

Please do not circulate the paper without the permission from the authors.

Short Sellers Are the Vanguard of the SEC Investigation: with Evidence from Controlled Experiment[†]

Yang Bai

Xiaohu Guo

Ruixiang Wang

(Preliminary Draft)

August 14, 2023

Abstract

We study the coordination between short sellers and the Securities Exchange Commission (SEC) in initiating the SEC investigations. Our results confirm that the SEC relies on the short sellers in selecting investigation targets. We establish the causal relation using an experiment, the Pilot Program of Reg SHO (2005-2007) that increased the short-selling interest in randomly selected stocks. During the program, we find that pilot firms' SEC investigation risk surged by 2.5%, or 73.5% relative to the average SEC investigation risk in the full sample and that such effect reverted after the program. Information asymmetry amplifies the SEC's reliance on the short sellers, while information certainty mitigates it. Short-selling interest predicts income-reducing, accounting, and error restatements, reflects poor accounting quality, and anticipates shareholder lawsuits. Consistent with our investigation results, the SEC also issues more comment letters to firms with high short-selling interest.

Keywords: Accounting Quality, Information Asymmetry, Comment Letters, SEC Investigation, Short Selling.

JEL Classification: G38, K22, M48.

[†] Yang Bai (yangbai@mail.missouri.edu) is from Robert J. Trulaske, Sr. College of Business, University of Missouri. Xiaohu Guo (xguo19@crimson.ua.edu) is from Culverhouse College of Business, University of Alabama. Ruixiang (Ray) Wang (ruixwang@clark.edu) is from the School of Management, Clark University.

Please do not circulate the paper without permission from the authors.

“The public benefit from the ‘short’ side of the market is less well understood, but no less valuable. As Edward Chancellor, the noted expert in the history of finance, wrote in 2001, ‘we need more, not less, shorting activity if, in the future, we are to avoid wasteful bubbles, such as the recent technology, media and telecoms boom.’ Many of the major corporate frauds and bankruptcies of the past quarter century were first exposed by short sellers doing fundamental research: Enron, Tyco, Sunbeam, Boston Chicken, Baldwin United, MicroStrategies, Conseco, ZZZZBest and Crazy Eddie are but a few examples of this phenomenon.”

— James Chanos, Prepared Statement, SEC Roundtable on Hedge Funds, May 15, 2003¹

1. Introduction

On January 31, 2020, Muddy Waters, a short-selling investment firm, publicized an anonymous report alleging that Luckin Coffee, a Chinese coffeehouse chain listed on NASDAQ at the time, inflated its sales.² The anonymous reporter hired 92 full-time and 1,418 part-time staff to run the investigation at the store level across the mainland China. They sent out field teams to collect store traffic information and obtained 11,260 hours of videos and granular order information, including 25,843 receipts from 10,119 customers in 2,213 stores in 45 cities. Shortly after the report, the Securities and Exchange Commission (SEC) opened an investigation into Luckin’s conduct, after which Luckin’s share plummeted by 81% on April 2, 2020. By the end of April, the firm admitted that more than 300 million USD was falsified in 2019. Luckin Coffee’s stock price collapsed soon after the start of the SEC investigation and was eventually delisted from NASDAQ.

In the stunning fall of Luckin Coffee, the SEC faced the most extreme example of its resource constraint and lack of information in the securities law enforcement due to the boundary of jurisdiction area. The economic cost was also expensive: the anonymous reporter spent a total of 981 store-days to collect the granular sales information, which is an impossible number of labor-days that the SEC can afford on a single case. However, the short sellers, driven by profits, stepped in for private scrutiny that in turn catered to the

¹ See the prepared statement from James Chanos for the SEC’s hedge fund roundtable in 2003 that provides details of how short sellers collect information and form their belief of firm value: <https://www.sec.gov/spotlight/hedgefunds/hedge-chanos.htm>.

² See Wall Street Journal article: <https://www.wsj.com/articles/coffees-for-closers-how-a-short-sellers-warning-helped-take-down-luckin-coffee-11593423002> for a general description of the fraud. The anonymous report can still be found online: https://cdn.gmtresearch.com/public-ckfinder/Short-sellers/Unknown%20author/Luckin%20Coffee_Anonymous.pdf. See the SEC’s complaint on the case of Luckin Coffee for more information: <https://www.sec.gov/litigation/complaints/2020/comp-pr2020-319.pdf>.

Please do not circulate the paper without permission from the authors.

regulator for their detailed evidence of corporate fraud. Such private scrutiny from short sellers is not rare. For example, James Chanos became famous for successfully predicting the downfall of Enron, and Andrew Left discovered the inflated sales of Valeant Pharmaceuticals. Collectively, these cases depict a possibility that the SEC, motivated by its resource constraint, can leverage short sellers as its vanguards of enforcement to initiate investigations according to the short-selling activities (Fang et al. 2016; Karpoff and Lou 2010). In this paper, we aim to understand the link between the initiation of SEC investigation and short seller interests..

The literature highlights the predictive ability of short sellers in many areas, such as detecting financial misrepresentations (Karpoff and Lou 2010). However, there is a missing link between the short sellers' activities and the initiation of the SEC enforcement. Using the SEC's nonpublic investigation records obtained through Freedom of Information Act (FOIA), we connect this missing link from the SEC's perspective and fill the gap in the literature. We generate a firm-year level short interest proxy to directly investigate its influence on the SEC's choice of investigation target. We show that, from 1996 to 2015, a one standard deviation increase in the short-selling interest is related to 0.5% increase in the SEC's investigation risk. In other words, firms with short-selling interest that is one standard-deviation higher than the average face a 14.3% increase in the SEC investigation risk.

We are cautious that our empirical finding is subject to endogeneity concerns. For example, perhaps there are some unobservable heterogeneities at firm, manager or board level that drove the SEC investigation beyond the short sellers' detection, causing the omitted variable bias. Or the short sellers had insider information within SEC to foresee an imminent investigation, thus a reverse causality issue. To reliably establish a causal relation, we leverage the SEC's regulatory experiment, the Rule 202T Pilot Program of the Regulation SHO (pilot program, hereafter), that took place between July 2004 and August 2007 (Fang et al. 2016; Li and Zhang 2015).³ During the period, the SEC adopted a new regulation governing the short-selling activities in the US stock markets. The commission randomly selected stocks into the Pilot Program that temporarily exempted them from the

³ The experiment effectively ran from May 2, 2005 to July 2, 2007.

Please do not circulate the paper without permission from the authors.

short-sale price tests, thus directly decreased the cost of short selling in the pilot stocks.⁴ Consequently, shorting on pilot stocks is more profitable compared to the non-pilot stocks. Mechanically, such change increases the short-sellers trading volume of the pilot stocks and hence raises the short-selling interest.

The pilot program represents an exogenous shock to the short selling activities with randomly chosen treatment and control groups and clearly identified experiment period (Fang et al. 2016). Therefore, the program provides technical convenience for the application of difference-in-differences (DID) analysis. We match the firm-year observations to the stock tickers and identify the members of the pilot group following the SEC's pilot.⁵

Following Fang et al. (2016) and Li and Zhang (2015), we model the SEC investigation with a double DID regression. We find supportive evidence on our main finding that the short sellers' focus can lead to SEC investigation. The pilot firms face significantly increased SEC investigation risk during the program, and such firms' SEC investigation risk became indistinguishable from nonpilot firms after the program concluded. Figure 1 demonstrates the changes of the SEC investigation risk by time periods of the pilot program and stock groups. The SEC investigation risk of the pilot stocks increased more than 3% during the pilot program compared to the risk level before the pilot program. The risk difference between the pilot and the nonpilot stocks magnified during the pilot program and reverted to a negligible level following the pilot program. Specifically, our double DID regression indicates that the pilot firms during the program face 2.5% higher SEC investigation risks, which is a 73.5% increase compared with the investigation risk of an average firm in the full sample from 1996 to 2016.

[Insert Figure 1 Here]

⁴ The SEC chose every third member of the Russell 3000 index ranked by average daily dollar volume from July 28, 2003 to July 28, 2004. The short-sale price tests comprise the uptick rule adopted by the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX) and the bid-price rule adopted by the National Association of Securities Dealers Automated Quotations Stock Market (NASDAQ). The uptick rule requires that the short-selling offers can only be placed at a price above the most recent trade and the bid-price rule mandates short-selling offers be placed at a price above the current bid price.

⁵ The SEC did not release the non-pilot list included in its 2007 study. However, the difference between our non-pilot list and the SEC's finalized non-pilot list is likely immaterial.

Please do not circulate the paper without permission from the authors.

We hypothesize that the information asymmetry, and thus search cost, between the firms and the SEC facilitates the SEC's reliance on the short sellers for initiating investigations. This effect is bidirectional. On the one hand, the SEC's limited resources can impede its ability to detect fraud because even the public information filed with the SEC can be deceiving. For example, Joseph Kennedy, the first SEC chair, once said that *"Before you can be asked to invest your money in a business, there must be a record in Washington of the important facts which should guide your judgment ... The truth of these facts cannot be guaranteed. There will always be people to whom an oath is meaningless. There will always be problems of bad management."* Specifically, we hypothesize that such under-detection can trouble the SEC the most during its busiest time, when the information asymmetry and search cost also become greatest. Therefore, the SEC will depend more on the short sellers when it is under pressure (Bonsall et al. 2021; Donelson et al. 2022; Ege et al. 2019; Kedia and Rajgopal 2011).

On the other hand, some firms' information itself can be complex and asymmetric, under which situation the SEC can face challenges in deciphering it. We hypothesize that such information asymmetry from the firm side can also direct the SEC to rely on the information from the experts in the private sector, i.e., short sellers.

Indeed, we show that the SEC's investigation depends more on the short sellers when the SEC faces higher investigation backlog. In particular, for a firm with an average level of short interest, a one-standard-deviation increase in the total number of SEC's investigation cases can lead to 0.3% increase in the SEC's investigation risk, which translates to an increase of 10% in the investigation risk relative to the average investigation risk of 3.4% in our sample.⁶

To proxy for the information asymmetry from the firm side, we use two sets of variables with respect to earning's uncertainty and stock return uncertainty. We find that information asymmetry at the firm level can amplify the SEC's dependence on the short sellers. For example, a one-standard-deviation increase in the volatility of return on assets (ROA) can lead to a 0.2% increase in the SEC investigation risk, which is a 5.3% increase

⁶ A one-standard-deviation increase in the number of SEC investigation cases is 85 cases in a fiscal year.

Please do not circulate the paper without permission from the authors.

relative to the average investigation risk. In addition, a one-standard-deviation increase in the absolute discretionary accrual can lead to a 0.2% in the probability of facing SEC investigation, and this increase is 6% relative to the unconditional average of the SEC investigation risk. The role of the information asymmetry at the firm level is also confirmed by the stock market proxies, including bid-ask spread and return volatility.

Next, we examine if the information certainty could mitigate the SEC's reliance on short sellers. We investigate the influence of bad corporate governance mechanisms that explicitly weaken shareholders' interest. We adopt the Governance Index (G-index), which sums up the number of corporate provisions that restrict the shareholders' rights (Gompers et al. 2003). We conjecture that in the control-enhancing mechanisms in corporate bylaws, which are easily accessible and understandable, could directly attract the SEC's attention while reducing its reliance on short sellers. This is particularly relevant since investor protection is one of the SEC's core missions.⁷ Our results support our hypothesis. The effect from short-selling interest on the SEC investigation risk is reduced by 0.5% relative to the average investigation risk with every one-standard-deviation increase in the G-index.

As a robustness check, we find consistent evidence with Fang et al. (2016) that the short-selling interest is positively related to the likelihood of earnings management and restatement. Specifically, our findings demonstrate that short-selling interest exhibits exceptional performance in detecting income-reducing restatements, restatements related to accounting principles, and restatements due to errors. However, it does not show the same effectiveness in detecting restatements related to financial fraud. This highlights short sellers' ability to analyze publicly available accounting information. In addition, short-selling interest also predicts class action lawsuits and the shareholder litigation. And the SEC is also more likely to issue comment letters to firms with short sellers' interest. A standard-deviation increase in the short position leads to 0.67% increase in the recipience of SEC comment letter, which is an 3% relative increase.

⁷ The SEC has three major missions, including protecting investors, facilitating capital formation, and maintaining fair, orderly and efficient markets. See the SEC's website for details: <https://www.sec.gov/about/what-we-do>.

Please do not circulate the paper without permission from the authors.

This paper primarily contributes to the regulatory enforcement literature and highlights the forerunner nature of the short sellers' trading activities in the SEC's enforcement. A strand of regulatory enforcement literature has explored the discretion in the SEC's enforcement decisions (Baugh et al. 2022; Do and Zhang 2022; Donelson et al. 2022; Ege et al. 2019; Kalmenovitz 2020). In particular, Correia (2014), deHann et al. (2015), and Heese (2019) identify several firm characteristics, including revolving doors and political election, that are related to the SEC's Accounting and Auditing Enforcement Releases (AAER) cases.

The literature, however, is relatively silent on SEC investigation as compared to AAER enforcement actions, even though the former is the *starting point* of any SEC enforcement. For example, Blackburne et al. (2021) and Solomon and Soltes (2021) highlight the consequences of the SEC investigations to firms, including the increased insider trading activities and a huge market penalty upon voluntary disclosure. Such serious consequences call for understanding of the determinants in the selection of the SEC's investigation targets. Holzman et al. (2023) take the first step and document that the likelihood of noncompliance, the private sector scrutiny, and the public triggers are associated with the SEC's investigation case selection. However, the specific influence from the short sellers on the SEC's investigation remains unclear.

Dechow et al. (1995) show that the short-selling interest increases in two months before the SEC's AAER. Fang et al. (2016) show that, during the pilot program of Regulation SHO, firms with prior fraud is more susceptible to the SEC's enforcement. Yet, the link between the financial misrepresentation and the detection of financial misrepresentation from the regulatory perspective is missing in the literature. Motivated by these studies including the predictive power of corporate misconduct documented in Karpoff and Lou (2010), we fill the gaps in the literature with a precise focus on the relation between short sellers' activities and the SEC's investigation case selection.

We explicitly detail a mechanism that increases the fraud detection in regulatory enforcement attributable to short-selling interest. Additionally, we document the interactions between the short-selling interest and other corporate environment variables,

Please do not circulate the paper without permission from the authors.

pointing out the importance of information and the effects from short-selling interest on the SEC enforcement risk.

Our findings also contribute to the literature of short sellers as informed investors. Short sellers are sophisticated investors who have superior access to information and only take actions with certainty, which adds new information to the market (Boehmer et al. 2008; Boehmer et al. 2010; Diamond and Verrecchia 1987; Miller 1977). For example, Callen and Fang (2015) document that the short-seller interest is positively related to the one-year-ahead stock crash risk. Bao et al (2018) show that the short-selling activities predict the bad news withheld by managers. This paper extends this strand of literature. We show that the SEC takes the advantage of short sellers' access to material nonpublic information and initiates investigations conditional on short sellers' actions.

The paper is organized as follows. Section 2 describes the empirical setting and data. Section 3 reports the empirical results for the relation between short-selling interest and the SEC investigation. Section 4 reports the amplifying and mitigating channels of the effect from short-selling interest on the investigation risk. Section 5 reports the reexamination of firm quality and the predictive actions of the short sellers. We also extend the study to include the comment letters in this section. Section 6 concludes this paper.

2. Background, Empirical Settings, and Data

In this section, we introduce our empirical methods. First, we briefly describe the SEC's process of enforcement and the duties of different SEC branches. Second, we introduce the main empirical setting which we leverage to establish causality between short-selling interests and the SEC's investigation case selection. Third, we discuss the variable definition and summary statistics.

2.1. Background of the SEC Enforcement

The SEC is an independent agency of the US federal government established based on the Securities Act of 1934 to regulate and oversee the secondary markets. The

Please do not circulate the paper without permission from the authors.

commission has broad powers to enforce the US federal securities law, investigate potential violations, and regulate trading of all non-exempt securities.⁸

The SEC consists of five divisions with a headquarter office in Washington, D.C. and 11 regional offices throughout the US.⁹ This paper focuses on two divisions, the Division of Enforcement and the Division of Corporation Finance. The Division of Enforcement investigates potential violations and litigates the SEC's enforcement actions. The Division of Enforcement does not monitor firms directly in terms of compliance but expects compliance. In contrast, the mission of the Division of Corporation Finance is "both to monitor and to enhance compliance with disclosure and accounting requirements." Specifically, the Division of Corporation Finance undertakes a review of each registrant's filings at least once every three years to monitor and enhance compliance with applicable disclosure and accounting rules. When the Division completes the review, it makes its comment letter(s) and registrant response(s) public on the SEC's EDGAR system.¹⁰

The investigation of Division of Enforcement at the SEC is nonpublic, i.e., no public information is released from the SEC during the investigation process. The Division of Enforcement staff first identifies a suspicious target with potential violation of the securities laws based on tips, complaints, and referrals submitted by the public, self-regulatory organizations, and others. The internal cooperation with cases recommended by the Division of Corporation Finance can also provide leads to the Division of Enforcement (Defond et al. 2018).

The enforcement staff then sends out an information inquiry known as the Matter Under Inquiry (MUI) program, after which the enforcement staff decides whether to initiate a formal investigation against a potential violation. Often, a formal investigation imposes legal responsibility on the target firm to comply with the information request.

⁸ Exempt securities are securities backed by the government that are required to register with the SEC.

⁹ The five main divisions are: the Division of Corporation Finance, the Division of Investment Management, the Division of Enforcement, the Division of Economic Risk Analysis, and the Division of Trading and Markets.

¹⁰ See the SEC's enforcement manual for the Division of Enforcement's mission and the guidelines for SEC investigations: <https://www.sec.gov/divisions/enforce/enforcementmanual.pdf>. Also, see the website of the Division of Corporation Finance for the detailed mission statement: <https://www.sec.gov/divisions/corpfin/cfabout>.

Please do not circulate the paper without permission from the authors.

After serious fact-finding, the Division of Enforcement staff decides whether to recommend the case to the commission for enforcement. If the staff finds solid evidence that warrants charges against the target firm or an individual within the firm, the staff issues a Wells Notice to inform the target firm of its decision to recommend the case to the commission for enforcement authorizations.¹¹ The chairperson and the commissioners of the SEC then vote on whether to authorize the Division of Enforcement to bring the case for an enforcement action. Since the commission almost never rejects any recommended case for enforcement, the initiation of an investigation becomes the most important statistical selection step for the SEC enforcement actions.¹² This paper precisely provides new insights of this selection process with respect to short seller activities.

2.2. Empirical Setting: Pilot Program of Regulation SHO

Our main empirical analyses rely on the regression with ordinary least squares (OLS) estimation and consequently are vulnerable to endogeneity concerns. To mitigate the endogeneity concerns, we situate our analyses in a controlled experiment and test the influence of short-selling interest on the SEC's investigation decisions.

Regulation SHO is a set of rules from the SEC that regulates short sale practices. The rules are designed to avoid price distortion associated with short-selling abuses, such as naked short selling. To assess the price efficiency under the situation without price tests, the SEC conducted a controlled experiment, i.e., the Pilot Program, from May 2, 2005 to August 6, 2007. Specifically, the SEC ranked the firms of the Russell 3000 index, essentially all US stocks, by 12-month average trading volume from June 2003 to May 2004, and selected one stock out of every three stocks into the pilot program that exempts the stocks from price tests in short-selling transactions. Such price tests include the uptick test and the bid price test.¹³ We obtain the pilot stock list from the SEC. Since the SEC did not reveal the nonpilot stocks, we keep the stocks that are included the Russell 3000 index

¹¹ The enforcement staff has the option but not obligation to issue the Wells Notice to inform the target firm of potential enforcement action.

¹² The vote process is symbolic. The approval rate is almost 100% on the recommended cases. See the SEC's public record for the voting results: <https://www.sec.gov/about/commission-votes>.

¹³ See footnote 3 for the details of the price tests.

Please do not circulate the paper without permission from the authors.

in May 2004 and that are not in the pilot stock list as the nonpilot stocks (Bhattacharya et al. 2022; Fang et al. 2016; Li and Zhang 2015).¹⁴

The experiment setting directly lowers the cost to the short selling in the pilot stocks relative to the nonpilot stocks and thus eliminates the need and the challenge to directly estimate the short selling costs in the analysis of short-selling activities (Fang et al. 2016; Lamont 2012; SEC 2007). Because the treatment group and the control group are conditionally randomly selected and the program has a clear time window, the pilot program thus provides technical convenience to the application of difference-in-differences (DID) analysis that can causally quantify the influence of the increase in the short sale and the SEC's selection of investigation cases (Fang et al. 2016; Roberts and Whited 2013).¹⁵

2.3. Data

We have several data components. Our main response variable, i.e., the SEC's nonpublic investigation, is from the SEC through the Freedom of Information Act (FOIA). We first obtain the digitized pictures in pdf format of the investigation records from calendar year 1996 to 2017. Then, we apply optical character recognition (OCR) to read the text from the pictures and fuzzy-match the firm names to find the top 20 candidates to the firm names from Compustat Fundamental database. Finally, we manually select the match for each record. Based the matched records, we create a dummy variable with a value of 1 indicating the investigation, and 0 otherwise. Since we focus on the predictive implication of short sellers' private information, we assume that the firm filings become publicly available three months after a fiscal period end date (Compustat item *datadate*). Based on the assumed date (Compustat item *datadate* plus three months) when the fiscal information becomes public, we estimate another 12-month trailing period to merge in the SEC investigation to the modeling variables. In other words, our main independent variable, i.e., short-selling interest, and the control variables are at least lagged for 3 months relative to the SEC investigation.

¹⁴ To be included in Russell Index of May 2004, a stock has to exist before the month.

¹⁵ For example, Fang et al. (2016) verify the randomness of the treatment assignment. They show that the pilot and nonpilot firms are similar in size, growth, investment, profitability, leverage, and dividend payout.

Please do not circulate the paper without permission from the authors.

We measure short-selling interest using the short-interest data from Compustat database. The short-interest data include the open short positions of stocks with settlements on the last business day on or before 15th of a month. Following Karpoff and Lou (2010), we first identify a firm's shares held short before the settlement date, adjusted for stock splits, in each month. Then we scale this variable by its current month's number of shares outstanding as reported in CRSP database. Moreover, since our empirical tests are performed at the annual level, we aggregate the monthly short interest variable to annual level to generate our main variable, *Short Interest*.

Our control variables are mainly constructed using the Compustat Fundamental database, while the M Score, restatement, shareholder litigation, and comment letter information are from Audit Analytics database. We supplement the shareholder litigation with the class action lawsuits from the Stanford Class Action Clearinghouse (SCAC). We count the analyst coverage using the I/B/E/S database. Our full sample covering the period from 1996 to 2016 include 58,033 firm-year observations. Due to additional limited control variable availability, some tests can have slightly fewer number of observations.

For our DID tests on the pilot program of Regulation SHO, we obtain the pilot list from the SEC's website which includes 986 pilot stocks. We follow the literature and construct the DID analysis data with equal windows for the three periods, i.e., before treatment (fiscal year end in 200204:200406), during treatment (fiscal year end in 200505:200707), and post treatment (fiscal year end in 200708:200910) (Bhattacharya et al. 2022; Fang et al. 2016; Li and Zhang 2015). Each of the three periods include 15 months. Since the SEC announced the first pilot list in July 2004 while the pilot program did not start until May 2005, we follow Fang et al. (2016) and drop the observations between July 2004 and May 2005 to avoid complications due to investors anticipations of the pilot stocks which could bias the DID estimation.¹⁶ Our DID sample identifies 865 pilot

¹⁶ We drop the year between the pilot list announcement and the effective date of the pilot program in our main DID analysis. In our untabulated results, we construct a similar sample following Bhattacharya et al. (2022) and Li and Zhang (2015) to include the year between the pilot list announcement and the effective date of the pilot program, our DID analysis results are consistent.

Please do not circulate the paper without permission from the authors.

stocks and 1,838 nonpilot stocks, totaling 12,093 observations with control variables.¹⁷ In Appendix Table A1, we detail the definition of our variables.

In Table 1, we report the summary stats for both our full sample and the DID analysis sample of the pilot program. Panel A details the summary stats of our main sample covering 1996:2015. During our sample period, the average short-selling interest of a firm-year observation is 4%, while the SEC investigation averages 3.4% of the firm-years. Panel B reports the summary stats of our pilot program sample for the DID analysis. The SEC investigation risk is around 6.3% on average in our DID sample, higher than the overall average of the SEC investigation risk. The reason is that our DID sampler period included several macro corporate events, such as the implementation of Sarbanes-Oxley (SOX) act which substantially increases the SEC's budget and thus investigation cases. The sample also covers the Global Financial Crisis. About one third of the observations are pilot stocks in our DID sample, i.e., the treatment group, while each of the pilot program periods, e.g., *During*, makes up around one-third of the observations.

[Insert Table 1 Here]

3. Short Sellers Are the Vanguard of the SEC Investigation

In this section, we detail our empirical methods and report our findings. We first focus on the entire sample covering from fiscal year 1996 to 2015. We analyze the relation between the short-selling interest and the SEC's selection of investigation target. Then, we leverage the controlled experiment to identify the causality.

3.1. Baseline Results

As the first step to understand the relation between short sellers' activities and the SEC's choice, we regress the dummy of the SEC's choice to short-selling interest

¹⁷ In our untabulated results, we also follow Bhattacharya et al. (2022), Fang et al. (2016), and Li and Zhang (2015) to perform DID analysis excluding financial (SIC 6000-6999) and utility industries (SIC 4900-4950). After excluding financial and utility industries, our sample for the DID analysis includes 684 pilot stocks and 1,434 nonpilot stocks, totaling 9,699 observations. When we exclude the financial and utility industries, the results from our main sample are consistent.

Please do not circulate the paper without permission from the authors.

controlling the common variables and fixed effects that could influence or signal financial misconduct. We estimate the following linear probability regression.

$$\text{SEC Investigation}_{i,t+1} = \beta * \text{Short-Selling Interest}_i + \text{Controls}'_i \Gamma + \varepsilon_{i,t+1}, \quad (1)$$

where subscripts i and t represent the specific firm and year, respectively. $\text{SEC Investigation}_{i,t}$ is an indicator variable equal to 1 if the firm-year is under the SEC's investigation, and 0 otherwise. $\text{Short-Selling Interest}_{i,t-1}$ is our main independent variable that proxies for the short-seller activities. Higher short-selling interest indicates greater market-wide short-selling position in the stock. β is the coefficient on the short-selling interest and the focus of this regression. It stands for the marginal increase in the SEC investigation risk in probability for firm i in time t . A positive and statistically significant β suggests that the short-selling interest are associated with higher SEC investigation probability, and vice versa.

$\text{Controls}'_{i,t-1}$ and Γ represents the vector of control variables and their coefficients. We define the control variables in Appendix Table A1. We consider a wide range of control variables that can signal financial misconduct or influence the SEC's enforcement (e.g., Fang et al. 2016; Holzman et al. 2023; Kedia and Rajgopal 2011). We group them into four different categories. First, we include governance environment variables including the number of analyst (*Analyst Coverage*), firms' numbers of years in Compustat database (*Firm Age*), and an indicator of whether the firm is included in *Fortune 500*. Second, we control for the firm fundamentals, including *Cash Holding*, the ratio between long-term debt and assets (*Leverage*), the ratio between market value and book value (*Market to Book*), annualized return volatility (*Volatility*), and the natural logarithm of market capitalization (*Market Cap*). Third, we control the investigation resource constraints. We the natural log of the great circle distance in miles between firm headquarters and their corresponding SEC regional offices (*SEC Distance*), and we include a lag of the dependent variable (*SEC Investigation*) to control for the potential effect of an earlier investigation on the same matter. Lastly, we include industry and fiscal year fixed effects to account for the unobservable industry and year variations.

3.2. We report the baseline results in Table 2. First, our main independent variable of short-selling interest (*SI*) indicates that the short sellers are indeed the pioneers of the regulatory enforcement. A one-standard-deviation increase in the short-selling interest is related to 0.5% increase in the SEC investigation likelihood, which is a 14% increase relative to the average SEC investigation likelihood. Next, we find that young firms or those with Fortune 500 designation are more likely to be investigated by the SEC. Additionally, SEC appears to be interested in large firms with more analyst coverage or those with high leverage or cash holdings. Pilot Program: A Controlled Experiment

In this section, we aim to confirm the causal relation between the short-selling interest and the SEC's investigation through a randomized experiment. The SEC implemented an experimental test to collect information of market reaction in the absence of the price tests in short sales. The SEC randomly selected every third stocks in the Russel 3000 index ranked by the average trading volume from June 2003 to May 2004 and assigned the selected stocks into the treatment group, i.e., pilot stock group, exempting the stocks from the price tests in short sales. The experiment was initially scheduled to run from May 2, 2005 to August 6, 2007 but was concluded earlier on July 6, 2007. Since the lift of the short price tests, the cost to short sales reduces and the size of the short position increases mechanically for the pilot stocks (Fang et al. 2016). Taking this exogenous variation in the policy, we conduct a double Difference-in-Differences analysis (DID). Specifically, we estimate the following regression.

$$\begin{aligned} \text{SEC Investigation}_{i,t+1} = & \beta_1 \text{During}_t \times \text{Pilot}_i + \beta_2 \text{Post}_t \times \text{Pilot}_i + \\ & \beta_3 \text{During}_t + \beta_4 \text{Post}_t + \beta_5 \text{Pilot}_i + \\ & \text{Controls}'_{i,t} \Gamma + \varepsilon_{i,t+1}, \end{aligned} \tag{2}$$

where *During* is a dummy variable of value 1 indicating the time is during the pilot program and value 0 indicating the time is not during the pilot program, while *Post* is a dummy

Please do not circulate the paper without permission from the authors.

variable of value 1 indicating the time is after the pilot program and value 0 indicating the time is not after the pilot program. Pilot is a dummy variable of value 1 or 0 indicating whether the stock i is a pilot stock or not. β_1 and β_2 are of the most importance. If the short sellers do lead the SEC to initiate more investigations, then we should expect to see that β_1 is positive and significant. Since the pilot program ends in 2007, we should also expect to see that the β_2 is *insignificant*. β_3 and β_4 captures the macro time trend. Their effects will be partially absorbed by the year fixed effects.¹⁸ β_5 is a dummy variable of value 1 and 0 indicating whether there is a direct influence on the SEC investigation from the treatment group or not. Because the random assignment, we should expect β_5 to be insignificant. We include the same control variables as in our baseline results.

Table 3 reports our results from the double DID analysis with our main DID sample excluding financial and utilities industries following Fang et al (2016) and Li and Zhang (2015). Our DID analysis show results consistent with our expectations that β_1 is positive and significant and that β_2 is insignificant. During the pilot program, pilot stocks compared with nonpilot stocks are significantly more likely to be investigated by the SEC. After the pilot program, the difference between the two groups of the stocks revert to zero. Specifically, the pilot stocks face 2.5% more SEC investigations, which is 73.5% more relative to the overall average of the SEC investigation in our sample from 1996 to 2015. Our results are robust across the models from column 1 to column 5 in Table 3, including or excluding control variables and fixed effects. We report the counterpart of our main DID analysis including the financial and utility industries in Appendix Table A2. Our results are robust.

4. Potential Channels

With the causal relation identified, we explore the possible channels that can affect the SEC's dependence on short sellers for initiating investigations. The SEC faces two obstacles in collecting and understanding firm information. First, from the SEC's perspective, it is bounded by its budget and manpower, which limit the federal agency's ability in investigating all potential fraudulent firms (Bonsal et al. 2021; Ege et al. 2019;

¹⁸ Because the time dummies are not exactly the same as the fiscal year fixed effects, the year fixed effect cannot fully absorb the effects from the dummies.

Please do not circulate the paper without permission from the authors.

Holzman et al. 2023; Kedia and Ragopal 2011). Since SEC only has limited resources, when the it faces great backlog, search costs between the SEC and the firms will be amplified. Due to the resource constraint, we hypothesize that the SEC’s reliance on the private sector for information will increase during such stressful time.

We proxy the SEC’s stress by the total number of investigations at the year level. We interact the SEC’s busyness with the short-selling interest in the following linear probability regression.

$$\begin{aligned} \text{SEC Investigation}_{i,t} = & \beta_1 \text{SI}_{i,t-1} \times \text{SEC Busyness}_t + \\ & \beta_2 \text{SI}_{i,t-1} + \beta_3 \text{SEC Busyness}_t + \\ & \text{Controls}'_{i,t-1} \Gamma + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where SEC Busyness_t stands for the number of investigation cases in a fiscal year. We conjecture that higher stress will increase the SEC’s dependence on the short sellers, and thus we expect that β_1 will be positive and significant. We report our analysis of the SEC’s resource constraints in Table 4.

[Insert Table 4 Here]

The results support our hypothesis. During the busy time when the SEC faces resource stress, it depends on the short sellers in initiating investigations. Moreover, our results indicate that the busyness together with the short-selling interest led to increased SEC investigation risk, while the standalone short-selling interest can have negative contribution to SEC investigation during other time which takes about two standard increases in the number of cases to offset the negative effect from the standalone short-selling interest. In general, for a firm with average short-selling interest, i.e., about 3.9%, a one-standard-deviation increase in the SEC’s resource stress measured by the total number of investigation cases, i.e., 85 cases, will lead to a 0.3% increase in the SEC investigation probability. Thus, we conclude that the information demand from the SEC due to its limited resources leads the federal agency to rely on the expert from the private sector for information collection and information processing.

Please do not circulate the paper without permission from the authors.

Second, we consider the firm perspective. Even if the SEC was granted with infinite resources, it would still face information asymmetry because of the complexity in the corporate conduct. The fundamental characteristics can lead to information uncertainty such that the interpretation for firms' accurate public information is challenging. When facing these firms of information uncertainty, the SEC can also rely on the private sector for their expertise. Specifically, we consider two sets of proxies to measure such information uncertainty. Table 5 reports the results from our first set of proxies measuring earnings quality and earnings uncertainty.

[Insert Table 5 Here]

We consider four aspects of the earning's quality and uncertainty. We include analyst's dispersion in earnings forecast (*Analyst Dispersion*), the absolute value of the modified Jones discretionary accrual (*Discretionary Accrual*), the volatility of the return on assets (*ROA Volatility*) and the R&D expenditure over assets (*R&D*). We regress the SEC investigation dummy on these variables with the following linear probability model.

$$\text{SEC Investigation}_{i,t+1} = \beta_1 \text{SI}_{i,t} \times \text{Info. Uncertain.}_{i,t} + \beta_2 \text{SI}_{i,t} + \beta_3 \text{Info. Uncertain.}_{i,t} + \text{Controls}'_{i,t} \Gamma + \varepsilon_{i,t+1}, \quad (4)$$

where we interact the short-selling interest and the information uncertainty variables.

Our results confirm that the earnings' quality and uncertainty can alter the relation between the short sellers and the SEC. Specifically, the SEC's dependence on the short sellers is indeed amplified by the uncertainty of earnings. For example, a firm with average level of short-selling interest and face 0.2% increase in the SEC investigation risk when the discretionary accrual increases for one standard deviation. This 0.2% increase is about 6% of the sample average of the SEC investigation risk. Comparably, together with the short interest, a one-standard-deviation increase in the analyst dispersion in earnings forecast, ROA volatility, or R&D expenditure can lead to an increase in the SEC investigation risk for 0.15%, 0.18%, and 0.2%, respectively. These numbers translate to an increase of 4.5%, 5% and 6%, respectively, relative to the sample average of the SEC investigation risk.

Please do not circulate the paper without permission from the authors.

Our second set of information uncertainty measures utilize firms' market characteristics to measure stock return uncertainty. We conduct the same regression of equation (4) and replace the information uncertainty variables with *Bid-Ask Spread*, *Idiosyncratic Volatility*, *Return Volatility*, and *Turnover*. Table 6 reports our results with the market characteristics as the proxies for the firm level information uncertainty. Our results using the stock return uncertainty to proxy information uncertainty is consistent with our results from the earnings quality and earnings uncertainty. Focusing on the interaction terms, in Table 6, our results suggest that the SEC's reliance on the short sellers increases when the stock return uncertainty is high. For example, a one-standard-deviation increase in idiosyncratic volatility together with an average firm short interest is related to an increase of 3.4% in the SEC investigation risk, which is a 100% increase in the SEC investigation risk relative to the sample average.

Our information channel tests suggest that the SEC is motivated by its resource constraint to condition its investigation decisions on the short sellers' information. Moreover, our information uncertainty tests from the firm perspective confirm that the SEC relies more on the short sellers' activities when the firm information is hard to interpret. Collectively, we conclude that the information asymmetry between the SEC and the firms is potentially a major channel, through which the short sellers cater to the SEC their expert testimony on the corporate conduct and influence the SEC's enforcement decisions.

[Insert Table 6 Here]

We further adopt an alternative logic to test the information asymmetry channel through which the short sellers may influence the SEC's investigation decisions. A potential mitigation of the short sellers' influence on the SEC's case selection can come from information *certainty*. In other words, we explore what would happen to the relation between the SEC and the short sellers, if a firm sends some clear information to the SEC.

To perform this test, we adopt the G index from Gompers et al. (2003). The merit of G index comes from its use of corporate bylaws, which are publicly accessible and easily understandable. The index sums up all the corporate provisions that can reduce the shareholders' power within the governance structure. In other words, the G index measures how explicitly a firm is poorly governed or how entrenched the management is. The firms

Please do not circulate the paper without permission from the authors.

that are obviously bad will have a higher score in G index, and vice versa. We conjecture that if a firm is obviously bad, then the SEC does not really need to depend much on the short sellers, because the information search is nearly costless.

We repeat our regression from equation (4) and replace the information uncertainty variable with G index. The results are reported in Table 8. The results from the G index confirms our hypothesis that clear information can reduce the SEC's reliance on the short sellers for investigation target selection. Focusing on the interaction term, for a firm with average short-selling interest of 3.9%, a one-standard-deviation increase in the G-index, i.e., 2.53, can reduce the influence of short selling interest by 0.2%, which is a 6% reduction relative the sample average of SEC investigation risk.

[Insert Table 7 Here]

5. Robustness Tests

In this section, we report our robustness analyses of the baseline results and the information channel. We also examine the short sellers' ability in discovering specific information related to the firm quality and the litigation.

5.1. Alternative Measure of the SEC's Selection of Target and Information Certainty

Specifically, we adopt SEC's issuance of comment letters related to 10K filings as an alternative measure of the SEC's enforcement. Our main dependent variable in the analysis before this section is SEC investigation. The SEC has five main divisions, and the investigation is an enforcement process overseen by the SEC's Division of Enforcement happening mostly in the SEC's regional offices.

The SEC's Division of Corporation Finance located in the SEC's DC headquarters is mainly tasked with reviewing firms' filings and communicating with the firms for compliance quality. When the SEC Division of Corporation Finance discovers a potential problem in firm filings, it can issue a question letter, i.e., comment letter, to the firm. Although the SEC claims that the comment letter issuance is random and that the agency reviews each firm at least once every three years, the literature has documented the

Please do not circulate the paper without permission from the authors.

determinants that can change the frequency of the SEC's issuance of comment letters (Heese et al. 2017).

The comment letter is thus an ideal alternative measure of the SEC's preference because it reflects the SEC's discretion in selecting target firms and the relative independence between the SEC Division of Enforcement and of Corporation Finance. Specifically, we repeat the baseline regression in equation (1) and replace the dependent variable with a dummy variable indicating whether a firm receives a 10K related comment letter or not.

[Insert Table 8 Here]

We report the results from the robustness analysis of our baseline results in Table 7. The robustness analysis shows that the short selling interest can lead to an increased frequency of comment letter issuance. Specifically, depending on the inclusion of fixed effects, for a one-standard-deviation increase in the short-selling interest, i.e., 6%, the likelihood of receiving comment letters can increase between 0.68% to 1.05%. Comparing to the sample average frequency of 10K-related comment letter issuance of 23% from 2005 to 2015, these numbers translate to relative increase of 3% to 4.6%. Therefore, our findings with comment letters are consistent with our baseline results. In other words, our findings of the SEC's reliance on short sellers in its enforcement is ubiquitous and not restricted to a specific enforcement program of the securities laws.

5.2. Short Sellers and the SEC: A Rational Coordination

The literature documents that the short sellers are informed investors and can reveal original information to the market. For example, Bao et al (2018) show that the short-selling activities predict the bad news withheld by managers. Therefore, the SEC's choice of relying on the short sellers is strategic and can be justified. In this section, we conduct important robustness analysis and present new evidence in line with the literature on short sellers' ability to anticipate financial misrepresentations, an area of common interest for the SEC. (Karpoff and Lou 2010). In particular, we focus on the relation between the short-selling interest and the firms' accounting quality and lawsuits. If short sellers can foresee

Please do not circulate the paper without permission from the authors.

accounting quality problems and stakeholders' disapproval, the SEC's choice of reliance is rationalized.

Therefore, first, we examine the accounting quality in two aspects, i.e., statement quality and restatements. Specifically, we regress the accounting quality variables, e.g., *M Score* and restatement (*Restate*), on the short-selling interest and control variables.

$$\text{Acct Quality}_{i,t+1} = \beta_1 \text{SI}_{i,t} + \text{Controls}'_{i,t} \Gamma + \varepsilon_{i,t+1}, \quad (5)$$

where $\text{Acct Quality}_{i,t+1}$ stands for accounting quality variables, including *M Score* and *Bog Index* for the statement quality and restatement variables. We report the results in Table 9. Panel A reports our regression results on the financial statement quality. Panel B reports the results for restatements by categories. We include five specific restatement types, including income-reducing restatement (*Income*), restatement related to accounting principle (*Accounting*), restatement related to errors (*Error*), restatement related to financial fraud (*Financial*), and restatement related to revenue recognition (*Revenue*).

Our results in Table 9 Panel A show that the short-selling interest is positively related to the *M Score*, a popular measure of earnings management, and negatively related to the *Bog Index*, a popular measure of statement readability and complexity. Consistent with the literature, we find that short sellers can foresee the restatement in the future. Particularly, Panel B demonstrates that short sellers excel in predicting income-reducing restatements, accounting restatements, and error restatements. However, they do not perform as well in predicting restatements related to financial fraud or revenue recognition. This comparison highlights the short sellers' ability in analyzing public information but not specifically financial fraud.

[Insert Table 9 Here]

In addition, we analyze corporate litigation and investigate whether the short sellers can also anticipate the corporate lawsuits. We consider two important litigation variables, i.e., class action lawsuits and shareholder lawsuits¹⁹. We perform the regression in equation

¹⁹ We obtain the class action lawsuit record from Stanford Class Action Clearinghouse (SCAC) and the shareholder litigation record from Audit Analytics.

Please do not circulate the paper without permission from the authors.

(5) and replace the dependent variable with the litigation variables. Table 10 reports our results.

[Insert Table 10 Here]

We show that the short-selling positions demonstrate strong predictive power in anticipating corporate lawsuits. Particularly, a one-standard-deviation increase in short-selling interest can lead to a 0.26% increase in class action lawsuit and 1.1% increase in shareholder lawsuit. These numbers translate to a 11% increase and a 12% increase in the frequencies of the class action lawsuit and the shareholder lawsuit, respectively. In other words, Table 10 confirms the short sellers' ability in discovering firm problems and foreseeing the litigation consequences. Together with our findings from Table 9, the short sellers' ability in discovering errors from public information and foresee future events provide a rational support for the SEC's choice of relying on the short sellers for information.

6. Conclusion

In this paper, we study the possibility that the short sellers can trigger the SEC investigation in secrecy. With our sample from fiscal year 1996 to 2015, we document that the short-selling interest increases the firm-level SEC investigation risk. We then exploit a randomized experiment to identify the causality of such relation. The SEC's Pilot Program of Regulation SHO randomly assigns the treatment status to every third stock in the Russell 3000 index ranked by trading volume, providing a true exogenous shock to the cost and thus the profit of the short-selling investment. Leveraging the randomized experiment, we document that during the pilot program, the pilot stocks that are exempt from the price tests for short sale face 50% increase in the SEC's investigation risk relative to the investigation risk of an average firm prior to the pilot program. The risk of these firms reverts to a level that is indistinguishable from the nonpilot stocks after the pilot program concludes. The randomized experiment confirms the critical role that the short sellers play in the SEC's decision-making for enforcement.

The literature shows that the short sellers are informed investors who can facilitate the detection of financial misconduct (Dechow et al. 1995; Fang et al. 2016; Karpoff and

Please do not circulate the paper without permission from the authors.

Lou 2010). However, we are the first to precisely document the relationship between the short sellers and the regulatory enforcement. Our findings highlight the regulator's dependence on the short sellers for the *initiation* of the enforcement process.

Our findings resonate with the literature and point out the motivating factor of the SEC's reliance on short sellers, i.e., information. The prior studies have identified two important characteristics of the SEC's regulatory enforcement. First, the SEC faces significant resource constraints. Second, the SEC's enforcement process depends on certain signals for violation likelihood of the securities laws, e.g., firm characteristics (Correia 2014; Heese 2019; Holzman et al. 2023; Kedia and Rajgopal 2011;). On the one hand, the federal agency does not have enough employees or budget to discover all the violations and regulate all fraudulent firms. On the other hand, the firms' information uncertainty makes the SEC's scrutiny process even more difficult, since the firms with information uncertainty can be hard to understand. As such, the SEC's dependence on short sellers seems a strategic choice. Indeed, our results show that the SEC's resource constraints and the firm level information uncertainty can both amplify the SEC's dependence on the short sellers in initiating investigation. In a robustness test, we examine the influence of information certainty on the SEC's relationship with short sellers and show that the information certainty indeed reduces the SEC's reliance on short sellers.

Finally, our findings are ubiquitous and not restricted to the SEC's investigation decisions. Taking the SEC comment letter as an alternative measure of the SEC's preference in selecting enforcement target, we find that the SEC also issues significantly more comment letters to the firms with higher short-selling interest. In our robustness test, we find that the short sellers can foresee future accounting quality, restatement, and corporate lawsuits. They are particularly good at discovering accounting-related and error-related restatements, which emphasize their ability in analyzing public information and discover problem. Therefore, the SEC's choice of relying on short sellers for initiating the investigations is rational. Collectively, from the SEC's perspective, this paper demonstrates a complete and clear link that explains the increased detection of misconduct attributable to the short sellers through the SEC's enforcement target selection (Fang et al. 2016). Our

Please do not circulate the paper without permission from the authors.

findings highlight the coordination between the investors and the regulator in the detection of fraud.

Please do not circulate the paper without permission from the authors.

References

- Ali, A, L Hwang, and MA Trombley, 2003, Arbitrage risk and the book-to-market anomaly, *Journal of Financial Economics* 69, 355–373.
- Amihud, Y, and H Mendelson, 1989, The effects of beta, bid-ask spread, residual risk, and size on stock returns, *The Journal of Finance* 44, 479–486.
- Anonymous, 2020, Luckin coffee: fraud + fundamentally broken business.
- Baugh, M, K Kim, and KJ Lee, 2021, The effect of sec reviewers on comment letters, *Contemporary Accounting Research* 39, 656–690.
- Beneish, MD, 1999, The Detection of Earnings Manipulation, *Financial Analysts Journal* 55, 24–36.
- Bhattacharya, N, TE Christensen, Q Liao, and B Ouyang, 2021, Can short sellers constrain aggressive non-GAAP reporting?, *Review of Accounting Studies* 27, 391–440.
- Blackburne, T, JD Kepler, PJ Quinn, and D Taylor, 2020, Undisclosed SEC investigations, *Management Science* 67, 3403–3418.
- Boehmer, E, ZR Huszar, and BD Jordan, 2010, The good news in short interest, *Journal of Financial Economics* 96, 80–97.
- Boehmer, E, CM Jones, and X Zhang, 2008, Which shorts are informed?, *The Journal of Finance* 63, 491–527.
- Bonsall, SB, E Holzman, and BP Miller, 2021, Wearing out the watchdog: the impact of sec case backlog on the formal investigation process, *papers.ssrn.com*.
- Chanos, James, 2003, Comments of James Chanos for the SEC Roundtable on Hedge Funds, *www.sec.gov*.
- Correia, MM, 2014, Political connections and SEC enforcement, *Journal of Accounting and Economics* 57, 241–262.
- Dechow, PM, RG Sloan, and AP Sweeney, 1995, Detecting Earnings Management, *The Accounting Review* 70, 193–225.
- Defond, ML, JR Francis, and NJ Hallman, 2018, Awareness of sec enforcement and auditor reporting decisions, *Contemporary Accounting Research* 35, 277–313.
- deHaan, E, S Kedia, K Koh, and S Rajgopal, 2015, The revolving door and the SEC's enforcement outcomes: Initial evidence from civil litigation, *Journal of Accounting and Economics* 60, 65–96.
- Diamond, DW, and RE Verrecchia, 1987, Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277–311.
- Diether, KB, CJ Malloy, and A Scherbina, 2002, Differences of opinion and the cross section of stock returns, *The Journal of Finance* 57, 2113–2141.
- Do, TPT, and H Zhang, 2022, Styles of Regulators: Evidence from the SEC's Comment Letters, *Contemporary Accounting Research* 39.
- Donelson, DC, M Kubic, and S Toynbee, 2022, The SEC's September spike: regulatory inconsistency with the fiscal year, *SSRN Electronic Journal*.
- Ege, M, JL Glenn, and JR Robinson, 2019, Unexpected SEC resource constraints and comment letter quality, *Contemporary Accounting Research* 37, 33–67.

Please do not circulate the paper without permission from the authors.

Fang, VW, AH Huang, and JM Karpoff, 2016, Short selling and earnings management: a controlled experiment, *The Journal of Finance* 71, 1251–1294.

Francis, J, R LaFond, PM Olsson, and K Schipper, 2004, Costs of equity and earnings attributes, *The Accounting Review* 79, 967–1010.

Gompers, P, J Ishii, and A Metrick, 2003, Corporate governance and equity prices, *The Quarterly Journal of Economics* 118, 107–155.

Heese, J, M Khan, and K Ramanna, 2017, Is the SEC captured? Evidence from comment-letter reviews, *Journal of Accounting and Economics* 64, 98–122.

Holzman, Eric R., Nathan T. Marshall, and Brent A. Schmidt, 2023, When are firms on the hot seat? An analysis of SEC investigation preferences, *Journal of Accounting and Economics*.

Kalmenovitz, Joseph, 2020, Incentivizing financial regulators, *The Review of Financial Studies* 34, 4745–4784.

Karpoff, JM, and X Lou, 2010, Short sellers and financial misconduct, *The Journal of Finance* 65, 1879–1913.

Kedia, S, and S Rajgopal, 2011, Do the SEC’s enforcement preferences affect corporate misconduct?, *Journal of Accounting and Economics* 51, 259–278.

Li, Y, and L Zhang, 2015, Short selling pressure, stock price behavior, and management forecast precision: evidence from a natural experiment, *Journal of Accounting Research* 53, 79–117.

Miller, EM, 1977, Risk, uncertainty, and divergence of opinion, *The Journal of Finance* 32, 1151.

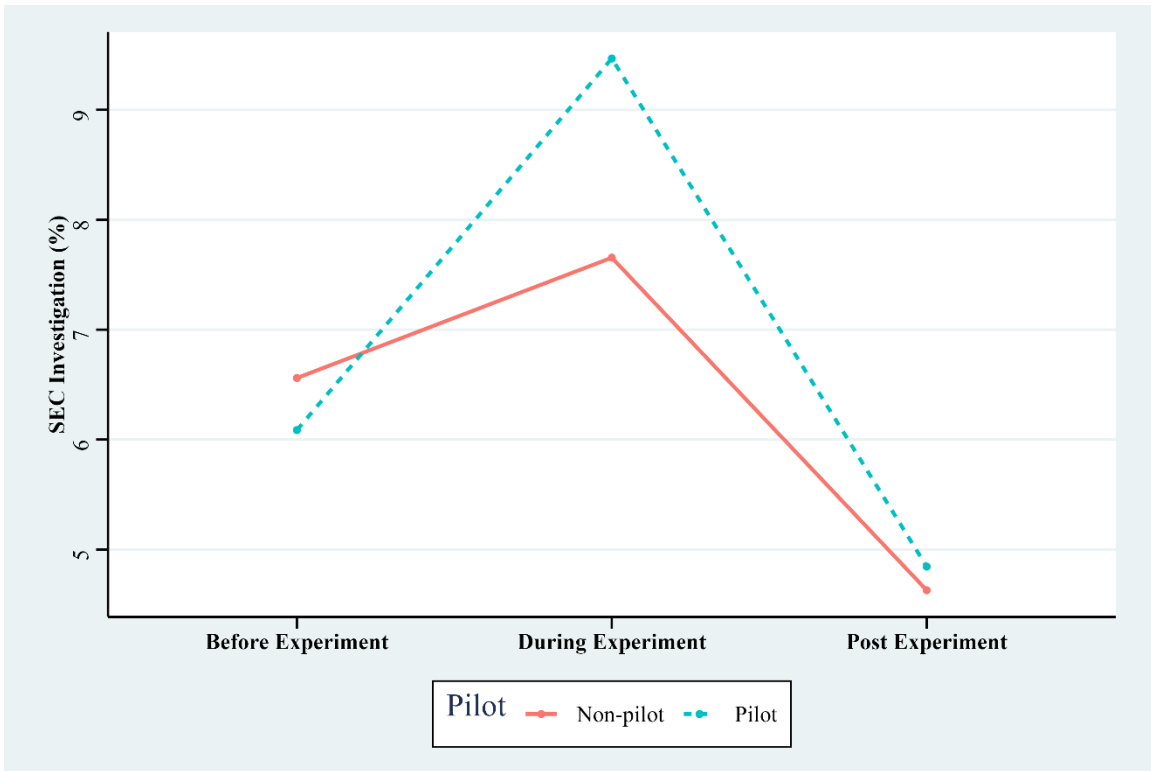
Securities and Exchange Commission (SEC), 2007, Economic analysis of the short sale price restrictions Under the regulation SHO pilot, *sec.gov*.

Securities and Exchange Commission (SEC) v. Luckin Coffee, Inc., No. 1:20-cv-10631 (S.D.N.Y. Dec. 16, 2020).

Solomon, DH, and E Soltes, 2021, Is “not guilty” the same as “innocent”? Evidence from sec financial fraud investigations, *Journal of Empirical Legal Studies* 18, 287–327.

Yang, J, J Chung, and J Steinberg, 2020, Coffee’s for closers: how a short seller’s warning helped take down Luckin Coffee, *Wall Street Journal*.

Figure 1 Investigation Risk Around Pilot Program Experiment



This figure plots the probabilities of the SEC investigation among the pilot stocks and the nonpilot stocks before, during, and after the pilot program. We replicate the figure in Fang et al. (2016) with our data. The difference between the pilot stocks and the nonpilot stocks during the pilot program experiment highlight the causal relation between the short-selling interest and the SEC's selection of investigation targets.

Table 1 Summary Statistics

This table reports the summary statistics of selected variables for the main sample and the difference-in-differences analysis sample. Panel A reports the descriptive statistics for the dependent, test, and control variables used to estimate Equation (1) and variables used in additional analysis, and Panel B reports the descriptive statistics for variables used in the DID analysis. Variable definitions are in Appendix Table A1.

Panel A: Main Sample						
Variable	N	Mean	SD	Q1	Median	Q3
Main Variables						
<i>SEC Investigation</i>	58033	0.034	0.18	0.00	0.000	0.00
<i>Short-Selling Interest</i>	58033	0.039	0.06	0.00	0.018	0.05
<i>Analyst Coverage</i>	58033	1.851	1.10	1.10	2.079	2.71
<i>Firm Age</i>	58033	2.686	0.81	2.08	2.708	3.30
<i>Fortune 500</i>	58033	0.084	0.28	0.00	0.000	0.00
<i>Cash</i>	58033	0.174	0.22	0.02	0.076	0.24
<i>Leverage</i>	58033	0.179	0.19	0.01	0.121	0.30
<i>MtB</i>	58033	3.240	6.06	1.15	1.825	3.16
<i>Mkt Cap</i>	58033	6.097	1.95	4.71	6.121	7.45
<i>SEC Distance</i>	58033	4.433	1.60	3.31	5.049	5.65
<i>SEC Investigation (t-1)</i>	58033	0.029	0.17	0.00	0.000	0.00
Additional Variables						
<i>10K Comment Letter</i>	58033	0.158	0.36	0.00	0.000	0.00
<i>Accounting Restatement</i>	58033	0.087	0.28	0.00	0.000	0.00
<i>Analyst Dispersion</i>	41884	0.157	0.28	0.02	0.050	0.14
<i>Bid-Ask Spread</i>	57803	0.039	0.02	0.02	0.033	0.05
<i>Bog Index</i>	55561	83.926	7.12	79.00	84.000	89.00
<i>Class Action</i>	58033	0.023	0.15	0.00	0.000	0.00
<i>Discretionary Accrual</i>	47562	0.078	0.11	0.02	0.046	0.09
<i>Error Restatement</i>	58033	0.004	0.06	0.00	0.000	0.00
<i>Financial Restatement</i>	58033	0.002	0.04	0.00	0.000	0.00
<i>Governance Index</i>	20203	9.253	2.53	7.00	9.000	11.00
<i>Idiosyncratic Vol</i>	54785	0.058	0.03	0.04	0.050	0.07
<i>Income-reducing Restatement</i>	58033	0.081	0.27	0.00	0.000	0.00
<i>Ln (SEC Busyness)</i>	58033	5.142	0.50	4.84	5.236	5.46
<i>M Score</i>	52514	0.000	0.00	0.00	0.000	0.00
<i>R&D/AT</i>	58029	0.039	0.09	0.00	0.000	0.03
<i>Restatement</i>	58033	0.091	0.29	0.00	0.000	0.00
<i>Return Vol (Daily)</i>	57795	0.028	0.02	0.02	0.024	0.04
<i>Revenue Restatement</i>	58033	0.015	0.120	0.000	0.000	0.000
<i>ROA Vol</i>	52889	0.032	0.06	0.01	0.013	0.03
<i>SEC Busyness</i>	58033	191.864	85.42	126.00	188.000	234.00
<i>Shareholder Lawsuit</i>	49469	0.089	0.28	0.00	0.000	0.00
<i>Turnover (Monthly)</i>	44269	0.700	0.25	0.52	0.757	0.92

Table 1 (Continues)

Panel B: Pilot Program Experiment Sample						
Variable	N	Mean	SD	Q1	Median	Q3
<i>SEC Investigation</i>	12093	0.063	0.24	0.00	0.00	0.00
<i>Pilot</i>	12093	0.324	0.47	0.00	0.00	1.00
<i>During</i>	12093	0.335	0.47	0.00	0.00	1.00
<i>Post</i>	12093	0.292	0.45	0.00	0.00	1.00
<i>Analyst Coverage</i>	12093	2.354	0.84	1.95	2.48	2.94
<i>Firm Age</i>	12093	2.850	0.70	2.40	2.77	3.47
<i>Fortune 500</i>	12093	0.118	0.32	0.00	0.00	0.00
<i>Cash</i>	12093	0.180	0.21	0.03	0.09	0.26
<i>Leverage</i>	12093	0.167	0.17	0.01	0.12	0.27
<i>MtB</i>	12093	2.782	2.75	1.31	1.98	3.18
<i>Mkt Cap</i>	12093	6.952	1.51	5.92	6.81	7.86
<i>SEC Distance</i>	12093	4.487	1.61	3.32	5.15	5.73
<i>SEC Investigation (t-1)</i>	12093	0.054	0.23	0.00	0.00	0.00

Table 2 Baseline Results

This table reports the baseline results from the linear probability model. The dependent variable is a dummy variable that takes the value 1 if a firm is under SEC investigation, 0 otherwise. The coefficients are directly interpretable as marginal percentage influence on the likelihood of an SEC investigation. Robust standard errors clustered at the firm level are reported in parentheses. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>SEC Investigation</i>			
	(1)	(2)	(3)	(4)
<i>Short-Selling Interest</i>	0.101*** (0.019)	0.096*** (0.019)	0.086*** (0.020)	0.082*** (0.020)
<i>Analyst Coverage</i>	0.007*** (0.001)	0.005*** (0.001)	0.009*** (0.001)	0.007*** (0.001)
<i>Firm Age</i>	-0.003*** (0.001)	-0.002** (0.001)	-0.002* (0.001)	-0.001 (0.001)
<i>Fortune 500</i>	0.016*** (0.004)	0.018*** (0.004)	0.014*** (0.004)	0.015*** (0.004)
<i>Cash</i>	0.028*** (0.004)	0.013*** (0.005)	0.027*** (0.004)	0.012*** (0.005)
<i>Leverage</i>	-0.009* (0.005)	-0.004 (0.005)	-0.003 (0.005)	0.001 (0.005)
<i>MtB</i>	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
<i>Market Cap</i>	0.004*** (0.001)	0.005*** (0.001)	0.004*** (0.001)	0.005*** (0.001)
<i>SEC Distance</i>	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
<i>SEC Investigation (t-1)</i>	0.021*** (0.006)	0.019*** (0.006)	0.014** (0.006)	0.012* (0.007)
<i>Constant</i>	-0.008* (0.005)	-0.009* (0.005)	-0.014*** (0.005)	-0.015*** (0.005)
N	58,033	58,033	58,033	58,033
Adj. R2	0.013	0.014	0.021	0.022
Industry	N	Y	N	Y
Year	N	N	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 3 Evidence from Controlled Experiment: An Identification Strategy

This table reports the results from our double difference-in-differences (DID) analysis based on the SEC's pilot program of Reg SHO. The sample include firms that exist in the Russell 3000 index with control variables. Our DID sample include three periods, i.e., before the pilot program, during the pilot program (200505:200707), and post the pilot program (200708:200910). The treatment group include the randomly selected pilot stocks. From column (1) to column (5), we fit the double DID regression in equation (2) and alternate the combination of control variables and fixed effects. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>SEC Investigation</i>				
	(1)	(2)	(3)	(4)	(5)
<i>Pilot X During</i>	0.026*** (0.010)	0.026** (0.011)	0.026** (0.011)	0.025** (0.011)	0.025** (0.011)
<i>Pilot X Post</i>	0.011 (0.010)	0.010 (0.011)	0.009 (0.011)	0.009 (0.011)	0.009 (0.011)
<i>Pilot</i>	-0.010 (0.007)	-0.006 (0.008)	-0.007 (0.008)	-0.006 (0.008)	-0.007 (0.008)
<i>During</i>	0.001 (0.006)	-0.006 (0.007)	-0.006 (0.007)	-0.024 (0.059)	-0.027 (0.059)
<i>Post</i>	-0.017*** (0.006)	-0.020*** (0.006)	-0.021*** (0.006)	0.004 (0.062)	0.002 (0.062)
<i>Analyst Coverage</i>		0.016*** (0.003)	0.014*** (0.003)	0.016*** (0.003)	0.014*** (0.003)
<i>Firm Age</i>		-0.005 (0.004)	-0.002 (0.004)	-0.004 (0.004)	-0.002 (0.004)
<i>Fortune 500</i>		0.005 (0.009)	0.008 (0.009)	0.005 (0.009)	0.008 (0.009)
<i>Cash</i>		0.070*** (0.013)	0.041*** (0.015)	0.069*** (0.013)	0.040*** (0.015)
<i>Leverage</i>		0.009 (0.014)	0.018 (0.015)	0.009 (0.014)	0.018 (0.015)
<i>MtB</i>		0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)
<i>Market Cap</i>		0.014*** (0.003)	0.014*** (0.003)	0.013*** (0.003)	0.014*** (0.003)
<i>SEC Distance</i>		0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
<i>SEC Investigation (t-1)</i>		0.013 (0.012)	0.010 (0.012)	0.014 (0.012)	0.011 (0.012)
<i>Constant</i>	0.070*** (0.004)	-0.069*** (0.018)	-0.071*** (0.019)	-0.071* (0.041)	-0.072* (0.041)
N	15,065	12,093	12,093	12,093	12,093
Adj. R2	0.001	0.020	0.022	0.022	0.025
Industry	N	N	Y	N	Y
Year	N	N	N	Y	Y
Error Cluster	Firm	Firm	Firm	Firm	Firm

Please do not circulate the paper without permission from the authors.

Table 4 The SEC's Reliance on Short Sellers with Respect to Resources Constraints

This table reports the results of linear probability model with interaction terms between the short-selling interest and the SEC's busyness. We proxy the SEC's backlog with the fiscal year total number of investigation cases (Bansall et al. 2021). We include results with the plain count (*SEC Busyness*) and the natural logarithm scale of the plain count (*SEC Busyness*). Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	SEC Investigation			
	(1)	(2)	(3)	(4)
<i>Short-Selling Interest X SEC Busyness</i>	0.001*** (0.000)	0.001*** (0.000)		
<i>SEC Busyness</i>	0.000*** (0.000)	0.000*** (0.000)		
<i>Short-Selling Interest X Ln (SEC Busyness)</i>			0.152*** (0.032)	0.158*** (0.033)
<i>Ln (SEC Busyness)</i>			0.027*** (0.002)	0.026*** (0.002)
<i>Short-Selling Interest X Analyst Coverage</i>	-0.120*** (0.040)	-0.130*** (0.041)	-0.709*** (0.164)	-0.743*** (0.165)
<i>Firm Age</i>	0.009*** (0.001)	0.007*** (0.001)	0.008*** (0.001)	0.007*** (0.001)
<i>Fortune 500</i>	-0.002* (0.001)	-0.001 (0.001)	-0.002* (0.001)	-0.002 (0.001)
<i>Cash</i>	0.014*** (0.004)	0.016*** (0.004)	0.014*** (0.004)	0.016*** (0.004)
<i>Leverage</i>	0.026*** (0.004)	0.012** (0.005)	0.027*** (0.004)	0.012*** (0.005)
<i>MtB</i>	-0.003 (0.005)	0.002 (0.005)	-0.003 (0.005)	0.002 (0.005)
<i>Market Cap</i>	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
<i>SEC Distance</i>	0.004*** (0.001)	0.005*** (0.001)	0.004*** (0.001)	0.005*** (0.001)
<i>SEC Investigation (t-1)</i>	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
<i>Constant</i>	0.012* (0.006)	0.011* (0.007)	0.013** (0.006)	0.012* (0.006)
	-0.044*** (0.005)	-0.045*** (0.005)	-0.151*** (0.010)	-0.149*** (0.010)
Observations	58,033	58,033	58,033	58,033
Adjusted R-squared	0.022	0.023	0.021	0.022
Industry	Y	Y	Y	Y
Year	Y	Y	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 5 Channel Analysis I: Information Asymmetry Around Earnings

This table reports the results of linear probability model with interaction terms between the short-selling interest and the proxy variables of earnings quality and uncertainty. We proxy the earnings quality and uncertainty with four variables: analyst dispersion in earnings forecast (*Analyst Dispersion*), absolute value of modified Jones *Discretionary Accrual*, volatility of return on assets (*ROA Volatility*), and R&D expenditure to assets (*R&D*). Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>SEC Investigation</i>			
	(1)	(2)	(3)	(4)
<i>SI X Analyst Dispersion</i>	0.139*			
	(0.082)			
<i>Analyst Dispersion</i>	0.008*			
	(0.005)			
<i>SI X Discretionary Accrual</i>		0.487**		
		(0.212)		
<i>Discretionary Accrual</i>		0.007		
		(0.009)		
<i>SI X ROA Volatility</i>			0.765**	
			(0.354)	
<i>ROA Volatility</i>			0.014	
			(0.019)	
<i>SI X R&D</i>				0.587***
				(0.218)
<i>R&D</i>				-0.014
				(0.013)
<i>SI</i>	0.052**	0.041*	0.055**	0.057***
	(0.023)	(0.024)	(0.022)	(0.020)
<i>Analyst Coverage</i>	0.013***	0.007***	0.008***	0.007***
	(0.002)	(0.001)	(0.001)	(0.001)
<i>Firm Age</i>	-0.003*	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
<i>Fortune 500</i>	0.011**	0.010**	0.014***	0.016***
	(0.005)	(0.005)	(0.004)	(0.004)
<i>Cash</i>	0.015**	0.007	0.007	0.011**
	(0.006)	(0.005)	(0.005)	(0.005)
<i>Leverage</i>	0.001	0.005	0.002	0.001
	(0.006)	(0.006)	(0.005)	(0.005)
<i>MtB</i>	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
<i>Market Cap</i>	0.007***	0.006***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SEC Distance</i>	0.001	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SEC Investigation (t-1)</i>	0.009	0.006	0.012*	0.012*
	(0.007)	(0.006)	(0.007)	(0.006)
<i>Constant</i>	-0.042***	-0.020***	-0.019***	-0.014***
	(0.007)	(0.006)	(0.006)	(0.005)
N	41,884	47,562	52,889	58,029
Adj. R2	0.024	0.023	0.023	0.022
Industry	Y	Y	Y	Y
Year	Y	Y	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 6 Channel Analysis II: Information Asymmetry Around Stock Returns

This table reports the results of linear probability model with interaction between the short-selling interest and the proxy variables of stock return uncertainty. We proxy the stock return uncertainty with four variables: annual average of *Bid-Ask Spread* by Amihud and Mendelson (1989), annual average of idiosyncratic volatility by Ali Hwang and Trombley (2003) (*Idiosyncratic Volatility*), daily *Return Volatility*, and annualized daily stock *Turnover*. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>SEC Investigation</i>			
	(1)	(2)	(3)	(4)
<i>SI X Bid-Ask Spread</i>	1.601*			
	(0.898)			
<i>Bid-Ask Spread</i>	0.237***			
	(0.048)			
<i>SI X Idiosyncratic Volatility</i>		2.892***		
		(0.714)		
<i>Idiosyncratic Volatility</i>		0.177***		
		(0.037)		
<i>SI X Return Volatility</i>			2.991**	
			(1.300)	
<i>Return Volatility</i>			0.491***	
			(0.072)	
<i>SI X Turnover</i>				0.376***
				(0.146)
<i>Turnover</i>				0.027***
				(0.006)
<i>SI</i>	0.016	-0.100**	-0.011	-0.304**
	(0.042)	(0.045)	(0.043)	(0.132)
<i>Analyst Coverage</i>	0.006***	0.007***	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.002)
<i>Firm Age</i>	-0.000	-0.000	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
<i>Fortune 500</i>	0.014***	0.014***	0.013***	0.012***
	(0.004)	(0.004)	(0.004)	(0.005)
<i>Cash</i>	0.008	0.004	0.007	0.011*
	(0.005)	(0.005)	(0.005)	(0.006)
<i>Leverage</i>	-0.000	-0.000	-0.000	0.006
	(0.005)	(0.005)	(0.005)	(0.006)
<i>MtB</i>	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
<i>Market Cap</i>	0.007***	0.007***	0.008***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SEC Distance</i>	0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SEC Investigation (t-1)</i>	0.011*	0.010	0.010	0.007
	(0.006)	(0.007)	(0.006)	(0.007)
<i>Constant</i>	-0.035***	-0.040***	-0.044***	-0.038***
	(0.006)	(0.007)	(0.006)	(0.006)
N	57,803	54,785	57,795	44,269
Adj. R2	0.023	0.025	0.024	0.024
Industry	Y	Y	Y	Y
Year	Y	Y	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 7 Channel Analysis III: Information Certainty with Disclosed Governance Condition

This table reports the robustness analysis results for the information channel that links the SEC's initiation of investigation and the short-selling interest. Instead of measuring information asymmetry between the firms and the SEC, we proxy the information certainty with G index, which sums up corporate provisions that harm the shareholders' interest that is publicly accessible and easy to understand. Firms with high G index signal publicly its governance problems with no uncertainty. We report the results from the linear probability model specified in equation (4). Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>SEC Investigation</i>			
	(1)	(2)	(3)	(4)
<i>SI X Governance Index</i>	-0.021*	-0.022**	-0.022**	-0.023**
	(0.011)	(0.011)	(0.011)	(0.011)
<i>Governance Index</i>	0.000	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SI</i>	0.252**	0.250**	0.239**	0.238**
	(0.107)	(0.107)	(0.107)	(0.108)
<i>Analyst Coverage</i>	0.008***	0.006**	0.012***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
<i>Firm Age</i>	-0.007**	-0.005	0.000	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
<i>Fortune 500</i>	0.011**	0.013**	0.009*	0.010*
	(0.005)	(0.005)	(0.005)	(0.005)
<i>Cash</i>	0.068***	0.052***	0.060***	0.044***
	(0.013)	(0.014)	(0.013)	(0.014)
<i>Leverage</i>	-0.015	-0.009	-0.011	-0.004
	(0.010)	(0.011)	(0.010)	(0.011)
<i>MtB</i>	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
<i>Market Cap</i>	0.009***	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
<i>SEC Distance</i>	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
<i>SEC Investigation (t-1)</i>	0.014	0.013	0.005	0.004
	(0.010)	(0.010)	(0.010)	(0.010)
<i>Constant</i>	-0.031**	-0.038***	-0.052***	-0.062***
	(0.014)	(0.015)	(0.015)	(0.016)
N	20,203	20,203	20,203	20,203
Adj. R2	0.012	0.013	0.022	0.023
Industry	N	Y	N	Y
Year	N	N	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 8 Robustness Test: The SEC's Selection of Enforcement Target

This table reports our robustness analysis results using an alternative measure to proxy the SEC's preference in selecting enforcement target. Instead of the SEC investigation, we adopt the 10K-related comment letter issuance as our dependent variable. We repeat our linear probability regression in equation (1) with the dummy variable of value 1 or 0 indicating whether a firm receives a 10K-related comment letter in the 12-month trailing period 3 months after fiscal period end. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

	<i>10K Comment Letter</i>			
	(1)	(2)	(3)	(4)
<i>SI</i>	0.176*** (0.043)	0.161*** (0.043)	0.131*** (0.043)	0.113*** (0.044)
<i>Analyst Coverage</i>	0.020*** (0.003)	0.019*** (0.003)	0.008*** (0.003)	0.006* (0.003)
<i>Firm Age</i>	0.018*** (0.003)	0.015*** (0.003)	0.013*** (0.003)	0.009*** (0.003)
<i>Fortune 500</i>	0.011 (0.012)	0.013 (0.013)	0.001 (0.012)	0.003 (0.012)
<i>Cash</i>	-0.019** (0.010)	-0.019* (0.011)	-0.012 (0.010)	-0.018 (0.011)
<i>Leverage</i>	0.086*** (0.014)	0.094*** (0.014)	0.085*** (0.014)	0.094*** (0.014)
<i>MtB</i>	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
<i>Market Cap</i>	0.027*** (0.002)	0.027*** (0.002)	0.036*** (0.002)	0.036*** (0.002)
<i>SEC Distance</i>	-0.005*** (0.001)	-0.006*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
<i>SEC Investigation (t-1)</i>	0.023* (0.013)	0.022 (0.013)	0.029** (0.013)	0.028** (0.013)
<i>Constant</i>	-0.016 (0.013)	-0.003 (0.013)	-0.034*** (0.013)	-0.019 (0.013)
N	37,401	37,401	37,401	37,401
Adj. R2	0.042	0.044	0.059	0.060
Industry	N	Y	N	Y
Year	N	N	Y	Y
Error Cluster	Firm	Firm	Firm	Firm

Table 9 Robustness Test: Short Sellers' Ability in Forecasting Accounting Quality

This table reports our robustness analysis results of the short seller's ability in discovering firm quality. We include on two sets of dependent variables. In Panel A, we report the results for the one-year-ahead statement quality proxied with M Score and Bog Index. In Panel B, we report the results for the restatement outcomes in the trailing 12-month period 3 months after the fiscal period end. The statement quality variables are continuous variables, and the restatement variables are dummy variables taking a value of 1 if the restatement happens and 0 if the restatement does not happen. We perform linear probability regression defined in equation (5) and report the results for the statement quality in Panel A and the results for the restatements in Panel B. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Panel A: Statement Quality		
Dependent Variable	(1)	(2)
	<i>M Score</i>	<i>Bog Index</i>
<i>SI</i>	0.045*** (0.013)	-3.921*** (1.133)
<i>Discretionary Accrual</i>	0.003 (0.005)	1.186*** (0.330)
<i>Analyst Coverage</i>	0.002** (0.001)	0.482*** (0.106)
<i>Firm Age</i>	-0.001 (0.001)	-0.536*** (0.101)
<i>Fortune 500</i>	0.003 (0.003)	0.334 (0.334)
<i>Cash</i>	-0.002 (0.003)	2.977*** (0.372)
<i>Leverage</i>	0.024*** (0.004)	3.470*** (0.406)
<i>MtB</i>	0.000* (0.000)	-0.047*** (0.011)
<i>Market Cap</i>	-0.001* (0.001)	0.232*** (0.061)
<i>SEC Distance</i>	0.000 (0.000)	-0.152*** (0.051)
<i>SEC Investigation (t-1)</i>	0.003 (0.003)	0.327** (0.164)
<i>Constant</i>	0.007** (0.003)	83.872*** (0.466)
N	41,403	39,703
Adj. R2	0.039	0.335
Industry	Y	Y
Year	Y	Y
Error Cluster	Firm	Firm

Table 9 (Continues)

Dependent Variable	Panel B: Restatement					
	(1) Restatement	(2) Income	(3) Accounting	(4) Error	(5) Financial	(6) Revenue
<i>SI</i>	0.102*** (0.029)	0.093*** (0.028)	0.092*** (0.028)	0.018*** (0.007)	-0.004 (0.002)	0.002 (0.011)
<i>Analyst Coverage</i>	0.004* (0.002)	0.002 (0.002)	0.003* (0.002)	0.000 (0.000)	0.000 (0.000)	0.001 (0.001)
<i>Firm Age</i>	-0.002 (0.002)	-0.002 (0.002)	-0.001 (0.002)	-0.000 (0.000)	-0.000 (0.000)	-0.002*** (0.001)
<i>Fortune 500</i>	-0.008 (0.006)	-0.008 (0.006)	-0.007 (0.006)	-0.001 (0.001)	0.001 (0.001)	-0.004* (0.002)
<i>Cash</i>	-0.024*** (0.008)	-0.021*** (0.008)	-0.021*** (0.008)	-0.001 (0.002)	-0.003*** (0.001)	-0.009*** (0.003)
<i>Leverage</i>	0.048*** (0.009)	0.040*** (0.008)	0.047*** (0.009)	-0.000 (0.001)	-0.000 (0.001)	0.003 (0.003)
<i>MtB</i>	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000* (0.000)
<i>Market Cap</i>	-0.001 (0.001)	0.001 (0.001)	-0.001 (0.001)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
<i>SEC Distance</i>	-0.002*** (0.001)	-0.002*** (0.001)	-0.002** (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.001** (0.000)
<i>SEC Investigation (t-1)</i>	0.018** (0.008)	0.016** (0.008)	0.019** (0.008)	0.000 (0.002)	0.001 (0.001)	0.011*** (0.004)
<i>Constant</i>	0.101*** (0.009)	0.078*** (0.008)	0.092*** (0.008)	0.006*** (0.002)	0.004*** (0.001)	0.022*** (0.003)
N	58,033	58,033	58,033	58,033	58,033	58,033
Adj. R2	0.020	0.022	0.020	0.002	0.000	0.005
Industry	Y	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y	Y
Error Cluster	Firm	Firm	Firm	Firm	Firm	Firm

Please do not circulate the paper without permission from the authors.

Table 10 Robustness Test: Short Sellers' Ability in Forecasting Corporate Litigation

This table reports our results for short sellers' ability in anticipating corporate litigation. We include two dependent variables to proxy the corporate lawsuits, i.e., class action lawsuits and shareholder lawsuits. The dependent variable are dummy variables taking a value of 1 if the lawsuit happens and a value of 0 if the lawsuit does not happen. Variable definitions are in the Appendix Table A1. ***, **, and * denote significance at the 1%, 5%, and 10% level (two-tailed), respectively.

Dependent Variable	(1)	(2)
	<i>Class Action</i>	<i>Shareholder</i>
<i>SI</i>	0.044*** (0.016)	0.183*** (0.051)
<i>Analyst Coverage</i>	0.007*** (0.001)	0.039*** (0.003)
<i>Firm Age</i>	-0.007*** (0.001)	-0.011*** (0.003)
<i>Fortune 500</i>	0.007* (0.004)	0.057*** (0.014)
<i>Cash</i>	0.005 (0.004)	0.068*** (0.015)
<i>Leverage</i>	-0.006 (0.004)	-0.001 (0.013)
<i>MtB</i>	0.000*** (0.000)	-0.000 (0.000)
<i>Market Cap</i>	0.004*** (0.001)	0.001 (0.002)
<i>SEC Distance</i>	0.000 (0.000)	0.000 (0.002)
<i>SEC Investigation (t-1)</i>	0.005 (0.005)	0.210*** (0.013)
<i>Constant</i>	0.001 (0.004)	0.009 (0.015)
N	50,704	47,802
Adj. R2	0.014	0.075
Industry	Y	Y
Year	Y	Y
Error Cluster	Firm	Firm

Appendix

Table A1 Variable Definition

The table below describes the main variable definitions we use in this paper, including their possible values and calculations.

Variables	Description	Variable Value and Calculation
<i>10K Comment Letter</i>	This variable is the SEC comment letter record from Audit Analytics. The SEC publicizes the comment letter records since 2005.	1 if the firm-year is associated with a comment letter related to 10K form, 0 otherwise.
<i>Accounting Restatement</i>	This variable is a dummy of restatements related to accounting principles.	1 if the firm-year is associated with an accounting restatement in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Analyst Coverage</i>	This variable is the number of distinct analysts who follow the firm from the I/B/E/S database.	Log of the number of equity research analysts covering the firm.
<i>Analyst Dispersion</i>	This variable is the analyst dispersion in the earnings forecast.	Standard deviation of analyst forecasts in month prior to fiscal period end data divided by the absolute value of the mean forecast (Diether et al. 2002). Then, the outcome is averaged to produce the fiscal year value.
<i>Bid-Ask Spread</i>	This variable is the bid-ask spread proxy at the annual level.	The annual average of the bid-ask spread defined at the monthly level (Amihud and Mendelson 1987s).
<i>Bog Index</i>	This variable is the Bog Index of statement readability.	The bog index is from Bonsall et al. (2017). See https://host.kelley.iu.edu/bpm/activities/bogindex.html .
<i>Cash Holdings</i>	This variable is the cash holding amount estimated with Compustat database.	Ratio between cash holding (Compustat item che) and total assets (Compustat item at).
<i>Class Action Lawsuit</i>	This variable is a dummy of class action lawsuit.	1 if the firm-year is associated with a class action lawsuit in the Stanford Class Action Clearinghouse (SCAC) in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Discretionary Accrual</i>	This variable is the discretionary accruals estimated using the modified Jones method (see Dechow et al. 1995).	The calculation is detailed in Table 2A.
<i>During</i>	This variable is a dummy of time.	1 if the year is during the implementation of Reg SHO, and 0 otherwise.
<i>Error Restatement</i>	This variable is a dummy of restatements due to errors.	1 if the firm-year is associated with an error restatement in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Financial Restatement</i>	This variable is a dummy of restatements due to financial fraud.	1 if the firm-year is associated with an restatement due to financial fraud in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Firm Age</i>	This variable is the number of years since a firm appeared in the Compustat database.	Log of firm age in years.
<i>Fortune 500</i>	This variable is a dummy of whether a firm is a Fortune 500 firm from the Compustat database.	1 if the firm is a Fortune 500 company, 0 otherwise.
<i>Governance Index</i>	This variable is the governance index (G index).	Summation of corporate governance provisions for individual firms in corporate takeover defenses from IRCC database (See Gompers et al. 2003).

Please do not circulate the paper without permission from the authors.

Table TA1 (Continues)

Variables	Description	Variable Value and Calculation
<i>Idiosyncratic Volatility</i>	This variable is the annual average of the idiosyncratic volatility (see Ali et al. 2003).	Annual average of standard deviation of residuals of weekly returns on weekly equal weighted market returns for 3 years prior to month end.
<i>Income-reducing Restatement</i>	This variable is a dummy of restatements with negative changes on earnings from Audit Analytics.	1 if the firm-year is associated with an income-reducing restatement in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Leverage</i>	This variable is the leverage level estimated with Compustat database.	Ratio of long-term debt (Compustat item ltd) to total assets (Compustat item at).
<i>Market Cap</i>	This variable is the market capitalization.	The market capitalization in the last month of the fiscal year.
<i>MtB</i>	This variable is the ratio between market capitalization and book equity.	The ratio is defined as the market capitalization divided by the book equity. We define book equity as the summation of shareholders' equity (Compustat item seq) and deferred taxes (Compustat item txdb) net the preferred stock (calculated with Compustat item pstkrv, pstkl, and pstk).
<i>M Score</i>	This variable is the measure of potential earnings management (see Beneish 1999).	This variable is obtained from Audit Analytics database.
<i>Pilot</i>	This variable is a dummy indicating pilot stock.	1 if the firm is included in the pilot program, 0 otherwise.
<i>Post</i>	This variable is a dummy indicating the time period after the regulation SHO pilot program experiment.	1 if the fiscal year end is included in the 15-month during period of SHO experiment.
<i>R&D</i>	This variable is the ratio between research and development expenditure and assets.	Ratio between research and development expenditure and assets.
<i>Return Volatility</i>	This variable measures the daily return volatility.	Standard deviation of the daily return for the fiscal year.
<i>Restatement</i>	This variable is a dummy of restatements.	1 if the firm-year is associated with an restatement in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>Revenue Restatement</i>	This variable is a dummy of restatements due to revenue recognition.	1 if the firm-year is associated with an restatement because of revenue recognition in the 12-month trailing period 3 months after the fiscal year end, 0 otherwise.
<i>ROA Volatility</i>	This variable is the volatility of return on assets (see Francis et al.).	The variable is first defined as the standard deviation for 16 quarters of income before extraordinary items (ibq) dividend by total assets (atq) and then the outcome is averaged to create the fiscal year variable.
<i>SEC Busyness</i>	This variable is the measure of SEC investigation case backlog.	The variable is defined as the total number of investigation cases in the fiscal year. A natural logarithm version of the variable is also included in the analysis.

Please do not circulate the paper without permission from the authors.

Table TA1 (Continues)

Variables	Description	Variable Value and Calculation
<i>SEC Distance</i>	This variable is the distance between the firm headquarter in Compustat database and the SEC office of the jurisdiction.	<p>The distance between a firm headquarter and the SEC regional office is defined as the great-circle distance using latitudes and longitudes based on the 5-digit zip code information. Specifically, the distance is calculated as</p> $3949.99 \times \arccos \left[\begin{array}{l} \sin \left(\frac{\arctan(1)}{45} \times \text{firm latitude} \right) \times \\ \sin \left(\frac{\arctan(1)}{45} \times \text{SEC latitude} \right) \\ + \\ \cos \left(\frac{\arctan(1)}{45} \times \text{firm latitude} \right) \times \\ \cos \left(\frac{\arctan(1)}{45} \times \text{SEC latitude} \right) \times \\ \cos \left(\frac{\arctan(1)}{45} \times \text{firm longitude} - \right. \\ \left. \frac{\arctan(1)}{45} \times \text{SEC longitude} \right) \end{array} \right]$ <p>and the log scale is applied to the distance calculation.</p>
<i>SEC Investigation</i>	This variable is the SEC undisclosed investigation record obtained through Freedom of Information Act (FOIA) request.	1 if the firm-year is under the SEC's undisclosed investigation, 0 otherwise.
<i>Short-selling Interest</i>	This variable is the short selling positions in the number of shares scaled by the total number of outstanding shares.	Ratio between the number of short-selling shares and the total number of outstanding shares.
<i>Shareholder Litigation</i>	This variable is a dummy variable of shareholder litigation from Audit Analytics database.	1 if the firm-year is associated with a shareholder lawsuit, 0 otherwise.
<i>Turnover</i>	This variable is the total annual turnover.	This variable is calculated as the fiscal year annual cumulative turnover (CRSP item vol) in percentage of the total outstanding shares (CRSP item shrou).

Please do not circulate the paper without permission from the authors.

Table 2A Discretionary Accrual Calculation

This table details the calculation of discretionary accrual in this paper using the Compustat database following the Jones model as modified by Dechow et al. (1995).

Step 1 Data Preparation	The discretionary accrual is calculated following the modified Jones model. To obtain robust estimation of the average two-digit SIC code-level discretionary accrual, three types of firms are excluded from the estimation process: Firms with total assets and lag 1 total asset smaller than 1 million USD, Firms associated with a 2-digit SIC code of less than ten fiscal year observations in the Compustat database, and Firms with missing values in total assets, lag 1 total assets, sales, lag 1 sales, income, operating net cash flow, and plant, property, and equipment.
Step 2 Raw Input Preparation	$\%Total\ Accrual_t = \frac{Total\ Income_t - Operating\ Net\ Cash\ Flow_t}{Total\ Assets_{t-1}}$ $Revenue\ Growth_t = Sales_t - Sales_{t-1}$ $Receivable\ Change_t = Total\ Receivable_t - Total\ Receivable_{t-1}$ $\%Sales_t = \frac{Revenue\ Growth_t - Receivable\ Change_t}{Total\ Assets_{t-1}}$ $\%Property\ Plant\ and\ Equipment_t = \frac{Property\ Plant\ and\ Equipment_t}{Total\ Asset_{t-1}}$
Step 3 Missing Value Substitution in %Total Accruals	If total accrual calculation in step 2 is not viable, $\%Total\ Accrual_t = [(Total\ Current\ Assets_t - Total\ Current\ Assets_{t-1}) - (Total\ Current\ Liability_t - Total\ Current\ Liability_{t-1}) + (Debt\ in\ Current\ Liability_t - Debt\ in\ Current\ Liability_{t-1}) - Depreciation_t] / Total\ Assets_t$
Step 4 Winsorization	To ensure the robustness of the industry-year benchmark, the input variables are Winsorized in fiscal-year groups at the 1% and 99% levels.
Step 5 Calculation	For each two-digit SIC code and fiscal year combination, we conduct the following regression: $\%Total\ Accrual_{it} = \beta_0 + \beta_1 \%Sales_{it} + \beta_2 \frac{1}{Total\ Assets_{it-1}} + \beta_3 \%Property\ Plant\ and\ Equipment_t + Discretionary\ Accrual_{it}$ <p>where the residual term <i>Discretionary Accrual_{it}</i> is taken as the discretionary accrual for firm <i>i</i> in fiscal year <i>t</i>. In the empirical analysis, the absolute value is adopted to focus the analysis only on the magnitude of the discretionary accrual without consideration of the direction.</p>