readability^{PCA}_{i,t}$	(1) Low Read.	(2) High Read.	(3) Low Read.	(4) High Read.	(5) Low Read.	(6) High Read.
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.287***	–0.119***	–1.545***	–0.559***	–2.552***	–1.152***
	(–13.89)	(–4.76)	(–12.95)	(–3.95)	(–13.97)	(–5.21)
$EDGAR_{i,t-1}$	0.352***	0.132***	1.955***	0.645***	3.071***	1.230***
	(18.28)	(5.59)	(17.30)	(4.82)	(18.12)	(5.94)
$P_{i,t-1}^{52h}$	–0.267***	–0.481***	–1.625***	–2.727***	–2.207***	–4.007***
	(–16.16)	(–23.64)	(–16.90)	(–23.42)	(–15.38)	(–22.36)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	304,239	292,609	304,239	292,609	304,239	292,609
R-Squared	0.297	0.303	0.298	0.304	0.282	0.290
$P^{52h} \times EDGAR$	–0.167***		–0.986***		–1.400***	
Low Read. – High Read.	(–6.15)		(–6.11)		(–5.72)	

Panel B: Sell orders	$Sell_{i,i,t}$		$SellShr_{i,i,t}$		$SellValue_{i,i,t}$	
Readability measure $Readability^{PCA}_{i,t}$	(1) Low Read.	(2) High Read.	(3) Low Read.	(4) High Read.	(5) Low Read.	(6) High Read.
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.287***	0.119***	1.524***	0.491***	2.640***	1.106***
	(13.89)	(4.76)	(12.54)	(3.38)	(14.30)	(4.94)
$EDGAR_{i,t-1}$	–0.352***	–0.132***	–2.004***	–0.657***	–3.138***	–1.101***
	(–18.28)	(–5.59)	(–17.64)	(–4.76)	(–18.40)	(–5.24)
$P_{i,t-1}^{52h}$	0.267***	0.481***	1.520***	2.616***	2.545***	4.386***
	(16.16)	(23.64)	(15.53)	(22.00)	(17.54)	(24.21)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	304,239	292,609	304,239	292,609	304,239	292,609
R-Squared	0.297	0.303	0.292	0.301	0.285	0.292
$P^{52h} \times EDGAR$	0.167***		1.033***		1.533***	
Low Read. – High Read.	(6.15)		(6.40)		(6.27)	

Panel C: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
Readability measure $Readability^{Rank}_{i,t}$	(1) Low Read.	(2) High Read.	(3) Low Read.	(4) High Read.	(5) Low Read.	(6) High Read.
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.261***	–0.111***	–1.407***	–0.536***	–2.325***	–1.067***
	(–13.09)	(–4.64)	(–12.25)	(–3.97)	(–13.20)	(–5.09)
$EDGAR_{i,t-1}$	0.324***	0.132***	1.807***	0.695***	2.832***	1.217***
	(17.32)	(5.92)	(16.47)	(5.47)	(17.22)	(6.23)
$P_{i,t-1}^{52h}$	–0.275***	–0.479***	–1.660***	–2.711***	–2.279***	–3.989***
	(–16.86)	(–24.65)	(–17.50)	(–24.39)	(–16.09)	(–23.30)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	318,751	320,407	318,751	320,407	318,751	320,407
R-Squared	0.293	0.294	0.294	0.296	0.279	0.281
$P^{52h} \times EDGAR$	–0.150***		–0.871***		–1.258***	
Low Read. – High Read.	(–6.08)		(–5.74)		(–5.40)	
Panel D: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
Readability measure $Readability^{Rank}_{i,t}$	(1) Low Read.	(2) High Read.	(3) Low Read.	(4) High Read.	(5) Low Read.	(6) High Read.
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.261***	0.111***	1.383***	0.460***	2.407***	1.059***
	(13.09)	(4.64)	(11.83)	(3.32)	(13.52)	(4.98)
$EDGAR_{i,t-1}$	–0.324***	–0.132***	–1.853***	–0.682***	–2.896***	–1.137***
	(–17.32)	(–5.92)	(–16.77)	(–5.22)	(–17.49)	(–5.73)
$P_{i,t-1}^{52h}$	0.275***	0.479***	1.572***	2.607***	2.613***	4.355***
	(16.86)	(24.65)	(16.27)	(22.96)	(18.23)	(25.15)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	318,751	320,407	318,751	320,407	318,751	320,407
R-Squared	0.293	0.294	0.289	0.292	0.282	0.283
$P^{52h} \times EDGAR$	0.150***		0.923***		1.348***	
Low Read. – High Read.	(6.08)		(5.42)		(5.21)	

Table 4. Results around information events.

This table examines whether the effect of EDGAR implementation on retail investors' trading is stronger around the release of 10-K/Q filings. Retail transactions within 10 days after the release of stock *i*'s 10K/Q reports are considered as surrounding the information event. Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) <=10 days	(2) > 10 days	(3) <=10 days	(4) >10 days	(5) <=10 days	(6) >10 days
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.202***	–0.124***	–1.087***	–0.640***	–1.843***	–1.141***
	(–6.24)	(–9.05)	(–5.67)	(–8.10)	(–6.27)	(–9.44)
$EDGAR_{i,t-1}$	0.199***	0.164***	1.085***	0.894***	1.775***	1.453***
	(6.54)	(13.29)	(5.99)	(12.34)	(6.49)	(13.35)
$P_{i,t-1}^{52h}$	–0.267***	–0.359***	–1.521***	–2.112***	–2.215***	–2.965***
	(–9.81)	(–32.83)	(–9.65)	(–34.00)	(–9.15)	(–31.33)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	157,114	919,750	157,114	919,750	157,114	919,750
R-Squared	0.517	0.229	0.515	0.238	0.493	0.219
$P^{52h} \times EDGAR$	–0.078***		–0.447***		–0.702***	
Within 10 days –After 10 days	(–2.86)		(–2.74)		(–2.98)	
Panel B: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) <=10 days	(2) > 10 days	(3) <=10 days	(4) > 10 days	(5) <=10 days	(6) > 10 days
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.202***	0.124***	1.024***	0.573***	1.976***	1.187***
	(6.24)	(9.05)	(5.27)	(7.24)	(6.64)	(9.83)
$EDGAR_{i,t-1}$	–0.199***	–0.164***	–1.048***	–0.877***	–1.847***	–1.465***
	(–6.54)	(–13.29)	(–5.69)	(–12.07)	(–6.65)	(–13.47)
$P_{i,t-1}^{52h}$	0.267***	0.359***	1.435***	1.990***	2.506***	3.328***
	(9.81)	(32.83)	(8.82)	(31.20)	(10.10)	(34.90)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	157,114	919,750	157,114	919,750	157,114	919,750
R-Squared	0.517	0.229	0.512	0.232	0.498	0.222
$P^{52h} \times EDGAR$	0.078***		0.451***		0.789***	
Within 10 days –After 10 days	(2.86)		(2.67)		(2.87)	

Table 5. Internet vs. non-Internet traders.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation conditional on investors' access to the Internet before the trading date. Panels A and B report results for buy and sell orders respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1)Internet	(2)Non-Internet	(3)Internet	(4)Non-Internet	(5)Internet	(6)Non-Internet
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.188***	–0.134***	–1.069***	–0.684***	–1.691***	–1.238***
	(–6.70)	(–10.00)	(–5.92)	(–9.00)	(–6.50)	(–10.59)
$EDGAR_{i,t-1}$	0.217***	0.173***	1.211***	0.936***	1.890***	1.536***
	(8.75)	(14.41)	(7.50)	(13.53)	(8.20)	(14.70)
$P_{i,t-1}^{52h}$	–0.193***	–0.364***	–1.216***	–2.134***	–1.558***	–2.991***
	(–7.55)	(–34.53)	(–7.99)	(–35.93)	(–6.88)	(–32.93)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	148,748	936,118	148,748	936,118	148,748	936,118
R-Squared	0.220	0.227	0.236	0.236	0.210	0.216
$P^{52h} \times EDGAR$	–0.055**		–0.385***		–0.453**	
Internet – Non-internet	(–2.36)		(–2.75)		(–2.17)	
Panel B: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1)Internet	(2)Non-Internet	(3)Internet	(4)Non-Internet	(5)Internet	(6)Non-Internet
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.188***	0.134***	1.072***	0.604***	1.817***	1.281***
	(6.70)	(10.00)	(5.78)	(7.89)	(6.90)	(10.92)
$EDGAR_{i,t-1}$	–0.217***	–0.173***	–1.289***	–0.902***	–1.998***	–1.536***
	(–8.75)	(–14.41)	(–7.82)	(–12.92)	(–8.61)	(–14.65)
$P_{i,t-1}^{52h}$	0.193***	0.364***	1.009***	2.031***	1.841***	3.392***
	(7.55)	(34.53)	(6.37)	(33.45)	(7.99)	(37.12)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	148,748	936,118	148,748	936,118	148,748	936,118
R-Squared	0.220	0.227	0.229	0.230	0.211	0.221
$P^{52h} \times EDGAR$	0.055**		0.468***		0.536**	
Internet – Non-internet	(2.36)		(3.30)		(2.55)	

Table 6. Sophisticated vs. unsophisticated investors.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation conditional on investor sophistication. Investors with the number of trades executed above the sample 70th percentile (below the 30th percentile) before EDGAR implementation in February 1991 are grouped as sophisticated (non-sophisticated) investors. Panels A and B report results for buy and sell orders respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) Sophisticated	(2) Unsophisticated	(3) Sophisticated	(4) Unsophisticated	(5) Sophisticated	(6) Unsophisticated
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.136***	–0.080***	–0.715***	–0.402***	–1.240***	–0.803***
	(–9.14)	(–3.34)	(–8.07)	(–3.00)	(–9.30)	(–3.88)
$EDGAR_{i,t-1}$	0.167***	0.172***	0.917***	0.919***	1.478***	1.496***
	(12.49)	(7.54)	(11.37)	(7.14)	(12.40)	(7.66)
$P_{i,t-1}^{52h}$	–0.278***	–0.451***	–1.699***	–2.514***	–2.274***	–3.708***
	(–25.12)	(–18.41)	(–26.58)	(–18.56)	(–23.73)	(–17.85)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	720,571	195,051	720,571	195,051	720,571	195,051
R-Squared	0.139	0.440	0.161	0.434	0.137	0.414
$P^{52h} \times EDGAR$	–0.056**		–0.314**		–0.437**	
Sophis.–Unsophis.	(–2.27)		(–2.24)		(–2.06)	
Panel B: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) Sophisticated	(2) Unsophisticated	(3) Sophisticated	(4) Unsophisticated	(5) Sophisticated	(6) Unsophisticated
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.136***	0.080***	0.672***	0.224*	1.320***	0.751***
	(9.14)	(3.34)	(7.53)	(1.66)	(9.91)	(3.63)
$EDGAR_{i,t-1}$	–0.167***	–0.172***	–0.924***	–0.852***	–1.517***	–1.453***
	(–12.49)	(–7.54)	(–11.42)	(–6.58)	(–12.77)	(–7.41)
$P_{i,t-1}^{52h}$	0.278***	0.451***	1.580***	2.468***	2.660***	4.149***
	(25.12)	(18.41)	(24.12)	(17.95)	(27.43)	(19.71)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	720,571	195,051	720,571	195,051	720,571	195,051
R-Squared	0.139	0.440	0.150	0.436	0.137	0.424
$P^{52h} \times EDGAR$	0.056**		0.448***		0.569***	
Sophis.–Unsophis.	(2.27)		(3.13)		(2.65)	

Table 7. Assessment of trading performance.

This table examines the relation between of 52-week high and stocks' post-trading performance after EDGAR implementation. Buy orders are examined in Panel A, where the dependent variable is the cumulative market-adjusted return for stock *i* 10-, 20-, 30-, 40-, 50- or 60 days after it is bought by investor *j* on day *t*. Sell orders are examined in Panel B, where the dependent variable is negative one multiplied by the cumulative market-adjusted return for stock *i* 10-, 20-, 30-, 40-, 50- or 60 days after it is sold by investor *j* on day *t*. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $CAR_{i,t,t+10}$	(2) $CAR_{i,t,t+20}$	(3) $CAR_{i,t,t+30}$	(4) $CAR_{i,t,t+40}$	(5) $CAR_{i,t,t+50}$	(6) $CAR_{i,t,t+60}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	0.034***	0.064***	0.096***	0.120***	0.140***	0.168***
	(6.08)	(8.08)	(10.44)	(11.66)	(12.05)	(13.51)
$EDGAR_{i,t}$	–0.021***	–0.037***	–0.055***	–0.065***	–0.076***	–0.092***
	(–4.00)	(–5.28)	(–6.69)	(–6.96)	(–7.18)	(–8.10)
$P_{i,t-1}^{52h}$	–0.001	–0.005	–0.004	0.001	–0.001	0.002
	(–0.31)	(–0.71)	(–0.54)	(0.10)	(–0.08)	(0.18)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	584,553	584,251	583,831	583,375	582,929	582,330
R-Squared	0.436	0.450	0.466	0.480	0.495	0.501
Panel B: Sell orders	(1) $CAR_{i,t,t+10}$	(2) $CAR_{i,t,t+20}$	(3) $CAR_{i,t,t+30}$	(4) $CAR_{i,t,t+40}$	(5) $CAR_{i,t,t+50}$	(6) $CAR_{i,t,t+60}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	–0.037***	–0.075***	–0.108***	–0.133***	–0.151***	–0.171***
	(–6.76)	(–9.46)	(–11.93)	(–13.08)	(–13.87)	(–14.62)
$EDGAR_{i,t}$	0.027***	0.050***	0.074***	0.092***	0.102***	0.113***
	(4.91)	(6.84)	(9.07)	(9.93)	(10.26)	(10.32)
$P_{i,t-1}^{52h}$	–0.009**	–0.012**	–0.015**	–0.017**	–0.026***	–0.035***
	(–2.07)	(–2.03)	(–2.18)	(–2.41)	(–3.45)	(–4.26)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	500,956	500,337	499,452	498,668	497,939	497,156
R-Squared	0.423	0.449	0.539	0.558	0.538	0.529

Internet Appendix A

Table IA1. Timetable for Implementation of EDGAR Division of Corporation Finance Filings

Phase-in Group	Preliminary phase-in date (SEC Release 33-6977)	Revised phase-in date (SEC Release 33-7122)	Free online date
CF-01	April 26,1993	April 26,1993	January 17,1994
CF-02	July 19,1993	July 19,1993	January 17,1994
CF-03	October 4,1993	October 4,1993	January 17,1994
CF-04	December 6,1993	December 6,1993	January 17,1994
CF-05	August 1994	January 30,1995	January 30,1995
CF-06	November 1994	March 6,1995	March 6,1995
CF-07	May 1995	May 1,1995	May 1,1995
CF-08	August 1995	August 7,1995	August 7,1995
CF-09	November 1995	November 6,1995	November 6,1995
CF-10	May 1996	May 6,1996	May 6,1996

Table IA2. Information uncertainty measured using market reaction surrounding earnings announcements.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation conditional on abnormal trading volume ($AVol_{i,t-1}$) and abnormal return volatility ($AVar_{i,t-1}$), which is calculated over a three-day window around the most recent earnings announcements. $AVol_{i,t-1}$ or $AVar_{i,t-1}$ above sample 70th percentile (below 30th percentile) is considered to be high (low) information uncertainty group. Panels A (C) and B (D) report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) High $AVol_{i,t-1}$	(2) Low $AVol_{i,t-1}$	(3) High $AVol_{i,t-1}$	(4) Low $AVol_{i,t-1}$	(5) High $AVol_{i,t-1}$	(6) Low $AVol_{i,t-1}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.173***	–0.108***	–0.919***	–0.558***	–1.618***	–0.972***
	(–8.00)	(–4.92)	(–7.39)	(–4.24)	(–8.39)	(–5.03)
$EDGAR_{i,t-1}$	0.172***	0.198***	0.908***	1.122***	1.558***	1.742***
	(8.60)	(9.60)	(7.86)	(8.94)	(8.78)	(9.68)
$P_{i,t-1}^{52h}$	–0.323***	–0.283***	–1.837***	–1.770***	–2.650***	–2.292***
	(–19.23)	(–16.46)	(–19.12)	(–17.04)	(–18.10)	(–15.52)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	328,731	334,776	328,731	334,776	328,731	334,776
R-Squared	0.345	0.370	0.345	0.375	0.326	0.353
$P^{52h} \times EDGAR$	–0.065***		–0.362**		–0.647***	
High $AVol_{i,t-1}$ – Low $AVol_{i,t-1}$	(–2.62)		(–2.48)		(–2.98)	
Panel B: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) High $AVol_{i,t-1}$	(2) Low $AVol_{i,t-1}$	(3) High $AVol_{i,t-1}$	(4) Low $AVol_{i,t-1}$	(5) High $AVol_{i,t-1}$	(6) Low $AVol_{i,t-1}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.173***	0.108***	0.898***	0.395***	1.698***	1.006***
	(8.00)	(4.92)	(7.16)	(2.99)	(8.73)	(5.20)
$EDGAR_{i,t-1}$	–0.172***	–0.198***	–0.977***	–1.020***	–1.580***	–1.719***
	(–8.60)	(–9.60)	(–8.34)	(–8.05)	(–8.81)	(–9.50)
$P_{i,t-1}^{52h}$	0.323***	0.283***	1.743***	1.711***	2.988***	2.688***
	(19.23)	(16.46)	(17.77)	(16.17)	(20.19)	(18.10)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	328,731	334,776	328,731	334,776	328,731	334,776
R-Squared	0.345	0.370	0.339	0.363	0.330	0.356
$P^{52h} \times EDGAR$	0.065***		0.503***		0.692***	
High $AVol_{i,t-1}$ – Low $AVol_{i,t-1}$	(2.62)		(3.39)		(3.15)	

Panel C: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) High $AVar_{i,t-1}$	(2) Low $AVar_{i,t-1}$	(3) High $AVar_{i,t-1}$	(4) Low $AVar_{i,t-1}$	(5) High $AVar_{i,t-1}$	(6) Low $AVar_{i,t-1}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.202***	–0.095***	–1.062***	–0.474***	–1.885***	–0.875***
	(–9.32)	(–4.50)	(–8.68)	(–3.84)	(–9.79)	(–4.76)
$EDGAR_{i,t-1}$	0.208***	0.181***	1.133***	0.998***	1.883***	1.601***
	(9.94)	(9.19)	(9.49)	(8.53)	(10.18)	(9.39)
$P_{i,t-1}^{52h}$	–0.396***	–0.293***	–2.169***	–1.808***	–3.307***	–2.408***
	(–21.59)	(–16.74)	(–20.81)	(–17.44)	(–20.59)	(–15.98)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	322,812	338,385	322,812	338,385	322,812	338,385
R-Squared	0.357	0.364	0.357	0.366	0.337	0.346
$P^{52h} \times EDGAR$	–0.107***		–0.588***		–1.010***	
High $AVar_{i,t-1}$ – *Low* $AVar_{i,t-1}$	(–3.14)		(–3.08)		(–4.03)	
Panel D: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) High $AVar_{i,t-1}$	(2) Low $AVar_{i,t-1}$	(3) High $AVar_{i,t-1}$	(4) Low $AVar_{i,t-1}$	(5) High $AVar_{i,t-1}$	(6) Low $AVar_{i,t-1}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.202***	0.095***	1.016***	0.358***	1.938***	0.885***
	(9.32)	(4.50)	(8.20)	(2.87)	(10.00)	(4.77)
$EDGAR_{i,t-1}$	–0.208***	–0.181***	–1.149***	–0.953***	–1.920***	–1.578***
	(–9.94)	(–9.19)	(–9.45)	(–8.08)	(–10.25)	(–9.19)
$P_{i,t-1}^{52h}$	0.396***	0.293***	2.108***	1.736***	3.632***	2.830***
	(21.59)	(16.74)	(19.72)	(16.35)	(22.28)	(18.61)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	322,812	338,385	322,812	338,385	322,812	338,385
R-Squared	0.357	0.364	0.351	0.358	0.342	0.348
$P^{52h} \times EDGAR$	0.107***		0.658***		1.053***	
High $AVar_{i,t-1}$ – *Low* $AVar_{i,t-1}$	(3.14)		(3.87)		(3.15)	

Table IA3. Information uncertainty measured based on firm characteristics.

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation conditional on firm characteristics. Stocks with firm size or age above the sample 70th percentile (below the 30th percentile) are considered to be large (small) or old (young). The subgroup results are reported in Panels A & B and C & D, respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) Large firm	(2) Small firm	(3) Large firm	(4) Small firm	(5) Large firm	(6) Small firm
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.015	–0.158***	–0.255	–0.891***	–0.228	–1.356***
	(–0.43)	(–8.25)	(–1.35)	(–7.14)	(–0.71)	(–8.15)
$EDGAR_{i,t-1}$	0.037	0.235***	0.206	1.342***	0.468	1.981***
	(0.98)	(14.14)	(1.02)	(12.37)	(1.36)	(13.83)
$P_{i,t-1}^{52h}$	–0.781***	–0.145***	–3.919***	–1.102***	–6.834***	–1.099***
	(–25.88)	(–11.88)	(–24.67)	(–14.47)	(–25.01)	(–10.82)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	297,118	363,436	297,118	363,436	297,118	363,436
R-Squared	0.295	0.303	0.282	0.310	0.279	0.291
Panel B: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) Large firm	(2) Small firm	(3) Large firm	(4) Small firm	(5) Large firm	(6) Small firm
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.015	0.158***	–0.045	0.838***	0.213	1.457***
	(0.43)	(8.25)	(–0.23)	(6.72)	(0.65)	(8.78)
$EDGAR_{i,t-1}$	–0.037	–0.235***	–0.126	–1.335***	–0.340	–2.039***
	(–0.98)	(–14.14)	(–0.60)	(–12.38)	(–0.97)	(–14.34)
$P_{i,t-1}^{52h}$	0.781***	0.145***	3.940***	1.030***	7.228***	1.474***
	(25.88)	(11.88)	(24.27)	(13.53)	(25.90)	(14.47)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	297,118	363,436	297,118	363,436	297,118	363,436
R-Squared	0.295	0.303	0.284	0.298	0.283	0.292

Panel C: Buy orders	$Buy_{i,j,t}$		$BuyShr_{i,j,t}$		$BuyValue_{i,j,t}$	
	(1) Old firm	(2) Young firm	(3) Old firm	(4) Young firm	(5) Old firm	(6) Young firm
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	–0.036	–0.136***	–0.240	–0.757***	–0.455*	–1.248***
	(–1.31)	(–6.67)	(–1.58)	(–6.26)	(–1.88)	(–7.08)
$EDGAR_{i,t-1}$	0.200***	0.183***	1.084***	1.065***	1.714***	1.613***
	(5.75)	(10.25)	(5.51)	(10.02)	(5.68)	(10.53)
$P_{i,t-1}^{52h}$	–0.683***	–0.209***	–3.587***	–1.419***	–5.853***	–1.659***
	(–29.58)	(–14.85)	(–28.43)	(–16.68)	(–28.47)	(–13.76)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	306,358	329,962	306,358	329,962	306,358	329,962
R-Squared	0.308	0.283	0.303	0.291	0.295	0.268
Panel D: Sell orders	$Sell_{i,j,t}$		$SellShr_{i,j,t}$		$SellValue_{i,j,t}$	
	(1) Old firm	(2) Young firm	(3) Old firm	(4) Young firm	(5) Old firm	(6) Young firm
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$	0.036	0.136***	0.067	0.796***	0.358	1.309***
	(1.31)	(6.67)	(0.44)	(6.48)	(1.48)	(7.32)
$EDGAR_{i,t-1}$	–0.200***	–0.183***	–1.018***	–1.109***	–1.777***	–1.618***
	(–5.75)	(–10.25)	(–5.02)	(–10.24)	(–5.83)	(–10.40)
$P_{i,t-1}^{52h}$	0.683***	0.209***	3.612***	1.268***	6.226***	2.010***
	(29.58)	(14.85)	(27.27)	(14.73)	(29.72)	(16.50)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	306,358	329,962	306,358	329,962	306,358	329,962
R-Squared	0.308	0.283	0.302	0.284	0.296	0.273

Table IA4. Isolate of the influence of investor attention

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation with additional control for stock *i*'s attention-grabbing characteristic. In Panels A and B, we consider $Attention_vol_{i,t}$ in our regressions, which is an indicator that equals one if stock *i* has experienced an abnormal trading volume higher than the sample 90th percentile and zero otherwise. In Panels C and D, we use an alternative measure $Attention_ret_{i,t}$, which is a dummy that equals one if stock *i* has experienced an extreme one-day return higher than the sample 95th percentile or lower than the sample 5th percentile and zero otherwise. Panels A (C) and B (D) report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $BuyShr_{i,j,t}$	(3) $BuyValue_{i,j,t}$
$P_{i.t-1}^{52h} \times EDGAR_{i.t}$	–0.128***	–0.661***	–1.180***
	(–10.05)	(–8.99)	(–10.49)
$EDGAR_{i,t}$	0.164***	0.889***	1.451***
	(14.42)	(13.36)	(14.51)
$P_{i,t-1}^{52h} \times Attention_vol_{i,t}$	0.145***	0.719***	1.328***
	(11.06)	(9.24)	(11.63)
$Attention_vol_{i,t}$	0.001	0.131**	–0.027
	(0.13)	(2.08)	(–0.29)
$P_{i,t-1}^{52h}$	–0.365***	–2.144***	–3.014***
	(–36.70)	(–37.88)	(–35.04)
Controls	YES	YES	YES
Wave-specific Account, stock, and date	YES	YES	YES
# of Obs.	1,086,330	1,086,330	1,086,330
R-Squared	0.213	0.223	0.203
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $SellShr_{i,j,t}$	(3) $SellValue_{i,j,t}$
$P_{i.t-1}^{52h} \times EDGAR_{i.t}$	0.128***	0.591***	1.232***
	(10.05)	(7.99)	(10.94)
$P_{i,t-1}^{52h}$	–0.164***	–0.869***	–1.465***
	(–14.42)	(–12.97)	(–14.63)
$P_{i,t-1}^{52h} \times Attention_vol_{i,t}$	–0.145***	–0.750***	–1.353***
	(–11.06)	(–9.47)	(–11.75)
$Attention_vol_{i,t}$	–0.001	–0.041	0.081
	(–0.13)	(–0.65)	(0.88)
$P_{i,t-1}^{52h}$	0.365***	2.023***	3.403***
	(36.70)	(34.82)	(39.08)
Controls	YES	YES	YES
Wave-specific Account, stock, and date	YES	YES	YES
# of Obs.	1,086,330	1,086,330	1,086,330
R-Squared	0.213	0.216	0.207

Panel C: Buy orders	(1) $Buy_{i,j,t}$	(2) $BuyShr_{i,j,t}$	(3) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	–0.134***	–0.703***	–1.229***
	(–10.41)	(–9.51)	(–10.85)
$EDGAR_{i,t}$	0.171***	0.933***	1.508***
	(14.77)	(13.84)	(14.86)
$P_{i,t-1}^{52h} \times Attention_ret_{i,t}$	0.183***	1.004***	1.639***
	(16.24)	(14.92)	(16.78)
$Attention_ret_{i,t}$	–0.066***	–0.330***	–0.606***
	(–8.05)	(–6.60)	(–8.52)
$P_{i,t-1}^{52h}$	–0.351***	–2.061***	–2.891***
	(–34.30)	(–35.37)	(–32.61)
Controls	YES	YES	YES
Wave-specific Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.208	0.217	0.198
Panel D: Sell orders	(1) $Sell_{i,j,t}$	(2) $SellShr_{i,j,t}$	(3) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	0.134***	0.625***	1.273***
	(10.41)	(8.39)	(11.21)
$EDGAR_{i,t}$	–0.171***	–0.908***	–1.517***
	(–14.77)	(–13.36)	(–14.90)
$P_{i,t-1}^{52h} \times Attention_ret_{i,t}$	–0.183***	–0.991***	–1.659***
	(–16.24)	(–14.46)	(–16.93)
$Attention_ret_{i,t}$	0.066***	0.340***	0.636***
	(8.05)	(6.72)	(8.94)
$P_{i,t-1}^{52h}$	0.351***	1.938***	3.278***
	(34.30)	(32.38)	(36.51)
Controls	YES	YES	YES
Wave-specific Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.208	0.211	0.202

Table IA5. Control for changing investor characteristics

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation additionally considering the EDGAR*Investor fixed effect. Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $BuyShr_{i,j,t}$	(3) $BuyValue_{i,j,t}$
$P_{i\,t-1}^{52h} \times EDGAR_{i\,t}$	–0.134***	–0.755***	–1.230***
	(–9.30)	(–9.08)	(–9.63)
$P_{i,t-1}^{52h}$	–0.318***	–1.881***	–2.614***
	(–30.42)	(–31.43)	(–28.86)
Controls	YES	YES	YES
Account-EDGAR, Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.257	0.260	0.243
Panel B: Sell orders	(4) $Sell_{i,j,t}$	(5) $SellShr_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	0.134***	0.693***	1.297***
	(9.30)	(8.25)	(10.14)
$P_{i,t-1}^{52h}$	0.318***	1.767***	3.003***
	(30.42)	(28.91)	(32.91)
Controls	YES	YES	YES
Account-EDGAR, Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.257	0.255	0.247

Table IA6. Control for changing return behavior

This table examines the influence of the 52-week high on retail investors' trading after EDGAR implementation additionally considering several interaction terms of EDGAR with past momentum effects. Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock, and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $BuyShr_{i,j,t}$	(3) $BuyValue_{i,j,t}$
$P_{i\,t-1}^{52h} \times EDGAR_{i\,t}$	–0.140***	–0.661***	–1.310***
	(–9.08)	(–7.58)	(–9.70)
$Runup_{i,t-5\sim t-1} \times EDGAR_{i,t}$	–0.072***	–0.482***	–0.671***
	(–3.23)	(–3.78)	(–3.43)
$Runup_{i,t-21\sim t-6} \times EDGAR_{i,t}$	–0.041***	–0.324***	–0.350***
	(–2.68)	(–3.64)	(–2.62)
$Runup_{i,m-6\sim m-2} \times EDGAR_{i,t}$	0.023***	0.098**	0.231***
	(3.48)	(2.55)	(3.92)
$Runup_{i,m-12\sim m-7} \times EDGAR_{i,t}$	0.016***	0.098***	0.145***
	(3.43)	(3.66)	(3.48)
$EDGAR_{i,t}$	0.170***	0.874***	1.523***
	(13.25)	(11.81)	(13.58)
$P_{i,t-1}^{52h}$	–0.332***	–1.976***	–2.713***
	(–32.47)	(–33.82)	(–30.70)
Controls	YES	YES	YES
Wave-specific Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.209	0.219	0.199
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $SellShr_{i,j,t}$	(3) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t}$	0.140***	0.558***	1.382***
	(9.08)	(6.32)	(10.22)
$Runup_{i,t-5\sim t-1} \times EDGAR_{i,t}$	0.072***	0.502***	0.602***
	(3.23)	(3.79)	(3.09)
$Runup_{i,t-21\sim t-6} \times EDGAR_{i,t}$	0.041***	0.328***	0.390***
	(2.68)	(3.58)	(2.89)
$Runup_{i,m-6\sim m-2} \times EDGAR_{i,t}$	–0.023***	–0.076*	–0.246***
	(–3.48)	(–1.90)	(–4.08)
$Runup_{i,m-12\sim m-7} \times EDGAR_{i,t}$	–0.016***	–0.080***	–0.168***
	(–3.43)	(–2.90)	(–4.00)
$EDGAR_{i,t}$	–0.170***	–0.831***	–1.549***
	(–13.25)	(–11.17)	(–13.82)
$P_{i,t-1}^{52h}$	0.332***	1.870***	3.095***
	(32.47)	(31.30)	(34.83)
Controls	YES	YES	YES
Wave-specific Account, stock, and date FE	YES	YES	YES
# of Obs.	1,086,337	1,086,337	1,086,337
R-Squared	0.209	0.213	0.203

Table IA7. Alternative EDGAR inclusion dates: preliminary phase-in dates.

This table repeats the tests in Table 2 but defines $EDGAR_{i,t-1}$ based on preliminary phase-in EDGAR inclusion dates (SEC Release 33-6977). Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are *t*-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		–0.141***		–0.747***		–1.306***
		(–13.12)		(–12.05)		(–13.85)
$EDGAR_{i,t-1}$		0.147***		0.803***		1.315***
		(15.74)		(14.69)		(16.08)
$P_{i,t-1}^{52h}$	–0.373***	–0.326***	–2.189***	–1.943***	–3.083***	–2.652***
	(–44.25)	(–36.02)	(–45.26)	(–37.27)	(–42.09)	(–33.95)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402
R-Squared	0.247	0.247	0.250	0.250	0.235	0.235
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		0.141***		0.680***		1.350***
		(13.12)		(10.74)		(14.17)
$EDGAR_{i,t-1}$		–0.147***		–0.778***		–1.331***
		(–15.74)		(–14.01)		(–16.16)
$P_{i,t-1}^{52h}$	0.373***	0.326***	2.050***	1.827***	3.488***	3.042***
	(44.25)	(36.02)	(41.56)	(34.31)	(47.08)	(38.50)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402
R-Squared	0.247	0.247	0.246	0.246	0.239	0.239

Table IA8. Alternative EDGAR inclusion dates: revised phase-in dates.

This table repeats the tests in Table 2 but defines $EDGAR_{i,t-1}$ based on revised phase-in EDGAR inclusion dates (SEC Release 33-7122). Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		–0.159***		–0.831***		–1.452***
		(–14.20)		(–12.92)		(–14.79)
$EDGAR_{i,t-1}$		0.175***		0.931***		1.552***
		(17.85)		(16.29)		(18.06)
$P_{i,t-1}^{52h}$	–0.373***	–0.326***	–2.189***	–1.946***	–3.083***	–2.657***
	(–44.25)	(–36.45)	(–45.26)	(–37.73)	(–42.09)	(–34.38)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402
R-Squared	0.247	0.247	0.250	0.250	0.235	0.236
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		0.159***		0.772***		1.493***
		(14.20)		(11.77)		(15.06)
$EDGAR_{i,t-1}$		–0.175***		–0.934***		–1.570***
		(–17.85)		(–16.12)		(–18.19)
$P_{i,t-1}^{52h}$	0.373***	0.326***	2.050***	1.828***	3.488***	3.048***
	(44.25)	(36.45)	(41.56)	(34.70)	(47.08)	(39.03)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402	1,252,402
R-Squared	0.247	0.247	0.246	0.246	0.239	0.239

Table IA9. Exclude transitional filers.

This table repeats the tests in Table 2 after excluding stocks in Group CF-01. Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		–0.146***		–0.763***		–1.341***
		(–11.91)		(–10.70)		(–12.42)
$EDGAR_{i,t-1}$		0.176***		0.946***		1.557***
		(16.22)		(14.82)		(16.40)
$P_{i,t-1}^{52h}$	–0.344***	–0.308***	–2.065***	–1.876***	–2.822***	–2.487***
	(–39.36)	(–33.50)	(–40.99)	(–35.43)	(–37.22)	(–31.44)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,105,096	1,105,096	1,105,096	1,105,096	1,105,096	1,105,096
R-Squared	0.234	0.234	0.240	0.241	0.222	0.222
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		0.146***		0.740***		1.397***
		(11.91)		(10.21)		(12.81)
$EDGAR_{i,t-1}$		–0.176***		–0.981***		–1.573***
		(–16.22)		(–15.18)		(–16.46)
$P_{i,t-1}^{52h}$	0.344***	0.308***	1.913***	1.733***	3.231***	2.881***
	(39.36)	(33.50)	(37.26)	(32.07)	(42.18)	(36.06)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,105,096	1,105,096	1,105,096	1,105,096	1,105,096	1,105,096
R-Squared	0.234	0.234	0.236	0.236	0.226	0.226

Table IA10. Stacked regressions with all clean controls.

This table repeats the tests in Table 2 using stacked regressions with all clean controls. Panels A and B report results for buy and sell orders, respectively. Following Cengiz et al. (2019) and Baker et al. (2022), we combine nine event-specific clean datasets, which only include treatment cohorts and not-yet-treated firms as the clean controls. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		–0.187***		–0.964***		–1.683***
		(–15.23)		(–13.60)		(–15.49)
$EDGAR_{i,t-1}$		0.208***		1.125***		1.820***
		(19.77)		(18.28)		(19.60)
$P_{i,t-1}^{52h}$	–0.287***	–0.245***	–1.777***	–1.560***	–2.329***	–1.947***
	(–29.31)	(–24.96)	(–33.52)	(–29.30)	(–27.04)	(–22.80)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,432,109	1,432,109	1,432,109	1,432,109	1,432,109	1,432,109
R-Squared	0.202	0.203	0.214	0.215	0.192	0.193
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		0.187***		0.888***		1.714***
		(15.23)		(12.50)		(15.83)
$EDGAR_{i,t-1}$		–0.208***		–1.094***		–1.828***
		(–19.77)		(–17.74)		(–19.74)
$P_{i,t-1}^{52h}$	0.287***	0.245***	1.613***	1.415***	2.825***	2.436***
	(29.31)	(24.96)	(29.07)	(25.38)	(37.52)	(32.62)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	1,432,109	1,432,109	1,432,109	1,432,109	1,432,109	1,432,109
R-Squared	0.202	0.203	0.206	0.207	0.198	0.199

Table IA11. Propensity-score matching.

This table repeats the tests in Table 2 using the propensity-score matched sample. Following Chang et al. (2022), treated firms are those included in EDGAR; control firms are nearest-neighbor propensity-scored matched on equity market capitalization (in levels and logs) and fiscal quarter using a 0.05 caliper. Panels A and B report results for buy and sell orders, respectively. All variables are as defined in Table 2. Wave-specific account, stock and date fixed effects are all controlled. Numbers in parentheses are t-statistics based on standard errors two-way clustered by wave-specific account and date. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Buy orders	(1) $Buy_{i,j,t}$	(2) $Buy_{i,j,t}$	(3) $BuyShr_{i,j,t}$	(4) $BuyShr_{i,j,t}$	(5) $BuyValue_{i,j,t}$	(6) $BuyValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		–0.107***		–0.556***		–0.980***
		(–9.47)		(–9.08)		(–9.84)
$EDGAR_{i,t-1}$		0.108***		0.574***		0.980***
		(11.58)		(11.14)		(11.93)
$P_{i,t-1}^{52h}$	–0.312***	–0.272***	–1.959***	–1.747***	–2.441***	–2.064***
	(–40.41)	(–31.98)	(–46.04)	(–37.51)	(–36.18)	(–28.47)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	910,949	910,949	910,949	910,949	910,949	910,949
R-Squared	0.164	0.164	0.171	0.171	0.157	0.157
Panel B: Sell orders	(1) $Sell_{i,j,t}$	(2) $Sell_{i,j,t}$	(3) $SellShr_{i,j,t}$	(4) $SellShr_{i,j,t}$	(5) $SellValue_{i,j,t}$	(6) $SellValue_{i,j,t}$
$P_{i,t-1}^{52h} \times EDGAR_{i,t-1}$		0.107***		0.431***		0.997***
		(9.47)		(6.80)		(10.09)
$EDGAR_{i,t-1}$		–0.108***		–0.509***		–0.940***
		(–11.58)		(–9.53)		(–11.51)
$P_{i,t-1}^{52h}$	0.312***	0.272***	1.582***	1.420***	3.036***	2.650***
	(40.41)	(31.98)	(35.70)	(29.29)	(44.55)	(36.14)
Controls	YES	YES	YES	YES	YES	YES
Wave-specific Acc., stk, & date FE	YES	YES	YES	YES	YES	YES
# of Obs.	910,949	910,949	910,949	910,949	910,949	910,949
R-Squared	0.164	0.164	0.164	0.164	0.163	0.164

Internet Appendix B: Proof

Proof of Proposition 2

As we show before,

$$\begin{aligned}\frac{\partial L^C}{\partial \theta} &= \pi 2p_B(1-p_B)\left((1-2p)\frac{1}{p_B+\frac{1-\theta}{2\theta-1}}+[p+(1-2p)\theta]\frac{1}{\left(\frac{1-\theta}{2\theta-1}+p_B\right)^2}\frac{1}{(2\theta-1)^2}\right)\\ &=\frac{\pi 2p_B(1-p_B)}{p_B+\frac{1-\theta}{2\theta-1}}\left((1-2p)+[p+(1-2p)\theta]\frac{1}{(1-\theta)(2\theta-1)+p_B(2\theta-1)^2}\right)\\ &=-\frac{\pi 2p_B(1-p_B)}{p_B+\frac{1-\theta}{2\theta-1}}\frac{1}{(1-\theta)(2\theta-1)+p_B(2\theta-1)^2}\\ &\quad\times\left((2p-1)(4p_B-2)\theta^2+(2p-1)(2-4p_B)\theta+(2p-1)(p_B-1)-p\right).\end{aligned}$$

Let $g=(2p-1)(4p_B-2)\theta^2+(2p-1)(2-4p_B)\theta+(2p-1)(p_B-1)-p$. As we can see, g is a quadratic function of θ. Meanwhile, we have the following properties.

$$g(0)=(2p-1)(p_B-1)-p<0.$$
$$g\left(\frac{1}{2}\right)<0.$$
$$g'(0)<0.$$

We can quickly conclude that when

$$\theta>\frac{-(2p-1)(2-4p_B)+\sqrt{(2p-1)^2(2-4p_B)^2+4(2p-1)(4p_B-2)[(2p-1)(1-p_B)+p]}}{2(2p-1)(4p_B-2)}=\theta^*$$

or $\theta<\frac{-(2p-1)(2-4p_B)-\sqrt{(2p-1)^2(2-4p_B)^2+4(2p-1)(4p_B-2)[(2p-1)(1-p_B)+p]}}{2(2p-1)(4p_B-2)}$.

However, since $\frac{-(2p-1)(2-4p_B)-\sqrt{(2p-1)^2(2-4p_B)^2+4(2p-1)(4p_B-2)[(2p-1)(1-p_B)+p]}}{2(2p-1)(4p_B-2)}<0$, it suggests that when $\frac{1}{2}\leq\theta<\theta^*$, we always have $g(\theta)<0$. $g(\theta)>0$ only if $\theta>\theta^*$.

Internet Appendix C: Experimental Evidence

1. Methods

We conduct an experiment to investigate how more information coupled with individuals' confirmation bias would affect their inherent anchoring bias. This experiment is hosted on Credamo, one of China's largest and most professional platforms for online experiments. It has a sample base of three million participants and is comparable to platforms like Amazon Mechanical Turk or Prolific in the U.S.

We request Credamo to randomly recruit two groups of participants, each with 125 subjects, and end up with 246 subjects in total. Group 1, the treated group, has 125 subjects with an average age of 28.46 and 58.40% females. Group 2, the control group, has 121 subjects with an average age 28.66 and 57.85% females. In each group, subjects are asked to predict the future price movement of a stock based on (1) the stock's simulated historical price chart over the past 52 weeks and (2) fundamental information about the stock. The price chart has two versions: one where the current price is close to its 52-week high (Figure IA1) and the other where the current price is still far from the 52-week high (Figure IA2). Figures IA1 and IA2 are designed in a way that even though the current price's proximity to its 52-week high differs, the most recent price trend is comparable. Thus, the difference in subjects' predictions of future price movements between Groups 1 and 2 is less likely to be influenced by the stock's recent performance.

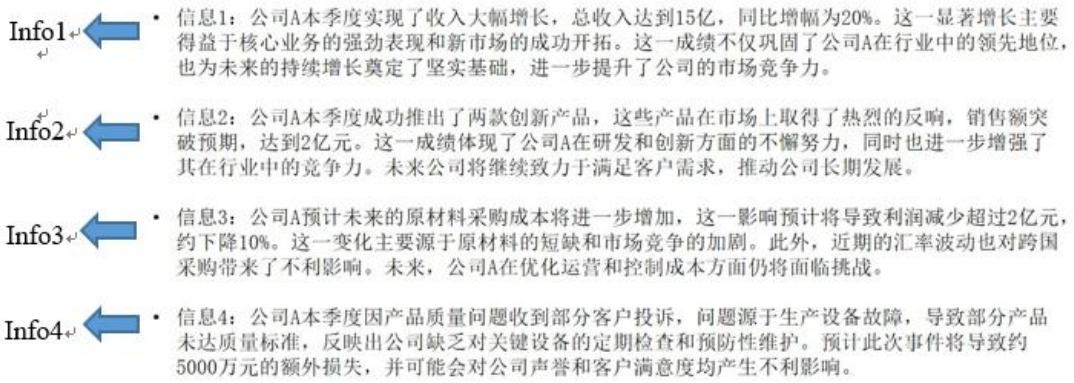


Figure IA1. Price Chart for Group 1



Figure IA2. Price Chart for Group 2

We further present to subjects financial information of the stock. The information set (Figure IA3) provides two pieces of positive information and two pieces of negative information, all containing detailed descriptions about the firm's business conditions. In designing these four pieces of information, we ensure that the degrees of their positive and negative tones are approximately equal as reflected in the sentiment scores. As a result, these four pieces of information combined provide a neutral view on the stock's fundamental performance overall.



The four pieces of information (displayed in Chinese on the subjects' screens) translated into English are as follows:

Info 1: Company A achieved rapid revenue growth this quarter, with total revenue reaching $1.5 billion, representing a 20% year-over-year increase. This significant growth is primarily driven by strong performance in core business areas and successful expansion into new markets. This achievement not only reinforces Company A's leading position in the industry but also lays a solid foundation for future sustainable growth, further enhancing the company's market competitiveness. **(Positive News Sentiment Score generated by ChatGPT: 9)**

Info 2: Company A successfully launched two innovative products this quarter, which received an enthusiastic market response and exceeded sales expectations, reaching 200 million yuan. This achievement reflects Company A's relentless efforts in research and development and innovation, while also enhancing its competitiveness in the industry. Moving forward, the company will continue to focus on meeting customer needs and driving long-term growth. **(Positive News Sentiment Score: 8)**

Info 3: Company A anticipates a further increase in raw material procurement costs in the future, which is expected to result in a reduction in profits of over 200 million yuan, representing a decrease of approximately 10%. This change is primarily driven by raw material shortages and increased market competition. Additionally, recent currency fluctuations have negatively impacted international procurement. Going forward, Company A will continue to face challenges in optimizing operations and controlling costs. **(Negative News Sentiment Score: –8)**

Info 4: This quarter, Company A received customer complaints due to product quality issues, which stemmed from equipment malfunctions that resulted in some products failing to meet quality standards. This situation highlights the company's lack of regular inspections and preventive maintenance for critical equipment. The incident is expected to incur additional losses of approximately 50 million yuan. and may negatively impact the company's reputation and customer satisfaction. **(Negative News Sentiment Score: –9)**

Figure IA3. Information Presented to Subjects in both Groups 1 and 2

Subjects in both groups undergo a 2-stage process. For Group 1, the treated group, the first stage involves presenting Figure IA1 to subjects and asking them to select the direction they believe the stock price is most likely to move: increase, decrease, or remain similar. They

are also required to rate the level of confidence in their prediction on a scale from 1 (lowest level of confidence) to 7 (highest level of confidence). After submitting their responses, the second stage begins with the display of the information set, shown in Figure IA3, on their computer screens. After reviewing this information, the subjects are required to make another prediction regarding the future price movement (increase, decrease, or remain similar) and indicate which of the four pieces of information (two positive and two negative) is most relevant for making the second prediction. Finally, they are asked to assess their confidence in second prediction as well.

The subjects in Group 2, the control group, go through the same 2-stage process. However, in the first stage, they are shown Figure IA2 instead of Figure IA1 before being asked to predict the direction of future price movement. In Figure IA2, the current stock price level is far from its 52-week high but has a similar recent price trend.

We collect three sets of information during the experiment: (1) the subjects' predictions of future price movement directions in the first and second stages, (2) the subjects' self-reported confidence scores about their predictions in both stages, and (3) the subjects' choices of information that they feel is the most relevant for their second prediction.

2. Results

2.1 Anchoring Bias

Table IA12 presents the distribution of subjects' predictions on future stock price movements after viewing the price charts in the first stage. In Groups 1 and 2, we retain only the subjects with a clear prediction of the directional price movement. In other words, the subjects who predict that the price would remain similar (0 out of 125 in Group 1, and 3 out of 121 in Group 2) are excluded from the follow-up analysis.

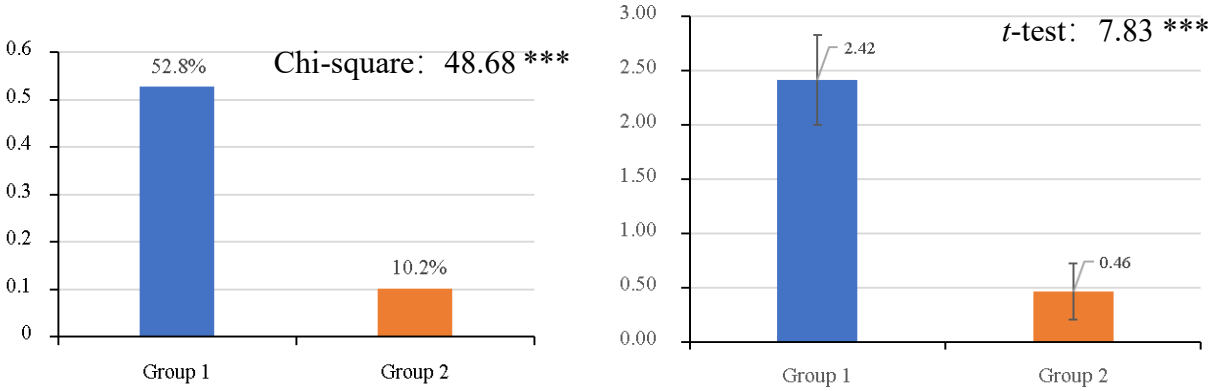
In Group 1, where subjects are presented with the price chart in Figure IA1, the percentage of subjects predicting a price decline (52.8%) is dominating. Conversely, in Group 2, where the subjects are presented with the price chart in Figure IA2, the percentage of subjects predicting a price increase (89.8%) is substantially higher than that predicting a price decline (10.2%). Panel A of Figure IA4 graphically illustrates the percentage of subjects predicting a price decline in Groups 1 and 2. The chi-square test indicates that the difference between the two groups is statistically significant at the 1% level (chi-square 48.68). Panel B of Figure IA4

further displays the average of subjects' self-reported confidence scores regarding this prediction. The difference between Groups 1 and 2 is also significantly positive. The evidence confirms that 52-week high serves as a robust anchor for individuals in predicting future stock price movements, consistent with the finding using Odean's account-level data.

Table IA12: Prediction of stock price movement direction in the first stage

Group	Prediction of stock price movement direction			
	Increase		Decline	
	N	Pct	N	Pct
1 (N=125)	59	47.20%	66	52.80%
2 (N=118)	106	89.83%	12	10.17%
Difference (1–2)	–47	–42.63%	54	42.63%

Figure IA4: Prediction of price movement direction and self-reported confidence scores in the first stage



Panel A. % of subjects predicting a price decline

Panel B. Average of self-reported confidence scores

2.2 Confirmation Bias

In stage 2, subjects are given four pieces of hypothetical information as shown in Figure IA3 - two positive ones (Info. 1 and 2) and two negative ones (Info. 3 and 4). They are asked to identify the piece of information they consider the most relevant for making the price change prediction. The four pieces of information are designed to ensure that their sentiment scores are comparable. Table IA13 provides strong evidence that the subjects who predict a price increase in the first stage tend to consider positive information as the most relevant (71.19% in Group 1 and 66.04% in Group 2) in the second stage, while those predicting a price decline tend to prioritize negative information (77.27% in Group 1 and 75.00% in Group 2). This evidence confirms that our subjects tend to focus on information that aligns with their prior

predictions.

Table IA13: Selection of the most relevant information in the second stage

		Group 1 (N=125)		Group 2 (N=121)	
		Price prediction, 1st stage			
		Increase	Decline	Increase	Decline
# of subjects selecting the info. as the most relevant	Info. 1 (positive)	28	11	47	2
	Info. 2 (positive)	14	4	23	1
	Info. 3 (negative)	10	32	22	4
	Info. 4 (negative)	7	19	14	5
% selecting pos. info. (piece 1 or 2)		71.19%	22.73%	66.04%	25.00%
% selecting neg. info. (piece 3 or 4)		28.81%	77.27%	33.96%	75.00%
Difference (Pos% – Neg%)		42.38%	–54.54%	32.08%	–50.00%

2.3 When confirmation bias met anchoring bias

Table IA12 shows that 66 out of 125 subjects in Group 1 predict a price decline after viewing the price chart showing the stock price approaching its 52-week high, supporting the conjecture that the 52-week high acts as an anchor (a resistance level). As we focus on how confirmation bias affects subjects' inherent anchoring bias, we analyze the behavior only for these 66 subjects in this step. And as 3 out these 66 subjects predict the price to stay similar in the second stage, we further exclude them from this analysis to concentrate on subjects who not only exhibit an anchoring tendency in the first stage but also make clear price movement predictions in the second stage.

Table IA14 shows that the majority of the subjects in Group 1 who predict a price decline in the first stage tend to consider negative information as the most relevant (50 out of 63) in the second stage, indicating their confirmation bias. Within this subgroup, shown in row (1), the vast majority (49 out of 50) continue to predict a price decline after selectively choosing to focus on information that confirms their initial prediction. Moreover, their average self-reported confidence score increases from 4.57 in stage 1 (where no information was provided) to 5.0 in stage 2 (where they could selectively choose to focus on information aligning with their initial prediction). This increase in confidence score is statistically significant at the 1% level.

In contrast, for those who predict a price decline in the first stage but select the positive information as the most relevant in the second stage, as shown in row (4), the change in their confidence scores from the first to the second stage is not statistically significant. This result suggests that for these subjects focusing on information that contradicts their prior beliefs, which is indicative of their weak confirmation bias, additional information is less likely to enhance their inherent anchoring bias. These pieces of evidence combined support our argument that when more information is available, information cherry-picking due to confirmation bias could exacerbate individuals' inherent behavioral biases such anchoring.

3. Supplementary Study

In a supplementary study, we recruit an additional control group through Credamo to investigate the impact of additional information on individuals' pre-existing anchoring bias. Specifically, the experiment aims to provide direct evidence on whether more information amplifies this anchoring tendency compared to the absence of additional information. This experimental study mirrors our large-scale empirical setting, replicating its key elements within a fully controlled environment.

The treated group is the same as Group 1 above, where subjects first observe the stock's price charts with a current price close to the past 52-week high (Figure IA1), then review the stock's fundamental information (Figure IA3). The newly collected control group (Group 3) consists of 123 subjects, with an average age of 28.38 and 56.10% females. These subjects only observe the same price chart as the treated group (Figure IA1) without browsing any fundamental information. Both groups are asked to make a price prediction and rate confidence level on a scale from 1 (lowest) to 7 (Highest). By comparing the confidence levels in anchoring tendency between Group 1 and Group 3, we can directly assess the impact of additional information on individuals' anchoring propensity, consistent with our empirical setting.

We only keep predictions of price decreases and calculate the average confidence level to capture the extent of individuals' anchoring tendency. Table IA15 shows that the confidence level of the anchoring tendency for Group 1 is 4.92, significantly higher than the 4.63 in Group 3, indicating subjects in Group 1 are more "confident" in this anchoring bias. The difference of 0.29 is statistically significant at the 5% level. This supplementary evidence reinforces our findings that providing more information can exacerbate, rather than mitigate, pre-existing behavioral biases.

Table IA14. Changes in confidence scores conditional on subjects' prior price prediction and the consumption of information

Subgroups	Price prediction, 1st stage	Price prediction, 2nd stage	Selection of most relevant info.	Avg. confidence score, 1st stage	Avg. confidence score, 2nd stage	Change in confidence score
(1) N=49, Pct = 77.8%	Decrease	Decrease	Negative info	4.57	5.0	**+0.43*** (p=0.00)**
(2) N=3, Pct = 4.8%	Decrease	Decrease	Positive info.	4.67	3.67	–1.00 (p-value N.A.)
(3) N=1, Pct = 1.6%	Decrease	Increase	Negative info	4.0	5.0	+1.00 (p-value N.A.)
(4) N=10, Pct=15.9%	Decrease	Increase	Positive info.	4.5	4.8	+0.30 (p=0.34)

Table IA15. Confidence levels of anchoring tendency: with vs. without fundamental information

	Prediction of Price Decrease		
	Group 1	Group 3	Diff (1–3)
Avg. conf. score	4.92	4.63	0.29**
			(2.07)