Textual Analysis of Short-seller Research Reports*

Jules H. van Binsbergen[1], Xiao Han[2], and Alejandro Lopez-Lira[†3]

[1] *The Wharton School, University of Pennsylvania, NBER and CEPR*

[2] *Bayes Business School, City, University of London*

[3] *University of Florida*

First Version: August 2021; This Version: May 2023

Abstract

We document that a large percentage of text in short-sell research reports pertains to accounting fraud and earnings mismanagement, and 95% of the reports mention fraud-related words at least once. Using survey cash-flow forecasts as a counterfactual, we find that investors underreact to the cash-flow news contained in short-sell reports. On average, target firms earn abnormal returns of -4.9% on the publication day, and the subsequent price revisions equal -15% over a 12-month horizon. We introduce a novel text-based fraud measure and find that reports more related to fraud are associated with larger negative long-term abnormal returns. Using a variance decomposition of returns, we find that the revisions in expected cash flows account for the negative return surprises associated with the reports. Furthermore, short-sell research reports are associated with significant reductions in future real investment ($116 million) and stock issuances ($196 million), and we present several potential mechanisms that can explain this correlation.

*The paper was previously circulated with the title "Textual Analysis of Short-seller Research Reports, Stock Prices, and Real Investment". We are grateful for helpful comments and suggestions from Jiafu An, Giacinta Cestone, Ricard De la O, Vasso Ioannidou, Tim Jenkinson, Manish Jha (discussant), Tong Liu, Alexander Ljungqvist, Asaf Manela (discussant), Sean Myers, Matthew Ringgenberg (discussant), Hongxun Ruan, Francisco Urzua, James Ryans (discussant), William Waller (discussant), and seminar participants at the King's College London, University of Bath, University of Florida, Future of Financial Information Conference, Swedish House of Finance FinTech Conference, EFA Annual Meeting 2022, and AFA Annual Meeting 2023. We thank the research assistance provided by Dong Xu.

†Emails: julesv@wharton.upenn.edu; xiao.han.3@city.ac.uk; alejandro.lopez-lira@warrington.ufl.edu

Investment firms that specialize in producing research reports on short-selling have recently gained significant attention in the media. While numerous studies have examined the potential benefits and drawbacks of short-selling, there has been little focus on the content and nature of these research reports and their impact on the targeted firms' expected cash flows, discount rates, and investments. In this paper, we aim to address this gap in the literature by examining the content of short-sell research reports. Our analysis reveals that a significant portion of these reports pertains to allegations of accounting fraud and earnings mismanagement. Moreover, using survey cash-flow forecasts as a benchmark, we show that investors may not thoroughly incorporate the information about long-term cash flows and real investment implications.

We collect research reports using a combination of web-scraping and human verification, ensuring the data's accuracy. We use text-based analysis to provide an in-depth look at the type of information that short-sell research reports contain. While short-selling has received plenty of negative attention in the popular press, we find that the reports are dedicated to uncovering fraud and earnings mismanagement. For example, 95% of the reports mention fraud-related words at least once, and around 80% of the text in those reports is dedicated to the allegations of accounting fraud and earnings mismanagement. We introduce a novel text-based fraud measure and find that reports more related to fraud are associated with larger negative long-term abnormal returns. We use word or sentence embeddings to project each word into a vector space and then compare the distance from each text to the vector of the word 'fraud.'

Moreover, the heterogeneity of research companies' reports' effect on stock returns appears to be driven by how much they focus on documenting fraud. Intuitively, only new sources of information, such as uncovering fraud, have a significant price impact, while reports discussing public data are less convincing. However, the full effect of the cash-flow news takes time to materialize, suggesting investors' underreaction to the information in the reports. We collect several facts related to how the information from short-seller reports incorporates slowly into stock prices.

First, we find that stocks with short-seller reports exhibit initial overpricing, particularly concerning inaccurate expectations about earnings. This initial overpricing can be explained by the high costs associated with short-selling. Second, the information in the reports is assimilated into stock prices with a large delay resulting in a sluggish sell-off by long-side investors. Third, reports with higher information content, as measured by their similarity to fraudulent cases, prompt more substantial price revisions. Fourth, analysts consistently update their earnings expectations slowly, consistent with news underreaction. Such underreaction is consistent with long-term downward price corrections, indicating that the shorts-sellers reports' negative news is only slowly incorporated into the stock prices. Finally, similarly to the information model of Coibion and Gorodnichenko (2015), we find expectations delays in cash flow forecasts.

An essential contribution of our paper is combining analysts' expectations regarding cash flows with the decline in the stock market price. This allows us to separately identify whether the price impact is through the cash flow or discount rate channel. We find that investors underreact to the cash-flow news contained in short-sell reports. Moreover, using a variance decomposition of returns, we find that the revisions in expected cash flows account for the negative return surprises associated with the reports. Despite the ex-post negative impact on cash flows, expectations about cash flows are updated sluggishly, and there are consistently negative earnings surprises and downward corrections. The average abnormal return equals −4.9% on the publication day, and prices continue to decline slowly, resulting in an average cumulative total effect of −15% in the 12-month horizon.[1]

We provide evidence demonstrating that short-sell research and its subsequent stock-market effects are associated with significant real economic impact. Companies reduce their capital and R&D expenditure by up to 2% of assets, translating to an average value of $116 million in corporate investment per report. Furthermore, firms mentioned in the reports seem to respond to the downward price correction by significantly reducing net stock issuances equal to an average reduction of $196 million.

The reduction in real activities must come from the sum of the selection and treatment

1. Our findings align with the existing literature, and we study longer horizons. See for example Ljungqvist and Qian (2016), Chen (2016), Zhao (2020), and Appel and Fos (2020).

effects. Intuitively, research firms could have private information and target stocks whose investment and net-stock issuances would decrease regardless of the report. This way, they would profit if this decrease in real activities is incorporated eventually into the price. Yet, writing short-seller research reports is costly. The costs include both monetary expenses incurred when researching a company and non-trivial monetary costs. Often, there are accusations of personal harassment from the investigated firms against short-seller researchers.[2] Hence, rational short-seller report writers would only incur the costs of writing the reports if they have an expectation that the reports would change the market's information set, since otherwise, they could free-ride by short-selling and not writing a report.

Alternatively, the report's publication has a causal impact on financing and investment decisions. This channel may appear similar to how relaxing short-sell constraints can decrease firms' real activity (Grullon, Michenaud, and Weston 2015). However, we show the effect is distinct from the short-sell constraints channel since firms with reports are costlier to short-sell on average, and the short-sell costs *increase* after the reports' publications. Instead, the considerable price pressure produced by the publication of the reports could induce an increase in the cost of capital. For instance, if the price substantially decreases but cash-flow expectations remain relatively stable, the cost of capital would significantly increase. The increase in the cost of capital would reduce the appeal of financing and investing.

When considering the treatment effect, we look into managers' future earnings and sales estimates to check whether target firms tend to have less planned investments and growth before the research reports.[3] We find no evidence indicating that analysts and firm managers lower their economic outlook before the publication of the short-seller research. Thus, the expectation surprise combined with the sharp decline in price makes a strong argument showing that the change in investment and stock issuance is unexpected. We further control for firm fixed effects, time-fixed effects, and observable characteristics that predict real activity.

Finally, we use firms that commit accounting fraud but have not been identified by the

2. See for example "Short Seller Accuses Kroll and Jones Day of Wirecard Harassment," at https://www.bloomberg.com/news/articles/2023-03-06/shortseller-accuses-kroll-and-jones-day-of-wirecard-harassment.

3. We obtain managers' earnings and sales forecasts from the I/B/E/S guidance database.

short-seller research reports as a 'control' group.[4] Importantly, our textual analysis shows that most short-sell research reports contain accounting fraud allegations. Only the firms whose accounting frauds are mentioned by short-seller research experience a downward change in real activities.

Because short-seller research firms could have access to private information that makes them more likely to report about a specific company, such as information about future investments, it is impossible to rule out the selection effect completely. Nevertheless, the evidence appears to be indicative of the treatment effect driving more of the variation in real effects. Any selection story would need to specify why short-seller research firms incur costly report writing if they do not expect the information to have an impact.

Related Literature

Our paper is related to the nascent literature on studying short-sale campaigns in the stock market. Ljungqvist and Qian (2016) find investors respond strongly to the information released in the short-sale campaigns, generating downward price corrections for target firms. Using US-listed Chinese firms, Chen (2016) shows that short sellers tend to target firms that have financial reporting red flags and that exhibit 'good' operating performance and stock valuations. Appel and Fos (2020) find increases in CEO turnover associated with short campaigns by hedge funds. Paugam, Stolowy, and Gendron (2021) analyze the short-sell reports for stylized narratives related to credibility-based, emotions-based, and logic-based rhetorical strategies. Finally, Brendel and Ryans (2021) study how and when firms respond to the short-seller research reports. We differ from these studies in two ways.

First, to the best of our knowledge, we are the first study to use machine learning techniques to study the topics in the shorts-seller research reports. Our textual analysis reveals that the research reports allocate a large percentage of the text commenting on accounting fraud and earnings mismanagement. More importantly, we develop a fraud similarity measure and show that firms with reports more related to fraud earn even larger negative

4. We obtain accounting fraud data from Audit Analytics' (AA) earnings restatement database.

long-term abnormal returns. Machine-learning methods avoid the subjectivity involved in manually classifying reports.

Second, in the spirit of our textual analysis, which reveals large adverse shocks to target firms' cash flows, we propose a novel decomposition to quantify the importance of cash flow news. Using revisions in survey cash-flow forecasts, we find that cash flow news accounts for a non-trivial fraction of the negative return surprises associated with the research reports. Our findings highlight the importance of cash flow news in driving returns.

Using survey cash-flow forecasts, we also contribute to the literature on the underreaction of news by providing novel evidence of how investors underreact to the negative cash-flow news in the short-sell research reports. Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), and Hirshleifer (2001) among others develop economic models that allow for underreaction. Poteshman (2001), Ikenberry and Ramnath (2002), Kadiyala and Rau (2004), Baker and Wurgler (2007), Milian (2015), Cen, Wei, and Yang (2017), Atilgan et al. (2020) document underreaction.

Finally, our study is also related to the literature on the real effects imposed by short-selling. In contrast to the literature focusing on the relaxation of short-selling constraints, short-selling costs increase in the case of the publication of short-seller reports. Hence, the channel of short-selling reports affecting real investment differs from the existing literature. For example, Grullon, Michenaud, and Weston (2015) find that relaxing short-selling constraints causes stock prices to fall. Small firms react to these price changes by reducing equity issuances and real investment. Massa, Zhang, and Zhang (2015) and Fang, Huang, and Karpoff (2016) find that relaxing short-selling constraints reduces earnings management. Using proprietary data from Seeking Alpha and Activist Shorts Research, Wong and Zhao (2017) studies the real consequences of activist short-selling. Wang et al. (2018) show that short sellers play a disciplinary role in reducing insiders' opportunistic selling. Our contribution is to show that short-sell research reports are associated with significant reductions in future real investment and stock issuances despite the increase in short-selling costs. We further present several potential mechanisms that can explain this correlation.

1 Background: Short-sellers' Research and Target Firms

1.1 Background

Short-seller research companies publish research on public companies and often take positions that reflect their research. In many cases, these research reports attract media attention that, at times, is polemic. For example, the Wall Street Journal documents that short-sell researchers often attract so much attention that they get personally threatened.[5]

As another illustration, Fortune magazine dedicates an entire article to short-seller reports. They feature a short-seller report by Hindenburg Research that divulged that Nikola, the electric vehicle company, misled the public by showing a video of a truck cruising through the high desert outside Salt Lake City. The short-seller research firm got a tip that the video was staged and "the truck wasn't traveling under its own power [...] it had been towed to the top of a hill. The person at the wheel then popped it into neutral and started it on its journey downhill—slowly at first, then accelerating."

On September 10, 2020, Hindenburg Research published a short-sell report about Nikola. Nikola defended itself the following day, saying that the article contained "false and misleading statements" and hired counsel to "evaluate potential legal recourse." Nevertheless, in a regulatory filing in November, the company revealed that the Justice Department had issued grand jury subpoenas against Nikola and its founder and executive chairman. Since the research release and until Fortune's article publication date, Nikola's stock fell by 30%, and the chairman resigned.[6] On October 14, 2022, the chairman was convicted for fraud.[7]

While these reports attract attention from the business media, relatively little academic work quantifies these reports' real and financial effects. We assemble a comprehensive database of publicly available short-sell research reports and study their transmission channels to fill this void.

5. https://www.wsj.com/articles/gamestop-stock-short-squeeze-ugly-side-11611750250
6. https://fortune.com/longform/short-selling-stock-market-bets-hindenburg-viceroy-muddy-waters/.
7. https://www.nytimes.com/2022/10/14/business/trevor-milton-nikola-fraud.html

1.2 Data Sample

Our short-sell research reports sample includes the reports publicly released by Bonitas Research, Citron Research, Viceroy Research, J Capital Research, Hindenburg Research, Blue Orca Capital, GMT Research, Spruce Point Capital Management, and Muddy Waters.

First, we use a combination of web-scraping and human verification to download all research reports from these short sellers' websites with the initial publication dates recorded. Second, we use tickers from these reports and merge them with the Center for Research in Security Prices (CRSP) database to match them with PERMNO codes. As a result, we can match 509 short-sale reports of 314 unique firms listed on NYSE, NASDAQ, and AMEX.[8] Our first observation is from January 2007, and the last is from December, 2021.

1.2.1 Firms Types

What types of firms are more likely to be targeted by short-seller research reports? We consider characteristics including firm age, market-cap, valuation ratio, profitability/earnings, and investments. We obtain stock-level price and return data from CRSP and firm-level accounting data from Compustat. The variables are as follows:

1. Age: The month difference between current month and the first month of a firm' stock appear in CRSP.

2. Firm Size (LNsize): Log of price times shares outstanding at the end of December.

3. Net Stock Issuance (NSI): The natural logarithm of the ratio of the split-adjusted shares outstanding at the fiscal year ending in t to the split-adjusted shares outstanding at the fiscal year ending in t-1,

4. Book-to-market Ratio (LNbeme): Log of book equity for the fiscal year ending in year t divided by the ME at the end of December of t.

5. Gross Profitability to Assets (GP): Total revenue minus cost of goods sold divided by total assets.

8. Some firms have multiple short-sale reports on different dates.

6. Return on Equity (ROE): Log of one plus income before extraordinary items divided by lagged book equity.

7. Investment-to-asset (IA): Total assets for the fiscal year ending in t divided by total assets for the fiscal year ending in t-1 minus 1.

8. Capital and R&D Expenditure (CAP_RD): Capital expenditure and R&D expenses divided by lagged total assets.

9. Sale Growth (SG): Net Sales for the fiscal year ending in t divided by net Sales for the fiscal year ending in t-1 minus 1.

10. Changes in PPE and Inventory (ΔPI/A): Changes in gross property, plant, and equipment plus changes in inventory scaled by lagged total assets.

11. Cost of Borrow Score (obtained from Markit): A number from 1 to 10 indicating the cost of borrowing a stock, where 1 is cheapest and 10 is most expensive.

Having obtained these characteristics, we measure the difference between the short firms' characteristics and the cross-sectional mean/median of other firms' to study which firms are more likely to be subject to short-sellers research. Specifically, for each target firm, we compare their characteristics in the year before the publication dates of the research reports to the cross-sectional mean/median of other firms'.[9] As such, we can infer what types of firms are more likely to be targeted by short-sellers.

Table 1 presents the summary statistics of the differences in these variables.[10] We find that, on average, the target firms are younger and larger than other firms, and are more costly to short. The target firms also issue more stocks and have higher profitability and real investment than other firms. The book-to-market ratios are also lower for target firms, indicating that growth firms are more likely to be mentioned in the short-seller research report. The cost of borrowing target firms' shares is also higher than other firms, which is

9. Since firm characteristics are annual accounting variables, we only count a firm once if it has multiple short-sellers research reports in a year and keeps the earliest report in a year.

10. The variables are winsorized at 1% to exclude outlier effects.

consistent with Ljungqvist and Qian (2016) who shows that shorting fees of target firms are high and thus are difficult to arbitrage.

[Insert Table 1 about here]

2 Textual Analysis

We start by conducting a textual analysis of the reports to understand their underlying content. We use machine-learning methods to avoid the subjectivity of manual classification. While the machine-learning algorithms are not flawless, any errors in the machine-learning methods should be orthogonal to the subsequent empirical analysis.[11] Moreover, machine learning provides continuous variables, which we use in the subsequent empirical analysis. In contrast, manual classification can only provide a discrete categorization.

We first train a word2vec model on the reports and construct a novel measure of fraud similarity. The word2vec algorithm projects each word into a vector space, and we represent each document in a 300-dimensional real-number space. We train the word2vec model on these texts to ensure the appropriate contexts. Using more sophisticated models such as BERT produces similar results.[12]

We find that 95% of the reports contain direct allegations of fraud or financial misconduct. We first construct a fraud dictionary by collecting a list of the words most related to the word 'fraud' as measured by the cosine distance. The closely related words in this dictionary are: 'accuse,' 'admission,' 'allegation,' 'allegations,' 'allege,' 'allegedly,' 'bribery,' 'commit,' 'conspiracy,' 'convict,' 'corruption,' 'crime,' 'delist,' 'felony,' 'fraudulent,' 'guilty,' 'impropriety,' 'indict,' 'intentional,' 'investigate,' 'investigation,' 'manipulation,' 'manipulative,' 'perpetrate,' 'perpetuate,' 'plead,' 'Ponzi,' 'prosecutor,' 'scandal,' and 'sue.' We say

11. For the textual analysis, we drop sample firms from GMT Research as most of the text from these reports is paywalled.

12. See Jha, Liu, and Manela 2022 for a recent application of BERT in the context of news. The FinBERT model of Araci (2019) does not seem to work well in this dataset. Finbert was trained on companies' annual reports and conference calls with a language different than what we observe in the reports. The (strategic) lack of discussion about fraud in official reports and earnings announcements may explain the discrepancy.

that a report addresses fraud or financial misconduct if it has any of the words in the fraud dictionary.

Next, we define our fraud similarity measure at the document level. We first calculate each report vector as the average of the individual word vectors in the report. In the case of BERT, we use the vector obtained from the entire text. Then we use one minus the cosine distance from the document vector to the 'fraud' vector as the similarity measure. The fraud similarity measure indicates how close the reports are to the word 'fraud' in the vector space, corresponding to reports alleging more fraud in different ways. Hence, our method focuses on the cross-sectional variation of the fraud-related content. In the following sections, we document the fraud similarity measure is related to the differences in realized returns. Further, the differences in returns across different research firms are well explained by the differences in fraud similarity.

Then, we use topic models to study the narratives the research reports provide. We use Latent Dirichlet Allocation (LDA), a standard Bayesian unsupervised algorithm. LDA describes each document as a distribution over topics and each topic as a distribution over words. Consequently, LDA provides a quick summary of the issues that the collection of documents describes and what percentage of each document discusses each topic.[13]

[**Insert Figure 1 about here**]

Further, the topic model reveals that companies spend a considerable percentage of the reports discussing accounting information and fraud. Figure 1 plots the result of running LDA with five topics. Each cluster of words corresponds to a different topic. Larger words correspond to a higher-than-average probability of occurring within that topic. Topics one, four, and five discuss accounting information. Topic two discusses fraud, particularly in connection with China. Topic three corresponds to disclaimer information. Figure 2 shows the average percentage of words (per report) that firms dedicate to each topic. We find that reports allocate, on average, around 80% in discussing accounting, fraud, or earnings mismanagement.

13. See Lopez-Lira (2019) for a detailed overview of the algorithm.

[Insert Figure 2 about here]

2.1 Post-publication Earnings for Target Firms

If the reports correctly identify the accounting frauds or earnings management of the target firms, then it is likely that the earnings of these firms decline following the report publications. To test this, we consider a window of eight years (with four years pre- and post-publications of the research reports) to study the change in earnings for target firms. That is, for each firm with a research report published in the fiscal year ending in year $t+0$, we measure the difference between its earnings in year $\{t-4, t-3, t-2, t-1, t+1, t+2, t+3, t+4\}$ and its earnings in year $t+0$. We consider two commonly used earnings measures: gross profitability to assets, and the return on equity, as defined above. Graphs (a) and (b) in Figure 3 show that both gross profitability and return on equity decline dramatically upon the report publications. The downward changes in profitability/earnings imply that direct allegations of fraud or financial misconduct in the reports are associated with negative shocks to target firms' cash flows.

[Insert Figure 3 about here]

3 Price Impacts of Short-seller Research

3.1 Abnormal Returns

What is the impact of short-seller research reports on stock prices? As discussed earlier, such reports often generate considerable media attention, so that one may expect price drops upon the revelation of this information. To study the price impact, we calculate the abnormal returns of stocks on the publication date of the short-seller research reports. The abnormal return relative to CAPM for stock i on day t ($abr_{i,t}$) is defined as,

$$abr_{i,t} = ret_{i,t} - (\hat{\beta}_i \times mktrf_t), \tag{1}$$

in which ret_t is the daily return at day t and $\hat{\beta}$ is the estimated exposure to the market excess return ($mktrf$) using daily returns from the most recent 200 trading days. As a robustness check, we also consider an alternative benchmark model, the Fama-French-Carhart four-factors model (FFC hereafter, Carhart (1997)), to measure the abnormal returns,

$$abr_{i,t} = ret_{i,t} - (\hat{\beta_{1,i}} \times mktrf_t + \hat{\beta_{2,i}} \times smb_t + \hat{\beta_{3,i}} \times hml_t + \hat{\beta_{4,i}} \times mom_t) \tag{2}$$

where smb, hml, and mom are size, value, and momentum factors.

Table 2 Panel A reports the summary statistics of abnormal returns (alphas hereafter) on the publication days of research reports. On average, stocks earn a negative CAPM and FFC alpha of 4.9% on the publication day, an economically large magnitude. Moreover, a t-statistic of -10.211 (-10.152) indicates that the negative average CAPM (FFC) alpha is statistically different from zero. Figure 4 plots the distribution of the alphas relative to the CAPM of stocks on the publication day of short-sell reports. The distribution is negatively skewed since most alphas are negative, suggesting a downward price correction imposed by short-seller research reports.

[Insert Table 2 and Figure 4 about here]

Figure 5 plots the average volatility for the targeted firms.[14] Volatility spikes on the event day, and the daily volatility reaches its highest point at around 0.0355 after one week of the publications of the short-seller research reports. Further, we can visually verify that a structural break occurs on the publication day for the constant level of return volatility. The mean of the return volatility is higher during the post-publication periods than during the pre-publication periods.

14. We measure return volatility as the standard deviation of daily returns in a rolling window of 15 trading days.

[Insert Figure 5 about here]

Figure 6 plots the average cumulative abnormal returns (relative to CAPM) for the post-event window of 250 trading days.[15] We find the cumulative abnormal returns after the publication become significantly negative with a lower bound (95 % confidence interval) below zero. More importantly, the confidence interval at day 0 is between around -4% and -6%; at day 150 between around -8% and -18%. The non-overlap between days 0 and 150 suggests a robust downward trend in price correction.

[Insert Figure 6 about here]

Table 2 Panel B reports the summary statistics of the one-year cumulative abnormal returns following the publication day of research reports for target firms.[16] We find that on average, each target firms earn one-year cumulative abnormal returns of -15.3% (-14.6%) relative CAPM (FFC) with t-statistic of -4.121 (-4.179). These findings suggest that the reports induce long-term price corrections on the reports. The difference between the long- and short-term price impacts is around 10% per report. Such a difference indicates that investors may underreact to the negative cash flow news in the reports on the publication day. We test in the next section how investors underreact to cash-flow news, resulting in long-term corrections.

Another possibility worth exploring is that targeted firms' stock prices start declining before the research reports' publication day. Combining such trends with a momentum-type continuation of the price decline could potentially provide a different interpretation of our results, where the research reports do not reveal information, and their publication coincides with an existing negative time trend.

We investigate the returns before pre-publication to alleviate this concern. Specifically, we sort sample stocks into two groups: stocks in the first group have positive cumulative alphas

15. To avoid overlaps in aggregating cumulative returns, we keep the date of the first report for a firm in a year if the firm has multiple reports in that year.

16. We require target firms to have at least 150 observations of daily returns following the publication day.

before the publication day, whereas stocks in the second group have negative alphas. Next, we estimate the cumulative alpha following the publication dates for each group. Figure A1 (a) and (b) plot the average of one-year cumulative abnormal returns of target firms whose cumulative abnormal returns before publication day are positive and negative, respectively. We find that regardless of positive or negative past abnormal returns, the cumulative alpha after following the publications is significantly negative, arriving at -15% in a one-year horizon. Therefore, our finding that stock prices are negatively impacted by short-seller research reports cannot be explained by momentum or long-term reversal effects.

3.2 Fraud Similarity and Abnormal Returns

As the textual analysis indicates, short-sellers' research reports allocate a large percentage of the text commenting on accounting fraud and earnings mismanagement. The extent to which short-sellers' reports negatively impact the stock prices of target firms should be driven by the strength of the fraud revealed by short-sellers' reports.

Table 3 Panel A looks at the abnormal returns of stock portfolios sorted on the text-based fraud measure. Specifically, we sort our sample stocks into two portfolios on the median of the fraud similarity measure based on the BERT model. Consistent with our hypothesis, the portfolio with higher fraud similarity has more negative one-year cumulative returns than the portfolio with low fraud similarity. The difference is 16.3% (13.6%) for alphas measured with CAPM (Fama-French-Carhart model), both economically meaningful and statistically significant. The results are consistent with the word2vec-based fraud similarity measure.

Table 3 Panel B looks at regressions of abnormal returns on fraud similarity. We find consistent results that reports more related to fraud are associated with larger negative long-term abnormal returns. In the next section, we explore this underreaction using survey cash-flow forecasts as a counterfactual.

[Insert Table 3 about here]

To explain why different research firms have different price effects, in Figure 7, we plot the average fraud similarity against the average of one-year cumulative CAPM alphas for reports issued by each short-sell research firm. The reports issued by different short-sell research companies are associated with different price impacts, e.g., the reports issued by "Hinderburg" on average generate about -36.6% one-year cumulative CAPM alpha whereas the reports published by "Spruce Point" on average generate about -2.85%. This return difference across firms is mostly driven by the difference in the fraud similarity measure, as shown by a negative correlation between fraud similarity and abnormal returns. Therefore, there will be a high correlation between research-firm fixed effects and fraud similarity. Still, our finding that reports with higher fraud similarities are associated with larger negative long-term abnormal returns remains consistent even when controlling for research-firm fixed effects.

[Insert Figure 7 about here]

4 Cash Flow News and Investor Reaction

4.1 Underreaction in Cash Flow Expectations

The long-term downward price corrections we find in the previous section show that the negative news in the shorts-sellers reports is slowly incorporated into the prices, indicating underreaction. To dissect how investors react to the short-seller research reports, we study revisions in analysts' cash flow forecasts, a widely used measure of market expectations in the literature. More importantly, the literature shows that the revisions in analysts' forecasts are related to a significant source of cash flow information in financial markets.[17]

We look at both short-term and long-term cash flow expectations: one is expected ROE, defined as the log of one plus analysts' one-year-ahead earnings forecasts to lagged book equity, and the other is long-term earnings growth forecasts (LTEG). Since analysts update earnings expectations monthly, we look at the monthly revisions.

17. See Kothari, So, and Verdi (2016) for a suvery of the related studies. More recently, Derrien et al. (2021) show that analysts revise forecasts downward following negative ESG news.

If investors underreact to the negative cash flow news in the reports, the revisions in their cash flow expectations would be gradual rather than dramatic. We consider a window of 12 months (post-publications of the research reports), which is in line with the return window we study above. For each firm with a research report published in the forecast month ending in time t, we measure the revision as the cash flow expectation in subsequent 12 months as the difference between expectations and expectations in month t, a month just before the report release. As a placebo test, we also measure the "revisions" as the difference between expectations in a 12-month pre-publication window and expectations in month t. In doing so, we can check whether analysts' downward revisions occur before the report publication.

Figure 8 plots the average time-varying revisions in analysts' cash flow expectations, with the blue bar indicating the 95% confidence interval. ROE expectations and long-term earnings growth expectations decrease slowly and significantly, and the decline takes at least one year to materialize fully with a downward revision of 4% for one-year expected ROE and of 3% for LTEG. For the 12-month window pre-publications, we do not find any significant downward revisions in cash flow expectations. Therefore, the downward revision we find following the publications of research reports is not driven by any past trends. The above findings suggest that investors tend to underreact to the negative cash flow news in short-seller research reports.

[**Insert Figure 8 about here**]

We conduct a placebo test to corroborate our finding of investor underreaction to reports. Specifically, for each target firm, we identify firms that can match the target firm's cash flow forecasts in a 12-month pre-publication window.[18] We then examine whether these matched firms experience similar downward revisions in analysts' cash flow forecasts as target firms. Figure A2 plots the revisions in both short- and long-term cash flow forecasts for matched firms. Overall, we do not find any significant and dramatic patterns of downward revisions

18. We identify matched firms by requiring their cash flow forecasts to be within 5% deviation from target firms' cash flow forecasts in the month before the publication month of short-seller research reports. We also require the change of cash flow forecasts in a 12-month pre-publication window to be within a 5% deviation to control for any downward trend in cash flow expectation in the pre-publication window.

in cash flow forecasts.

Thus far, an interesting question is whether investors change their discount rate expectations upon the research reports. Our textual analysis reveals rich cash flow news in these short-seller research reports, and investing in these firms may bring uncertainty, so investors may also update their discount rate expectations. We define the expected return as the log of price targets (for the 12-month horizon) plus the yearly dividend forecasts (if any) to current stock prices. For a firm whose reports are published in month t, we then define the revisions in discount rate expectations as the difference between discount rate expectation in a 24-month publication window and discount rate expectation in month $t-1$.

Figure 8 panel (c) plots the revision in the discount rate expectations. We find that revisions in expected returns are significantly positive, with an increase of around 7% upon the research reports' publications. The increase in the expected returns happens because publishing short-sell research reports do not change investors' expectations about firms' future cash flows quickly enough. The downward price correction, therefore, makes the discount rate higher. Hence, what should be a pure cash-flow effect, instead becomes a discount rate effect on the publication days. The increase in the expected returns may also happen because investors think target firms are riskier to hold after the attack by short-sellers and demand higher returns from them.

However, unlike revisions in the cash flow expectations, we do not find significant revisions in expected returns in the subsequent months following the publications, as indicated by a relatively flat line. This suggests that the one-year cumulative negative abnormal returns (negative return surprises) associated with the report publications tend to be explained by revisions in cash flow news instead of revisions in discount rate news.

4.1.1 Test Expectation Stickiness

Thus far, an interesting question is why investors underreact to the negative information in the short-sellers' research reports. To explain this, we shed light on the noise information model of Coibion and Gorodnichenko (2015), who show that forecasts are a weighted average of agents' prior beliefs and the new information received, where the weight on prior beliefs can be interpreted as the degree of expectation stickiness or information rigidity. The model predicts that ex-post mean forecast errors and ex-ante mean forecast holds,

$$X_{t+h} - F_t X_{t+h} = \frac{1-G}{G}(F_t X_{t+h} - F_{t-1} X_{t+h}) + v_{t+h,t}, \tag{3}$$

where $v_{t+h,t}$ is the rational expectation errors, and F_t denotes the average forecast across agents at time t. $(1-G)$ can be interpreted as the degree of expectation stickiness that measures to what extent agents are resistant to using the new information.

Following Bouchaud et al. (2019), we use analysts' consensus earnings expectations, a widely used measure of investors' expectations for firm fundamentals, to test the stickiness $(1-G)$ in earnings expectations for firms targeted by short-sellers' research reports. Specifically, we run a Panel OLS regression,

$$\frac{X_{i,t+1} - F_t X_{i,t+1}}{P_{i,t-1}} = \alpha + \beta(\frac{F_t X_{i,t+1} - F_{t-1} X_{i,t+1}}{P_{i,t-1}}) + \epsilon_{i,t} \tag{4}$$

in which $X_{i,t+1}$ denotes earnings per share for target firm i at year $t+1$. Following Bouchaud et al. (2019), we normalize the earnings revision and forecast errors by price $P_{i,t-1}$, which denotes the price per share for firm i at year $t-1$. The coefficient β can be interpreted as a function of the stickiness parameter, so that $1-G{=}\beta/(1{+}\beta)$.

Table 4 reports the regression of forecast error on the forecast revision for the firms targeted by short-sellers. We find a significantly positive beta of 0.19 (t-statistic is 3.309) in a pool OLS regression, which means $1-G = 0.16$. The results suggest that the weight on prior expectation at the quarterly frequency is given by $0.16^{\frac{1}{4}} = 0.63$, which is very similar to what Bouchaud et al. (2019) find for the quarterly revisions in analysts' earnings forecasts and Coibion and Gorodnichenko (2015) find for the quarterly revisions in inflation forecasts.

[**Insert Table 4 about here**]

Columns (2) and (3) in Table 4 look at the Panel OLS regressions with time-fixed effect and firm-fixed effect, respectively. Controlling for time-fixed effects (firm-fixed effects) suggests a cross-sectional (time-series) variation of expectation stickiness. Regardless of the fixed effects that we control for, we find a significant pattern of stickiness in the cash flow expectations.

4.2 Quantify Cash Flow News

The above discussion shows that due to expectation stickiness, the negative cash flow news in the reports is slowly incorporated into stock prices, resulting in long-term downward price corrections. In this section, we quantify the role of cash flow news in the return surprises associated with the reports.

Consider the Gordon growth model,

$$P_t = \frac{E_t[CF_{t+1}]}{E(r-g)}$$

The negative shocks to target firms' prices must come from either 1) a decrease in the short-term expected cash flow (CF_{t+1}) or in the long-term growth rate of cash flows (g) as reports reveal accounting frauds/earnings management, or 2) increases in the discount rate (r) if investors are more risk averse and require higher returns for target firms (though idiosyncratic).

To quantify the importance of cash flow shocks in driving the negative return surprises associated with the reports, we follow Vuolteenaho (2002) to decompose the return surprises into revisions in expected cash flows and revisions in expected returns,

$$r_t - E_{t-1}r_t \approx \Delta E \sum_{j=0}^{\infty} \rho^j roe_{t+j} - \Delta E \sum_{j=1}^{\infty} \rho^j r_{t+j} \tag{5}$$

where $roe_t = log(1+X_t/B_{t-1})$ with X and B denoting earnings and book equity, respectively. ΔE denotes the change of expectations between E_t and E_{t-1}. r_t indicates log stock return.

Equation (3) states that a negative shock to returns ($[r_t - E_{t-1}r_t]$ is negative) is attributed to either a downward revision in expected cash flow or an upward revision in expected returns. Unlike the standard methods of using a VAR model to measure the changes in predicted cash flows as the revisions in cash flow expectations, we use analysts' revisions in their forecast as a direct cash flow measure. Our motivations for using these direct cash flow measures come from two sources. First, VAR models impose restrictions by assuming that the dynamics of returns or cash flows are determined by the lags of themselves and other variables. In the cases where lags cannot predict returns (or cash flows) well, the discount rate (cash flow) would be poorly measured.[19] Second, our previous results show that the one-year cumulative abnormal returns are likely driven by investors' gradual revisions in their cash flow expectations, namely underreaction. We can test the link between return surprises and forecast revisions more quantitatively using the direct measures of revisions in expected cash flows from analysts' forecasts. Hence, we can test directly whether revisions in analysts' expectations can explain a large fraction of return surprises.

Since we study how revisions in expectation drive return surprises in a one-year event window, we denote E_{t-1} and E_t as the expectations formed one month before the publications of reports and 11 months after the publications of reports. Each target firm's return surprise $r_t - E_{t-1}r_t$ is defined as the difference between its log of one-year cumulative returns upon the report publication and the log of analysts' expected one-year returns one-month before the report publication. Due to the forecasts having a finite horizon, we construct revisions in cash flow expectations $\Delta E(roe_{t+j})$ up to a five-year forecast horizon using earnings and book value forecast data from I/B/E/S analysts. We detail the construction of cash flow revisions in Appendix A1.

Table 5 reports the summary statistics of return surprises and revisions in cash flow expectation for the target firms, with two findings emerging. First, consistent with the downward price corrections induced by the short-sellers' reports, we find an average annual log return surprise of -52.1% for target firms. The finding suggests that realized returns following the publication of short-sellers' reports are lower than expected returns before the reports. Second, revisions in cash flow expectations at most horizons are significantly negative. In particular, the revision at the one-year forecast horizon, defined as $roe_t - E_{t-1}(roe_t)$, is -0.148 (t-statistic is -5.723). Such a negative revision suggests that realized roe after the report's

19. See the discussion in Chen and Zhao (2009).

publications is -0.148 lower than the expected roe before the reports. The result shows that before the report publications, investors were too optimistic about target firms' earnings, which could be a consequence of fraud and earnings management.

[Insert Table 5 about here]

Having estimated these variables, we then decompose the variance of return surprises into its covariance with cash flow revisions and discount rate revisions,

$$Var(r_t - E_{t-1}r_t) \approx Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=0}^{\infty} \rho^j roe_{t+j}) - Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=1}^{\infty} \rho^j r_{t+j}) \quad (6)$$

where $Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=0}^{\infty} \rho^j roe_{t+j})$ can be further decomposed into,

$$\underbrace{Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=0}^{5} \rho^j roe_{t+j})}_{\text{Cash flow news in six-years forecast horizon, } CF_5} + \underbrace{Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=5}^{\infty} \rho^j roe_{t+j})}_{\text{Cash flow news beyond six-years forecast horizon, } CF_{5-\infty}}$$

Divide both sides in Equation (5) by the variance of the return surprise,

$$1 \approx \frac{CF_5}{Var(r_t - E_{t-1}r_t)} + \frac{CF_{5-\infty}}{Var(r_t - E_{t-1}r_t)} - \frac{Cov(r_t - E_{t-1}r_t, \Delta E \sum_{j=1}^{\infty} \rho^j r_{t+j})}{Var(r_t - E_{t-1}r_t)} \quad (7)$$

where $\frac{CF_5}{Var(r_t - E_{t-1}r_t)}$ represents the fraction of variance of the return surprise that cumulative revisions in expected cash flow at a five-years-horizon can explain. It is equivalent to the beta estimated from the cross-sectional regression of cash flow revisions on the return surprise,

$$\Delta E \sum_{j=0}^{4} \rho^j roe_{i,t+j} = \beta[r_{i,t} - E_{i,t-1}r_{i,t}] + \epsilon_i \quad (8)$$

One can also further decompose $\Delta E \sum_{j=0}^{4} \rho^j roe_{i,t+j}$ into cash flow revisions at individual forecast horizons to estimate β,

$$\Delta E roe_{i,t} = \beta_1 [r_{i,t} - E_{i,t-1} r_{i,t}] + \epsilon_i \tag{9}$$

$$\Delta E \rho^1 roe_{i,t+1} = \beta_2 [r_{i,t} - E_{i,t-1} r_{i,t}] + \epsilon_i \tag{10}$$

$$\Delta E \rho^2 roe_{i,t+2} = \beta_3 [r_{i,t} - E_{i,t-1} r_{i,t}] + \epsilon_i \tag{11}$$

$$\Delta E \rho^3 roe_{i,t+3} = \beta_4 [r_{i,t} - E_{i,t-1} r_{i,t}] + \epsilon_i \tag{12}$$

$$\Delta E \rho^4 roe_{i,t+4} = \beta_5 [r_{i,t} - E_{i,t-1} r_{i,t}] + \epsilon_i \tag{13}$$

where β_1, β_2, β_3, β_4, and β_5 measure the cash flow news at the forecast horizon of one to five years, respectively.[20]

Table 6 Panel A reports the decomposition results for revision in cash flow forecasts at the individual horizon with two findings emerging. First, almost all estimated βs are significantly positive, indicating that firms with more downward revisions in cash flow expectations are associated with negative shocks to returns. This is consistent with our findings of negative price impacts of short-sellers' research reports, which bring negative cash news such as accounting fraud and earnings management, as implied by our textual analysis.

[Insert Table 6 about here]

Second, the explanation power of revisions in cash flow forecasts for the variance of return surprise is sizable. The revision at the one-year forecast horizon accounts for 12% of the return surprise variance. Table 6 Panel B reports the decomposition results using cumulative revisions in cash flow forecasts. Cumulative cash-flow news explains 23.6% of the variation of return surprise up to a five-year forecast horizon.

20. Following Vuolteenaho 2002, we choose a ρ of 0.967.

5 Short-seller Research and Real Activities

5.1 Real Activities

Grullon, Michenaud, and Weston (2015) find that relaxing short-selling constraints cause stock prices to fall. Small firms react to these price changes by reducing equity issuances and real investment. Their findings suggest that short-selling activity has a causal impact on financing and investment decisions. Since the presence of short-seller research also results in a downward price correction, we explore in this section several theories related to stock issuances and real investment. However, the short-seller research report channel differs from the short-selling constraints channel since we find the shorting-cost are high for targeted firms and increase after the report's publication.

First, we explore whether firms have lower net stock issuances after short-sell research reports. Our empirical analysis is motivated by the literature on managerial market timing: firm managers tend to issue more stocks when their firms are overpriced (Baker and Wurgler 2013). Since the short-sell reports result in a downward price correction for the overpriced firms, we expect their managers to issue fewer stocks.

Second, we explore whether firms exhibit lower rates of real investment, thereby shedding light on the quantitative model provided in Binsbergen and Opp (2019) where equity prices have an important real feedback effect on real investment. More specifically, firms that are overvalued tend to invest more since their cost of equity is low, and a downward price correction could push up the cost of equity, leading to lower real investment rates.

To study the effects of the reports on real activity, we consider a window of eight years (with four years pre- and post-publications of the research reports). That is, for each firm with a research report published in the fiscal year ending in year $t + 0$, we measure the difference between its real activities in year $\{t-4, t-3, t-2, t-1, t+1, t+2, t+3, t+4\}$ and its real activities in year $t + 0$.

Figure 9 plots these differences in real activities for an 8-year window with a blue bar as the 95% confidence interval. We find downward trends in the changes of net stock issuance, investment-to-asset ratio, capital and R&D expenditure, changes in PPE and inventory, and sales growth. Moreover, we find statistical significance for changes in all real activity variables, with the exception of net stock issuance, where the effect disappears after one year. On average, firms impacted by short-sell reports significantly invest less and issue less stock even four years after the publication of the report.

[Insert Figure 9 about here]

One concern is that the reports may be written at times when firms that are not subject to short-seller research reports also invest less and issue fewer stocks. That would imply that the downward change in real activities occurs at the aggregate market level rather than at the firm level. In this case, it might not be the short-seller research that drives the declines in the target firms' investments and stock issuance.

To explore this potential explanation, we use a standard difference in differences regression with firm fixed effects.[21] This standard diff-in-diff approach better measures the research report's effect since it compares the values of companies with a report to those without reports while controlling for any aggregate variation in real activity using time-fixed effects. We also control for individual fixed effects for any individual unobserved variation that is constant through time. We remark that the short-seller reports are endogenous since they are the result of an optimization problem when selecting the companies, but adding standard controls does not seem to suggest that the selection effect is driving the results.

Hence, we run a panel OLS regression of real investment variables on four dummy variables: $D_{1,i,t}$ that takes the value of one if a firm has a short-sale report written about them in the past twelve months; $D_{2,i,t}$ that takes the value of one if such a report is written in the past 12-24 months; $D_{3,i,t}$ that takes the value one if such a report is written in the past 24-36 months; $D_{4,i,t}$ that takes the value one if such a report is written in the past 36-48 months.

$$Real\_activity_{i,t} = \beta_1 D_{1,i,t} + \beta_2 D_{2,i,t} + \beta_3 D_{3,i,t} + \beta_4 D_{4,i,t} + \gamma control_{i,t} + Year_t + Firm_i + \epsilon_{i,t}, \quad (14)$$

21. Naturally, because the writing and publication timing of the short-sale research reports is endogenous, we abstain from causal statements (see Wooldridge (2010)).

where $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. Since the dummy variable is zero for firms not subject to short-sell research reports, the coefficient estimate on this variable tells us how short-selling firms change their real activities relative to other firms. In this multivariate panel OLS regression, we control for short interest and firm size following Grullon, Michenaud, and Weston (2015). In addition to short interests and firm size, we also control for financial leverage and book-to-market ratios.

Table 7 reports the results of this panel OLS regression. We find significantly negative coefficients of -0.297, -0.022, -0.021, and -0.091 on the dummy variable $D_{1,i,t}$ for investment-to-assets, capital, and R&D expenditure, changes in PPE and inventory and sale growth, respectively. Companies reduce their capital and R&D spending by up to 2% as a percentage of assets, which translates to an average value of $116 million in corporate investment relative to other firms in the subsequent year. The coefficient estimate is -0.042 (with t-statistic of -3.63), indicating that firms with short-seller research reports experience 4.2% more negative changes in net stock issuance than other firms in the subsequent year value. By multiplying this percentage change with the average market value of target firms, we find an average decline of $196 million in stock issuances. The effect of the report in the two years is given by $D_{2,i,t}$, in the third year by $D_{3,i,t}$, and in the fourth year is $D_{4,i,t}$.

[Insert Table 7 about here]

Grullon, Michenaud, and Weston (2015) show a causal impact of short-selling activity on firms' financing and investment decisions. Notably, the effects of the reports on real activity are not driven by any change in short-selling activity for two reasons. First, the panel OLS controlling for short interest (the proxy for short-selling activity used by Grullon, Michenaud, and Weston (2015)) still features the effects of research reports on target firms' real activities.

Second, we find a relatively constant short-selling activity around the publication day of short-selling reports. Therefore, it is unlikely that the increase in short-selling activity (partially) explains our results. Following Wang, Yan, and Zheng (2020), we obtain daily aggregate short-sale volume at the individual stock level from Financial Industry Regulatory

Authority (FINRA). The daily aggregate short-sale volume is defined as the total short-sale volume divided by the total trading volume reported to FIRNA. Graph (a) in Figure A2 plots the average of the daily share-sale trading volume for the target firms in short-seller research reports, with day 0 indicating the publication day of short-seller research reports. The short-sale volume is only slightly higher on day -1 and has no notable difference from other days. The finding is consistent with Ljungqvist and Qian (2016), who show that the spike in new shorts on day -1 is not only small but also short-lived and, therefore, unlikely to drive share prices of target firms down.

However, investors who long the stocks are not constrained, and therefore, the selling pressure from their tradings contributes to the downward price corrections associated with the shorts. To study whether tradings increase substantially on the event day, we obtain daily share turnover (daily trading volume data scaled by total share outstanding) at the firm security level from The Center for Research in Security Prices (CRSP). Graph (b) in Figure A2 plots the average share turnover for the target firms in the short-sellers research reports. Consistent with our conjecture, the turnover increases to approximately 6% on the day of the publication date, compared to an average value of 2%. Interestingly, we observe an increase in trading before the report's publication date, consistent with either information leakage or the short seller firms timing the information release to benefit from short positions.

5.2 Additional Tests

While the results presented in Tables 4 and 5 indicate that target firms issue fewer stocks and reduce their investments upon the short-seller research, the concern still exists that short-sellers do not randomly select firms to write reports on. That says, the firms that draw the attention of short sellers are experiencing some problems, which cause their outcomes to slide down. Short sellers, therefore, could select these firms to report by anticipating them to reduce investments. In this Section, we provide more evidence to suggest that the change in the real activities is due to a treatment effect rather than a selection effect.

5.2.1 Analysts' forecasts of real investment and cash flow

In the first test, we look at whether analysts forecast downward changes in real investments of target firms before the publications of short-sellers' research reports. It is feasible that short-sellers have no information advantage over analysts. In this case, if analysts do not forecast the downward changes in real investments, it is plausible that short-sellers cannot predict as well. Therefore, short-sellers are unlikely to write research reports because they anticipate target firms to reduce investments.

We obtain analysts' capital expenditure forecasts at one-year and two-year horizons from the I/B/E/S database and scale the forecasts by total assets.[22] We then run the OLS panel regression of the scaled forecasts on a dummy variable $D_{i,t}$ that takes one if firms are targeted by short-seller research in the next year,

$$Analysts\_forecasts_{i,t} = \beta_1 D_{i,t} + Year_t + Firm_i + Control_{i,t} + \epsilon_{i,t}, \tag{15}$$

where $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. A negative coefficient estimation on $D_{1,i,t}$ implies that analysts forecast the downward changes in real investments for target firms at one-year and two-year horizons. Appendix Table A1 Panel A reports the regression results with controlling for variables that are the same as those controlled in 7. We find that the coefficient on the dummy is not significant. Further, the results remain consistent after controlling other variables.

[Insert Table A1 about here]

Appendix Table A1 Panel B reports the regression of analysts' forecasts of cash flows per share at one-year and two-year horizons on the dummy variable $D_{i,t}$. We study analysts' estimates of cash flows to examine whether short-sellers could anticipate target firms to have lower cash flows and therefore have more financial constraints, which would reduce target firms' investments. Consistent with our findings of analysts' capital expenditure forecasts,

22. We focus on one-year and two-year horizons as analysts' annual forecasts for other horizons have significantly fewer observations.

we do not find evidence that analysts expect target firms to have lower cash flows at one-year and two-year horizons. In contrast, we find that the coefficient estimation on the dummy variable is positive, which indicates optimistic expectations of target firms' future cash flows, though the estimation is weakly significant.

Overall, the above findings suggest that the real effects of short-sell research reports are unlikely to arise from short-sellers' expectations that target firms tend to have financial constraints and reduce their investments in the future.

5.2.2 Short-seller research and I/B/E/S Guidance

In the second test, we examine whether managers of target firms are looking down on their future earnings and sales. In doing so, we can check whether target firms tend to have lower investments and growth regardless of the short-sellers' research reports. We obtain the sales and earnings forecast data from the I/B/E/S Guidance database, which compares managers' forecasts with analysts' consensus forecasts. If the forecast of a firm's manager beats (shortfalls) the consensus, I/B/E/S assigns a value of 2 (1) for this firm's guidance code. If the forecast matches the consensus, the value of the guidance code is 3. To facilitate our test, we change the value of the guidance code to 2 if the manager forecast matches the consensus and three if the forecast beats the consensus. As such, a higher value of the guidance code implies that managers tend to make optimistic forecasts relative to analysts.

Having obtained the I/B/E/S guidance data, we then run the OLS panel regression of guidance code for earnings and sales forecasts on a dummy variable $D_{i,t}$ that takes one if firms are targeted by short-seller research in the next year,

$$Guidance\_code_{i,t} = \beta_1 D_{i,t} + Year_t + Firm_i + Control_{i,t} + \epsilon_{i,t}, \tag{16}$$

where $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. A negative coefficient estimation on $D_{1,i,t}$ implies that managers are pessimistic about their future earnings and sales. Table A2 reports the regression results with controlling variables that are the same as those controlled in 7. We find that the coefficient on the dummy is not significant. Therefore, there is no evidence indicating that managers are looking down on their future earnings and sales before the publication of the short-seller research. In this case, it is unlikely that the downward change in investment and stock issuance coincides with short-seller research.

[Insert Table A2 about here]

5.2.3 Accounting fraud and real activity

Finally, we examine whether firms with accounting frauds not identified by short-seller research would not issue fewer stocks and reduce investments. Only the firms whose accounting frauds are mentioned by short-seller research experience a downward change in real activities. We use accounting fraud as an instrument here because most target firms have accounting frauds identified by short-sellers, as shown by our textual analysis. If it is the short-seller research that drives the downturn in the target firms' real activities, then we should expect that firms with accounting frauds that have not been identified by short-seller research will not have any changes in their real activities.

To test this, we obtain accounting fraud data from Audit Analytics' (AA) earnings restatement database for the period between 2007 and 2020. AA earnings restatement database assigns a dummy variable "RES_FRAUD" that takes one for the public firms with accounting frauds reported by the SEC. One issue is an overlap between the accounting-fraud firms mentioned in short-seller research reports and those identified by SEC. We exclude the overlapped ones when testing the accounting fraud identified by SEC.[23]

Having obtained the accounting fraud data, we then run the panel OLS regressions of changes in real activities on four dummy variables that take one if firms are listed in the short-seller research reports in the past one ($DS_{1,i,t}$), two ($DS_{2,i,t}$), three ($DS_{3,i,t}$), and four years ($DS_{4,i,t}$), respectively, and on another four dummies if firms have accounting fraud but

23. Using Audit Analytics' (AA) earnings restatement database, we obtain 138 firms with accounting fraud for the period between 2007 and 2020. We then match PERMNO to these firms using the CIK identifier ("COMPANY_FKEY"), and we have 85 observations left at the firm-year level. Among these 85 firms, three are also mentioned in short-seller research reports. Therefore, we exclude the three firms and have 82 observations left at the firm-year level for accounting fraud.

are not listed in the short-seller research reports in the past one ($DF_{1,i,t}$), two ($DF_{2,i,t}$), three ($DF_{3,i,t}$), and four years ($DF_{4,i,t}$), respectively,

$$\begin{aligned} Real\_activity_{i,t} = \beta_1 DS_{1,i,t} + \beta_2 DS_{2,i,t} + \beta_3 DS_{3,i,t} + \beta_4 DS_{4,i,t} + \beta_5 DF_{1,i,t} + \beta_6 DF_{2,i,t} \\ + \beta_7 DF_{3,i,t} + \beta_8 DF_{4,i,t} + \gamma Control_{i,t} + Year_t + Firm_i + \epsilon_{i,t} \end{aligned} \tag{17}$$

where $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively.

[Insert Table A3 about here]

Table A3 reports the regression results with control variables considered in Table 7. Consistent with our findings in Table 7, we find that firms who are the targets of short-seller research issue fewer stocks and reduce investment, and the effect remains significant even after four years of the publication of research reports. In sharp contrast, none of the coefficients on the dummies representing firms with accounting fraud are significantly negative. Therefore, only the firms whose accounting frauds are mentioned by short-seller research experience a downward change in real activities.

6 Conclusion

This paper uses textual analysis to investigate the content of short-sell research reports. We construct a comprehensive database of short-sell research reports and find that many relate to exposing accounting fraud or other financial misconduct. Using survey cash-flow forecasts as a counterfactual, we find that investors underreact to the negative cash-flow news contained in short-sell reports. On average, target firms earn abnormal returns of -4.7% on the publication day, and the subsequent price revisions equal -13% over a 12-month horizon. Moreover, we introduce a novel text-based fraud measure and find that reports more related to fraud are associated with larger negative long-term abnormal returns.

We propose a novel decomposition to quantify the importance of cash flow news in explaining the negative return surprises associated with the research reports. Using revisions in survey cash-flow forecasts, we find that cash-flow news accounts for 70% of the return surprises, with the remaining attributed to discount rate news. Our decomposition highlights the importance of cash flow news, such as financial fraud, in the research reports that drive returns of target firms.

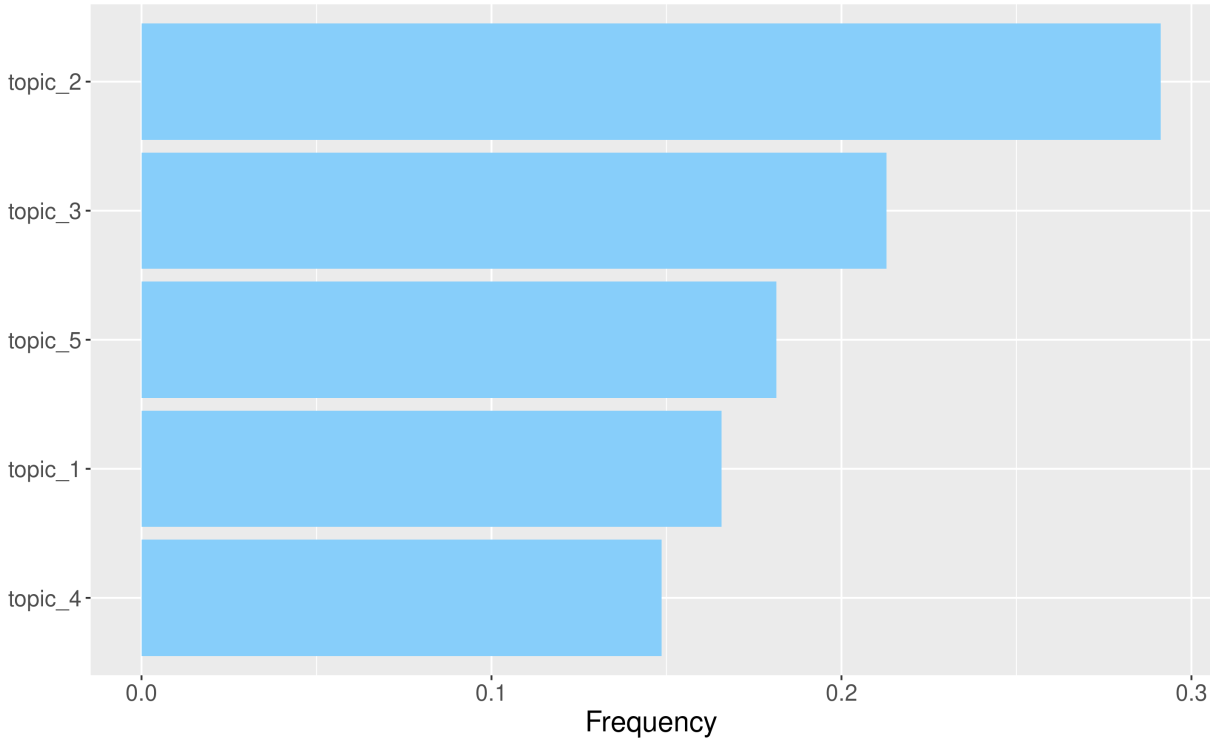
Finally, we provide evidence demonstrating that short-sell research and its subsequent stock-market effects are associated with significant real economic impact. Short-sell research reports are associated with substantial reductions in future real investment and stock issuances. While we cannot rule out the selection effect entirely, the evidence indicates the treatment effect drives more of the variation in real outcomes. We conclude that short-sell research affects financial market outcomes and significantly affects real economic allocations.

Figures and Tables

Figure 1: Word Cloud of Topics

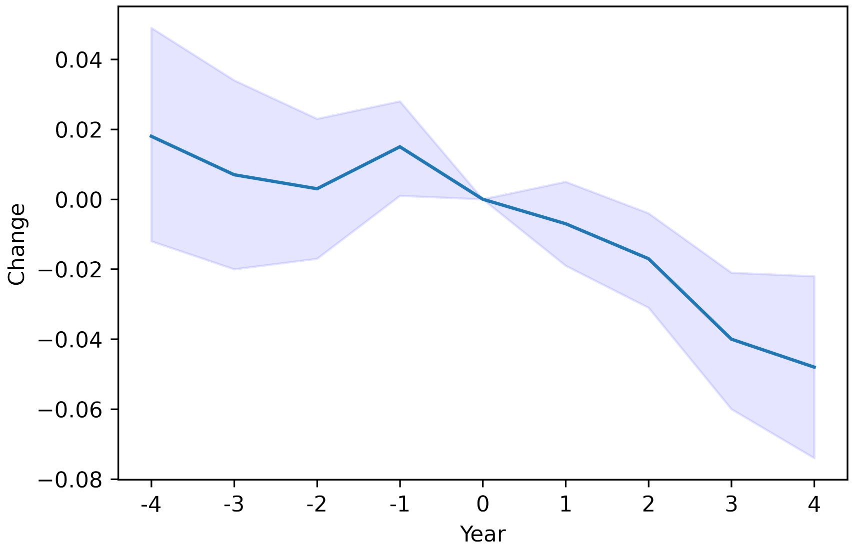
This figure plots the result of running LDA with five topics. Each cluster of words corresponds to each topic. Bigger words correspond to a higher-than-average probability of occurring within that topic.

Figure 2: Frequency of Topics

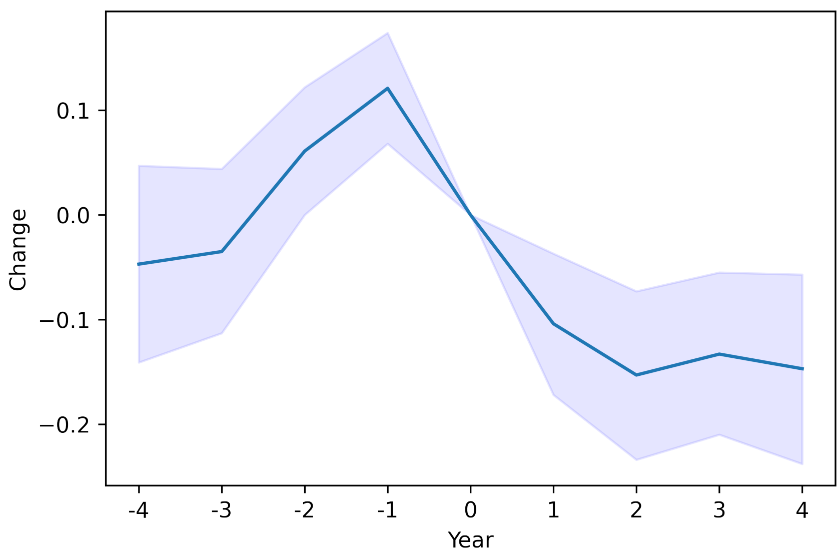


This figure plots the frequency of each topic. The frequency corresponds to the average percentage of the report that firms allocate to each topic.

Figure 3: Changes in Realized Profitability and Earnings



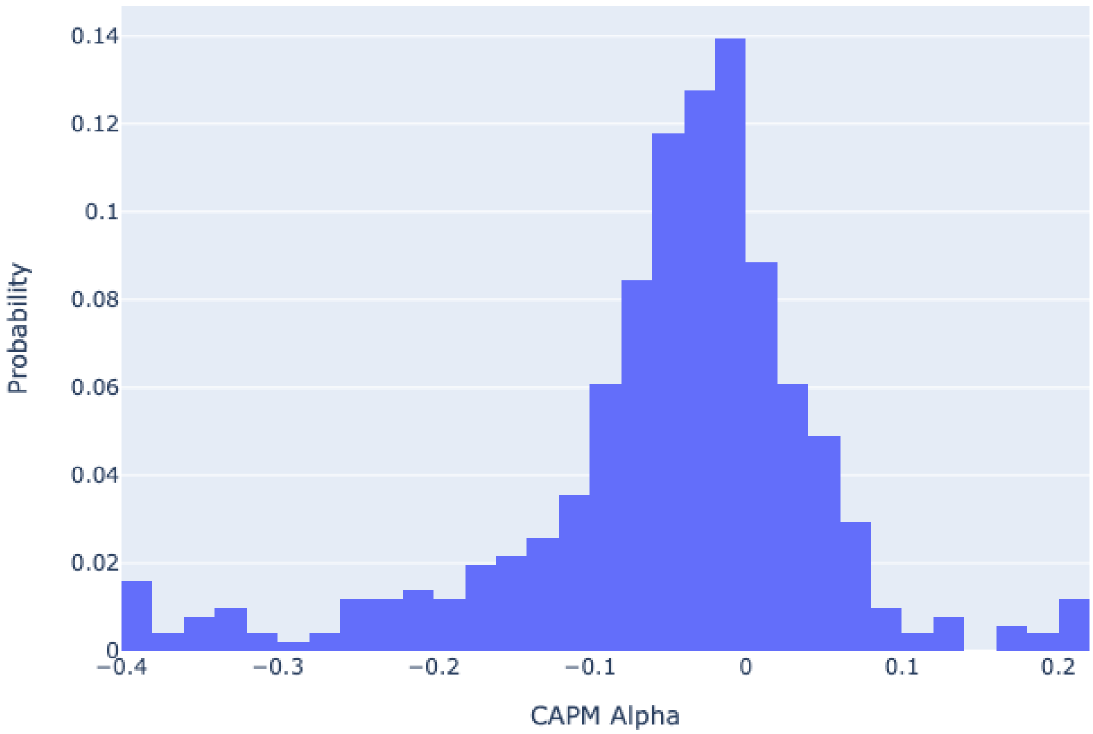
(a) Gross Profitability



(b) Return on Equity

This figure plots the time-varying changes in target firms' realized earnings and profitability. Year 0 is the year when short-seller research reports are published. The changes in earnings and profitability in each year t are measured as the differences between real activities in year t and real activities in year 0. The blue bar indicates the 95% confidence interval.

Figure 4: Distribution of Abnormal Returns on The Publication Day

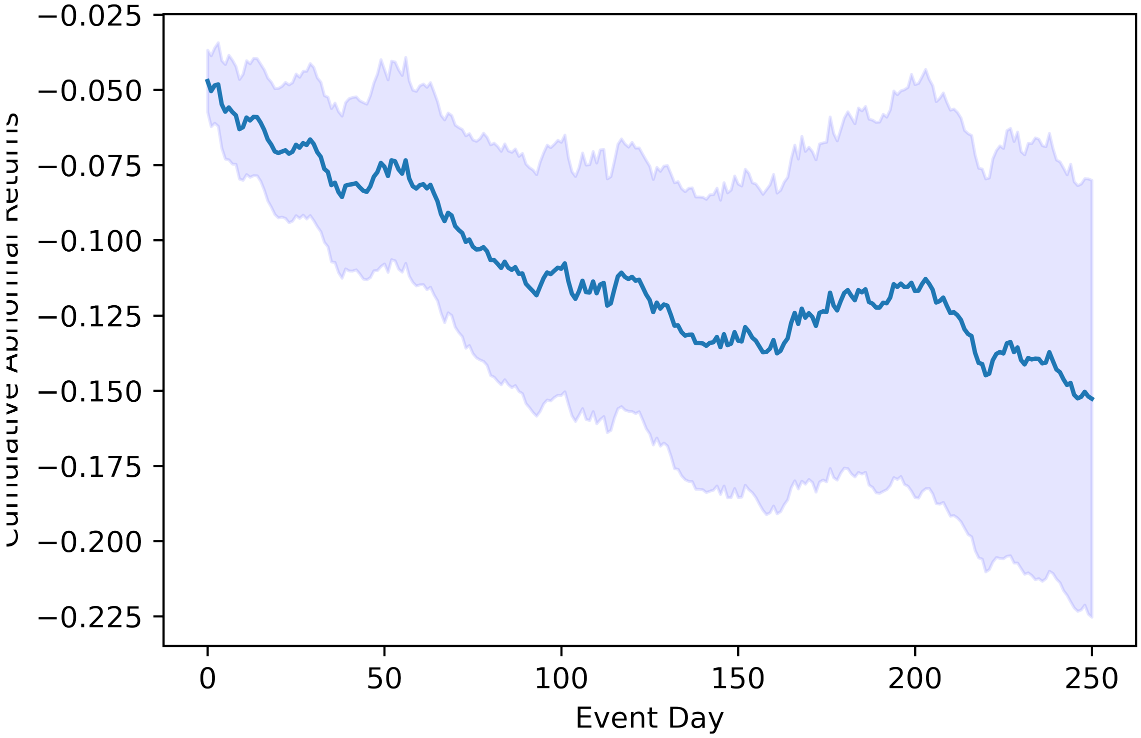


This figure plots the distribution of abnormal returns of target firms on the day when short-seller research reports are publicly available. We winsorize the abnormal returns at 1% in this figure.

Figure 5: Return Volatility

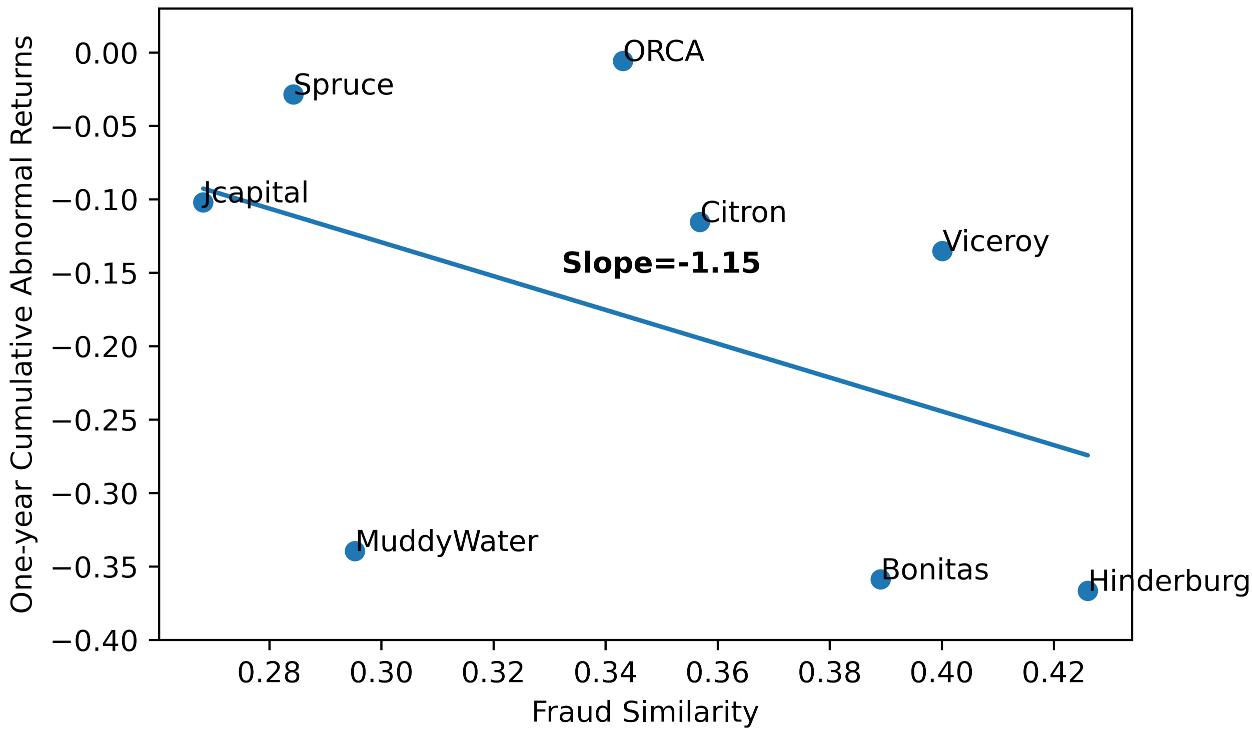
This figure plots the average return volatility of target firms during pre- and post-publication days. Event day 0 is the publication day of short-seller research reports.

Figure 6: One-year Cumulative Abnormal Return



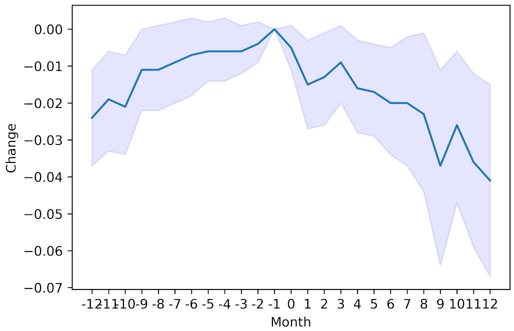
This figure plots the average of cumulative abnormal returns for the post-event window of 250 trading days. Event day 0 is the publication day of short-seller research reports. The blue bar indicates the 95% confidence interval.

Figure 7: Fraud Similarity and Alphas for Each Short-sell Research Firms

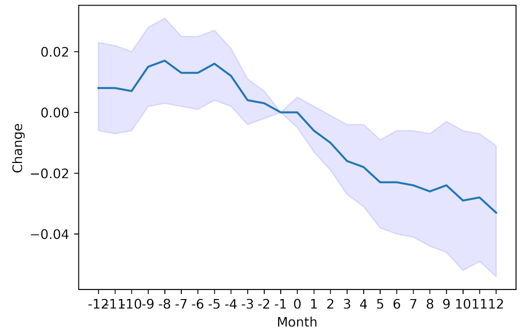


This figure plots the average of one-year cumulative CAPM alphas against the average of BERT fraud similarity for reports issued by each short-sell research firm.

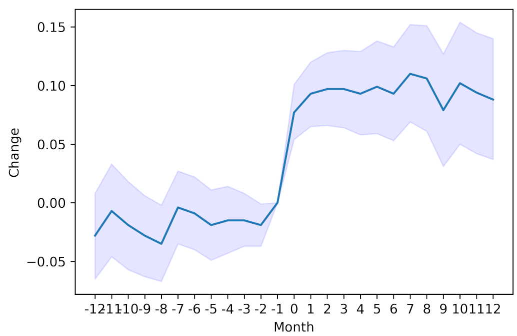
Figure 8: Subjective Cash Flow and Discount Rates Expectations



(a) One-year-ahead Forecast of Return on Equity



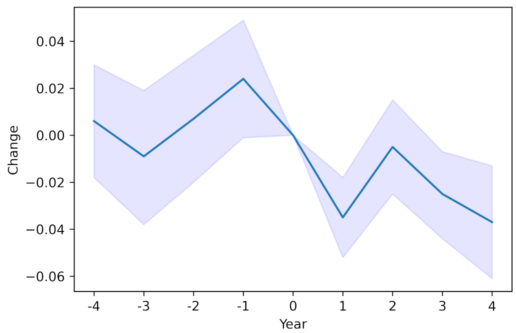
(b) Long-term Earnings Growth Expectation



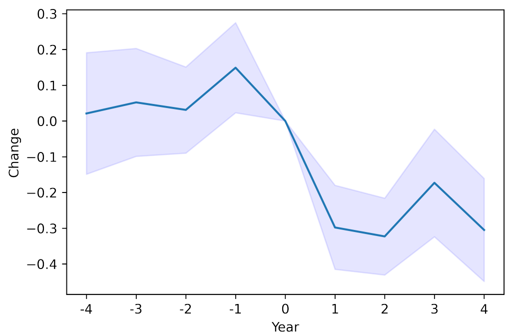
(c) One-Year-ahead Return Expectation

This figure plots the time-varying changes in target firms' a) return on equity, defined as the log of analysts' one-year-ahead earnings forecasts to lagged book equity, b) long-term earnings growth forecasts and c) the average of demeaned subjective discount rate, defined as the log of price targets (for the 12-month horizon) plus the annual dividend forecasts to current stock prices. Month 0 is the publication Month of short-seller research reports. The changes in subjective cash flow and discount rate expectations in month t are measured as the differences between values in month t and values in month 0. The blue bar indicates the 95% confidence interval.

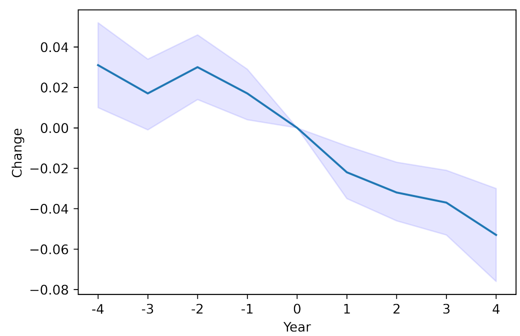
Figure 9: Changes in real activities



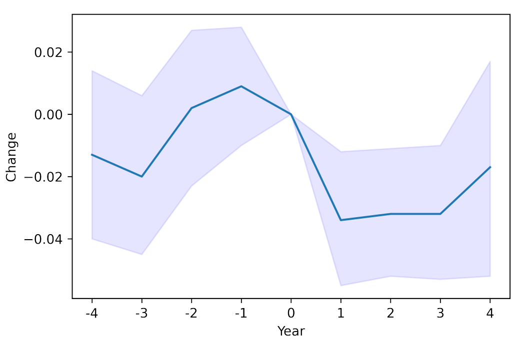
(a) Net Stock Issuance



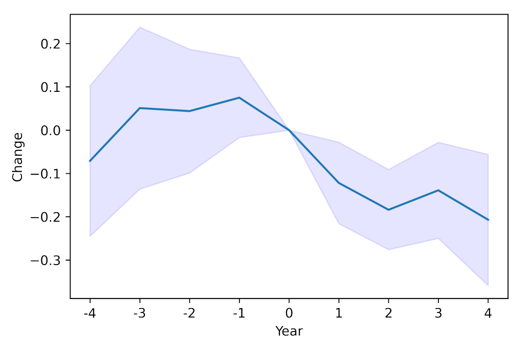
(b) Investment-to-Asset



(c) Capital and R&D Expenditure



(d) Changes in PPE and Inventory



(e) Sale Growth

This figure plots the time-varying changes in target firms' real activities in a window of eight years. Year 0 is the year when short-seller research reports are published. The changes in real activities in each year t are measured as the differences between real activities in year t and real activities in year 0. The blue bar indicates the 95% confidence interval.

Table 1: Firm Types

This table presents the summary statistics of the differences between the characteristics of firms that are subject to short-seller research reports and the cross-sectional mean/median of other firms.

Variable	Mean	Std	P5	P10	Q1	Median	Q3	P90	P95	*t*-stat
Panel A: Comparison to the mean										
Age	-88.271	134.641	-200.960	-194.663	-180.552	-144.497	-28.305	54.757	189.225	-12.508
LNsize	0.718	1.510	-1.772	-1.201	-0.345	0.674	1.774	2.767	3.193	9.062
NSI	0.036	0.196	-0.112	-0.095	-0.062	-0.039	0.069	0.226	0.419	3.053
GP	0.093	0.285	-0.215	-0.184	-0.091	0.032	0.182	0.490	0.742	5.830
ROE	0.115	0.438	-0.488	-0.213	-0.012	0.104	0.228	0.523	0.784	4.526
LNbeme	-0.640	1.040	-2.355	-1.921	-1.321	-0.594	0.010	0.536	0.945	-10.747
IA	0.412	1.096	-0.340	-0.279	-0.128	0.078	0.476	1.322	2.345	6.577
CAP_RD	0.026	0.174	-0.121	-0.106	-0.082	-0.043	0.078	0.228	0.384	2.581
PIA	0.053	0.161	-0.079	-0.059	-0.038	-0.007	0.066	0.245	0.400	5.604
SG	0.306	0.666	-0.323	-0.204	-0.053	0.107	0.443	1.026	1.454	7.819
Cost of Borrow Score	0.956	3.072	-1.677	-1.639	-1.554	-0.539	3.156	6.118	7.479	5.357
Panel B: Comparison to the median										
Age	-40.696	134.383	-156.000	-149.700	-132.250	-93.000	16.500	100.700	231.625	-5.778
LNsize	0.813	1.507	-1.645	-1.092	-0.227	0.782	1.867	2.819	3.323	10.277
NSI	0.095	0.199	-0.066	-0.028	-0.004	0.017	0.128	0.281	0.487	8.000
GP	0.126	0.285	-0.186	-0.156	-0.058	0.064	0.215	0.526	0.764	7.885
ROE	0.063	0.437	-0.544	-0.264	-0.069	0.055	0.169	0.467	0.721	2.483
LNbeme	-0.680	1.039	-2.404	-1.943	-1.353	-0.623	-0.018	0.504	0.931	-11.431
IA	0.560	1.109	-0.185	-0.113	0.010	0.194	0.629	1.509	2.566	8.850
CAP_RD	0.089	0.174	-0.048	-0.042	-0.020	0.020	0.142	0.287	0.438	8.904
PIA	0.081	0.162	-0.043	-0.032	-0.010	0.020	0.094	0.276	0.433	8.563
SG	0.397	0.666	-0.205	-0.105	0.034	0.199	0.523	1.122	1.546	10.152
Cost of Borrow Score	2.216	2.954	0.000	0.000	0.000	0.388	4.137	7.443	9.000	12.909

Table 2: Abnormal Returns

This table presents the summary statistics of abnormal returns of target firms on the date the short-seller research reports are publicly available. Panel A looks at the abnormal returns on the publication day. Panel B look at the one-year cumulative abnormal returns (CAR) following the publication day. "CAPM" and "FFC" indicate abnormal returns measured with CAPM and Fama-French-Carhart four factor model, respectively.

Summary	Mean	Std	P5	P10	Q1	Median	Q3	P90	P95	*t*-stat
Panel A: Alpha on the publication day										
CAPM Alpha	-0.049	0.109	-0.256	-0.179	-0.083	-0.032	0.003	0.052	0.073	-10.211
FFC Alpha	-0.049	0.109	-0.255	-0.182	-0.083	-0.033	0.003	0.052	0.071	-10.152
Panel B: One-year cumulative alpha										
CAPM CAR	-0.153	0.677	-0.992	-0.777	-0.589	-0.223	0.090	0.403	2.125	-4.121
FFC CAR	-0.146	0.639	-0.992	-0.774	-0.565	-0.210	0.084	0.422	2.455	-4.179

Table 3: Abnormal returns and fraud similarity

This table presents the abnormal returns conditional on text-based fraud measure. Panel A looks at the abnormal returns of stock portfolios sorted on fraud similarity. p-value indicates the difference in the mean abnormal returns between the portfolio with high fraud similarity and the portfolio with low fraud similarity. Panel B look at the regressions of abnormal returns on fraud similarity. We test both abnormal returns on the report-publication days and one-year cumulative abnormal returns following the report publications. "CAPM" and "FFC" indicate abnormal returns measured with CAPM and Fama-French-Carhart four-factors model, respectively. We look at the fraud similarity measured based on both BERT and word2vec model. Parentheses report t-statistics based on the heteroskedasticity-consistent standard errors of White (1980).

	BERT		word2vec	
	CAPM CAR	FFC CAR	CAPM CAR	FFC CAR
Panel A: Abnormal returns of portfolios sorted on fraud similarity				
High Fruad Similarity	-0.223	-0.212	-0.216	-0.21
	(-4.591)	(-4.201)	(-4.462)	(-4.223)
Low Fruad Similarity	-0.059	-0.076	-0.066	-0.078
	(-0.92)	(-1.474)	(-1.02)	(-1.496)
High- Low	-0.163	-0.136	-0.15	-0.131
p-value	0.044	0.061	0.064	0.071
Panel B: Regressing abnormal returns on fraud similarity				
Coefficient	-0.942	-0.804	-0.668	-0.645
	(-2.642)	(-2.292)	(-1.679)	(-1.734)

Table 4: Expectation stickiness

Column (1) presents the regression of forecast error on the forecast revision for the firms tagrted by short-sellers. Column (2) and (3) look at the regressions with time-fixed effect and firm-fixed effect, respectively. Parentheses report t-statistics based on the standard errors clustered at the firm and month level.

	(1)	(2)	(3)
Coeffi	0.19	0.186	0.122
	(3.309)	(3.244)	(2.065)
Time Fixed Effect	No	Yes	No
Firm Fixed Effect	No	No	Yes

Table 5: Summary statistics of return surprises and revisions in cash flow expectations

This table presents the summary statistics of return surprises and revisions in cash flow expectation for target firms.

	Mean	Std	P5	P10	Q1	Median	Q3	P90	P95	t-statistic
Return Surprise	-0.521	0.909	-2.316	-1.884	-0.990	-0.276	0.094	0.424	0.627	-9.184
CF Shock One-Year Horizon	-0.148	0.350	-0.776	-0.428	-0.152	-0.055	-0.005	0.050	0.122	-5.723
CF Shock Two-Year Horizon	-0.071	0.196	-0.354	-0.235	-0.099	-0.042	-0.000	0.053	0.110	-5.015
CF Shock Three-Year Horizon	-0.006	0.225	-0.244	-0.172	-0.069	-0.023	0.016	0.116	0.226	-0.376
CF Shock Four-Year Horizon	-0.027	0.161	-0.280	-0.208	-0.085	-0.018	0.021	0.120	0.180	-2.337
CF Shock Five-Year Horizon	-0.017	0.193	-0.264	-0.218	-0.095	-0.017	0.024	0.123	0.220	-1.188

Table 6: Return surprises and cash flow news

This table presents the results of the decomposition of return surprises using revisions in the cash-flow forecasts at a six-year horizon. The column "Cash Flow News" presents the beta estimated from the cross-sectional regression of cash-flow revisions at the individual horizon on the return surprise. The column "Cumulative Cash Flow News" presents the beta estimated from the cross-sectional regression of cumulative cash flow revisions on the return surprise. Parentheses report t-statistics based on the heteroskedasticity-consistent standard errors of White (1980).

	Cash Flow News	Cumulative Cash Flow News
One-years-ahead	0.12	0.12
	(2.52)	(2.52)
Two-years-ahead	0.04	0.159
	(2.243)	(2.858)
Three-years-ahead	0.032	0.189
	(3.633)	(3.341)
Four-year-ahead	0.03	0.217
	(2.452)	(3.888)
Five-years-ahead	0.021	0.236
	(1.528)	(4.17)

Table 7: Tests of real activities

This table reports the panel OLS regressions of real activities on four dummy variables that take one if firms are listed in the short-seller research reports in the past one ($D_{1,i,t}$), two ($D_{2,i,t}$), three ($D_{3,i,t}$), and four years ($D_{4,i,t}$), respectively. $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. Parentheses report t-statistics based on the standard errors clustered at the firm and year level.

$$Real\_activity_{i,t} = \beta_1 D_{1,i,t} + \beta_2 D_{2,i,t} + \beta_3 D_{3,i,t} + \beta_4 D_{4,i,t} + \gamma Control_{i,t} + Year_t + Firm_i + \epsilon_{i,t}$$

	NSI	IA	CAP_RD	ΔPI/A	SG
D_1	-0.042	-0.297	-0.022	-0.021	-0.091
	(-3.63)	(-4.191)	(-2.782)	(-2.013)	(-1.64)
D_2	-0.007	-0.261	-0.024	-0.001	-0.165
	(-0.513)	(-5.643)	(-2.482)	(-0.168)	(-3.923)
D_3	-0.024	-0.038	-0.007	-0.007	-0.081
	(-1.34)	(-0.272)	(-0.833)	(-0.972)	(-1.119)
D_4	-0.038	-0.17	-0.009	0.026	0.04
	(-3.555)	(-6.658)	(-0.695)	(1.639)	(0.239)
Adj R-sqr(%)	0.226	3.7	1.434	3.401	1.597
N	61597.0	64145.0	63740.0	53519.0	61338.0
Firm Fixed Effect	YES	YES	YES	YES	YES
Time Fixed Effect	YES	YES	YES	YES	YES
Control Variables	YES	YES	YES	YES	YES

References

Appel, Ian, and Vyacheslav Fos. 2020. "Active Short Selling by Hedge Funds." *SSRN Electronic Journal* (June). (Cited on pages 2, 4).

Araci, Dogu Tan. 2019. "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models" (August). (Cited on page 9).

Atilgan, Yigit, Turan G. Bali, K. Ozgur Demirtas, and A. Doruk Gunaydin. 2020. "Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns." *Journal of Financial Economics* 135, no. 3 (March): 725–753. ISSN: 0304405X. (Cited on page 5).

Baker, M., and J. Wurgler. 2007. "Investor sentiment in the stock market." *Journal of Economic Perspectives* 21:129–152. (Cited on page 5).

Baker, Malcolm, and Jeffrey Wurgler. 2013. "Behavioral corporate finance: An updated survey." In *Handbook of the Economics of Finance,* 2:357–424. Elsevier. (Cited on page 23).

Barberis, N., A. Shleifer, and R.W. Vishny. 1998. "A model of investor sentiment." *Journal of Financial Economics* 49:307–343. (Cited on page 5).

Binsbergen, Jules H. van, and Christian C. Opp. 2019. "Real Anomalies." *Journal of Finance* 74, no. 4 (August): 1659–1706. ISSN: 15406261. (Cited on page 23).

Bouchaud, Jean-philippe, Philipp Krueger, Augustin Landier, and David Thesmar. 2019. "Sticky expectations and the profitability anomaly." *The Journal of Finance* 74 (2): 639–674. (Cited on page 18).

Brendel, Janja, and James Ryans. 2021. "Responding to Activist Short Sellers: Allegations, Firm Responses, and Outcomes." *Journal of Accounting Research* 59, no. 2 (May): 487–528. ISSN: 1475-679X. (Cited on page 4).

Carhart, Mark M. 1997. "On persistence in mutual fund performance." *The Journal of finance* 52 (1): 57–82. (Cited on page 12).

Cen, Ling, K. C.John Wei, and Liyan Yang. 2017. "Disagreement, underreaction, and stock returns." *Management Science* 63, no. 4 (May): 1214–1231. ISSN: 15265501. (Cited on page 5).

Chen, Lei. 2016. "The Informational Role of Short Sellers: The Evidence from Short Sellers' Reports on US-Listed Chinese Firms." *Journal of Business Finance & Accounting* 43, nos. 9-10 (October): 1444–1482. ISSN: 1468-5957. (Cited on pages 2, 4).

Chen, Long, and Xinlei Zhao. 2009. "Return decomposition." *The Review of Financial Studies* 22 (12): 5213–5249. (Cited on page 20).

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review* 105 (8): 2644–2678. (Cited on pages 2, 17, 18).

Daniel, K.D, D. Hirshleifer, and A. Subrahmanyam. 1998. "Investor Psychology and Security Market under and overreactions." *Journal of Finance* 56:1839–1885. (Cited on page 5).

Derrien, François, Philipp Krueger, Augustin Landier, and Tianhao Yao. 2021. "ESG news, future cash flows, and firm value." *Swiss Finance Institute Research Paper,* nos. 21-84, (cited on page 15).

Fang, Vivian W, Allen H Huang, and Jonathan M Karpoff. 2016. "Short selling and earnings management: A controlled experiment." *The Journal of Finance* 71 (3): 1251–1294. (Cited on page 5).

Grullon, Gustavo, Sébastien Michenaud, and James P Weston. 2015. "The real effects of short-selling constraints." *The Review of Financial Studies* 28 (6): 1737–1767. (Cited on pages 3, 5, 23, 25).

Hirshleifer, D. 2001. "Investor psychology and asset pricing." *Journal of Finance* 56:1553–1597. (Cited on page 5).

Ikenberry, David L., and Sundaresh Ramnath. 2002. "Underreaction to Self-Selected News Events: The Case of Stock Splits." *Review of Financial Studies* 15, no. 2 SPEC. (January): 489–526. ISSN: 08939454. (Cited on page 5).

Jha, Manish, Hongyi Liu, and Asaf Manela. 2022. "Does Finance Benefit Society? A Language Embedding Approach." *SSRN Electronic Journal* (June). (Cited on page 9).

Kadiyala, Padmaja, and P. Raghavendra Rau. 2004. *Investor reaction to corporate event announcements: Underreaction or overreaction?,* 2, April. (Cited on page 5).

Kothari, Sagar P, Eric So, and Rodrigo Verdi. 2016. "Analysts' forecasts and asset pricing: A survey." *Annual Review of Financial Economics* 8:197–219. (Cited on page 15).

Ljungqvist, Alexander, and Wenlan Qian. 2016. "How Constraining Are Limits to Arbitrage?" *The Review of Financial Studies* 29, no. 8 (August): 1975–2028. ISSN: 0893-9454. (Cited on pages 2, 4, 9, 26).

Lopez-Lira, Alejandro. 2019. "Risk Factors That Matter: Textual Analysis of Risk Disclosures for the Cross-Section of Returns." *SSRN Electronic Journal* (September). ISSN: 1556-5068. (Cited on page 10).

Massa, Massimo, Bohui Zhang, and Hong Zhang. 2015. "The invisible hand of short selling: Does short selling discipline earnings management?" *The Review of Financial Studies* 28 (6): 1701–1736. (Cited on page 5).

Milian, Jonathan A. 2015. "Unsophisticated arbitrageurs and market efficiency: Overreacting to a history of underreaction?" *Journal of Accounting Research* 53, no. 1 (March): 175–220. ISSN: 1475679X. (Cited on page 5).

Pástor, L'uboš, Meenakshi Sinha, and Bhaskaran Swaminathan. 2008. "Estimating the intertemporal risk–return tradeoff using the implied cost of capital." *The Journal of Finance* 63 (6): 2859–2897. (Cited on page 53).

Paugam, Luc, Hervé Stolowy, and Yves Gendron. 2021. "Deploying Narrative Economics to Understand Financial Market Dynamics: An Analysis of Activist Short Sellers' Rhetoric*." *Contemporary Accounting Research* 38, no. 3 (September): 1809–1848. ISSN: 1911-3846. (Cited on page 4).

Poteshman, Allen M. 2001. "Underreaction, overreaction, and increasing misreaction to information in the options market." *Journal of Finance* 56, no. 3 (June): 851–876. ISSN: 00221082. (Cited on page 5).

Vuolteenaho, Tuomo. 2002. "What drives firm-level stock returns?" *The Journal of Finance* 57 (1): 233–264. (Cited on pages 19, 22).

Wang, Kemin, Rencheng Wang, Kuo-Chiang John Wei, Bohui Zhang, and Yi Zhou. 2018. "Insider Sales under the Threat of Short Sellers: New Hypothesis and New Tests." *SSRN Electronic Journal* (April). ISSN: 1556-5068. (Cited on page 5).

Wang, Xue, Xuemin Sterling Yan, and Lingling Zheng. 2020. "Shorting flows, public disclosure, and market efficiency." *Journal of Financial Economics* 135 (1): 191–212. (Cited on page 25).

White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: journal of the Econometric Society,* 817–838. (Cited on pages 43, 45).

Wong, Yu Ting Forester, and Wuyang Zhao. 2017. "Post-Apocalyptic: The Real Consequences of Activist Short-Selling." *SSRN Electronic Journal* (March). (Cited on page 5).

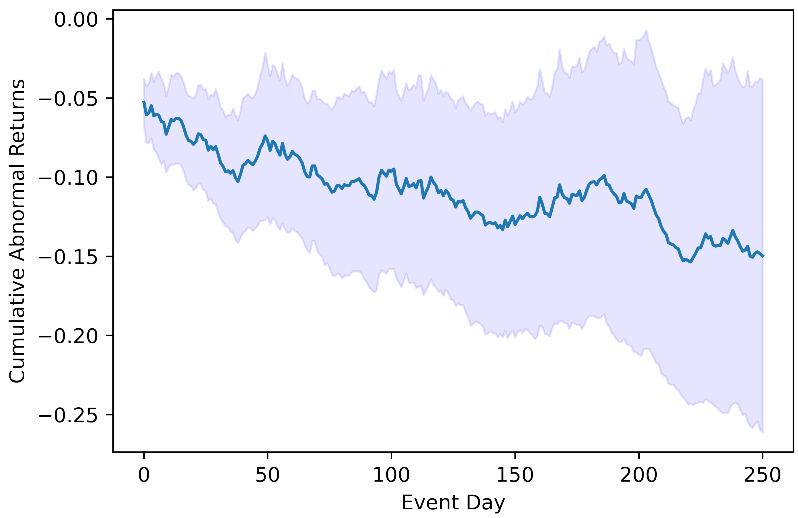
Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* 2nd ed., vol. 1. The MIT Press. (Cited on page 24).

Zhao, Wuyang. 2020. "Activist Short-Selling and Corporate Opacity." *SSRN Electronic Journal* (May). ISSN: 1556-5068. (Cited on page 2).

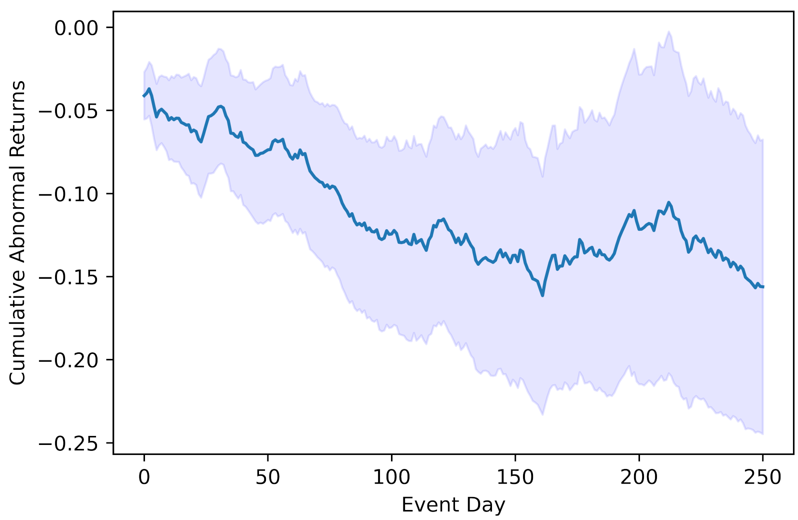
Appendix

A. Past Abnoraml Returns and Post-event Abnoraml Returns

Figure A1: Post-Event Cumulative Abnormal Return



(a) Positive Past Returns

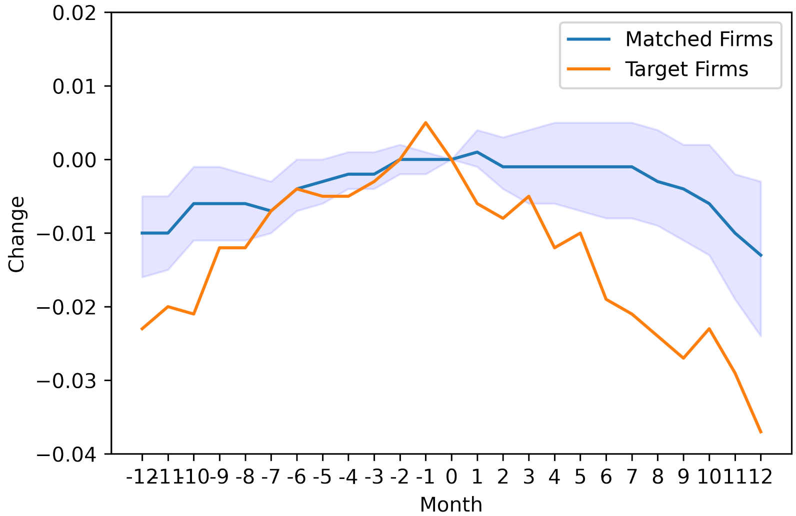


(b) Negative Past Returns

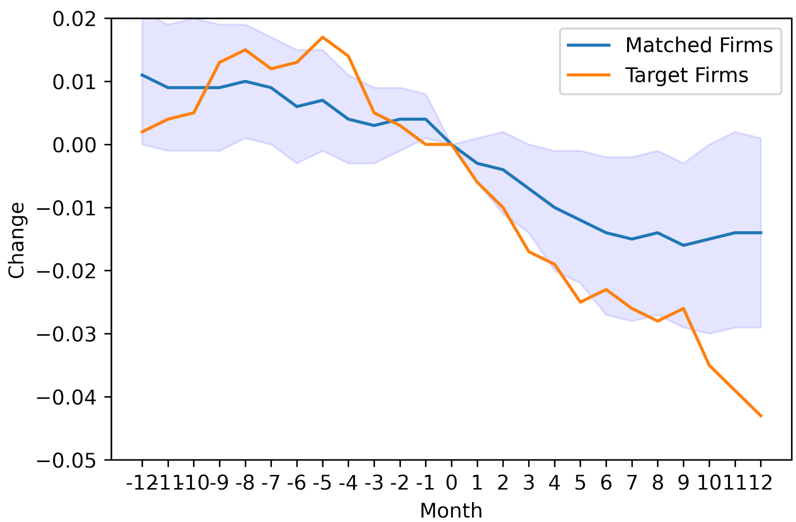
Figures (a) and (b) plot the average one-year cumulative abnormal returns of target firms whose cumulative abnormal returns before publication day are positive and negative, respectively. Event day 0 is the publication day of short-seller research reports. The blue bar indicates the 95% confidence interval.

B. Cash Low Forecsts Revision for Matched Firms

Figure A2: Cash Low Forecsts Revision for Matched Firms



(a) One-year-ahead Forecst of Return on Equity



(b) Long-term Earnings Growth Forecasts

This figure plots the time-varying changes in a) return on equity, defined as the log of analysts' one-year-ahead earnings forecasts to lagged book equity, b) long-term earnings growth forecasts for firms that match target firms' past cash flow forecasts. Month 0 is the publication month of short-seller research reports for target firms. The changes in subjective cash flow in month t are measured as the differences between values in month t and values in month 0. The blue bar indicates the 95% confidence interval.

C. Variables in Return Decomposition

I/B/E/S analysts provide one-year-, two-year-, and three-year-ahead forecasts for annual earnings per share. Analysts also provide long-term earnings growth forecasts (LTG) meant to forecast earnings growth over the next "three-to-five years".[24] Because of the present-value identity that requires forecast horizons with exact years, we interpolate across the different horizons to obtain a precise expectation over the next 12 months following the response of the analyst. For example, if the fiscal year of Firm A ends nine months after the forecast date, we may only have a nine-month earnings expectation and a 21-month earnings expectation for that firm. We interpolate these two measures to ensure that every expectation is exactly 12 months ahead (FE_1). We use an analogous procedure to construct two-year expectations (FE_2). To obtain exact three-year earnings expectations (FE_3), we multiply two-year expectations by one plus long-term earnings growth forecasts. We repeat this procedure to get exact four- (FE_4) and five-year earnings expectations (FE_5). To measure the revisions in expected roe, we obtain the forecast of book value per share from I/B/E/S. I/B/E/S analysts provide one-year- (FB_1), two-years- (FB_2), and three-years-ahead (FB_3) forecasts for annual book equity per share. We measure four-year book equity forecasts (FB_4) by assuming FB_3 grows at a rate between FB_3 and FB_2.

Based on these earnings and book value forecast data, we construct revision for expected

24. Following Pástor, Sinha, and Swaminathan (2008), we try to infer the missing two- and three-year forecasts from the available one-year forecasts and long-term growth forecasts. Specifically, if FE_{t+2} is not available but FE_{t+1} and LTG are available, we compute $FE_{t+2} = FE_{t+1} \times (1 + LTG)$. If FE_{t+3} is not available but FE_{t+2} and FE_{t+1} or LTG are available, we measure $FE_{t+3} = FE_{t+2} \times \frac{FE_{t+2}}{FE_{t+1}}$ and $FE_{t+3} = FE_{t+2} \times (1 + LTG)$, in this order. We also infer the missing LTG based on $FE_{t+3}/FE_{t+2} - 1$ and $FE_{t+2}/FE_{t+1} - 1$ (if available), in this order.

roe at forecast horizons of one-to-five years,

$$roe_t - E_{t-1}(roe_t) = log(1 + X_t/B_{t-1}) - log(1 + E_{t-1}(FY_1)/B_{t-1}) \tag{A1}$$

$$E_t(roe_{t+1}) - E_{t-1}(roe_{t+1}) = log(1 + E_t(FE_1)/B_t) - log(1 + E_{t-1}(FE_2/FB_1)) \tag{A2}$$

$$E_t(roe_{t+2}) - E_{t-1}(roe_{t+2}) = log(1 + E_t(FE_2/FB_1)) - log(1 + E_{t-1}(FE_3/FB_2)) \tag{A3}$$

$$E_t(roe_{t+3}) - E_{t-1}(roe_{t+3}) = log(1 + E_t(FE_3/FB_2)) - log(1 + E_{t-1}(FE_4/FB_3)) \tag{A4}$$

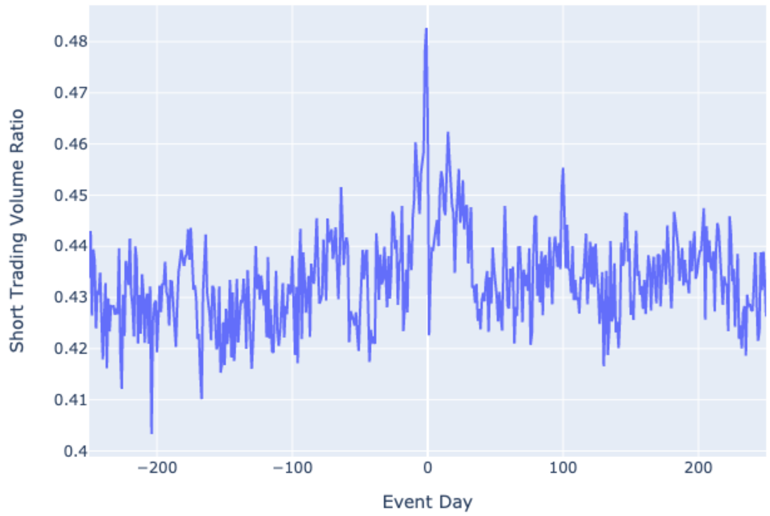
$$E_t(roe_{t+4}) - E_{t-1}(roe_{t+4}) = log(1 + E_t(FE_4/FB_3)) - log(1 + E_{t-1}(FE_5/FB_4)) \tag{A5}$$

where E_{t-1} and E_t denote the expectations formed one-month before the publications of reports and 11-month after the publications of reports. Variables without $E[\cdot]$ denote the realized values that can be observed, e.g., X_t are realized earnings per share 11-month after the publications of reports and B_{t-1} is the lagged book equity per share.[25] We winsorize the variables at 1% to exclude the effects of extreme values in our regression-based decomposition in Equation (13).

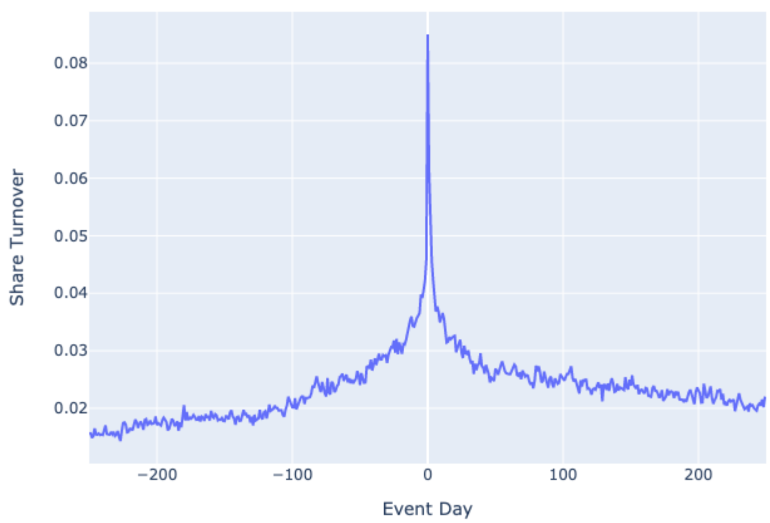
25. To account for possible data errors or extreme outliers, we exclude firms where the ratio of earnings to lagged book equity is less than -1 or greater than 1.

D. Short-sale Volume and Share Trading Volume

Figure A3: Post-Event Cumulative Abnormal Return



(a) Short-sale Volume



(b) Share Trading Volume

Figures (a) and (b) plot the average daily short-sale volume and share turnover of target firms, respectively. Event day 0 is the publication day of short-seller research reports.

E. Robustness tests of real activities

Table A1: Analysts' forecasts of investments and cash flows before the publications of short-seller reports

This table reports the panel OLS regressions of analysts' forecasts of capital expenditures and cash flow per share on an dummy variable ($D_{i,t}$) that takes one if firms are targeted by short-seller research in the next year. $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. Parentheses report t-statistics based on clustered standard errors.

$$Analysts\_forecasts_{i,t} = \beta_1 D_{i,t} + Year_t + Firm_i + Control_{i,t} + \epsilon_{i,t}$$

	Panel A: Capital Expenditures		Panel B: Cash Flows	
Horizon	One-year-ahead	Two-years-ahead	One-year-ahead	Two-years-ahead
$D_{i,t}$	0.013	0.027	1.187	1.83
	(0.509)	(0.597)	(1.33)	(1.724)
Adj R-sqr(%)	0.28	0.359	2.262	1.153
N	39157.0	35780.0	32764.0	31154.0
Firm Fixed Effect	YES	YES	YES	YES
Time Fixed Effect	YES	YES	YES	YES
Control Variables	YES	YES	YES	YES

Table A2: Guidance code for earnings and sales forecasts

This table reports the panel OLS regressions of guidance code for earnings and sales forecasts on an dummy variable ($D_{i,t}$) that takes one if firms are targeted by short-seller research in the next year. $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. t-statistic are measured based on clustered standard errors.

$$Guidance\_code_{i,t} = \beta_1 D_{i,t} + Year_t + Firm_i + Control_{i,t} + \epsilon_{i,t},$$

	Earnings	Sales
$D_{i,t}$	0.052	-0.026
	(0.436)	(-0.317)
Adj R-sqr(%)	0.205	0.126
N	13243.0	14307.0
Firm Fixed Effect	YES	YES
Time Fixed Effect	YES	YES
Control Variable	YES	YES

Table A3: Accounting fraud and changes in real activities

This table reports the panel OLS regressions of changes in real activities on four dummy variables that take one if firms are listed in the short-seller research reports in the past one ($DS_{1,i,t}$), two ($DS_{2,i,t}$), three ($DS_{3,i,t}$), and four years ($DS_{4,i,t}$), respectively, and on another four dummies if firms have accounting fraud but are not listed in the short-seller research reports in the past one ($DF_{1,i,t}$), two ($DF_{2,i,t}$), three ($DF_{3,i,t}$), and four years ($DF_{4,i,t}$), respectively, and on another four dummies if . $Year_t$ and $Firm_i$ denote time and firm fixed effects, respectively. Parentheses report t-statistics based on clustered standard errors.

$$Real\_activity_{i,t} = \beta_1 DS_{1,i,t} + \beta_2 DS_{2,i,t} + \beta_3 DS_{3,i,t} + \beta_4 DS_{4,i,t} + \beta_5 DF_{1,i,t} + \beta_6 DF_{2,i,t} \\ + \beta_7 DF_{3,i,t} + \beta_8 DF_{4,i,t} + \gamma Control_{i,t} + Year_t + Firm_i + \epsilon_{i,t}$$

	NSI	IA	CAP_RD	ΔPI/A	SG
DS_1	-0.042	-0.297	-0.022	-0.021	-0.091
	(-3.653)	(-4.916)	(-3.086)	(-2.268)	(-1.806)
DS_2	-0.007	-0.261	-0.024	-0.001	-0.166
	(-0.546)	(-7.898)	(-2.528)	(-0.167)	(-4.034)
DS_3	-0.024	-0.037	-0.007	-0.007	-0.081
	(-1.414)	(-0.34)	(-1.076)	(-1.424)	(-1.345)
DS_4	-0.038	-0.169	-0.009	0.026	0.04
	(-3.507)	(-6.144)	(-0.855)	(1.762)	(0.264)
DF_1	0.016	0.002	0.003	0.012	-0.03
	(0.723)	(0.03)	(0.587)	(0.615)	(-0.496)
DF_2	0.047	-0.046	-0.01	0.008	-0.063
	(0.992)	(-1.59)	(-0.883)	(0.857)	(-0.743)
DF_3	0.002	0.04	0.0	0.018	-0.028
	(0.095)	(0.736)	(0.006)	(2.584)	(-0.457)
DF_4	0.015	0.053	-0.002	0.003	-0.02
	(0.557)	(0.809)	(-0.098)	(0.171)	(-0.265)
Adj R-sqr(%)	0.225	3.695	1.429	3.396	1.592
N	61597.0	64145.0	63740.0	53519.0	61338.0
Firm Fixed Effect	YES	YES	YES	YES	YES
Time Fixed Effect	YES	YES	YES	YES	YES
Control Variable	YES	YES	YES	YES	YES