Financial Markets and Portfolio Management
https://doi.org/10.1007/s11408-023-00442-1

Short selling and firm investment efficiency

Chang Yu[1]

Accepted: 7 December 2023

Abstract
This paper investigates the informativeness of short sales on detecting firm investment inefficiency. Neoclassical and agency theory suggest that investment inefficiency destroys firm value by allocating resources to less-valued uses. This paper finds that short-sellers adjust their short positions before the announcement of a financial statement, to use their information advantage on firm investment inefficiency. The relation between the short positions in a firm and its future investment inefficiency is both statistically and economically significant, and robust to a broad set of control variables. Subsample analyses show that the informativeness of short sales positions about future investment inefficiency is concentrated on overinvestment firms, firms with little board independence, and firms with low CEO incentive pay.

Keywords Short selling · Investment efficiency · Capital investment · Corporate governance

JEL Classification G14 · G31

1 Introduction

The information content of short sales has attracted attention due to the tremendous growth of short-selling activities by hedge funds since 2004[1] . This research helps us understand whether short-sellers are informed and whether they improve price

[1] One reason is the boom of hedge funds that undertook short-selling activities on a large scale. As Jiao et al. (2016) document, the correlation between changes in hedge fund ownership and short interest was as high as 0.75, while the correlation between changes in mutual fund ownership and short interest was only 0.18; therefore, the emergence of hedge funds reshaped short-selling activities. Saffi and Sigurdsson (2011) also indicate that the short-selling market changed immensely during the global financial crisis. In January 2021, millions of retail investors rallied on the online community leading to the short squeeze of Gamestop's stock and other securities with a large short position. The short squeeze caused tremendous losses for short-sellers like hedge funds. This event drew the world's attention to short-selling activities and brought the debate on the real consequences of short-selling activities.

✉ Chang Yu
chang.yu@otago.ac.nz; yuc219797@gmail.com

1 University of Otago, Wellington, New Zealand

Published online: 07 February 2024

efficiency. The literature indicates that short-sellers can discern information about future stock returns (Boehmer et al. 2008; Diether et al. 2009; Chang et al. 2014; Jiao et al. 2016), as well as detect fraud and earnings manipulation (Desai et al. 2006; Karpoff and Lou 2010; Hirshleifer et al. 2011; Boehmer and Wu 2013; Massa et al. 2015; Fang et al. 2016; Park 2017). However, less attention has been devoted to the relation between short-selling and firm investment inefficiency, which has been documented as being value-destroying (Biddle and Hilary 2006; Biddle et al. 2009; Bushman et al. 2011; Balakrishnan et al. 2014; Goodman et al. 2014).

Despite the substantial growth of short-selling activities, short-selling remains a contentious practice and has been restricted in many countries during recent financial crises (Massa et al. 2015; Caby et al. 2020). Both regulators and researchers have intensified their focus on comprehending the impact of short-selling on various corporate behaviors. Proponents of short-selling argue that its primary benefits stem from its disciplinary effect, particularly its influence on curbing a firm's value-destructive investments (Chang et al. 2019). This effect hinges on the assumption that short-sellers will respond to a firm's investment inefficiencies.

Conversely, Grullon et al. (2015) contend that increased short-selling activities can lead to a decline in stock prices, prompting firms to react by reducing their investments. In this scenario, the reduction in investments is a response to a downward price manipulated by uninformed bear raiders rather than a result of the disciplinary effect. Consequently, the ongoing debate revolves around the question of whether short-sellers possess information about a firm's investment inefficiency, thereby warranting further investigation into the distinctive nature of investment inefficiency relative to other dimensions of operating performance.

In general, investment inefficiency constitutes a crucial aspect of investment performance. From a conceptual standpoint, inefficient investment entails the allocation of capital to less valuable uses, impeding the maximization of firm value (McNichols and Stubben 2008; Biddle et al. 2009; Bushman et al. 2011; Balakrishnan et al. 2014; Goodman et al. 2014; Gao and Yu 2020). Previous research has underscored the economic significance of investment inefficiency. For instance, Cooper et al. (2008) and Titman et al. (2004) have demonstrated that investment inefficiency, particularly in the form of overinvestment, is correlated with subsequent poor stock performance, which in turn attracts the attention of short-sellers.

Before delving deeper into the topic, it is crucial to understand the fundamental idea behind investment inefficiency. While prior literature has not explicitly identified a direct proxy for investment inefficiency (Benlemlih and Bitar 2018; Gao and Yu 2020), numerous studies have operationalized investment inefficiency as the deviation from the optimal level of investment (Biddle et al. 2009; Chen et al. 2011a; Gomariz and Ballesta 2014; Benlemlih and Bitar 2018; Yu 2023). This definition aligns with neoclassical theory, which posits that a firm's net present value is bound by certain constraints imposed by its production function and growth opportunities, consequently leading to an optimal and industry-specific operating size. Any deviations from this optimal size can give rise to inefficiencies (Gao and Yu 2020).

Value maximization is the key objective of a firm (Chen et al. 2011b). The agency theory explains how investment inefficiency destroys firm value, underpinning the importance of studying investment inefficiency. When information asymmetry exists, managers will possess superior information compared to shareholders. If the incentives of managers are not aligned with shareholders, managers may invest in their own interests rather than those of shareholders, either by rejecting positive net present value projects or by investing in negative net present value projects (Gao and Yu 2020). The theory of corporate investment suggests that value maximization requires a firm to invest based on the net present value principle (Chen et al. 2011b). Therefore, this scenario provides fundamental motivations for short-sellers to capture the information rent.

This paper investigates whether short-sellers can detect a firm's investment inefficiency before the release of its annual financial report. In particular, it focuses on whether short-sellers have an information advantage with which to make correct forecasts before the public can.[2] To do so, I follow Biddle et al. (2009) and McNichols and Stubben (2008) to use the deviations from the optimal investment as a measure of investment inefficiency. I then use the levels of abnormal short interest to measure short-selling activities.[3] I will address three main questions. First, does a firm's short-selling activity contain information about its investment inefficiency released in the upcoming annual report? Second, can short-sellers gain economic profits if they act according to such information? Third, does the information advantage of short-sellers regarding a firm's investment inefficiency depend on the characteristics of the firm? I document several interesting findings.

Firstly, short-sellers can identify investment inefficiencies before the annual financial report is released. When I sort firms based on their abnormal short interest before the annual report release,[4] the investment inefficiency measure in the top decile is significantly greater than that in the bottom decile. For instance, the disparity in the investment inefficiency measure from Biddle et al. (2009) between the top and bottom deciles of abnormal short interest is 0.078, with a *t*-value of 4.43. Multivariate panel regression yields consistent results. Using Biddle et al. (2009) measure, the coefficient for pre-annual report abnormal short interest is 0.003, significant at the 1% level. Coefficient dynamics highlight greater significance near annual report release. A case study on Texas Industries INC. provides anecdotal evidence. To tackle endogeneity, I employ a propensity score model to pair the sample firms with control firms sharing similar characteristics. The outcomes remain significant postmatching. Employing alternative investment inefficiency measures upholds the consistent results.

[2] Theoretically speaking, there are two sources of information advantage for informed investors. One is their superior ability to analyze publicly available information (Engelberg et al. 2012), while the other is their access to private information (Boehmer et al. 2020). I do not distinguish between these two in my analysis.

[3] I explain, in Appendix, how to calculate the abnormal short interest for a firm.

[4] I use the earnings report date of the fourth quarter in the Compustat Fundamentals Quarterly database as the proxy for the release date of the annual report because firms usually announce the last quarter report simultaneously with the annual report. The abnormal short interest is measured on the 15th of

Additionally, based on Jiao et al. (2016), I conduct a Fama–MacBeth regression annually. The empirical results demonstrate a notable economic significance in the information advantage. Specifically, employing the inefficiency measure from McNichols and Stubben (2008), the coefficient of investment inefficiency on the out-of-sample cumulative abnormal return is strongly negative at −0.017, with a highly significant *t*-value of −3.37, adjusted using the Newey–West method with three lags. Similar significant outcomes are observed with alternative investment inefficiency measures.

Thirdly, subsample analyses reveal that the correlation between short-selling activity and investment inefficiency focuses on overinvestment firms, those with limited board independence, and firms with low CEO incentive pay.[5] Initially, I categorize the sample into overinvestment and underinvestment firms, reflecting the two forms of investment inefficiency (Bebchuk and Stole 1993; Franzoni 2009). Upon forming these subsamples, the coefficient of abnormal short interest on investment inefficiency remains statistically significant for overinvestment firms, whereas it loses significance for underinvestment firms. This outcome aligns with the expectation that overinvestment is more easily identifiable externally (Bebchuk and Stole 1993) and is more connected to overvaluation (Titman et al. 2004). Further division of the sample based on board independence and CEO incentive pay demonstrates that the significant relationship between short-selling activity and investment inefficiency is limited to firms with scarce board independence and low CEO incentive pay. This could be attributed to the heightened susceptibility of such firms to short-selling activities. Rahman et al. (2021) argue that a lack of board independence facilitates short-selling due to increased information asymmetry, while lower CEO incentive pay aids short-sellers as managers are less cautious about stock price decline (De Angelis et al. 2017; Shi et al. 2021).

The present study contributes to the literature in several ways. First, it sheds light on the informational content of short sales by investigating a new perspective on firm performance. The literature has addressed the information content of short-selling activity for stock returns and firms' misconduct. The present paper fills some gaps by first documenting the relation between short-selling activity and firm investment inefficiency.

Secondly, the present research contributes to the literature on how short-selling impacts corporate investments. As discussed earlier, there are two competing views on how short selling reshapes a firm's investments. According to Grullon

Footnote 4 (continued)
each month or the preceding business day if the 15th is not a business day. The lag between the date of abnormal short interest and the date of the annual release of the financial report should not exceed one month.

5 While several measures exist to assess corporate governance, such as CEO-chairman duality, E-index, and financial slack, the current study concentrates on board independence and CEO incentive pay due to their substantial connection with short-selling activities, as emphasized in recent research (Rahman et al. 2021; De Angelis et al. 2017; Shi et al. 2021). It is important to acknowledge that other aspects of corporate governance may also influence corporate investments. However, the primary objective of this paper is to investigate the specific firm characteristics that impact the informativeness of short-sellers.

et al. (2015), short-sellers, acting as bear raiders, distort the firm's investment by artificially reducing the stock price. In this case, short-sellers manipulate the stock price and are uninformed (Goldstein and Guembel 2008; Grullon et al. 2015). On the other hand, some researchers argue that short-selling has a disciplinary effect on a firm's value-destroying investments (Chang et al. 2019). The underlying assumption behind such a disciplinary effect is that short-sellers are informed about the firm's inefficient investments. This paper supports the latter proposition by demonstrating that short-sellers possess information about a firm's investment inefficiency.

Thirdly, the present study is closely related to the works of Hirshleifer et al. (2011) and Park (2017), which examine short-selling activities following the release of annual financial reports. They discover that short-sellers possess a superior ability to analyze negative information related to the accruals anomaly or profit subsequent to the report. Trading strategies based on this enhanced information processing yield significant returns. In contrast to Hirshleifer et al. (2011) and Park (2017), I demonstrate that informativeness also arises from an information advantage preceding the public release of the report. These findings complement those documented in Hirshleifer et al. (2011) and Park (2017), further enhancing the comprehension of short-sellers.

Fourthly, it supports the proposition in the literature that differences exist between overinvestment and underinvestment firms (Bebchuk and Stole 1993; Franzoni 2009). The study delves deeply into the side of investment inefficiency for which short-sellers possess information. It also enhances the understanding of overinvestment, including its connection with overvaluation and external financial markets. Moreover, there has been a growing focus on uncovering the link between corporate governance and short-selling in recent studies (De Angelis et al. 2017; Shi et al. 2021; Rahman et al. 2021). The present study provides further evidence that a lack of board independence and CEO-incentive pay facilitates short-selling activities.

The remainder of this paper is organized as follows. In Sect. 2, I provide a literature review and develop my hypotheses. In Sect. 3, I explain how to construct the measures of the level of abnormal short interest and investment inefficiency. Section 4 presents the data and the main empirical results. I show the results of several robustness tests in Sect. 5 and the subsample analyses in Sect. 6. Section 7 presents the conclusions.

2 Literature review and main hypotheses

2.1 Literature review

2.1.1 Short-selling

Short-selling has been carried out in major financial markets around the world for years, and it continues to grow. This growth was especially pronounced between 2004 and 2008 due to booms in hedge funds (Saffi and Sigurdsson 2011; Jiao et al. 2016). Much effort continues to go into studying whether short-sellers are informed and how short-selling affects a firm's behavior.

One category of research studies whether short-sellers possess information by focusing on whether short sales can predict future movements in stock prices. For example, Boehmer et al. (2008) investigate the predictive power of short-selling for next month's stock returns on the NYSE from January 2000 to April 2004. The study finds that heavily shorted stocks significantly underperform lightly shorted stocks in their returns.

Diether et al. (2009) also find that short-sellers correctly predict future negative abnormal returns. The study documents that stocks in the highest quintile of short-selling significantly underperform stocks in the lowest quintile during a window of two to five days in both the NYSE and NASDAQ. Similar studies are conducted by Chang et al. (2014) and Jiao et al. (2016).

Another category of the literature concerning whether short-sellers are informed investigates whether they can detect firms' misconduct. The literature demonstrates that managers have incentives to manipulate a firm's financial performance for personal compensation (Bergstresser and Philippon 2006), gains from stock sales (Beneish and Vargus 2002), job security, and control-related motives (DeFond and Park 1997). Short-sellers, as sophisticated informed investors, can identify these manipulated earnings or deviations from the firm's fundamental value.

For example, Hirshleifer et al. (2011) use operating accruals as a proxy for earnings manipulation and finds that stocks in the highest decile of accruals have significantly higher short interest than stocks in the lowest decile. Furthermore, stocks with high accruals earn significantly negative abnormal returns. Park (2017) employs abnormal cash flow from operations, abnormal production costs, and abnormal discretionary expenses as three alternative measures of real earnings management. They discover that firms with more real earnings management subsequently attract significantly higher short interest. This positive relationship is more pronounced for firms with high accrual-based earnings management costs, such as firms with low accounting flexibility or firms subject to stricter scrutiny by high-quality auditors.

Several findings have also been documented in the literature on how short-selling affects a firm's behavior. Grullon et al. (2015) show that short-selling reduces a small firm's equity issues and capital expenditures. De Angelis et al. (2017) indicate that short-selling can also affect a firm's compensation behavior using the Regulation SHO's pilot program, which removed various short-selling constraints for randomly selected stocks in the pilot group. By running a difference-in-difference test, they document that short-selling significantly increases the convexity of compensation payoffs, and the increase in convexity is due to using more stock options to compensate managers.

Moreover, the literature suggests that short-selling benefits the firm in terms of a disciplinary effect. Fang et al. (2016) use discretionary accrual as a proxy for earnings management. They find that short-selling disciplines a firm's earnings manipulations by running a difference-in-difference test on Regulation SHO's pilot program. Massa et al. (2015) confirm the finding by using lendable shares as the proxy for the short-selling threat. They also extend the study by using an international sample of 33 countries. Another example of the disciplinary effect of short-selling is on a firm's value-destroying investments. Chang et al. (2019) find that short selling improves the announcement returns of M &As. They also find that this effect is more pronounced when managers' wealth is more linked to the stock price and when deals are prone to agency problems.

2.1.2 Investment inefficiency

The disciplinary impact of short-selling on a firm's value-diminishing investments implies that short-selling can help eliminate overinvestment, which is often associated with a firm's investment inefficiency (Biddle and Hilary 2006). As investment inefficiency tends to decrease firm value, the literature predominantly focuses on its measurement.

Biddle and Hilary (2006) utilize cash flow sensitivity as a measure of investment inefficiency: The higher a firm's investment sensitivity to internally generated cash flow, the greater the investment inefficiency. Meanwhile, McNichols and Stubben (2008), Biddle et al. (2009), and Goodman et al. (2014) employ different models to determine optimal investment levels and use deviations from these levels as proxies for investment inefficiency: The greater the deviation, the higher the investment inefficiency. Other relevant studies in this area include those by Bushman et al. (2011), Balakrishnan et al. (2014), and Gao and Yu (2020).

The literature on investment inefficiency also examines the factors that can influence it. For example, Biddle et al. (2009) use accruals quality, modifications to accruals quality, and financial disclosure transparency as proxies for a firm's financial reporting quality. They find that financial reporting quality reduces both overinvestment and underinvestment. They argue that this effect results from the reduced information asymmetry between firms and external suppliers of capital.

In a similar vein, Goodman et al. (2014) measure managerial forecasting quality using the accuracy of managers' external forecasts. They discover a negative association between the ability of managers to estimate payoffs of potential projects and investment inefficiency.

Furthermore, Chen et al. (2011b) reveal that investment inefficiency is significantly higher for state-owned enterprises compared to non-state-owned enterprises. Moreover, when top executives in state-owned enterprises have a government background, investment inefficiency is significantly higher. These findings suggest that government intervention increases a firm's investment inefficiency, adding to the challenges posed by information asymmetry and agency problems.

2.2 Main hypotheses

The question of interest in this paper is whether short-sellers have an information advantage about a firm's investment inefficiency before the public. Generally speaking, if short-sellers are informed about the investment inefficiency before it becomes publicly available, it should have a significant relationship between short-selling activities preceding the release of the financial report and the information about inefficiency contained in the report. The level of abnormal short interest will increase as the investment inefficiency worsens. The primary hypothesis is as follows:

Hypothesis 1 If short-sellers are informed about firm investment inefficiency, then a firm's short-selling activity will be positively related to its future investment inefficiency.

Investment inefficiency manifests itself mainly in two ways: either overinvestment or underinvestment. There are three main differences documented in the literature regarding overinvestment and underinvestment.

First, Bebchuk and Stole (1993) show that underinvestment occurs when the market has incomplete information about the level of investment undertaken, while overinvestment occurs when the market can observe the level of investment but not its productivity. Following this signaling theory, if investment inefficiency can be detected by short-sellers, it should be more related to overinvestment rather than underinvestment.

Second, Gilchrist et al. (2005) and Grullon et al. (2015) document that overvaluation leads to overinvestment by artificially reducing the cost of capital. Therefore, overinvestment reflects some information about overvaluation. As a result, short-sellers search for overvalued stocks to make a profit, so they pay more attention to overinvesting stocks.

Third, according to Franzoni (2009), the incentives behind underinvestment and overinvestment are different. Overinvestment occurs when management has empire-building incentives and when corporate governance for the firm is poor, both of which provide negative signals. On the other hand, the motivation behind underinvestment is not clear. It might happen when firms are financially constrained, leading to insufficient funds for investment. Under this case, the signal is negative, and market reactions tend to be significant. However, underinvestment could also be due to other reasons, with market reactions being indistinguishable from zero. These two effects make the response to underinvestment unclear.

Following these arguments, the second hypothesis is formulated as follows:

Hypothesis 2 The positive relationship between short-selling activities and future investment inefficiency is stronger for "overinvestment firms".

Recently, researchers find that board independence can significantly impact short-selling activities. For instance, Rahman et al. (2021) argue that a lack of board independence facilitates short-selling due to increased information asymmetry. Additionally, studies suggest that insufficient board independence leads to a deterioration in disclosure quality (Byrd and Hickman 1992; Fama and Jensen 1983; Beasley 1996; Yekini et al. 2015). Short-sellers are known as traders who exploit their informational advantage arising from the information asymmetry. Consequently, a lack of board independence may enable short-sellers to capitalize on their informational edge concerning investment inefficiencies.

Furthermore, board independence can mitigate a firm's investment inefficiency (Liu et al. 2015), indicating that firms lacking board independence are more susceptible to such inefficiencies. Souther (2021) finds that firms with less board independence are associated with lower firm value, making them more attractive to short-sellers. In other words, the lack of board independence enhances the informativeness of short-sellers regarding a firm's investment inefficiencies.

Hypothesis 3 The positive relation between short-selling and future investment inefficiency is stronger for 'firms with little board independence'.

Short-selling is closely related to managers' equity compensations (De Angelis et al. 2017; Shi et al. 2021) because more equity-based compensation can mitigate short-selling threats by aligning managers' personal interests with corporate stock prices. Firms with less equity-based compensation are more vulnerable to short-selling, as their managers are more likely to engage in value-destroying investments (Mehran 1995; Morck et al. 1988; McConnell and Servaes 1990; Jensen and Murphy 1990; Minnick et al. 2011). As short-sellers profit from downward stock price movements, they tend to avoid firms whose managers are cautious and diligent in stock price movements. Therefore, short-sellers are more likely to pay attention to the investment inefficiency of firms whose managers are not bound by stock performances.

Hypothesis 4 The positive relation between short-selling and future investment inefficiency is stronger for 'firms with low CEO incentive pay'.

3 Measures of short interest and investment inefficiency

3.1 Two measures of investment inefficiency

I employ two measures proposed by Biddle et al. (2009) and McNichols and Stubben (2008) to capture a firm's investment inefficiency. These measures align with neoclassical theory, which posits that a firm's primary objective is to maximize its net present value while adhering to the constraints of the production function. This pursuit of profit maximization shapes a firm's optimal investment policy, and any deviation from the optimal level of investment can be considered inefficient.

The two measures diverge in their approach to determining the optimal investment level. Drawing from the acceleration theory, Biddle et al. (2009) assert that the level of capital is directly proportional to the output level. When an economy operates at full capital utilization, firms adapt their investment in response to changes in demand. Biddle et al. (2009) model current investment as a function of past output growth, proxied by sales growth.

$$TI_{i,t} = \alpha + \beta_1 SG_{i,t-1} + \epsilon_{i,t}. \quad (1)$$

In this equation, $TI_{i,t}$ represents firm i's total investment,[6] and SG denotes firm i's percentage sales growth from year $t-1$ to year t. $\epsilon_{i,t}$ represents the residual in the regression, capturing deviations from the expected level of investment.

I utilize its absolute value, standardized by the standard deviation of its industry and year distribution, as the first measure of investment inefficiency.

$$Inefficiency^{B}_{i,j,t} = \left| \frac{\epsilon_{i,t}}{\sigma_{j,t}} \right|. \quad (2)$$

A higher level of $Inefficiency^B$ indicates lower efficiency in the firm's investment. I estimate Eq. (1) for each industry-year using the Fama and French 48-industry classification. Consistent with Biddle et al. (2009), I exclude industries with fewer than 20 observations in a given year.

Similarly, in line with Biddle et al. (2009), I calculate the total investment in a given year as the sum of capital expenditures, R &D expenditures, and acquisitions minus the sales of property, plant, and equipment (PPE). This total is then scaled by the lagged total assets. This approach employs an accounting-based framework to estimate total investment as the difference between total investment and asset sales (Richardson 2006). It also takes into account various types of investments, including capital expenditures and acquisitions.

The second measure of investment inefficiency, based on McNichols and Stubben (2008), assumes that investment can be predicted by investment opportunities. The use of Tobin's Q as a measure of investment opportunities is prevalent in the literature. For example, Goodman et al. (2014) use it as the proxy for investment opportunities and also control for cash flows, growth in assets, and past investment to allow for variations in the relationship between investment and Tobin's Q. McNichols and Stubben (2008) augment the model by further adding three interaction terms, which individually interact the lagged Tobin's Q with an indicator variable that equals 1 if the lagged Tobin's Q is in the second, third, or fourth quartile of its industry and year distribution. The model adopted from McNichols and Stubben (2008) is as follows:

$$\begin{aligned} TI_{i,t} = \alpha &+ \beta_1 Q_{i,t-1} + \beta_2 Q * QRT2_{i,t-1} \\ &+ \beta_3 Q * QRT3_{i,t-1} + \beta_4 Q * QRT4_{i,t-1} + \beta_5 CF_{i,t} \\ &+ \beta_6 AG_{i,t-1} + \beta_7 TI_{i,t-1} + \epsilon_{i,t}, \end{aligned} \quad (3)$$

where $Q_{i,t-1}$ is the lagged Tobin's Q.[7] $Q * QRT2_{i,t-1}$ (or $Q * QRT3_{i,t-1}$, $Q * QRT4_{i,t-1}$) equals $Q_{i,t-1}$ times an indicator variable that equals 1 if $Q_{i,t-1}$ is in the second (third, fourth) quartile of its industry and year distribution. $CF_{i,t}$ stands for net cash flows from operations (Compustat data item 308) scaled by the total book value of assets (Compustat data item 6). $AG_{i,t-1}$ stands for the lag of asset growth,

[6] Refer to Appendix for a detailed definition

[7] See Appendix for a detailed definition

which equals the natural log of the total book value of assets at the end of year $t-1$ divided by the total book value of assets at the end of year $t-2$. $TI_{i,t-1}$ is the lag of total investment.

I estimate Eq. (3) for each industry-year based on the Fama–French 48-industry classification and use the absolute value of the residuals from Eq. (3), standardized by the standard deviations of their corresponding industry and year distributions, as our second measure of investment inefficiency, $Inefficiency^{M}$:

$$Inefficiency^{M}_{i,j,t} = \left| \frac{\epsilon_{i,t}}{\sigma_{j,t}} \right|. \tag{4}$$

Table 2 reports the regression results from the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I calculate the cross-sectional means of the coefficient estimates for all industries in a given year and report the time-series average of the cross-sectional means of the coefficient estimates. We can see that all the factors in the models of Biddle et al. (2009) and McNichols and Stubben (2008) are significantly correlated with the firm's investment level, which is consistent with the assumptions behind these models. Besides the total investment, I also use the investment excluding R &D expenditure, TIW, to estimate Eqs. (1) and (3) and calculate these measures of investment inefficiency. They will be denoted by $Inefficiency^{BW}$ and $Inefficiency^{MW}$, respectively, and used for a robustness check.

3.2 Abnormal short interest

The primary independent variable is abnormal short interest. I employ the method of Karpoff and Lou (2010) to calculate it. First, I independently sort stocks into three groups by size, market-to-book ratio, and momentum, so that each stock is assigned to one of 27 constructed portfolios ($3 \times 3 \times 3$). I further categorize each stock into industry groups based on Fama and French's 48-industry classifications. In particular, I define the model as follows:

$$\begin{aligned} SI_{i,t} = {} & \sum_{g=low}^{medium} s_{gt} Size_{igt} + \sum_{g=low}^{medium} b_{gt} MB_{igt} \\ & + \sum_{g=low}^{medium} m_{gt} Momentum_{igt} + \sum_{k=1}^{K} \phi_{kt} Ind_{ikt} + u_{i,t}, \end{aligned} \tag{5}$$

where $SI_{i,t}$ represents the short interest level of firm i in month t. The first three independent variables are dummy variables that collectively determine the 27 size, market-to-book, and prior month return portfolios. For example, if firm i is sorted into the lowest market-to-book portfolio in month t, then $MB_{i,low,t} = 1$ and $MB_{i,medium,t} = 0$. The industry dummy $Ind_{ikt} = 1$ if firm i is assigned to industry k in month t, and K is the total number of industries in the sample.

I conduct monthly cross-sectional regressions of Eq. (5) and then calculate the time-series averages of the coefficient estimates. I use these time-series averages in conjunction with the dummy variables to derive the expected short interest $E(SI_{i,t})$ and calculate the abnormal short interest, $ASI_{i,t}$:

$$ASI_{i,t} = SI_{i,t} - E(SI)_{i,t}. \tag{6}$$

In an unpublished table,[8] I find that the smallest firms have the lowest short interest, growth stocks (with the highest market-to-book ratio) are more heavily shorted, and momentum has a U-shaped relationship with short interest as the coefficients on $Momentum_{low}$ and $Momentum_{medium}$ have different signs. These findings are consistent with Dechow et al. (2001) and Karpoff and Lou (2010).

4 Data and main results

4.1 Data source

The sample consists of stocks listed on NYSE, AMEX, and NASDAQ for the period from 1996 to 2018. Following Fang et al. (2016), I exclude firms in the financial services industry (SIC 6000–6999) and utilities industry (SIC 4900–4949) as the accounting rules, investment purposes, and processes significantly differ for these industries. Several steps are taken to clean the data. I require firms to have available information about investment and various variables needed for the calculation of investment inefficiency to remain in the sample. In particular, I exclude firms without capital expenditure information since it is a major component of total investment. The data used in the calculation of investment inefficiency are retrieved from the Compustat-Fundamentals Annual database. Additionally, following Biddle et al. (2009), I exclude industries with less than 20 observations in a given year. After these screenings, there are 61,947 firm-year observations, which are then used to calculate investment inefficiency.

Next, I match the measure of investment inefficiency with the short interest data and require firms to have available information on short interest before the release date of the annual report to remain in the sample. I retrieve short interest data, defined as the number of all open short positions scaled by total common shares outstanding, from the Compustat-Supplement Short Interest File. Short interest is measured on the 15th of each month or the preceding business day if the 15th is not a business day. The lag between the date of the short interest and the date of the annual release of the financial report should not exceed one month. I impose this lag between the release date of the annual report and the short sale date so that I can test whether short-sellers can detect investment inefficiency ex-ante, in other words, whether they are informed traders who possess superior information.

Finally, I require firms to have available information about various control variables. Data for the control variables are retrieved from the Compustat-Fundamentals

[8] It can be obtained upon request.

Table 1 Descriptive statistics

Variable	*N*	Mean	STD	25%	Median	75%
Variables for the inefficiency model						
TI (%)	61,947	16.469	19.247	4.845	10.380	20.356
TIW (%)	61,947	9.280	13.191	2.084	4.879	10.680
SG (%)	61,947	20.778	62.370	− 2.026	8.885	24.769
AG (%)	61,947	10.179	30.501	− 3.494	6.261	19.136
Q	61,947	2.218	1.816	1.177	1.617	2.507
CF (%)	61,947	4.322	18.983	1.819	8.011	13.364
Firm characteristics						
$Inefficiency^{B}$	40,594	0.633	0.642	0.257	0.487	0.763
$Inefficiency^{BW}$	40,594	0.601	0.704	0.248	0.434	0.637
$Inefficiency^{M}$	40,594	0.618	0.677	0.203	0.425	0.768
$Inefficiency^{MW}$	40,594	0.603	0.707	0.205	0.411	0.704
SI (%)	40,594	4.344	5.904	0.520	2.181	5.650
ASI (%)	40,594	0.061	5.526	− 3.044	− 1.431	1.087
ROA (%)	40,594	2.200	5.114	1.436	3.035	4.558
Leverage	40,594	0.296	0.294	0.021	0.250	0.453
MB Ratio	40,594	3.099	4.759	1.306	2.159	3.694
Ln_MV	40,594	6.545	2.040	5.107	6.513	7.917
Ln_assets	40,594	6.431	2.003	4.995	6.400	7.794
Operating Accruals	40,594	− 0.035	0.078	− 0.064	− 0.030	0.001
Board independence	15,713	0.743	0.151	0.667	0.778	0.875
CEO incentive pay	20,118	0.652	0.263	0.522	0.742	0.853
Variables for return predictability						
CAR	37,690	− 0.010	0.327	− 0.177	− 0.020	0.138
Price	37,690	29.542	72.891	7	18.225	36.870
Volatility	37,690	0.134	0.110	0.083	0.116	0.162
Turnover	37,690	0.167	0.297	0.058	0.114	0.201
Age	37,690	20.442	13.636	9	17	30
Dividend	37,690	0.011	0.023	0	0	0.013
SP500	37,690	0.134	0.341	0	0	0

This table presents the summary statistics of variables in the whole sample. For each variable, I report the number of observations (*N*), mean, standard deviation (STD), 25 percentile, median, and 75 percentile. All variables are winsorized at the 1% and 99% to mitigate the impact of outliers. The sample period is from January 1996 to December 2018. See Appendix for detailed descriptions of the variables

Quarterly database because some control variables are measured from the closest quarterly report released before the release date of the annual report. After data cleaning, there are 40,594 firm-year observations with available information about investment inefficiency, short interest, and various control variables, making up the main sample. Appendix shows the variables used in my empirical analysis and how to construct them. Table 1 reports the summary statistics of variables in the whole sample. I report the results of the main dependent and independent variables and other controlling variables. For each variable, I report the number of observations (*N*), mean, standard deviation (STD), and the 25th percentile, median, and 75th percentile values. I winsorize all variables at the 1 and 99% levels to mitigate the impact of outliers. These variables are used in the following analysis.

In the economic significance analysis, I collect stock return information from the Center for Research on Security Prices (CRSP) database. I require firms to have available stock return information for the past two years since we need to calculate the volatility of the stock return, which equals the standard deviation of monthly stock returns over the past 24 months. Therefore, we have 37,690 firm-year observations in the economic significance analysis.

For the subsample analyses, I use information about board independence and CEO incentive pay. Information about board independence is retrieved from the RiskMetrics database, which covers Standard & Poor's (S &P) 500, S &P MidCaps, and S &P SmallCap firms. Information about CEO incentive pay is obtained from the ExecuComp database.

Table 2 presents the regression results of the investment inefficiency model. Both models are run for each industry-year based on the Fama-French 48-industry classification. Columns (1) and (2) report the results of the models by Biddle et al. (2009) and McNichols and Stubben (2008), respectively.

As shown in column (1) of Table 2, the coefficient of sales growth is highly significant. Therefore, investment is significantly correlated with output growth, which is consistent with the assumption of acceleration theory. The results in column (2) show that the coefficients of various variables related to investment opportunities are highly significant. These results indicate that investment is strongly correlated with investment opportunities.[9]

4.2 Univariate test

In this subsection, I present the results of a univariate test to examine the relation between short-selling and investment inefficiency. If short-sellers are informed about a firm's investment inefficiency, then short-selling should be more severe for firms with higher investment inefficiency. Following Hirshleifer et al. (2011), I sort these sample firms into deciles based on the abnormal short interest closest to, but before, the annual release of the financial report for each year. The abnormal short interest is measured on the 15th of each month or the preceding business day if the 15th is not a business day. The lag between the date of abnormal short interest and the date of the annual release of the financial report should not exceed one month. I then test whether the mean of investment inefficiency in the top decile of abnormal short interest is significantly higher than the mean of investment inefficiency in the bottom decile. To do this, I first calculate the cross-sectional means of investment

[9] Biddle and Hilary (2006), Francis and Martin (2010), and Bushman et al. (2011) also study capital expenditures and R &D expenditures separately. To test whether these results are robust to the inclusion of R &D expenditure, I use investment excluding R &D expenditure to calculate the investment inefficiency measures as one robustness check, and I present the results in Sect. 5.3.

Table 2 Investment inefficiency

	(1) Biddle et al. (2009)	(2) McNichols and Stubben (2008)
	$TI_{i,t}$	$TI_{i,t}$
$SG_{i,t-1}$	0.048***	
	(16.67)	
$AG_{i,t-1}$		− 0.041***
		(−9.55)
$Q_{i,t-1}$		− 1.307**
		(−2.31)
$Q*QRT2_{i,t-1}$		1.958***
		(6.03)
$Q*QRT3_{i,t-1}$		3.010***
		(7.06)
$Q*QRT4_{i,t-1}$		3.195***
		(6.40)
$CF_{i,t}$		− 0.030*
		(−1.74)
$TI_{i,t-1}$		0.289***
		(16.28)
R^2	0.052	0.323
Observations	61,947	61,947

Regression results from the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008), run for each industry-year based on the Fama–French 48-industry classification. The cross-sectional means of the coefficient estimates for all industries in a given year are calculated and then the time-series averages of these cross-sectional means. Column (1) reports the results of the investment inefficiency model of Biddle et al. (2009),

$$TI_{i,t} = \alpha + \beta_1 SG_{i,t-1} + \epsilon_{i,t}$$

where $TI_{i,t}$ is firm i's total investment at year t scaled by its lagged total assets, and SG is firm i's sales growth in percentage from year $t-1$ to year t. Column (2) reports the results of the investment inefficiency model of McNichols and Stubben (2008),

$$TI_{i,t} = \alpha + \beta_1 Q_{i,t-1} + \beta_2 Q*QRT2_{i,t-1} + \beta_3 Q*QRT3_{i,t-1} + \beta_4 Q*QRT4_{i,t-1} + \beta_5 CF_{i,t} + \beta_6 AG_{i,t-1} + \beta_7 TI_{i,t-1} + \epsilon_{i,t}$$

where $Q_{i,t-1}$ is the lag of Tobin's q. $Q*QRT2_{i,t-1}(Q*QRT3_{i,t-1}, Q*QRT4_{i,t-1})$ equals $Q_{i,t-1}$ times an indicator variable that equals 1 if $Q_{i,t-1}$ is in the second (third, fourth) quartile of its industry and year distribution. $CF_{i,t}$ stands for net cash flows from operations (Compustat data item 308) scaled by total assets. $AG_{i,t-1}$ equals the natural log of total assets at the end of year $t-1$ divided by assets at the end of year $t-2$. The sample period is from January 1996 to December 2018. Newey–West corrected t-statistics of the time series averages in parentheses. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

Table 3 Univariate test for the relation between short-selling and investment inefficiency

	(1) Low	(2) Decile 2	(3) Decile 9	(4) High	(5) Difference (High–Low)	(6) Difference (Decile 9–Decile 2)
$Inefficiency^B$	0.611	0.592	0.657	0.689	0.078***	0.066***
					(4.43)	(4.69)
$Inefficiency^M$	0.609	0.602	0.658	0.684	0.075***	0.056***
					(4.45)	(4.39)

This table reports the results from univariate test for the relation between historical abnormal short interest and investment inefficiency. Following Hirshleifer et al. (2011), I sort these sample firms into deciles based on their abnormal short interest before the annual release of the financial report for each year. The abnormal short interest is measured on the 15th of each month or the preceding business day if the 15th is not a business day. The lag between the date of abnormal short interest and the date of annual release of the financial report should not exceed one month. Then, I first calculate the cross-sectional means of investment inefficiency for the top and the bottom deciles of abnormal short interest in each year and then their differences. After that, I calculate the time-series average of these differences and test whether the mean difference is significantly different from zero. This table reports the time-series average of investment inefficiency for the top and bottom deciles of abnormal short interest and the differences in means. The values in brackets are the *t*-values of the mean differences. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

inefficiency for the top and the bottom abnormal short interest deciles in each year. I then calculate their difference. After that, I calculate the time-series average of the difference and check whether the mean difference is significantly different from zero.

Table 3 reports the results of the univariate test. I consider two measures of investment inefficiency, $Inefficiency^B$ and $Inefficiency^M$, calculated by Eqs. (1) and (3), respectively. Table 3 shows that the means of $Inefficiency^B$ in the top and bottom deciles of abnormal short interest are 0.689 and 0.611, respectively. The difference in mean is 0.078 and is significant at the 1% level (t-value = 4.43). It accounts for 12.15% of the difference in the standard deviations of $Inefficiency^B$. The difference in mean of $Inefficiency^M$ between the top and bottom deciles is 0.075, which is also significant at the 1% level (t-value = 4.45). It accounts for 10.65% of the difference in the standard deviations of $Inefficiency^M$. These results support the hypothesis that short-sellers are informed about firms' investment inefficiency. There is a positive relation between abnormal short-selling activity and investment inefficiency.

4.3 Panel regression: baseline model

In this subsection, I present a baseline panel regression using the whole sample data to investigate the relation between short-selling and future investment inefficiency. In particular, I follow Jiao et al. (2016) to run the following panel regressions using all data:

$$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}, \tag{7}$$

where $ASI_{;i,t-1}$ represents the abnormal short interest before the release of the annual financial statement. $Inefficiency_{i,t}$ serves as the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables, including firm size, market-to-book ratio, leverage, and return on assets, all measured with a lag of one quarter. Furthermore, prior research finds a significant correlation between short interest and operating accruals (Hirshleifer et al. 2011; Fang et al. 2016; Massa et al. 2015; Park 2017). I further add operating accruals as control variables.[10] $Y_{i,t-1}$ refers to the list of independent variables used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). According to Chen et al. (2018), it leads to biased coefficients and standard errors if the researcher decomposes a dependent variable into its predicted and residual components in the first regression and uses the residuals as the dependent variable in a second regression. As discussed above, the investment inefficiency measures take advantage of the residuals in the models of Biddle et al. (2009) and McNichols and Stubben (2008). Hence, there is the concern of biased coefficients and standard errors here because the equation above uses investment inefficiency as the dependent variable. One of the approaches to address this problem is to regress the residuals on the combination of all the independent variables in the first regression and the independent variables in the second regression (Chen et al. 2018). Following this approach, I further add regressors in the models of Biddle et al. (2009) and McNichols and Stubben (2008) as independent variables in the above model.

Table 4 reports the results of the baseline panel regressions. I consider both firm and year fixed effects and cluster the standard errors at the firm level. Columns (1) and (2) report the results of $Inefficiency^B$ and $Inefficiency^M$, respectively. The coefficient of abnormal short interest on $Inefficiency^B$ is 0.003, as shown in column (1), which is significant at the 1% level (t-value = 3.58). The coefficient of abnormal short interest on $Inefficiency^M$ is 0.003, as shown in column (2), which is also significant at the 1% level (t-value = 3.37). Therefore, the result of the baseline regressions demonstrates that short-selling reflects information about a firm's investment inefficiency. If a firm is becoming less efficient in investment, then short-sellers react by holding a larger short position in the firm's stock before this information is made public. In summary, the result of the baseline regression is consistent with that of the univariate test, and both of them support Hypothesis 1.

To spell out why I choose to measure the short-selling activities just before the release of the annual report, I draw the dynamic changes of the coefficient on ASI. I run the baseline model of Eq. (7) but replace ASI_{t-1} with the abnormal short interest levels 2 months, 4 months, 6 months, and 8 months before the release of the annual financial report. I include the same controlling variables and consider the year and firm fixed effects. I also cluster the standard errors at the firm level. Figure 1 shows that the relationship between abnormal short interest and investment inefficiency is stronger when it is closer to the release of the annual financial report. As short-sellers can observe more information and have a more comprehensive understanding of

[10] Appendix explains how to construct these variables.

Table 4 Baseline regression result

	(1) $Inefficiency_t^B$	(2) $Inefficiency_t^M$
ASI_{t-1}	0.003***	0.003***
	(3.58)	(3.37)
$MB\ Ratio_{t-1}$	0.002**	0.002**
	(2.20)	(2.15)
$Leverage_{t-1}$	0.187***	0.211***
	(6.78)	(8.40)
ROA_{t-1}	− 0.007***	0.002
	(− 4.52)	(0.93)
Ln_MV_{t-1}	0.065***	0.017*
	(7.04)	(1.68)
Ln_Assets_{t-1}	− 0.008	0.072***
	(− 0.56)	(5.15)
$Operating\ Accruals_{t-1}$	− 0.083	− 0.123*
	(− 1.45)	(− 1.94)
SG_{t-1}	0.000	
	(0.89)	
AG_{t-1}		− 0.056***
		(− 2.93)
Q_{t-1}		− 0.108***
		(− 6.52)
$Q * QRT2_{t-1}$		0.101***
		(10.05)
$Q * QRT3_{t-1}$		0.145***
		(11.85)
$Q * QRT4_{t-1}$		0.146***
		(10.24)
CF_t		− 0.003***
		(− 5.14)
TI_{t-1}		0.003***
		(9.36)
R^2	0.262	0.232
Firm fixed effect	Yes	Yes
Year fixed effect	Yes	Yes
Cluster standard errors	Yes	Yes
Observations	40,594	40,594
Number of firms	4919	4919

This table reports the panel regression results for the relation between historical abnormal short interest and investment inefficiency. I run the following panel regressions using all data,

$$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t},$$

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. The abnormal short interest is measured on the 15th of each month or the preceding business day if the 15th

Table 4 (continued) is not a business day. The lag between the date of abnormal short interest and the date of annual release of the financial report should not exceed one month. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables including firm size (*Ln_MV* and *Ln_Asset*), market-to-book ratio (*MB Ratio*), leverage (*Leverage*), and return on assets (*ROA*), which are all measured at a lag of one quarter. Furthermore, previous research has found significant correlations between short interest and operating accruals (Hirshleifer et al. 2011; Fang et al. 2016; Massa et al. 2015; Park 2017), I further add operating accruals (*Operating Accruals*) as another control variable. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I further add these variables because it is one of the approaches to address the biased coefficients and standard errors when using residuals as the dependent variable in a second regression (Chen et al. 2018). I consider both firm and year fixed effects and cluster the standard errors at the firm level. Columns (1) and (2) report the results of $Inefficiency^B$ and $Inefficiency^M$, respectively. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

a firm's investment inefficiency when it is closer to the fiscal year end, I investigate whether short-sellers are informed about a firm's investment inefficiency at the time point when the information is considered sufficient and comprehensive.

Figure 1 also shows that the lower bound of the 95% confidence interval for the coefficient on *ASI* crosses over zero at the 4-month mark before the annual report release. This suggests that the relationship between short-selling activities and investment inefficiency becomes significant approximately 4 months prior to the annual report release, coinciding with the release of the 3rd quarter report. As discussed in previous sections, the investment inefficiency measures are derived from information disclosed in the annual financial report. This implies that short-sellers gather and analyze information about a firm's investment behaviors along the way, enabling them to make informed conjectures about a firm's investment inefficiency one quarter before it becomes evident in the annual financial report.

To address the concern of omitted variables, I further include additional control variables in the model, as shown in Appendix. These variables encompass dividend cuts, competition, momentum, and misstatements in the financial reports. However, the inclusion of these additional variables does not alter the significant relationship between short-selling activities and investment inefficiency.

4.4 Anecdotal evidence

Texas Industries Inc. (Ticker Symbol: TXI) is a publicly traded company listed on the New York Stock Exchange (NYSE). We analyze key financial metrics, including the mean values of the natural logarithm of market value (Ln_{MV}), the natural logarithm of total book value of assets (Ln_{Assets}), the market-to-book ratio (MB_{Ratio}), leverage (*Leverage*), and return on assets (*ROA*), over the period from 2005 to 2009. These metrics exhibit the following means during this period: Ln_{MV} (7.081), Ln_{Assets} (7.146), MB_{Ratio} (2.364), *Leverage* (0.398), and *ROA* (2.436%).

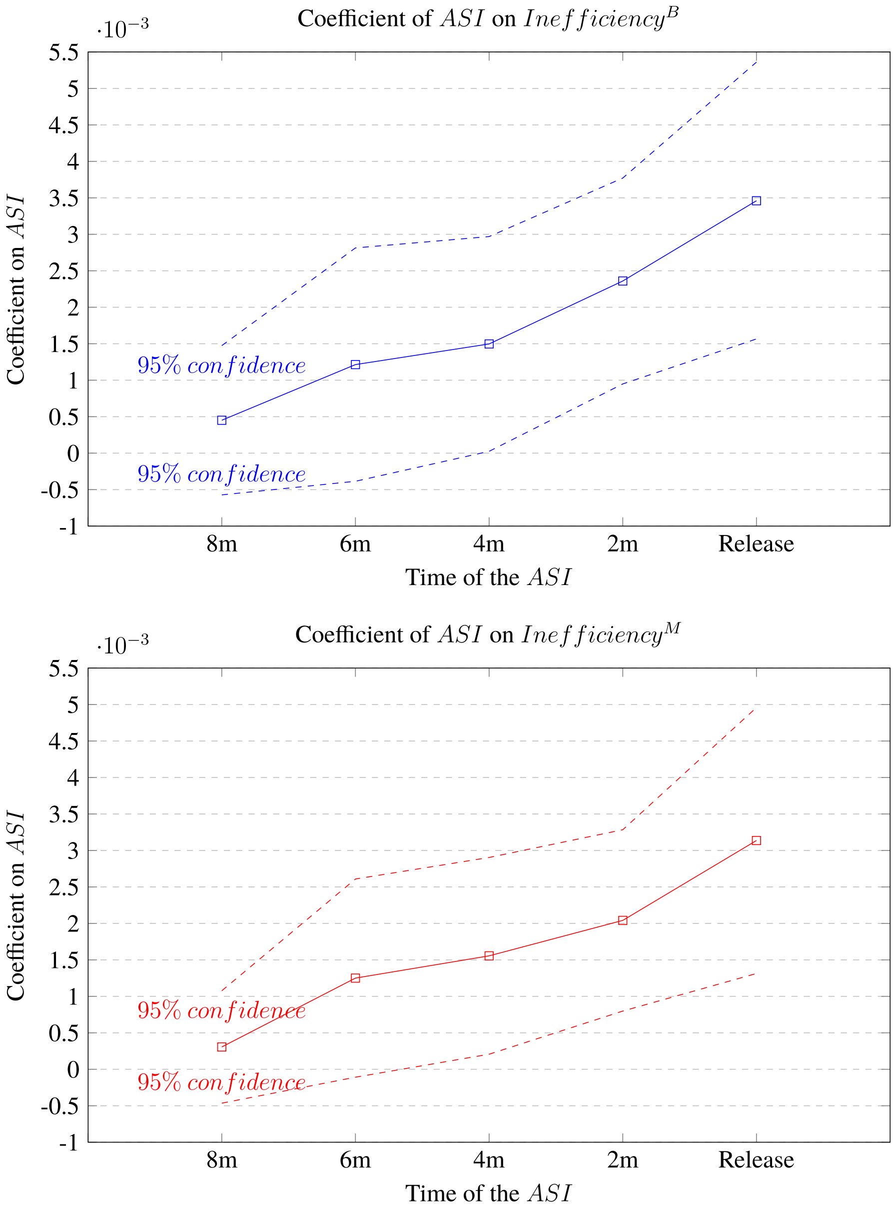

Fig. 1 Dynamic changes of the coefficient. This graph plots the dynamic changes of the coefficient of *ASI*. The *ASI* is measured on the closest short interest date, 2 months, 4 months, 6 months, and 8 months before the release of the annual financial report. The top and bottom panels plot the results of $Inefficiency^B$ and $Inefficiency^M$, respectively

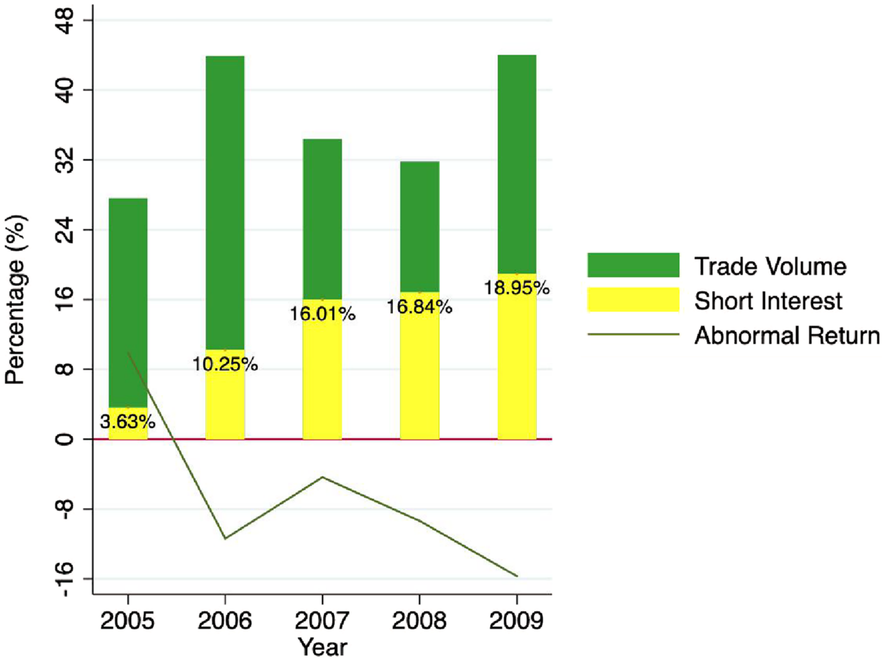

Fig. 2 Case study of Texas Industries INC. This graph plots the trading volume, short interest, and abnormal return around the release of the annual financial report for Texas Industries INC. The trading volume represents the monthly shares traded when the annual report is released, scaled by the total shares outstanding. The abnormal return is the monthly abnormal stock return after the release of the annual financial report

These values collectively suggest that Texas Industries Inc. does not display extreme financial characteristics during this timeframe, indicating a relatively stable performance. However, in the year 2006, the investment inefficiency of the company worsens. Specifically, the measures of investment inefficiency ($Inefficiency^B$ and $Inefficiency^M$) increase from 0.560 and 0.345 in 2005 to 0.573 and 0.748, respectively. Consequently, I observe an increase in short-selling activities before the annual report release. For instance, the levels of short interest and abnormal short interest rise from 3.627% and − 2.646% in 2005 to 10.251% and 3.932%, respectively, in 2006.

After 2006, the investment inefficiency of Texas Industries Inc. continued to deteriorate until 2009. Consequently, the short-selling activities before the annual report release also increased. Figure 2 depicts the short interest before the annual report release,[11] clearly illustrating an ascending pattern from 2005 to 2009. Consistent with expectations, as short-selling activities before the annual report release intensified for Texas Industries Inc. after 2005, the abnormal stock return around the annual report release also turned negative, with a clear decreasing trend evident in Fig. 2

[11] The trend of abnormal short interest remains consistent. For the sake of maintaining a sense of dignity, I have chosen not to plot the levels of abnormal short interest.

4.5 Economic significance

I find a statistically significant positive correlation between short-selling before the annual release of financial information and investment inefficiency. In this subsection, I study whether short-sellers can also use this information to generate economic profits.

To measure the abnormal returns of a stock, I deduct the individual stock returns by the returns of an equal-weighted benchmark portfolio, constructed following Daniel et al. (1997). Specifically, I sort all firms traded on NYSE, AMEX, and NASDAQ, excluding American Depository Receipts (ADRs),[12] into quintiles based on market capitalization. In each quintile of market capitalization, I sort firms into book-to-market ratio quintiles to generate 5×5 groups. Then, in each of these 25 groups, I sort firms into momentum quintiles using the stock's past one-year returns while skipping the most recent month. This results in 125 ($5 \times 5 \times 5$) groups, and then I construct the equal-weighted portfolio for each group. The benchmark portfolio for a stock is that of the group to which the stock belongs.

Next, I use the model in Jiao et al. (2016) to run a Fama–MacBeth regression at a yearly frequency:

$$CAR_{i,[t+1,t+6]} = \alpha + \beta_1 Inefficiency_{i,t} + \beta_2 X_{i,t} + \epsilon_{i,t} \quad (8)$$

where $CAR_{i,[t+1,t+6]}$ represents the accumulation over 6 months of abnormal returns. $Inefficiency_{i,t}$ serves as the measure of investment inefficiency. $X_{i,t}$ refers to a list of control variables used in Jiao et al. (2016), including the natural logarithm of the stock price ($Ln\_Price_{i,t}$), the natural logarithm of the standard deviation of stock returns over the past 24 months ($Ln\_Volatility_{i,t}$), the natural logarithm of firm age ($Ln\_Age_{i,t}$), the natural logarithm of stock turnover ($Ln\_Turnover_{i,t}$), the dividend ($Dividend_{i,t}$), and a dummy variable equal to one if the firm is in the S &P 500 index ($SP500_{i,t}$). I also add operating accruals ($Operating;Accruals_{i,t}$) as a control variable since it is well-documented in the literature to have return predictability (Hirshleifer et al. 2011).

Table 5 reports the time-series average of the coefficient estimates from the yearly cross-sectional regressions with their *t*-values (in brackets) adjusted using the Newey–West method with three lags. The result in Table 5 shows that the coefficients of the four measures of investment inefficiency are significantly negative, above the 10% level. This means that investment inefficiency forecasts negative abnormal returns. It is also worth noting that the coefficients of operating accruals are significantly negative, which is consistent with the previous literature, which shows that accruals forecast negative returns. Therefore, the results suggest that short-selling based on information inefficiency information can generate significant profits. Short-sellers increase their short position in the stocks with a higher level of investment inefficiency and receive significant returns from these activities.

[12] The abnormal return is calculated using 1,192,888 firm-month return observations on CRSP from January 1996 to December 2019, where the shares are traded on NYSE, AMEX, and NASDAQ, excluding American Depository Receipts (ADRs). The average number of stocks per month is 4142.

Table 5 Economic significance

	(1) $CAR_{i,[t+1,t+6]}$	(2) $CAR_{i,[t+1,t+6]}$	(3) $CAR_{i,[t+1,t+6]}$	(4) $CAR_{i,[t+1,t+6]}$
$Inefficiency_t^B$	− 0.010***			
	(− 3.75)			
$Inefficiency_t^M$		− 0.017***		
		(− 3.37)		
$Inefficiency_t^{BW}$			− 0.007*	
			(− 2.00)	
$Inefficiency_t^{MW}$				− 0.010**
				(− 2.16)
Ln_Price_t	− 0.037***	− 0.036***	− 0.036***	− 0.036***
	(− 5.54)	(− 5.48)	(− 5.54)	(− 5.52)
$Ln\_Volatility_t$	− 0.043***	− 0.042***	− 0.043***	− 0.043***
	(− 3.71)	(− 3.74)	(− 3.75)	(− 3.79)
Ln_Age_t	0.025***	0.024***	0.025***	0.024***
	(6.57)	(6.52)	(6.52)	(6.55)
$Ln\_Turnover_t$	− 0.007	− 0.007	− 0.007	− 0.006
	(− 1.39)	(− 1.37)	(− 1.36)	(− 1.36)
$Dividend_t$	− 0.81***	− 0.80***	− 0.81***	− 0.80***
	(− 4.85)	(− 4.73)	(− 4.86)	(− 4.80)
$Operating\ Accruals_t$	− 0.082**	− 0.085**	− 0.080**	− 0.082**
	(− 2.46)	(− 2.50)	(− 2.39)	(− 2.42)
$SP500_t$	− 0.011	− 0.011	− 0.010	− 0.011
	(− 1.35)	(− 1.40)	(− 1.34)	(− 1.38)
Intercept	− 0.075***	− 0.071***	− 0.078***	− 0.077***
	(− 2.84)	(− 2.60)	(− 2.89)	(− 2.81)
R^2	0.045	0.047	0.045	0.046
Observations	37,690	37,690	37,690	37,690
Number of firms	4575	4575	4575	4575

Testing the economic significance of the relation between short sales and investment inefficiency. I use the following model, incorporated in Jiao et al. (2016), to run a Fama–MacBeth regression at yearly frequency:

$$CAR_{i,[t+1,t+6]} = \alpha + \beta_1 Inefficiency_{i,t} + \beta_2 X_{i,t} + \epsilon_{i,t}$$

where $CAR_{i,[t+1,t+6]}$ is the accumulated abnormal returns over six months. The abnormal returns are determined by reducing the individual stock returns by the returns of an equal-weighted benchmark portfolio constructed using the firms within the same size, book-to-market ratio, and momentum groups. The abnormal returns are calculated using 1,192,888 observations of firm-month returns in the CRSP from January 1996 to December 2019, for shares traded on NYSE, AMEX, and NASDAQ excluding American Depository Receipts (ADRs). The average number of stocks per month is 4142. $Inefficiency_{i,t}$ is our measure of investment inefficiency: it is one of $Inefficiency_{i,t}^B$, $Inefficiency_{i,t}^M$, $Inefficiency_{i,t}^{BW}$ and $Inefficiency_{i,t}^{MW}$. $X_{i,t}$ refers to a list of control variables used in Jiao et al. (2016) including the natural logarithm of stock price ($Ln\_Price_{i,t}$), the natural logarithm of the standard deviation of stock returns over the past 24 months ($Ln\_Volatility_{i,t}$), the natural logarithm of firm age ($Ln\_Age_{i,t}$), the natural logarithm of stock turnover ($Ln\_Turnover_{i,t}$), dividend ($Dividend_{i,t}$), and a dummy variable equal to one if the firm is in the S &P 500 index ($SP500_{i,t}$). I further add operating accruals ($Operating\ Accruals_{i,t}$) as a control variable as in the literature it is well-documented to have return predictability (Hirshleifer et al. 2011). The *t*-statistics (in brackets) of our four measures of investment inefficiency are calculated using

Table 5 (continued)

the Newey–West method with three lags. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

5 Robustness tests

5.1 Endogeneity

I document a significantly positive correlation between short-selling and upcoming firm investment inefficiency information. However, this finding is subject to an endogeneity concern, as unknown factors influence both abnormal short interest and firm investment inefficiency. In other words, abnormal short interest may be correlated with the error term in the regression. To address this potential endogeneity issue, I follow Brav et al. (2018) and Park (2017) in adopting a propensity score matching procedure. This procedure reduces the potential bias that can result either from the omission of observable variables or the specification of an improper functional form for the relation between the observable variables and the outcome variable of interest (Armstrong et al. 2010; Park 2017). In the matched design, each treatment observation is matched with a control observation that does not receive that treatment but is indistinguishable from it in terms of other relevant dimensions. Therefore, any difference in the outcome between the treatment and control groups is attributed to the treatment effect (Armstrong et al. 2010; Park 2017).

To run the propensity score matching, I first rank my sample firms into quintiles based on abnormal short interest before the annual release of the financial statement for each year. I consider the top quintile of abnormal short interest as the treatment group because firms in the top quintile of abnormal short interest are the ones most heavily shorted. The firms in the other groups are treated as control groups. Second, I establish a dummy variable, $DASI_{i,t}$, that equals one if it is in the top quintile of abnormal short interest and zero otherwise. Third, I run a logistic regression that sets $DASI_{i,t}$ as the dependent variable:

$$logit(DASI_{i,t} = 1) = \phi(\alpha + \beta X_{I,t}). \quad (9)$$

Here, $X_{i,t}$ constitutes the vector of variables considered as determinants of short-selling activities in previous literature, including firm size, return on assets, market-to-book ratio, momentum, leverage, and operating accruals (Hirshleifer et al. 2011; Karpoff and Lou 2010; Grullon et al. 2015; Diether et al. 2009). Industry and year fixed effects are also controlled for in the regression. The predicted value in the logit regression serves as the propensity score for matching purposes, and one-to-one matching with replacement is employed. In other words, each observation in the top quintile of abnormal short interest is matched with one observation in the control group with the closest score. Observations in the control group may be used multiple times in the matched sample. Additionally, an additional "maximum 1%" constraint is imposed, meaning that matched pairs should have a propensity score difference of less than 1

Table 6 Comparison between treatment group and control group

	(1)	(2)	(3)	(4)	(5)	(6)
	Before matching			After matching		
	Treatment	Control	Difference	Treatment	Control	Difference
ROA (%)	2.206	2.199	0.007	2.213	2.137	0.076
			(0.11)			(0.84)
Momentum (%)	17.199	16.332	0.867	17.175	15.819	1.356
			(1.22)			(1.49)
Ln_MV	6.819	6,477	0.343***	6.816	6.801	0.015
			(17.74)			(0.50)
Ln_Assets	6.568	6.397	0.170***	6.569	6.593	− 0.024
			(8.42)			(− 0.77)
MB_ratio	3.438	3.014	0.424***	3.464	3.393	0.071
			(6.49)			(0.82)
Operating Accruals	− 0.038	− 0.035	− 0.003***	− 0.038	− 0.038	0.000
			(− 3.31)			(0.36)
Leverage	0.313	0.291	0.022***	0.311	0.310	0.000
			(5.72)			(0.08)

Comparison of various firm characteristics between the firms in the treatment and control groups. We will regard the firms in the top quintile of abnormal short interest as the treatment group and the firms in the other quintiles as controls. I define a dummy variable $DASI_{i,t}$ that equals one if it is in the top quintile of abnormal short interest and zero otherwise, and then run a logistic regression that sets $DASI_{i,t}$ as the dependent variable as follows:

$logit(DASI_{i,t} = 1) = \phi(\alpha + \beta X_{I,t})$.

where $X_{i,t}$ is the vector of the set of variables that have been regarded in the literature as determinants of short selling, including firm size, return on assets, market-to-book ratio, momentum, leverage and operating accruals (Hirshleifer et al. 2011; Karpoff and Lou 2010; Grullon et al. 2015; Diether et al. 2009). I also account for industry and year fixed effects. I use the predicted value in the logit regression as the propensity score for matching purposes and employ a one to one match with replacement. I also impose an additional "maximum 1% " constraint, which means the matched pairs should have a propensity score difference less than 1%. The left and right columns report the results before and after the matching, respectively. The numbers in brackets are the *t*-values of the mean differences. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

To assess the effectiveness of the propensity score matching, I compare the means of various firm fundamentals and characteristics between the treatment and control groups before and after matching. Table 6 reports the results. Before matching, significant differences exist between the treatment and control groups in terms of *Ln_MV*, *Ln_Asset*, MB_{ratio}, and *Leverage*. When all sample firms are considered, larger firms, higher market-to-book ratios, and higher leverages are associated with more short-selling. For example, the difference in *Ln_MV* between the treatment and control groups before matching is 0.343 and significant at the 1% level (*t*-value = 17.74). The differences in *Ln_Asset*, MB_{ratio}, and *Leverage* are 0.170, 0.424, and 0.022, respectively, all of which are significant at the 1% level. However, after matching, differences in firm characteristics between the treatment and control groups become indistinguishable. None of the differences are significant after matching. These results suggest that this propensity score matching effectively mitigates endogeneity concerns.

Next, I present the results of the following panel regression using the propensity score matching sample to study the relationship between the level of abnormal short interest and firm investment inefficiency.

$$Inefficiency_{i,t} = \alpha + \beta_1 DASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}, \tag{10}$$

Here, $X_{i,t-1}$ constitutes the vector of variables considered as determinants of short-selling activities in previous literature, including firm size, market-to-book ratio, leverage, and return on assets, all measured with a lag of one quarter. $Y_{i,t-1}$ refers to the list of independent variables used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). A positive coefficient of β_1 suggests that a higher level of abnormal short interest before the release date is associated with a higher level of investment inefficiency. Similarly, I control for industry and year fixed effects and cluster standard errors at the firm level.

Table 7 reports the regression results. Columns (1) and (2) report the results of $Inefficiency^B$ and $Inefficiency^M$, respectively. The results show that for the matched samples, the coefficients of $DASI_{i,t}$ are both significantly positive. The coefficients of $DASI$ for $Inefficiency^B$ and $Inefficiency^M$ are 0.050 and 0.046, respectively. Both of them are significant at the 1% level. These results suggest the positive relationship between abnormal short interest and investment inefficiency is robust after controlling for endogeneity concerns.

5.2 Alternative explanation

Some other alternative explanations for the relationship between short-selling activities and investment inefficiency bring up endogeneity concerns. There are two major alternative explanations. First, weak corporate governance is associated with investment inefficiency due to increased agency problems (Shleifer and Vishny 1997). If short-sellers target firms with weak corporate governance, then corporate governance tends to become endogenous. Second, certain events prior to the annual financial report release may be simultaneously related to short-selling activities and investment inefficiency. However, I suggest that these concerns are considerably minimal.

First, although short-sellers can identify firms with significant internal control weaknesses, the evidence primarily focuses on material internal governance failures (Singer et al. 2022). Considering the large sample size in this paper, the relationship should not be driven solely by these rare instances of material internal governance failure.

More importantly, if short-sellers target firms with material internal governance failure, the significant relationship should appear around the discovery of such failures rather than the financial report release. Furthermore, since material internal

Table 7 Regression results using the matched sample

	(1)	(2)
	$Inefficiency_t^B$	$Inefficiency_t^M$
$DASI_{t-1}$	0.050***	0.046***
	(2.98)	(2.67)
$MB\ Ratio_{t-1}$	0.004*	0.002
	(1.94)	(1.06)
$Leverage_{t-1}$	0.206***	0.181***
	(5.02)	(4.62)
ROA_{t-1}	− 0.005	0.002
	(− 1.65)	(0.58)
Ln_MV_{t-1}	0.074***	0.006
	(4.60)	(0.36)
Ln_Assets_{t-1}	− 0.030	0.070***
	(− 1.28)	(2.98)
$Operating\ Accruals_{t-1}$	− 0.185*	− 0.272**
	(− 1.72)	(− 2.42)
SG_{t-1}	0.000	
	(0.29)	
AG_{t-1}		− 0.049
		(− 1.57)
Q_{t-1}		− 0.103***
		(− 3.64)
$Q * QRT2_{t-1}$		0.104***
		(5.24)
$Q * QRT3_{t-1}$		0.152***
		(6.66)
$Q * QRT4_{t-1}$		0.151***
		(6.02)
CF_t		− 0.003***
		(− 2.90)
TI_{t-1}		0.004***
		(4.71)
R^2	0.388	0.364
Firm fixed effect	Yes	Yes
Year fixed effect	Yes	Yes
Cluster standard errors	Yes	Yes
Observations	14790	14790
Number of firms	3797	3797

Regression results for the matched sample. Specifically, I ran the following regression using the matched sample,

$Inefficiency_{i,t} = \alpha + \beta_1 DASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}$,

where *DASI* is a dummy variable that equals one if it is in the top quintile of abnormal short interest and zero otherwise. $X_{i,t-1}$ is the vector of the set of variables that have been regarded in the literature as determinants of short selling, including firm size, market-to-book ratio, leverage, and return on assets, which are all measured at a lag

Table 7 (continued) of one quarter. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I also control for industry and year fixed effects and cluster standard errors at the firm level. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

governance failures tend to be enduring, the significant relationship should not be limited to the period around the financial report release.

The concern regarding events prior to the financial report release is also considered minimal. First, events preceding the financial report release are reporting-specific, and short-selling activities related to these events occur within a very short period surrounding them. For example, Dai et al. (2021) find that the sharp increase in short-selling activities for firms delaying financial report release focuses on the month prior to the scheduled release.

However, Fig. 1 suggests that short-sellers have begun to detect investment inefficiency four months prior to the financial report release. The short-selling activities related to these reporting-specific events demonstrate a different pattern from that documented in the present paper.

Second, the number of such reporting-specific events is relatively scarce compared to the sample size in the present paper. For instance, the average number of late reporting events each year in Dai et al. (2021) is 127, while the average number of firms each year in the present paper is 1765. The late-reporting events account for only around 7% of the firms in the present study. For the subsequent analysis, I have identified 502 severe misstatement events, which constitute only about 1% of the firms each year in the present study. It is highly unlikely that the significant relationship between short-selling activities and investment inefficiency is biased by these rare cases.

To support this point, I have conducted an additional analysis as shown in Appendix. Short-selling activities are related to serious financial misstatements (Karpoff and Lou 2010), and there is concern that these misstatements are related to investment inefficiency. I have implemented the following model to investigate whether the filing of serious financial misstatements will affect the relationship between short-selling activities and investment inefficiency:

$$\begin{aligned} Inefficiency_{i,t} = \alpha &+ \beta_1 \cdot ASI_{i,t-1} \times Restatement_{i,t} \\ &+ \beta_2 \cdot ASI_{i,t-1} + \beta_3 \cdot Restatement_{i,t} \\ &+ X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}, \end{aligned} \tag{11}$$

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. $Restatement_{i,t}$ is a dummy variable that indicates whether a material inadvertent (unintentional) or fraudulent (intentional) error exists in a firm's financial statements. $ASI_{i,t-1} \times Restatement_{i,t}$ is the interaction term between $ASI_{i,t-1}$ and $Restatement_{i,t}$. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables included in Table 4. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009)

and McNichols and Stubben (2008). I consider both firm and year fixed effects and cluster the standard errors at the firm level.

The result in Appendix shows that the coefficient of $ASI_{;i,t-1} \times Restatement_{;i,t}$ is insignificant, indicating that the filing of serious financial misstatement does not affect the relationship between short-selling activities and investment inefficiency.

5.3 Investment inefficiency measure excluding R &D

I use a firm's total investment (*TI*) to calculate the measures $Inefficiency^{B}$ and $Inefficiency^{M}$, which equal the sum of capital expenditures, R &D expenditures, and acquisitions minus the sale of property, plant, and equipment. The literature documents several reasons for separating capital expenditures and R &D expenditures. First, according to the accounting principles generally accepted in the USA (US GAAP), all R &D expenditures must be expensed when incurred, while capital expenditures must be capitalized and depreciated over their economic useful lives (Amir et al. 2007). Therefore, R &D expenses could be used to manage earnings (Baber et al. 1991; Dechow and Sloan 1991; Graham et al. 2005; McNichols and Stubben 2008). For example, Graham et al. (2005) indicate that 80% of executives are willing to decrease R &D to meet an earnings target. As previous literature has highlighted the relation between short sales and earnings management (Desai et al. 2006; Karpoff and Lou 2010; Hirshleifer et al. 2011; Boehmer and Wu 2013; Massa et al. 2015; Fang et al. 2016; Park 2017), to test whether the relation between short sales and investment inefficiency is robust, it is worthwhile to exclude R &D expenditures and focus on physical investments that are unrelated to earnings management. Second, according to Amir et al. (2007), R &D expenditures contribute more to earnings volatility than capital expenditures, especially in R &D-intensive industries. There are fundamental differences between R &D expenditures and capital expenditures, which make it more challenging to investigate the subsequent economic benefits of R &D expenditures than those of capital expenditures.

To test the robustness of the relationship between abnormal short interest and firm investment inefficiency, I use total investment excluding R &D expenditures (TIW) to construct a measure of investment inefficiency. I run the same models as in Eqs. (1) and (3) but use *TIW* as the dependent variable. I define the measures of investment inefficiency as $Inefficiency^{BW}$ and $Inefficiency^{MW}$ using Biddle et al. (2009) and McNichols and Stubben (2008), respectively.

Table 8 reports the results. Panel A of Table 8 reports the results of the univariate test. The results show that the means of the measures of investment inefficiency for the top deciles of abnormal short interest are significantly higher than those of the low decile groups. The differences between High and Low (H–L) and between the decile 9 and decile 2 portfolios (Decile 9–Decile 2) are all significantly positive. Panel B of Table 8 reports the results of the panel regressions. The left columns report the results using all firms, while the right columns report the results using the matched sample. When I use all firms, the coefficients of *ASI* for $Inefficiency^{BW}$ and $Inefficiency^{MW}$ are 0.004 and 0.003, respectively. Both of them are significant at the

Table 8 Measure of investment inefficiency that excludes R &D expenditure

Panel A: Univariate test

	(1)	(2)	(3)	(4)	(5)	(6)
	Low	Decile 2	Decile 9	High	Difference (High–Low)	Difference (Decile 9–Decile 2)
$Inefficiency^{BW}$	0.607	0.569	0.639	0.676	0.069***	0.071***
					(2.99)	(3.78)
$Inefficiency^{MW}$	0.618	0.597	0.646	0.671	0.053**	0.049**
					(2.65)	(2.76)

Panel B: Panel regression results

	(1)	(2)	(3)	(4)
	Full sample		Matched sample	
	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
ASI_{t-1}	0.004***	0.003***		
	(4.14)	(2.87)		
$DASI_{t-1}$			0.050***	0.038**
			(2.67)	(2.13)
$MB\ Ratio_{t-1}$	0.003***	0.001	0.004***	0.002
	(3.39)	(1.35)	(2.91)	(1.40)
$Leverage_{t-1}$	0.253***	0.249***	0.278***	0.250***
	(9.34)	(9.55)	(6.30)	(6.13)
ROA_{t-1}	− 0.003**	0.001	− 0.002	− 0.000
	(− 2.08)	(0.37)	(− 0.74)	(− 0.07)
Ln_MV_{t-1}	0.061***	0.012	0.064***	− 0.001
	(6.56)	(1.22)	(3.72)	(− 0.07)

Table 8 (continued)

Panel B: Panel regression results

	(1)	(2)	(3)	(4)
	Full sample		Matched sample	
	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
Ln_Assets_{t-1}	0.039***	0.090***	0.036	0.098***
	(3.00)	(6.29)	(1.59)	(4.02)
$Operating\ Accruals_{t-1}$	0.208***	0.122**	0.199**	0.057
	(4.24)	(2.15)	(2.11)	(0.56)
SG_{t-1}	0.000***		0.000	
	(2.86)		(1.64)	
AG_{t-1}		0.112***		0.114***
		(6.76)		(3.87)
Q_{t-1}		− 0.115***		− 0.123***
		(− 6.67)		(− 3.99)
$Q * QRT2_{t-1}$		0.111***		0.115***
		(10.42)		(5.28)
$Q * QRT3_{t-1}$		0.153***		0.161***
		(11.92)		(6.53)
$Q * QRT4_{t-1}$		0.152***		0.164***
		(10.32)		(6.07)
CF_{t-1}		− 0.000		0.000
		(− 0.59)		(0.23)
TI_{t-1}		− 0.000		− 0.002**
		(− 1.07)		(− 2.06)

Table 8 (continued)

Panel B: Panel regression results

	(1)	(2)	(3)	(4)
	Full sample		Matched sample	
	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
R^2	0.198	0.213	0.334	0.347
Firm fixed effect	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes
Observations	40,594	40,594	14,790	14,790
Number of Firms	4919	4919	3797	3797

Results using the measures of investment inefficiency that exclude R &D expenditures: $Inefficiency^{BW}$ and $Inefficiency^{MW}$. Panel A reports the results of an univariate test, while panel B reports the results of panel regressions. In panel B, the left and right columns report the results using all firms and the matched sample, respectively. I also control for industry and year fixed effects and cluster standard errors at the firm level. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

1% level. When I use the matched sample, the coefficients of *DASI* for $Inefficiency^{BW}$ and $Inefficiency^{MW}$ are 0.050 and 0.038, respectively. They are significant above the 5% level. These results suggest that the positive relationship between the abnormal short interest level and firm investment inefficiency is robust when excluding R &D expenditures.

I also conduct the economic significance analysis using $Inefficiency^{BW}$ and $Inefficiency^{MW}$ and report the results in columns (3) and (4) of Table 5. The results continue to demonstrate significant economic profits from short-selling that utilizes information about investment inefficiency. The coefficients of $Inefficiency^{BW}$ and $Inefficiency^{MW}$ on CAR are significantly negative at the 10% level or above.

6 Subsample analysis

In this section, I am presenting the results of subsample analyses to test Hypothesis 2, 3, and 4. Specifically, I am investigating whether the relation between historical abnormal short interest and investment inefficiency is different between overinvestment or underinvestment (Hypothesis 2), the firms with more or less board independence (Hypothesis 3), and the firms with high or low CEO incentive pay (Hypothesis 4).

6.1 Overinvestment and underinvestment

First, I test the relationship between historical abnormal short interest and investment inefficiency for the subsamples of overinvestment firms and underinvestment firms. Initially, I define overinvestment and underinvestment firms following Chen et al. (2011a). I sort all firms into two groups each year based on the standardized residuals ($\frac{\epsilon_{i,t}}{\sigma_{j,t}}$) from Eqs. (1) and (3). Firms in the top group are classified as overinvestment firms, and firms in the bottom group are categorized as underinvestment firms. Then, I run Eq. (7) for both subsamples and present the results in Table 9. The coefficients of *ASI* on $Inefficiency^{B}$, $Inefficiency^{M}$, $Inefficiency^{BW}$, and $Inefficiency^{MW}$ are 0.006, 0.005, 0.006, and 0.004, respectively, for the overinvestment firms. All of them are statistically significant at 5% or above. Meanwhile, none of the four coefficients for the underinvestment firms is significant, except for $Inefficiency^{M}$. Although the coefficient of *ASI* on $Inefficiency^{M}$ is significant in the subsample of underinvestment firms, the magnitude and the significance level of the coefficient are lower than those in the subsample of overinvestment firms. The results indicate that the relationship between historical abnormal short interest and investment inefficiency is significant for the overinvestment subsample but insignificant for the subsample of underinvestment firms.

The coefficients of historical abnormal short interest for the subsample of overinvestment firms are also larger than those of the whole sample. For example, the coefficient of *ASI* on $Inefficiency^{B}$ is 0.003 ($t = 3.58$) for the whole sample, while it is 0.006 ($t = 3.47$) for the overinvestment firms. The results using other measures are similar. Overall, the significantly positive relationship between historical

Table 9 Subsample analysis: overinvestment and underinvestment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overinvestment				Underinvestment			
	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
ASI_{t-1}	0.006***	0.005***	0.006***	0.004**	0.000	0.002**	0.001	0.001
	(3.47)	(3.01)	(3.48)	(2.34)	(0.69)	(2.48)	(1.36)	(0.94)
R^2	0.354	0.321	0.309	0.328	0.505	0.449	0.474	0.409
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,302	20,302	20,302	20,302	20,292	20,292	20,292	20,292
Number of firms	4361	4361	4361	4361	4210	4210	4210	4210

This table reports the results of subsample analysis based on overinvestment and underinvestment. I sort all firms into two groups each year based on the standardized residuals ($\frac{\epsilon_{i,t}}{\sigma_{j,t}}$) from Eqs. (1) and (3). Firms in the top group are regarded as overinvestment firms and firms in the bottom group are regarded as underinvestment firms (Chen et al. 2011a). Then, I run the following equation for both subsamples.

$$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t},$$

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables including firm size (*Ln_MV* and *Ln_Asset*), market-to-book ratio (*MB Ratio*), leverage (*Leverage*), return on assets (*ROA*), and operating accruals (*Operating Accruals*), which are all measured at a lag of one quarter. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I further add these variables because it is one of the approaches to address the biased coefficients and standard errors when using residuals as the dependent variable in a second regression (Chen et al. 2018). I also introduce firm and year fixed effects and cluster the standard errors at the firm level. For simplicity, I only report the results of the coefficients of *ASI* and the regression R^2. The numbers in the brackets are the *t*-values of the coefficients. Columns (1) to (4) report the results of over-investment firms using $Inefficiency^B$, $Inefficiency^M$, $Inefficiency^{BW}$, and $Inefficiency^{MW}$, respectively, while columns (5) to (8) report the results of under-investment firms the four inefficiency measures, respectively. ***, **, and * denote significance at the 1, 5, and 10% level, respectively abnormal short interest and investment inefficiency only exists for the overinvestment firms, supporting Hypothesis 2.

6.2 Board independence

I test the relationship between historical abnormal short interest and investment inefficiency for the subsamples of firms with little board independence and firms with a high degree of board independence. As an important component of internal governance, board independence can mitigate investment inefficiency (Liu et al. 2015), which makes firms lacking board independence more susceptible to investment inefficiency and more difficult to correct such investment inefficiency.

More importantly, compared with other general governance measures, board independence is linked with short-selling activities (Rahman et al. 2021). Since a lack of board independence will damage disclosure quality and increase information asymmetry, it will facilitate short-selling activities by enhancing the information advantage of short-sellers (Byrd and Hickman 1992; Fama and Jensen 1983; Beasley 1996; Yekini et al. 2015; Rahman et al. 2021).

First, I sort all firms into two groups each year based on board independence. A firm with a board that has a fraction of independent directors above the median is regarded as a firm with high board independence, and a firm with a board that has a fraction of independent directors below the median is regarded as a firm with little board independence. Then, I run Eq. (7) for both subsamples and report the results in Table 10. The results indicate that the relationship between historical abnormal short interest and investment inefficiency is significant for firms with little board independence, as shown in columns (1) to (4), but insignificant for firms with a high degree of independence, as shown in columns (5) to (8). The coefficients of *ASI* using $Inefficiency^{B}$, $Inefficiency^{M}$, $Inefficiency^{BW}$, and $Inefficiency^{MW}$ are 0.006, 0.005, 0.006, and 0.004, respectively, for firms with little board independence. All of them are significant at 5% or above. Meanwhile, none of the four coefficients for firms with a high degree of board independence is significant.

The coefficients of historical abnormal short interest for the subsample of firms with little board independence are also larger than those of the whole sample. For example, the coefficient of *ASI* using $Inefficiency^{B}$ is 0.003 ($t = 3.58$) for the whole sample, while it is 0.006 ($t = 3.31$) for firms with little board independence. The results using other measures are similar. Overall, there is only a significantly positive relationship between historical abnormal short interest and investment inefficiency for firms with little board independence, which supports Hypothesis 3.

6.3 CEO incentive pay

I will next examine whether the relationship between historical abnormal short interest and investment inefficiency varies with CEO incentive pay. Linking managers' compensation to firm performance motivates managers to make more

Table 10 Subsample analysis: board independence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Little board independence				High board independence			
	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
ASI_{t-1}	0.006***	0.005***	0.006***	0.004**	0.001	0.000	0.001	− 0.000
	(3.31)	(2.68)	(2.96)	(2.22)	(0.37)	(0.12)	(0.69)	(− 0.02)
R^2	0.263	0.285	0.248	0.265	0.255	0.260	0.225	0.237
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7808	7808	7808	7808	7905	7905	7905	7905
Number of firms	1318	1318	1318	1318	1177	1177	1177	1177

This table reports the results of a subsample analysis based on board independence. I sort all firms into two groups each year based on board independence. A firm with a board that has a fraction of independent directors above the median is regarded as a firm with a high degree of board independence. A firm with a board that has a fraction of independent directors below the median is regarded as a firm with little board independence. This table presents the results of running the following equation for both subsamples.

$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}$,

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables including firm size (*Ln_MV* and *Ln_Asset*), market-to-book ratio (*MB Ratio*), leverage (*Leverage*), return on assets (*ROA*), and operating accruals (*Operating Accruals*), which are all measured at a lag of one quarter. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I further add these variables because it is one of the approaches to address the biased coefficients and standard errors when using residuals as the dependent variable in a second regression (Chen et al. 2018). I also introduce firm and year fixed effects and cluster the standard errors at the firm level. For simplicity, I only report the results of the coefficients of *ASI* and the regression R^2. The numbers in brackets are the *t*-values of the coefficients. Columns (1) to (4) report the results for firms with little board independence using $Inefficiency^B$, $Inefficiency^M$, $Inefficiency^{BW}$, and $Inefficiency^{MW}$, respectively, while columns (5) to (8) report the results of firms with a high degree of board independence using the four measures of inefficiency. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

Table 11 Subsample analysis: CEO incentive pay

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low CEO_incentive_pay				High CEO_incentive_pay			
	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$	$Inefficiency_t^B$	$Inefficiency_t^M$	$Inefficiency_t^{BW}$	$Inefficiency_t^{MW}$
ASI_{t-1}	0.005***	0.004**	0.006***	0.004***	0.002	0.002	0.003*	0.002
	(3.09)	(2.39)	(3.14)	(2.66)	(1.37)	(0.94)	(1.90)	(0.92)
R^2	0.263	0.260	0.232	0.244	0.303	0.288	0.260	0.260
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster standard errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,054	10,054	10,054	10,054	10,064	10,064	10,064	10,064
Number of firms	1697	1697	1697	1697	1604	1604	1604	1604

This table reports the results of subsample analysis based on CEO incentive pay. I sort all firms into two groups each year based on CEO incentive pay. The CEO incentive pay is calculated by

$$CEO\_incentive\_pay_{i,t} = 1 - \frac{Salary_{i,t} + Bonus_{i,t}}{Total\_CEO\_compensation_{i,t}},$$

where *Total_CEO_compensation* is item TDC1 in the ExecuComp database, which equals the sum of salary, bonus, other annual, the total value of restricted stock granted, the total value of stock options granted calculated using the Black–Scholes option-pricing formula, long-term incentive payouts, and all other totals. Firms with a CEO incentive pay above the median are regarded as having high CEO incentive pay, and firms with a CEO incentive pay below the median are regarded as having low CEO incentive pay. Then, I run the following equation for both subsamples.

$$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t},$$

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables including firm size (*Ln_MV* and *Ln_Asset*), market-to-book ratio (*MB Ratio*), leverage (*Leverage*), return on assets (*ROA*), and operating accruals (*Operating Accruals*), which are all measured at a lag of one quarter. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I further add these variables because it is one of the approaches to address the biased coefficients and standard errors when using residuals as the dependent variable in a second regression (Chen et al. 2018). I also introduce firm and year fixed effects and cluster the standard errors at the firm level. For simplicity, I only report the results of the coefficients of *ASI* and the regression R^2. The numbers in brackets are the *t*-values of the coefficients. Columns (1) to (4) report the results of firms with low CEO incentive pay using $Inefficiency^B$, $Inefficiency^M$, $Inefficiency^{BW}$, and $Inefficiency^{MW}$, respectively, while columns (5) to (8) report the results of firms with high CEO incentive pay using the same four measures. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively value-maximizing decisions (Mehran 1995; Morck et al. 1988; McConnell and Servaes 1990; Jensen and Murphy 1990; Minnick et al. 2011). Short-sellers' threats are significantly lower if a manager's compensation is closely related to stock performance, as managers will be diligent and pay attention to market movements.

To investigate whether the relationship between historical abnormal short interest and investment inefficiency depends on CEO incentive pay, I will sort all firms into two groups each year based on CEO incentive pay. CEO incentive pay is calculated by

$$CEO\_incentive\_pay_{i,t} = 1 - \frac{Salary_{i,t} + Bonus_{i,t}}{Total\_CEO\_compensation_{i,t}},$$

where *Total_CEO_compensation* is item TDC1 in the ExecuComp database, which equals the sum of salary, bonus, other annual, the total value of restricted stock granted, the total value of stock options granted calculated using the Black–Scholes option-pricing formula, long-term incentive payouts, and all other totals. A firm with CEO incentive pay above the median is regarded as having high CEO incentive pay, and a firm with CEO incentive pay below the median is regarded as having low CEO incentive pay. Then, Eq. (7) is run for both subsamples. I also introduce firm and year fixed effects and cluster the standard errors at the firm level.

The results in Table 11 indicate that there is only a significant relationship between historical abnormal short interest and investment inefficiency for firms with low CEO incentive pay. For firms with high CEO incentive pay, the relationship is insignificant. Moreover, the coefficients of *ASI* for firms with low CEO incentive pay are larger than those of the whole sample. For example, the coefficient of *ASI* using $Inefficiency^B$ is 0.003 (t = 3.58) for the whole sample, while it is 0.005 (t = 3.09) for firms with low CEO incentive pay. Short-sellers focus more on the investment inefficiency of firms with low CEO incentive pay when making short selling decisions. Overall, the results in Table 11 support Hypothesis 4.

7 Conclusion

I have researched whether short-sellers possess information about firm investment inefficiency before it becomes publicly available. Utilizing the two inefficiency measures proposed by Biddle et al. (2009) and McNichols and Stubben (2008), as well as measures considering total investment or investment excluding R &D expenditures, I identify a significant relationship between short-selling and investment inefficiency. This relationship holds significance not just statistically but also economically. Furthermore, I determine that the significant relationship between short-selling and investment inefficiency is specific to overinvestment firms, firms with limited board independence, and firms with low CEO incentive pay.

The literature documents that short-sellers can predict downward price movements, but the information advantage they possess remains an empirical question.

Moreover, the effect of short-selling on corporate investment is controversial. It raises the question of whether the reduction in investments is attributable to the disciplinary impact of informed short-sellers or if it arises from the feedback effect of reduced stock prices manipulated by uninformed short-sellers. These debates render investment inefficiency unique in comparison with other instances of poor performance in firms targeted by short-sellers.

I present, for the first time, evidence indicating that short-sellers possess an information advantage concerning investment inefficiency. This insight elucidates the disciplinary impact of short-sellers on investments that erode value, as illuminated by Chang et al. (2019). Given their informed perspective on investment inefficiency, particularly excessive investments, short-sellers exert a disciplinary influence on firms, discouraging their involvement in such value-destroying investments.

The present study provides insights and implications for policymakers. Although short-selling activities have increased tremendously in recent years, they are concentrated in the US market and underrepresented in the financial markets of developing countries. While some developing countries have aimed to relax restrictions on short-selling, the prior research focuses on implications for price efficiency and stock liquidity (Chang et al. 2014; Li et al. 2018). However, the current study offers implications from a different perspective. I find that the informativeness of short-selling activities regarding investment inefficiency centers on firms with overinvestment, limited board independence, and CEO incentive pay. Developing countries that aspire to ease restrictions on short-selling are grappling with significant overinvestment issues (Shi 2019) due to a higher prevalence of government-controlled firms and lower levels of board independence (Liu et al. 2015). Consequently, these nations could potentially benefit more from the disciplinary impact of short-selling activities. The limited board independence necessitates greater external market discipline, prompting policymakers in these nations to further advance efforts in promoting short-selling. While the present study focuses on the US context, the subsample analysis of firms experiencing overinvestment, limited board independence, and CEO incentive pay still offers insights and implications applicable to developing countries grappling with similar issues.

As mentioned earlier, the information advantage of short-sellers falls into two broad categories: either from public information or from private information (Engelberg et al. 2012; Boehmer et al. 2020). If the advantage comes from public information, it implies that short-sellers have better analytical skills than the general public. If the advantage comes from private information, it is evident since the general public does not possess such private information. However, I have not distinguished between these two channels in the present paper. It is expected that there will be some future research in this area regarding whether the information advantage of short-sellers over investment inefficiency arises from their superior analytical skills or their unique information that the general public does not possess.

Appendix: Definitions of the Variables

Variables	Definition
Inefficiency^{B}	Investment inefficiency measure calculated using the model of Biddle et al. (2009)
Inefficiency^{M}	Investment inefficiency measure calculated using the model of McNichols and Stubben (2008)
Inefficiency^{BW}	Investment inefficiency measure calculated using the model of Biddle et al. (2009) with the total investment excluding R &D as the dependent variable
Inefficiency^{MW}	Investment inefficiency measure calculated using the model of McNichols and Stubben (2008) with the total investment excluding R &D as the dependent variable
SI	Short interest level
ASI	Abnormal short interest, which is defined as the difference between the short interest and expected short interest calculated using the model proposed in Karpoff and Lou (2010)
DASI	A dummy variable equal to one if abnormal short interest is in the top quintile and zero otherwise
CAR	Cumulative abnormal returns relative to size, MB ratio, and momentum benchmark portfolio over a period of 6 months after the release of the annual report
TI	Total investment, calculated as the sum of capital expenditures (Compustat data item 128), R &D expenditures (Compustat data item 46) and acquisitions (Compustat data item 129) minus the sale of property, plant and equipment (Compustat data item 107), scaled by total book value of assets (Compustat data item 6) at the beginning of the fiscal year
TIW	Total investment excluding R &D expenditures. It equals the sum of capital expenditures (Compustat data item 128) and acquisitions (Compustat data item 129) minus the sale of property, plant and equipment (Compustat data item 107), scaled by total book value of assets (Compustat data item 6) at the beginning of the fiscal year
SG	Sales growth, which equals sales at the end of current year minus sales at the end of previous year, divided by sales at the end of previous year
AG	Assets growth, equal to the natural log of total book value of assets (Compustat data item 6) at the end of current year divided by total book value of assets at the end of previous year
CF	Operating cash flow, equal to net cash flow from operating activities (Compustat data item 308) divided by total book value of assets (Compustat data item 6) at the beginning of the fiscal year
Q	Tobin's Q, equal to (MV + AT − CEQ)/AT, where MV is the market value of equity (Compustat data item 25 x Compustat data item 199), AT is the total book value of assets (Compustat data item 6), and CEQ is the book value of common equity (Compustat data item 60)
QRT2	An indicator variable that equals 1 if $Q_{i,t-1}$ is in the second quartile of its industry-year distribution and zero otherwise
QRT3	An indicator variable that equals 1 if $Q_{i,t-1}$ is in the third quartile of its industry-year distribution and zero otherwise
QRT4	An indicator variable that equals 1 if $Q_{i,t-1}$ is in the fourth quartile of its industry-year distribution and zero otherwise
$Size_{low}$	Dummy variable that indicates whether a firm is sorted into the lowest size portfolio in month *t*. Firms are ranked into three size-groups, lowest, medium and highest, using their market values

Variables	Definition
$Size_{medium}$	Dummy variable that indicates whether a firm is sorted into the medium size portfolio in month *t*. Firms are ranked into three size-groups, lowest, medium, and highest, using their market values
MB_{low}	Dummy variable that indicates whether a firm is sorted into the lowest market-to-book ratio portfolio in month *t*. Firms are ranked into three market-to-book ratio groups, lowest, medium, and highest, using their ratios of market value of equity over book value of equity
MB_{medium}	Dummy variable that indicates whether a firm is sorted into the medium market-to-book ratio portfolio in month *t*. Firms are ranked into three market-to-book ratio groups, lowest, medium, and highest, using their ratios of market value of equity over book value of equity
$Momentum_{low}$	Dummy variable that indicates whether a firm's momentum is sorted into the lowest portfolio in month *t*. Firms are ranked into three momentum groups, lowest, medium, and highest, using their returns in month $t-1$
$Momentum_{medium}$	Dummy variable that indicates whether a firm's momentum is sorted into the medium portfolio in month *t*. Firms are ranked into three momentum groups, lowest, medium, and highest, using their returns in month $t-1$
MB Ratio	Firm's market-to-book ratio in the quarter prior to end of fiscal year, calculated as market value of equity divided by book value of equity (Compustat data item 59)
Leverage	Firm's leverage in the quarter prior to the end of fiscal year, calculated as (DLTTQ + DLCQ)/(DLTTQ + DLCQ + SEQQ), where DLTTQ and DLCQ are the long-term debt (Compustat data item 51) and short-term debt (Compustat data item 45) of a firm, respectively. SEQQ is the total shareholders' equity (Compustat data item 60)
ROA	Firm's return on assets in the quarter prior to the end of fiscal year, calculated as OIBDPQ/ATQ, where OIBDPQ is the income before depreciation and amortization (Compustat data item 21) and ATQ is the total book value of assets (Compustat data item 44)
Ln_MV	The natural logarithm of firm's market value in the quarter prior to the end of fiscal year
Ln_Assets	The natural logarithm of firm's total book value of assets (item 44) in the quarter prior to the end of fiscal year
Operating Accruals	Operating accruals in the quarter prior to the end of fiscal year, calculated as IBCY-(OANCFY-XIDOCY), where IBCY is the income before extraordinary items (Compustat data item 76). OANCFY is the net cash flow from operating activities (Compustat data item 108). XIDOCY is the cash flow from extraordinary items and discontinued operations (Compustat data item 78). All variables are standardized by the average total book value of assets (Compustat data item 44), which is the average between the total book value of assets in the current quarter and that in the previous quarter
Board independence	Number of independent directors divided by total number of directors on the board
CEO_incentive_pay	$CEO\_incentive\_pay_{i,t} = 1 - \frac{Salary_{i,t}+Bonus_{i,t}}{Total\_CEO\_compensation_{i,t}}$, where *Total_CEO_compensation* is item TDC1 in the ExecuComp database, which is the sum of the salary, bonus, other annual, the total value of restricted stock granted, the total value of stock options granted calculated using the Black–Scholes option-pricing formula, long-term incentive payouts, and all other totals
Dividend Cut	Dummy variable that indicates whether a firm's dividend has decreased compared to the previous year

Variables	Definition
HHI	Competition is measured by the Herfindahl–Hirschman Index. The Herfindahl Index is defined as the sum of squared market shares: $HHI_{i,t} = \sum_{i=1}^{N} S_{i,j,t}^{2}$, where $S_{i,j,t}$ is the market share of firm *i* in industry *j* in year *t*, N_j is the number of firms in industry *j* in year *t*, and $HHI_{j,t}$ is the Herfindahl Index of industry *j* in year *t*. The market share of an individual firm is calculated by using the firm's net sales divided by the total sales value of the entire industry (Gu 2016; Giroud and Mueller 2010)
Restatement	Dummy variable that indicates whether a material inadvertent (unintentional) or fraudulent (intentional) error exists in a firm's financial statements. The information is sourced from the Audit Analytics (AA) database. Material restatements are identified based on the filing of Form 8-K Item 4.02, beginning on August 23, 2004 (Bartov et al. 2021)
Momentum	Cumulative monthly stock return over the past 12 months

Appendix: Baseline regression result with additional control variables

	(1) $Inefficiency_t^B$	(2) $Inefficiency_t^M$
ASI_{t-1}	0.004***	0.003***
	(3.76)	(3.40)
$MB\ Ratio_{t-1}$	0.002**	0.002*
	(1.98)	(1.68)
$Leverage_{t-1}$	0.182***	0.195***
	(6.58)	(7.65)
ROA_{t-1}	− 0.008***	0.001
	(− 4.96)	(0.37)
Ln_MV_{t-1}	0.051***	− 0.020
	(5.45)	(− 1.64)
Ln_Assets_{t-1}	0.010	0.111***
	(0.69)	(7.14)
$Operating\ Accruals_{t-1}$	− 0.068	− 0.133**
	(− 1.18)	(− 2.05)
$Dividend\ Cut_t$	0.051***	0.050***
	(5.11)	(4.62)
HHI_t	0.047	0.034
	(0.63)	(0.45)
$Momentum_t$	0.042***	0.063***
	(7.17)	(6.34)
$Restatement_t$	0.010	− 0.027
	(0.34)	(− 0.85)
SG_{t-1}	0.000	
	(1.29)	

	(1) $Inefficiency^B_t$	(2) $Inefficiency^M_t$
AG_{t-1}		− 0.035*
		(− 1.79)
Q_{t-1}		− 0.110***
		(− 6.68)
$Q * QRT2_{t-1}$		0.113***
		(11.03)
$Q * QRT3_{t-1}$		0.161***
		(12.75)
$Q * QRT4_{t-1}$		0.159***
		(11.05)
CF_t		− 0.003***
		(− 5.54)
TI_{t-1}		0.003***
		(9.17)
R^2	0.267	0.235
Firm fixed effect	Yes	Yes
Year fixed effect	Yes	Yes
Cluster standard errors	Yes	Yes
Observations	40,594	40,594
Number of firms	4919	4919

This table reports the panel regression results for the relation between historical abnormal short interest and investment inefficiency. I run the following panel regressions using all data,

$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t}$,

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. The abnormal short interest is measured on the 15th of each month or the preceding business day if the 15th is not a business day. The lag between the date of abnormal short interest and the date of annual release of the financial report should not exceed one month. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables included in Table 4. I have also included dividend cut ($Dividend\ Cut$), competition (HHI), momentum ($Momentum$), and material restatement ($Restatement$) as control variables. Detailed explanations for these additional control variables are provided in Appendix: Definition of the Variables. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I consider both firm and year fixed effects and cluster the standard errors at the firm level. Columns (1) and (2) report the results of $Inefficiency^B$ and $Inefficiency^M$, respectively. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

Appendix: Alternative explanation-financial misstatement

	(1) $Inefficiency_t^B$	(2) $Inefficiency_t^M$
$ASI_{t-1} \times Restatement_t$	− 0.006	− 0.002
	(− 0.74)	(− 0.34)
ASI_{t-1}	0.004***	0.002***
	(3.66)	(3.41)
$Restatement_t$	0.015	− 0.028
	(0.47)	(− 0.87)
R^2	0.265	0.235
Control variables	Yes	Yes
Firm fixed effect	Yes	Yes
Year fixed effect	Yes	Yes
Cluster standard errors	Yes	Yes
Observations	40,594	40,594
Number of firms	4919	4919

This table examines whether the relationship between short-selling activities before the financial report release and investment inefficiency will be affected by a significant financial misstatement in the financial report. I have executed the following model:

$$Inefficiency_{i,t} = \alpha + \beta_1 \cdot ASI_{i,t-1} \times Restatement_{i,t} + \beta_2 \cdot ASI_{i,t-1} + \beta_3 \cdot Restatement_{i,t} + X_{i,t-1} \cdot \eta + Y_{i,t-1} \cdot \gamma + \epsilon_{i,t},$$

where $ASI_{i,t-1}$ is the abnormal short interest before the release of the annual financial statement. $Restatement_{i,t}$ is a dummy variable that indicates whether a material inadvertent (unintentional) or fraudulent (intentional) error exists in a firm's financial statements. $ASI_{i,t-1} \times Restatement_{i,t}$ is the interaction term between $ASI_{i,t-1}$ and $Restatement_{i,t}$. $Inefficiency_{i,t}$ is the measure of investment inefficiency. $X_{i,t-1}$ refers to a list of control variables included in Table 4. $Y_{i,t-1}$ refers to the list of independent variables that are used in the investment inefficiency models of Biddle et al. (2009) and McNichols and Stubben (2008). I consider both firm and year fixed effects and cluster the standard errors at the firm level. Columns (1) and (2) report the results of $Inefficiency^B$ and $Inefficiency^M$, respectively. ***, **, and * denote significance at the 1, 5, and 10% levels, respectively

Acknowledgements I thank Professor Hai Lin and Professor Jianfeng Shen for their support and comments on this paper. I thank Dr.Griffin Geng, Dr. Peter Huang, and Prof. Alfred Yawson for their comments on this paper in the examiner's report of my Ph.D. thesis. I also thank the anonymous referee for the helpful comments. All remaining errors are my own. I appreciate Victoria University of Wellington for a 6000 NZD research fund that can be used for purchasing software, attending conferences, journal submissions, and other research-related purposes. Besides the above, I have no relevant financial or non-financial interests to disclose. The data that support the findings of this study are available from the corresponding author upon reasonable request

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

References

Amir, E., Guan, Y., Livne, G.: The association of R &D and capital expenditures with subsequent earnings variability. J. Bus. Finance Account. **34**, 222–246 (2007)

Armstrong, C.S., Jagolinzer, A.D., Larcker, D.F.: Chief executive officer equity incentives and accounting irregularities. J. Account. Res. **48**, 225–271 (2010)

Baber, W.R., Fairfield, P.M., Haggard, J.A.: The effect of concern about reported income on discretionary spending decisions: the case of research and development. Account. Rev. **66**, 818–829 (1991)

Balakrishnan, K., Core, J.E., Verdi, R.S.: The relation between reporting quality and financing and investment: evidence from changes in financing capacity. J. Account. Res. **52**, 1–36 (2014)

Bartov, E., Marra, A., Momenté, F.: Corporate social responsibility and the market reaction to negative events: evidence from inadvertent and fraudulent restatement announcements. Account. Rev. **96**, 81–106 (2021)

Beasley, M.S.: An empirical analysis of the relation between the board of director composition and financial statement fraud. Account. Rev. 443–465 (1996)

Bebchuk, L.A., Stole, L.A.: Do short-term objectives lead to under-or overinvestment in long-term projects? J. Finance **48**, 719–729 (1993)

Beneish, M.D., Vargus, M.E.: Insider trading, earnings quality, and accrual mispricing. Account. Rev. **77**, 755–791 (2002)

Benlemlih, M., Bitar, M.: Corporate social responsibility and investment efficiency. J. Bus. Ethics **148**, 647–671 (2018)

Bergstresser, D., Philippon, T.: CEO incentives and earnings management. J. Financial Econ. **80**, 511–529 (2006)

Biddle, G.C., Hilary, G.: Accounting quality and firm-level capital investment. Account. Rev. **81**, 963–982 (2006)

Biddle, G.C., Hilary, G., Verdi, R.S.: How does financial reporting quality relate to investment efficiency? J. Account. Econ. **48**, 112–131 (2009)

Boehmer, E., Wu, J.: Short selling and the price discovery process. Rev. Financial Stud. **26**, 287–322 (2013)

Boehmer, E., Jones, C.M., Zhang, X.: Which shorts are informed? J. Finance **63**, 491–527 (2008)

Boehmer, E., Jones, C.M., Wu, J., Zhang, X.: What do short sellers know? Rev. Finance **24**, 1203–1235 (2020)

Brav, A., Jiang, W., Ma, S., Tian, X.: How does hedge fund activism reshape corporate innovation? J. Financial Econ. **130**, 237–264 (2018)

Bushman, R.M., Piotroski, J.D., Smith, A.J.: Capital allocation and timely accounting recognition of economic losses. J. Bus. Finance Account. **38**, 1–33 (2011)

Byrd, J.W., Hickman, K.A.: Do outside directors monitor managers?: evidence from tender offer bids. J. Financial Econ. **32**, 195–221 (1992)

Caby, J., et al.: The impact of short selling on firms: an empirical literature review. J. Bus. Account. Finance Perspect. **2**, 1–9 (2020)

Chang, E.C., Luo, Y., Ren, J.: Short-selling, margin-trading, and price efficiency: evidence from the Chinese market. J. Bank. Finance **48**, 411–424 (2014)

Chang, E.C., Lin, T.-C., Ma, X.: Does short-selling threat discipline managers in mergers and acquisitions decisions? J. Account. Econ. **68**, 101223 (2019)

Chen, F., Hope, O.-K., Li, Q., Wang, X.: Financial reporting quality and investment efficiency of private firms in emerging markets. Account. Rev. **86**, 1255–1288 (2011)

Chen, S., Sun, Z., Tang, S., Wu, D.: Government intervention and investment efficiency: evidence from China. J. Corp. Finance **17**, 259–271 (2011)
Chen, W., Hribar, P., Melessa, S.: Incorrect inferences when using residuals as dependent variables. J. Account. Res. **56**, 751–796 (2018)
Cooper, M.J., Gulen, H., Schill, M.J.: Asset growth and the cross-section of stock returns. J. Finance **63**, 1609–1651 (2008)
Dai, R., Li, T., Zaiats, N., Zhao, X.: Do short sellers anticipate late filings? J. Corp. Finance **69**, 102045 (2021)
Daniel, K., Grinblatt, M., Titman, S., Wermers, R.: Measuring mutual fund performance with characteristic-based benchmarks. J. Finance **52**, 1035–1058 (1997)
De Angelis, D., Grullon, G., Michenaud, S.: The effects of short-selling threats on incentive contracts: evidence from an experiment. Rev. Financial Stud. **30**, 1627–1659 (2017)
Dechow, P.M., Sloan, R.G.: Executive incentives and the horizon problem: an empirical investigation. J. Account. Econ. **14**, 51–89 (1991)
Dechow, P.M., Hutton, A.P., Meulbroek, L., Sloan, R.G.: Short-sellers, fundamental analysis, and stock returns. J. Financial Econ. **61**, 77–106 (2001)
DeFond, M.L., Park, C.W.: Smoothing income in anticipation of future earnings. J. Account. Econ. **23**, 115–139 (1997)
Desai, H., Krishnamurthy, S., Venkataraman, K.: Do short sellers target firms with poor earnings quality? Evidence from earnings restatements. Rev. Account. Stud. **11**, 71–90 (2006)
Diether, K.B., Lee, K.-H., Werner, I.M.: Short-sale strategies and return predictability. Rev. Financial Stud. **22**, 575–607 (2009)
Engelberg, J.E., Reed, A.V., Ringgenberg, M.C.: How are shorts informed?: short sellers, news, and information processing. J. Financial Econ. **105**, 260–278 (2012)
Fama, E.F., Jensen, M.C.: Separation of ownership and control. J. Law Econ. **26**, 301–325 (1983)
Fang, V.W., Huang, A.H., Karpoff, J.M.: Short selling and earnings management: a controlled experiment. J. Finance **71**, 1251–1294 (2016)
Francis, J.R., Martin, X.: Acquisition profitability and timely loss recognition. J. Account. Econ. **49**, 161–178 (2010)
Franzoni, F.: Underinvestment vs. overinvestment: evidence from price reactions to pension contributions. J. Financial Econ. **92**, 491–518 (2009)
Gao, R., Yu, X.: How to measure capital investment efficiency: a literature synthesis. Account. Finance **60**, 299–334 (2020)
Gilchrist, S., Himmelberg, C.P., Huberman, G.: Do stock price bubbles influence corporate investment? J. Monet. Econ. **52**, 805–827 (2005)
Giroud, X., Mueller, H.M.: Does corporate governance matter in competitive industries? J. Financial Econ. **95**, 312–331 (2010)
Goldstein, I., Guembel, A.: Manipulation and the allocational role of prices. Rev. Econ. Stud. **75**, 133–164 (2008)
Gomariz, M.F.C., Ballesta, J.P.S.: Financial reporting quality, debt maturity and investment efficiency. J. Bank. Finance **40**, 494–506 (2014)
Goodman, T.H., Neamtiu, M., Shroff, N., White, H.D.: Management forecast quality and capital investment decisions. Account. Rev. **89**, 331–365 (2014)
Graham, J.R., Harvey, C.R., Rajgopal, S.: The economic implications of corporate financial reporting. J. Account. Econ. **40**, 3–73 (2005)
Grullon, G., Michenaud, S., Weston, J.P.: The real effects of short-selling constraints. Rev. Financial Stud. **28**, 1737–1767 (2015)
Gu, L.: Product market competition, R &D investment, and stock returns. J. Financial Econ. **119**, 441–455 (2016)
Hirshleifer, D., Teoh, S.H., Yu, J.J.: Short arbitrage, return asymmetry, and the accrual anomaly. Rev. Financial Stud. **24**, 2429–2461 (2011)
Jensen, M.C., Murphy, K.J.: Performance pay and top-management incentives. J. Polit. Econ. **98**, 225–264 (1990)
Jiao, Y., Massa, M., Zhang, H.: Short selling meets hedge fund 13f: an anatomy of informed demand. J. Financial Econ. **122**, 544–567 (2016)
Karpoff, J.M., Lou, X.: Short sellers and financial misconduct. J. Finance **65**, 1879–1913 (2010)
Li, Z., Lin, B., Zhang, T., Chen, C.: Does short selling improve stock price efficiency and liquidity? Evidence from a natural experiment in China. Eur. J. Finance **24**, 1350–1368 (2018)

Liu, Y., Miletkov, M.K., Wei, Z., Yang, T.: Board independence and firm performance in China. J. Corp. Finance **30**, 223–244 (2015)
Massa, M., Zhang, B., Zhang, H.: The invisible hand of short selling: does short selling discipline earnings management? Rev. Financial Stud. **28**, 1701–1736 (2015)
McConnell, J.J., Servaes, H.: Additional evidence on equity ownership and corporate value. J. Financial Econ. **27**, 595–612 (1990)
McNichols, M.F., Stubben, S.R.: Does earnings management affect firms' investment decisions? Account. Rev. **83**, 1571–1603 (2008)
Mehran, H.: Executive compensation structure, ownership, and firm performance. J. Financial Econ. **38**, 163–184 (1995)
Minnick, K., Unal, H., Yang, L.: Pay for performance? CEO compensation and acquirer returns in BHCs. Rev. Financial Stud. **24**, 439–472 (2011)
Morck, R., Shleifer, A., Vishny, R.W.: Management ownership and market valuation: an empirical analysis. J. Financial Econ. **20**, 293–315 (1988)
Park, K.: Earnings quality and short selling: evidence from real earnings management in the United States. J. Bus. Finance Account. **44**, 1214–1240 (2017)
Rahman, A., Talukdar, B., Bhuyan, R.: Board independence and short selling. Finance Res. Lett. **39**, 101616 (2021)
Richardson, S.: Over-investment of free cash flow. Rev. Account. Stud. **11**, 159–189 (2006)
Saffi, P.A., Sigurdsson, K.: Price efficiency and short selling. Rev. Financial Stud. **24**, 821–852 (2011)
Shi, M.: Overinvestment and corporate governance in energy listed companies: evidence from China. Finance Res. Lett. **30**, 436–445 (2019)
Shi, W., Ndofor, H.A., Hoskisson, R.E.: Disciplining role of short sellers: evidence from M &A activity. J. Manag. **47**, 1103–1133 (2021)
Shleifer, A., Vishny, R.W.: A survey of corporate governance. J. Finance **52**, 737–783 (1997)
Singer, Z., Wang, Y., Zhang, J.: Can short sellers detect internal control material weaknesses? Evidence from section 404 of the Sarbanes-oxley Act. J. Account. Auditing Finance **37**, 3–38 (2022)
Souther, M.E.: Does board independence increase firm value? Evidence from closed-end funds. J. Financial Quant. Anal. **56**, 313–336 (2021)
Titman, S., Wei, K.J., Xie, F.: Capital investments and stock returns. J. Financial Quant. Anal. **39**, 677–700 (2004)
Yekini, K.C., Adelopo, I., Andrikopoulos, P., Yekini, S.: Impact of board independence on the quality of community disclosures in annual reports. Account. Forum **39**, 249–267 (2015)
Yu, C.: Board gender diversity and investment inefficiency. J. Econ. Bus. **124**, 106–107 (2023)

Chang Yu holds a Ph.D. in finance from Victoria University of Wellington. During his Ph.D., his research interests focused on short selling, board gender diversity, and corporate investment. After graduation, he worked as a data analyst at ESR, one of the crown research institutes in New Zealand, for one year. Currently, he is a research fellow at the University of Otago where he focuses on the linkage between economics and public health as a health economist.