Disclosure of private in-house meetings and stock crash risk: evidence from China

Hang Zhou (corresponding author)

University of Edinburgh Business School, University of Edinburgh, Edinburgh, EH8 9JS, UK

Rong Ding

Warwick Business School, University of Warwick, Coventry, CV4 7AL UK

Yifan Li

Alliance Manchester Business School, University of Manchester, Manchester, UK

Yuxin Sun

Queen's University Belfast, Belfast, UK

Abstract

In July 2012, Shenzhen Stock Exchange mandated that all listed firms electronically publish a standard summary report for each private in-house meeting through its web portal. This study develops a model to analyze the relation between disclosure of in-house meetings and stock crash risk. Our model predicts a negative association between them, and that the textual characteristics of in-house meeting reports (readability and sentiment) significantly impact on crash risk. Using data collected from Chinese listed firms between 2012 and 2017, we find results consistent with the model's predictions.

Keywords: in-house meeting; crash risk; readability; sentiment; informed trading

JEL: D82 G15 G19 M48

1. Introduction

Previous literature suggests that corporate managers can often access private information about their firms' operations, asset value, and future growth opportunities that is unavailable to external investors. However, managers may strategically withhold or delay the disclosure of bad news, expecting that it can be offset by subsequent good news. Doing so enables managers to protect their future career, maintain the esteem of peers, and maximize their equity-based compensation (Core et al. 2003; Kothari et al. 2009; Kim et al. 2011). Once the amount of accumulated bad news reaches a certain tipping point, managers are unable to continue withholding. Consequently, a large amount of negative information will be simultaneously released to the market, leading to a sudden decline in stock price, also termed a stock crash (Jin and Myers 2006).

In-house meetings allow attendees to visit a firm's headquarters, interact directly with the senior management team, and observe plant operations (Solomon and Soltes 2015). There is a consensus that in-house meetings inform investors about the firm's future operations and challenges, and so decrease the probability of unfavorable news being accumulated and eventually released to the market en masse. However, because only a select group of investors are permitted to participate in in-house meetings, the benefit is not equally shared among all market participants; this makes it difficult to explore how in-house meeting disclosures impact on crash risk.

We seek to fill this research gap by exploiting a unique mandatory disclosure requirement for firms listed on the Shenzhen Stock Exchange (SZSE) in China. Since July 2012, SZSE has mandated that all listed firms publish an electronic copy of a standard summary report for each in-house meeting through its web portal (http://irm.cninfo.com.cn/szse). As these reports are publicly accessible, they help to reduce information asymmetry between corporate insiders and external investors, leading to lower future stock crash risk. By contrast, such disclosure requirements do not apply to firms listed on Shanghai Stock Exchange (SHSE), which creates a natural control group for us to implement a difference-in-difference (DiD) test.[1]

We suggest that the disclosure requirement has two potential effects on future crash risk: first, because information released during in-house meetings receives higher public attention when accessible on the web portal, it is increasingly difficult for managers to withhold negative information, thus decreasing the probability of unfavorable news being disclosed to the market en masse. Second, as investors have timely access to in-house meeting reports, they can trade on the information released therein, facilitating the incorporation of both positive and negative information into the stock price. The overall consequence is lower future crash risk. Therefore, we expect to find that the crash risk of SZSE-listed firms has decreased since July 2012 relative to that of SHSE-listed firms.[2]

To further investigate the mechanism, we develop a model based on Jin and Myers's (2006) framework to show how the cross-sectional variation in in-house meeting disclosure affects future crash risk. The main takeaway of the model is that the number of in-house meetings and the textual characteristics of the released in-house meeting reports (i.e., readability and tone) are associated with future crash risk (details of the model are discussed in section 3).

[1] Before 2012, the only disclosure mandated of SZSE-listed firms was their annual reports. Consequently, the issues discussed during in-house meetings were unknown to non-attending investors before July 2012.

[2] Private interactions between investors and managers could occur at a variety of venues including conferences, investors' offices, and firms' headquarters (Solomon and Soltes 2015). Due to data availability, we narrow our attention to in-house meetings at firms' headquarters.

We design a program with Python to download and categorize information from these meeting reports released on the SZSE web portal, and use the data to test our model's predictions. First, we count the annual number of each firm's in-house meetings and expect that a higher number facilitates the flow of more extensive firm-specific information to the market, thus reducing future crash risk. Second, we follow the literature on readability (i.e., Li 2008) by computing the average readability of each firm's in-house meeting reports on an annual basis; we predict that higher average readability makes it easier for readers to understand and consequently trade on the firm-specific information released in these meeting reports, enabling the information to be better incorporated into the stock price and become public knowledge. This effectively decreases firm-level information asymmetry, resulting in reduced future crash risk. Finally, we follow the literature on "sentiment" (Tetlock 2007; Loughran and McDonald 2011) by calculating the average sentiment of all in-house meeting reports for each firm on an annual basis; we predict that negative sentiment helps the release of unfavorable news to the market and decreases the likelihood of bad news being accumulated and released to the market en masse, which implies lower future crash risk.[3]

Based on the analysis of 35,976 in-house meeting reports published by SZSE-listed firms between 2012 and 2017, we report the following findings. First, the crash risk becomes lower for SZSE-listed firms relative to SHSE-listed firms after 2012, which suggests that the mandatory disclosure of in-house meeting reports increases corporate transparency and reduces the probability of a future stock price crash. Second, the number

[3] Researchers typically count the number of words associated with a particular sentiment word list *scaled* by the total number of words in the document to measure the tone or sentiment of accounting and financial documents (Tetlock 2007, 2008; Loughran and McDonald 2014).

of in-house meetings is negatively associated with future crash risk, which is consistent with the general public being aware of the firm-specific information disclosed after in-house meetings, thus effectively decreasing the future crash risk. Third, the (average) negative sentiment of in-house meeting reports can inform the market of the difficulties and challenges confronting the firm, which reduces the likelihood of unfavorable news being continuously accumulated and released en masse, resulting in lower crash risk.

We further explore the channel though which in-house meetings disclosure can affect crash risk. We show that the effect is concentrated in firms with higher informed trading and lower institutional ownership. These findings suggest that investors who did not have the opportunity to attend the in-house meeting become aware of the issues discussed therein, and may trade on the information released from the disclosure of in-house meeting reports, which mitigates information asymmetry and decreases the crash risk.

Our contribution to the literature is threefold. First, we extend the accounting literature on disclosure by showing that private in-house meetings are an important disclosure channel that has received limited attention. Data availability enables us to investigate this type of corporate disclosure by analyzing publicly available reports of listed firms' in-house meetings, with incremental value for investors. Prior research on in-house meetings largely focuses on the potential benefits received by a select group of investors (i.e., in-house meeting attendees).[4] Our study differs by unraveling one of the capital

[4] Solomon and Soltes (2015) use a proprietary dataset that records all private meetings between senior management and investors for a NYSE-listed firm over a six-year period (2004-2010), and find that: 1) investors who attended the meetings make more correlated trades relative to others who did not meet management privately; 2) the trades of meeting attendees have greater predictive power of future stock returns (relative to those of non-attendees). These findings suggest that investors receive valuable information for making informed trades from private interactions with management. Bushee et al. (2018) use a unique sample of corporate jet flights as a proxy for managers attending non-deal roadshows, and report evidence of greater abnormal stock reactions during the roadshow period.

market consequences of in-house meeting issues being disclosed to market participants in China: lower stock crash risk. Second, we contribute to the literature on stock crash risk by clarifying the impact of in-house meetings on future crash risk. Finally, we add to the burgeoning research on in-house meetings by focusing on the unexplored textual characteristics of in-house meeting reports—an alternative source of value-relevant information.[5]

The remainder of the paper is structured as follows. Section 2 reviews the literature and section 3 presents the model. The sample and research design are described in section 4, and section 5 presents the empirical results. Finally, section 6 concludes.

2. Literature review

2.1 Literature on stock crash risk

Jin and Myers (2006) posit that when a firm is not completely transparent, its managers can capture a portion of cash flow that cannot be observed by outsiders. To protect themselves, managers may partially absorb downside risk and loss caused by temporarily unfavorable firm performance by delaying the release of bad news until a certain point. Benmelech et al. (2010) show that stock-based compensation induces managers to conceal bad news about future growth, thereby inflating the stock price and, ultimately, triggering a stock crash. However, incentive alone would not suffice to allow management to withhold bad news. Earnings management, tax avoidance, and voluntary

[5] Bowen et al. (2018) examine inside trading around private in-house meetings using data for SZSE-listed firms between 2012 and 2014. They find that corporate insiders time their trading based on the nature of news released in the in-house meetings and realize substantial gains. Our study differs by investigating the association between disclosure of in-house meetings and stock crash risk, and our sample covers a longer period (2012-2017), which is expected to increase the power of our test.

disclosure have been identified as some of the mechanisms used by managers to conceal unfavorable news (Hutton et al. 2009; Kim et al. 2011). For example, Hutton et al. (2009) find that stock crash risk is associated with financial reporting transparency, measured by the three-year average value of discretionary accruals. DeFond et al. (2015) report that the adoption of International Financial Reporting Standards (IFRS) reduced crash risk for non-financial firms, with a more pronounced effect for firms with severe information asymmetry prior to IFRS adoption and those for which local generally accepted accounting principles more credibly changed after IFRS adoption. In Kim and Zhang's (2015) study, firms that are conservative in financial reporting (more timely recognition of losses) have reduced stock crash risk. Meanwhile, Kim et al. (2016) find that overconfident managers are likely to misperceive projects with negative net present value as value-creating projects, which eventually leads to a stock price crash.

2.2 Literature on readability

In an influential and widely-cited paper, Li (2008) develops techniques to measure the readability of 10-K reports. Analyzing a large sample of annual reports of US public firms between 1994 and 2004, he finds that firms with lower earnings tend to have 10-Ks that are difficult to read (high Fog index value or high word count), which is consistent with managers intentionally limiting the readability of financial statements to conceal unfavorable information. Li (2008) also reports that firms with less readable annual reports have lower earnings persistence. Lehavy (2011) shows that firms with less readable 10-K reports are followed by more analysts and the reports these analysts produce have higher informativeness. Such findings suggest that the demand for analyst services is associated with the lexical properties of financial reports.

Regarding the capital market consequences of financial report readability, Miller (2010) shows that small investors reduce their trading when the length of 10-K reports increases, and Loughran and McDonald (2009) find a significant increase in the proportion of small trades when firms show improvement in plain English compliance (which leads to more readable financial reports). Callen et al. (2010) use the Fog index to measure the quality of investors' existing information set and find that lower quality information (i.e., less readable financial reports) is associated with greater stock price delay. You and Zhang (2009) find that there is price momentum after 10-K filing dates for firms with long annual reports that are hard to read, suggesting that investors under-react to information provided in longer and less readable 10-K filings.

A burgeoning stream of research is examining the readability of alternative sources of investment-relevant information: analyst reports. Based on analyzing over 350,000 analyst reports downloaded from Investext for US public companies between 2002 and 2009, De Franco et al. (2015) show that more readable reports are issued by "high-ability" experts (i.e., those who issue more timely forecasts, revise forecasts more frequently, and issue more consistent forecasts and recommendation revisions). They also find that more readable analyst reports are associated with higher trading volumes over a three-day window surrounding the analyst report date, which is consistent with the theoretical prediction that more precise information induces investors to trade.

2.3 Literature on tone/sentiment

Tetlock (2007) conducted the first investigation of the capital market consequences of tone in financial news. The author relies on the Harvard General Inquirer Word List to measure the tone of *Wall Street Journal's* daily column "Abreast of the market" between 1884 and 1999, and shows that high journalistic pessimism is related to lower subsequent stock returns and higher subsequent stock market volatility. Extending the work of Tetlock (2007), Garcia (2013) uses Loughran and McDonald's (2011) positive and negative word list to measure the sentiment of two financial columns in *New York Times* between 1905 and 2005, and concludes that newspaper sentiment plays a crucial role in predicting future stock returns during recession periods. Tetlock et al. (2008) examine *Wall Street Journal* and *Dow Jones News Service* stories on S&P 500 firms between 1980 and 2004, and report that a higher frequency of negative words in firm-specific news stories predicts lower subsequent earnings; this effect is also more pronounced if the negative words concern company fundamentals. Furthermore, negative information embedded in negative tone seems to be incorporated into stock price with a delay. Using Harvard IV-4 negative and positive word categories, Kothari et al. (2009) analyze the content of disclosure by firms, analysts, and the media. They posit that disclosure tone is associated with both stock return volatility and analyst forecast dispersion. In particular, more positive firm-related disclosure by the firm itself, analysts, and news outlets is related to lower volatility and smaller forecast dispersion.

Attempting to increase the power of their test, Davis et al. (2015) examine manager-specific optimism during earnings conference calls with multiple positive and negative word lists to gauge sentiment. They report that manager-specific tone is positively associated with future operating performance using the Henry (2008) and Loughran and McDonald (2011) word list, but tone generated with Diction word list is not associated with future performance.

3. A theoretical model for disclosure of in-house meeting and crash risk

Jin and Myers's (2006) theoretical model suggests that increased information asymmetry between corporate insiders and external investors enables managers to withhold bad news from the public for an extended period to maximize compensation, protect employment, and mitigate litigation risk. Based on their framework, we develop a model that directly links a firm's crash probability to the signals investors receive about its private cash flow (i.e., the in-house meeting reports). Our model demonstrates that an increase in the quantity or quality (e.g., greater readability) of in-house meeting reports leads to a decrease in future crash risk measures such as crash probability and return skewness, whereas an increase in the sentiment of in-house meeting reports has the opposite effect. We discuss the economic intuition of this model in this section; its technical details are presented in Appendix 2 for conciseness.

Following the setting of Jin and Myers (2006), we assume that the focal firm is owned by investors (outsiders) but operated by a manager (insider). The cash flow generated by the firm can be split into two parts: public cash flow, which is available to both parties, and private cash flow, which is only available to the manager (insider). Consequently, the manager can intercept private cash flow without being detected by the investors; in the model, the manager will do so as much as possible to maximize their personal wealth.

The investors can initiate a collective action to dismiss the manager by paying a cost proportional to the company's current value. However, they will allow the manager to continue operating the firm as long as they are indifferent between receiving dividend payments and firing the manager. Therefore, the manager's optimal strategy is to intercept all private cash flows and pay the exact amount of the required dividend. Paying a lower dividend will trigger their dismissal by the investors, whereas paying a higher dividend will increase the level of future dividend payments, thus jeopardizing the manager's ability to retain future private cash flow.

However, the required dividend is not necessarily smaller than the private cash flow captured by the manager. Where the required dividend is larger, the manager has to pay the dividend from their own pocket to conceal this negative cash flow. If the manager cannot afford the dividend payment, they will choose to "abandon" the firm by taking all private cash flow and refusing to pay the dividend. This leads to collective action by the investors, leading to the private cash flow, and hence the firm's true value, being revealed.

In Jin and Myers's (2006) model, firing the manager is an option for the investors to exercise. Though economically appealing, this construction makes it difficult to analyze the impact of abandonment on crash risk. Deviating from Jin and Myers (2006), we assume that the manager continues to conceal negative cash flow if the firm's true value is larger than a fraction of its investor-perceived value, and will abandon the firm immediately if this condition is breached. Intuitively, this fraction can be linked to the stringency of the firm's accounting policy, which limits the extent of creative accounting to inflate its perceived value. It could also be related to the manager's personal wealth, which affects their capacity to absorb negative cash flow. Upon abandonment, the investors discover that the firm's true value is less than what they perceived. Consequently, the market value shrinks accordingly, resulting in a stock price crash.

A major innovation of our model is allowing investors to learn from signals (i.e., in-house meeting reports) of the private cash flow,[6] and linking the sentiment and quality of signals to the crash risk. In our model, crash risk is measured as both the one-step ahead prediction of a crash and the conditional skewness of the one-period return. Our model generates the following predictions when the hidden private cash flow is largely negative:

(1) The availability of in-house meeting reports reduces future crash risk, via two mechanisms. On the one hand, investors can use the signals to learn about the otherwise-undetected negative private cash flow, thus reducing the discrepancy between the actual and investors' perceived value of the firm. On the other hand, the amount of concealed negative cash flow is partially reduced by the signals.

(2) Both the quantity and quality (measured by readability) of in-house meeting reports are negatively correlated with future crash risk. An increase in report quality enables more precise estimation of hidden negative information, which enhances the two mechanisms reported in (1), thus contributing to a lower future crash risk.

(3) The (average) negative sentiment of in-house meeting reports is negatively correlated with future crash risk when the sentiment underestimates the magnitude of the hidden private cash flow. This effect is unsurprising because more negative sentiment results in less negative private cash flow undetected by investors, which consequently contributes to a lower future crash risk.

[6] The signals in our model can also be interpreted as disclosure of private information, following Dasgupta et al. (2010). However, they do not explicitly consider the impact of private information on crash risk.

4. Hypotheses development

Since July 2012, SZSE has required all listed firms to electronically publish a standard summary report for each in-house meeting through its web portal. We expect that the mandatory disclosure requirement has two potential effects on future crash risk: first, given that the information released during in-house meetings receives higher public attention after reports become accessible on the web portal, managers should have greater difficulty withholding negative information, thus decreasing the probability of negative news being disclosed to the market en masse. Second, as investors have access to the in-house meeting reports, they can trade on the information released therein, facilitating the incorporation of both positive and negative information into the stock price. The overall consequence is a lower future crash risk. Therefore, we expect to find that since the implementation of mandatory disclosure in July 2012, there has been a systematic reduction in the stock crash risk of SZSE-listed firms. Since the disclosure regulation does not apply to SHSE-listed firms, we expect that reduction in crash risk is more salient in SZSE-listed firms than SHSE-listed firms. We thus formulate H1 as follows:

H1: *SZSE-listed firms have experienced a greater reduction in crash risk since the disclosure of in-house meeting reports became mandatory, compared to SHSE-listed firms.*

Next we analyze the cross-sectional variation in in-house meeting disclosure and its impact on crash risk. First, the number of in-house meetings held by a firm signals the amount of firm-specific information incremental to that published by the firm (in financial reports) and financial intermediaries (in analyst reports). In particular, in-house meeting reports provide details on attendees and questions answered during the meeting, enabling quantification of the scope of institutional investors with access to the information and the amount of value-relevant information released during the meeting. The accessibility of in-house meeting reports enables the general public to more effectively risk-assess the firm. For example, after reading an in-house meeting report discussing a firm's cash holdings in light of its current level of short-term debt, an investor can better judge the firm's default probability. This constrains manages' ability to conceal unfavorable information, leading to reduced future stock crash risk. Based on the discussion, we expect to find support for the following hypothesis:

H2: *Future stock price crash risk is negatively associated with the number of in-house meetings organized by a firm.*

We now turn to the textual characteristics of in-house meeting reports. Previous research shows that the readability of financial statements and analyst reports has an important effect on how efficiently value-relevant information is released to the market (Li 2008; Lee 2012; De Franco et al. 2015). For example, Lee (2012) investigates the association between readability of quarterly reports (10-Q) and post-earnings announcement drift using more than 60,000 quarterly reports published between 2001 and 2007. The results show there is less market response during the 10-Q filing window when 10-Q reports are difficult to read, suggesting that difficult-to-process information impedes investors' ability to react to earnings-related information. We expect the average readability of in-house meeting reports to affect the efficiency of readers' reaction to the information they contain. Specifically, investors who read reports with higher readability should find it reasonably easy to comprehend and, consequently, trade on the information, which suggests that the information is more likely to be capitalized into the stock price. For firms with more firm-specific information being disseminated to the market, the information asymmetry between corporate insiders and external investors decreases, resulting in lower probability of managers being able to withhold unfavorable news and, thus, a lower future crash risk. This leads to H3a:

H3a: *Future stock price crash risk is negatively associated with the average readability of the firm's in-house meeting reports.*

Prior studies report that negative sentiment or a substantial increase in negative sentiment, whether in the business press (Tetlock 2007; Garcia 2013), corporate disclosure (Durnev and Mangen 2011; Rogers at al. 2011), or analyst reports (Rees and Twedt 2012), exerts downward pressure on the stock price. For example, Rees and Twedt (2012) conclude that textual sentiment in analyst reports gives information about firms' future performance beyond that conveyed by quantitative content. We predict that the sentiment of in-house meeting reports has incremental value in predicting the firm's crash risk. In particular, if the average sentiment in these reports is negative, this implies poor expectations of the firm's future performance. Once this information is released to the market and incorporated into the stock price, the probability of bad news being released en masse declines, indicating decreased stock crash risk.[7] H3b is formally stated as follows:

H3b: *Future stock price crash risk is reduced when the firm releases in-house meeting reports with more negative sentiment.*

[7] We focus on negative sentiment because it is less ambiguous to identify the negative sentiment by computing the percentage of negative words in a document. Managers may use negation of positive words to frame a negative statement, so calculating positive sentiment could be problematic (Loughran and McDonald 2014). Both Tetlock (2007) and Loughran and McDonald (2011) find little incremental information in positive word list.

4. Data and research design

4.1 Data

Our data cover all Chinese A-share firms listed on the main-board markets of SHSE and SZSE between 2012 and 2017. We use a computer program to extract all in-house meeting reports published on the SZSE web portal. Our sample includes 35,976 in-house meeting reports for 1,787 SZSE-listed firms (SHSE-listed firms are used as the control group). To alleviate skewness, we construct our main measure of in-house meeting occurrences for each firm-year using the natural log of 1 plus the number of annual in-house meetings. Appendix 1 shows a typical example of in-house meeting reports. Table 1 illustrates sample characteristics of the in-house meetings of SZSE-listed firms. The number of in-house meeting reports increased from 667 in 2012 (July-December) to 1,228 in 2017.[8] We also observe that a typical SZSE-listed firm in our sample hosted 5.7 in-house meetings annually. In a typical in-house meeting, there were 6.2 institutional investors and 2.2 corporate insiders.

We collected firm-level accounting and trading data from China Stock Market & Accounting Research. We further obtained institutional ownership data from the Wind Financial Database (WindDB). We then excluded financial sector firms and any firms missing values for key variables in our regression model (details of the variables are provided later in this section). All continuous variables are winsorized at 1% (both tails) to mitigate the influence of extreme observations.

4.2 Measurement of stock crash risk

[8] This sample is larger than the actual in-house meeting report sample used in our analysis as we only include firms listed on the main-board markets.

To calculate firm-specific crash risk, we first estimate weekly returns for each firm-year. The firm-specific weekly return is defined as the natural log of 1 plus the residual return from the following regression:

$$r_{j,t} = \alpha_j + \beta_{1j} r_{m,t-2} + \beta_{2j} r_{m,t-1} + \beta_{3j} r_{m,t} + \beta_{4j} r_{m,t+1} + \beta_{5j} r_{m,t+2} + \varepsilon_{j,t} \quad (1)$$

where $r_{j,t}$ is the return on stock *j* in week *t*, and $r_{m,t}$ is the return on the value-weighted market index in week *t*. We include the lead and lag terms for the market index return to account for nonsynchronous trading (Dimson 1979). The firm-specific weekly return for firm *j* in week *t*, $W_{j,t}$ is measured by the natural logarithm of 1 plus the residual return in Eq. (1), so $W_{j,t} = \ln(1+\varepsilon_{j,t})$.

We define crash weeks in a fiscal year for a given firm as those weeks during which it experiences a firm-specific weekly return 3.2 standard deviations below its mean specific weekly return over the entire fiscal year. We chose 3.2 to generate a frequency of 0.1% in the normal distribution. Following Hutton et al. (2009) and Chen et al. (2001), our first firm-level measure of crash risk is the negative conditional return skewness (*NCSKEW*) because negative values for skewness indicate that data are skewed to the left. We calculate *NCSKEW* for a given firm in a fiscal year by taking the negative of the third moment of firm-specific weekly returns for each sample year and dividing it by the standard deviation of firm-specific weekly returns raised to the third power. Specifically, for each firm *j* in year *t*, we compute *NCSKEW* as:

$$NCSKEW_{j,t} = -\left[n(n-1)^{3/2}\sum W^3{}_{j,t}\right] / \left[(n-1)(n-2)\left(\sum W^2{}_{j,t}\right)^{3/2}\right] \quad (2)$$

where *n* is the number of trading weeks for stock *j* in year *t*. A higher value of *NCSKEW* implies greater crash risk.

Our second measure of crash risk, *DUVOL*, is the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. For each stock *i* over a fiscal year period, we separate the days with firm-specific daily return above (below) the period's mean and classify them as the "up" ("down") sample. We then calculate the samples' standard deviations and compute the natural log of the ratio of variance in the "down" sample to variance of the "up" sample. *DUVOL* is calculated as follows:

$$DUVOL_{i,t} = \ln\left\{\frac{(n_u - 1)\sum_{down} R_{i,d}^2}{(n_d - 1)\sum_{up} R_{i,d}^2}\right\} \qquad (3)$$

where n_u and n_d are the number of up and down days over the entire fiscal year, respectively. A higher value of *DUVOL* indicates greater crash risk.

4.3 Control variable

The set of control variables includes $DTURN_{t-1}$, $NCSKEW_{t-1}$, $SIGMA_{t-1}$, RET_{t-1}, $SIZE_{t-1}$, MB_{t-1}, LEV_{t-1}, ROA_{t-1}, and $ACCM_{t-1}$, which are taken from previous research (Chen et al. 2001; Hutton et al. 2009). The variable *DTURNt–1* is the detrended average monthly stock turnover in year $t-1$. This is the key variable of interest in Chen et al. (2001). The authors find this detrended turnover variable to be positively related to future crash risk. The variable *NCSKEWt–1* is the negative skewness of firm-specific weekly returns in year $t-1$. Firms with high return skewness in year $t-1$ are likely to also have high return skewness in year *t*. The variable $SIGMA_{t-1}$ is the standard deviation of firm-specific weekly returns over year $t-1$. More volatile stocks are likely to experience stock price crashes in the future (Chen et al. 2001). The variable RET_{t-1} is defined as the arithmetic average of firm-specific weekly returns in year $t-1$. The variable $SIZE_{t-1}$ is defined as the log of the market value of equity in year $t-1$. Both Chen et al. (2001) and Hutton et al. (2009) report a positive relation between size and crash risk. The variable MB_{t-1} is the market value of equity divided by the book value of equity in year $t-1$. Prior research shows that growth stocks are more likely to experience future price crashes. The variable LEV_{t-1} is the total long-term debt divided by total assets. The variable ROA_{t-1} is income before extraordinary items divided by lagged total assets. Hutton et al. (2009) show that financial leverage and operating performance are both negatively related to crash risk. The variable $ACCM_{t-1}$ is the sum of the past three years' absolute values of discretionary accruals, calculated as the residual from the modified Jones model (Dechow et al. 1995). All the variables are defined in Appendix 3.

4.4 Research design

First, we test whether the mandatory disclosure of in-house meetings contributes to reduced stock crash risk in SZSE-listed firms relative to SHSE-listed firms. We employ a DiD approach by extending our sample back to 2010 and comparing change in crash risk during the three years before and after 2012 for SZSE-listed firms with the same change for the control group of firms not subject to the 2012 regulatory change (SHSE-listed firms). In particular, we expect the reduction in crash risk to be more salient for SZSE-listed firms than SHSE-listed firms. The DiD approach allows us to control for unobservable commonalities between the treatment and control groups.

$$\mathrm{Crash}_{i,t+1} = \alpha_0 + \beta_1(\mathrm{SZSE*POST}) + \sum_{q\text{-}2}^{m} \beta_q(\mathrm{q^{th}Control})_{i,t} + \mathrm{Firm}_i + \mathrm{Year}_t + \varepsilon_{i,t} \quad (4)$$

We also repeat the DiD tests using a propensity-score-matched sample for listed firms with in-house meeting records.[9] First, we estimate a logit model each year from 2012 to 2015: the dependent variable is coded 1 if a SZSE-listed firm has at least one in-house meeting record in a given year, and 0 otherwise; the firm-level control variables in Eq. (4) are used as independent variables. Second, for each year we match each SZSE-listed firm with in-house meetings with two matching SHSE-listed firms that have the closest propensity scores within a maximum distance of 5% (Caliper constraint 0.05).

Next we investigate whether the variation in number of in-house meeting reports for a given firm in year *t* is associated with variation in stock price crash risk in year *t+1*. For this analysis, the variable of interest is L_INH, calculated as the natural logarithm of 1 plus the number of in-house meeting reports for firm *j* in year *t*. The other variables are defined as before, and both year and industry fixed effects are included:

$$Crash_{i,t+1} = \alpha_0 + \beta_1 L\_INH + \sum_{q-2}^{m} \beta_q (q^{th}Control)_{i,t} + industry_i + Year_t + \varepsilon_{i,t} \tag{5}$$

Furthermore, we analyze the textual characteristics of in-house meeting reports and their impact on crash risk. Specifically, we test whether the average readability and sentiment of in-house meeting reports for a given firm in year *t* is associated with crash risk in year *t+1*. We use the negative value of the natural logarithm of average sentence length in in-house meeting reports as a proxy of readability, and use Baidu's deep learning natural language processing model for mandarin (accessible at

[9] Since in-house meeting records are not accessible for SHSE-listed firms, we construct a quasi-in-house meeting sample for them based on annual logit models with respect to the number of in-house meetings using our SZSE-listed sample.

https://cloud.baidu.com/doc/NLP/index.html) to compute average sentiment. The model is as follows:

$$Crash_{i,t+1} = \alpha_0 + \beta_1 Negative + \sum_{q-2}^{m} \beta_q (q^{th} Control)_{i,t} + Industry_i + Year_t + \varepsilon_{i,t} \quad (6)$$

5. Empirical results

5.1 Descriptive statistics

Table 2 presents descriptive statistics for the variables used in our analyses. The two metrics of future crash risk, *NCSKEW* and *DUVOL*, have means (medians) of −0.185 (−0.184) and −0.104 (−0.135), respectively. The value of *NCSKEW* varies from −0.871 (25th percentile) to 0.527 (75th percentile), whereas the 25th and 75th percentile values are −0.707 and 0.469, respectively, for *DUVOL*. The in-house meeting variable L_INH has a mean (median) of 0.480 (0.000). This value underestimates the average annual number of in-house meetings because our sample also contains SHSE-listed firms, which do not disclose in-house meetings. The mean (median) of DTURN is 0.358 (0.570). SIGMA has a mean (median) of 0.069 (0.061). The mean (median) of LEV is 0.394 (0.381). All variables have substantial variation.

<Insert Table 2 about here >

5.2 Correlation

Table 3 reports the Pearson correlation matrix of the variables. We observe a high correlation (0.928) between the two crash risk metrics, *NCSKEW* and *DUVOL*, which is consistent with the existing literature (Li et al. 2017). The number of in-house meetings is positively correlated with each of the two crash risk measures, contrary to H2. However, these correlations should be interpreted with caution since we do not control for the other determinants of future crash risks. Furthermore, we find that the number of in-house meetings is positively correlated with stock return, market capitalization, and ROA, and negatively related to LEV and ACCM.

<Insert Table 3 about here>

5.3 Results

Table 4 reports the results from DiD estimation. The variable of interest regarding H1 is the interaction term between *SZSE* and *POST*, which captures the difference in crash risk changes between SZSE- and SHSE-listed firms after the disclosure regulation was adopted in 2012. In both columns (1) (where crash risk is measured by *DUVOL*) and (2) (where crash risk is measured by *NCSKEW*), the coefficients on the interaction term are negative and statistically significant for both crash risk measures (−0.091, t= −2.993; −0.112, t = −2.931). These results lend credence to the hypothesis that SZSE-listed firms have lower crash risk relative to SHSE-listed firms after the adoption of the in-house meeting mandatory disclosure requirement. In columns (3) and (4), we show the repeated analysis using the PSM sample, and find qualitatively unchanged results on the coefficients of the interaction term (−0.092, t= −1.804; −0.136, t = −2.180). This further implies that the disclosure regulation's adoption may have contributed to improving the corporate information environment of SZSE-listed firms. Therefore, H1 is supported.

<Insert Table 4 about here>

Table 5 reports the results for H2. In columns (1) and (2), our full sample results (including both SZSE- and SHSE-listed firms) show that the coefficients of L_INS, our main variable of interest, are −0.025 and −0.034 when future crash risk is measured by *DUVOL* and *NCSKEW*, with respective t-statistics of −2.260 and −2.504. These results support H2's prediction that a firm's future stock price crash risk is negatively associated with its number of in-house meetings in the previous year. A higher number of in-house meetings likely increases the amount of firm-specific information released to the market and absorbed into the stock price, thereby reducing the information asymmetry and future crash risk. Therefore, H2 is supported.

We further test H2 using SZSE-listed firms only to alleviate the confounding effect of including SHSE-listed firms. The results reported in columns (3) and (4) also confirm the negative and significant relation between the number of in-house meetings and future crash risk. Regarding its economic significance, ceteris paribus, a one standard deviation increase in L_INH among SZEX-listed firms (1.00) is associated with 3.47% and 3.66% reductions in the respective standard deviations of *DUVOL* and *NCSKEW*. Therefore, we conclude that the negative effect of the number of in-house meetings on future stock price crash risk is both statistically and economically significant.

For control variables, the coefficients of negative return skewness, arithmetic average return, and ROA are positive and significant. We find no statistically significant relations between de-trended turnover, the standard deviation of firm-specific weekly returns, financial opacity, and future stock crash risk. These findings are largely consistent with prior studies such as Li et al. (2017).

<Insert Table 5 about here>

We next turn to the impact on crash risk of in-house meeting reports' textual characteristics. We begin with the average readability of in-house meeting reports. Our regression sample only includes those firm-year observations with in-house meeting reports, so as to exclude the confounding effect of the remaining observations without readability measures. As reported in columns (1) and (2) of Table 6, the coefficients on average readability are −0.111 and −0.131, both with statistical significance. These results confirm that more readable in-house meeting reports are likelier than less readable reports to disseminate information to investors and thus reduce future crash risk. Therefore, H3a is supported.

<Insert Table 6 about here>

Finally, we test the effect of (average) sentiment of in-house meeting reports on crash risk (H3b), and present the results in Table 7. As reported in columns (1) and (2), our full sample results show that the coefficients on Negative are −0.130 and −0.143 when crash risk is measured by *DUVOL* and *NCSKEW*, with t-statistics of −2.488 and −2.131, respectively. These results show that a firm's future crash risk is negatively associated with the negative sentiment of in-house meetings reports in the previous year, thus supporting H3b.[10] It is plausible that the negative sentiment informs investors of the firm's unfavorable prospects. Given that the information has been (partially) incorporated into the stock price through trading, this decreases the probability of negative news being accumulated and released to the market en masse (which leads to a stock crash). We further

[10] We also employ cumulative abnormal stock returns as a measure of in-house meeting report sentiment, which is based on a standard market model with window of stock returns between 255 and 43 days before the meeting date. Our results and inferences remain qualitatively unchanged.

run the regression with SZSE-listed firms to alleviate the confounding effect of including SHSE-listed firms. As reported in columns (3) and (4), the results confirm the negative and significant relation between the negative sentiment of in-house meetings and future crash risk. Overall, therefore, H3b is supported.

<Insert Table 7 about here>

5.4 Additional analyses

This section investigates the mechanism through which reported in-house meetings could impact stock crash risk. The first channel is informed trading. That is, we expect the negative effect of in-house meeting disclosure on crash risk to be concentrated in stocks with a higher level of informed trading. A stream of extant literature suggests that informed trading activities are an important channel to capitalize firm-specific information into stock prices (Roll 1988; Durnev et al. 2003; Vega 2006). Firm-specific information released by in-house meeting reports will also be absorbed by stock prices through informed trading activities. For instance, negative news about a firm's operations in an in-house meeting report will be impounded in the stock price through more selling activities from arbitrageurs, suggesting that a higher level of informed trading helps incorporate more firm-specific information from these reports into the current stock price, resulting in lower future crash risk. Following the literature (Roll 1988; Amihud 2002; Jin and Myers 2006; Ferreira et al. 2011; Jiang and Yuan 2018), we use stock return synchronicity (SYNCH) to reflect the perceived amount of informed trading,[11] calculated as the association between

[11] We also find qualitatively similar results using an alternative measure of informed trading volume based on the price impact of information asymmetry (IAPI) (Amihud 2002). According to Ferreira et al. (2011), IAPI serves as a positive signal of the perceived amount of informed trading. We calculate IAPI as the

a firm's stock return and market return. Past literature suggests a negative relation between SYNCH and the level of informed trading (Durnev et al. 2003; Vega 2006). Following Jin and Myers (2006), we first estimate the adjusted R^2 of the following equation for each firm-year:

$$RET_{i,t} = \alpha_0 + \beta_1 Ret\_mkt_{i,t-2} + \beta_2 Ret\_mkt_{i,t-1} + \beta_3 Ret\_mkt_{i,t} + \beta_4 Ret\_mkt_{i,t+1} + \beta_5 Ret\_mkt_{i,t+2} + \varepsilon_{i,t} \quad (7)$$

where RET is the daily return of individual firm *i* on day *t*; Ret_mkt is the weekly return calculated as the value-weighted return of A-share market return on day *t*. We then calculate the synchronicity measure following the log transformation to free the adjusted R^2 from its bounded nature within [0,1]:

$$SYNCH = \log(\frac{Adj.R^2}{1 - Adj.R^2}) \quad (8)$$

We expect that releasing in-house meeting reports plays a stronger role in reducing future crash risk for firms with a higher amount of informed trading. We partition our sample into high and low informed trading sub-samples based on the annual median values of *SYNCH*, and compare the effect of the number of reported in-house meetings on crash risk across two sub-samples. The results are reported in Table 8.

In columns (1)-(2) ((3)-(4)) of Table 8, the crash risk measure is *DUVOL* (*NCSKEW*). The coefficients of L_INH are significantly negative in the higher informed trading sub-sample based on *SYNCH* (columns (2) and (4)), but the coefficients are insignificant in the lower informed trading sub-sample. Overall, the results accord with our prediction that the

---

following for each firm-year: $IAPI_{i,t} = \frac{1}{D_i}\sum_{d=1}^{D_i}\frac{|r_{i,t}|}{VOLD_{i,t}}$, where D_i is the number of trading days for stock *i*; $r_{i,t}$ is the daily return of stock *i* of day *t*; $VOLD_{i,t}$ is the trading volume of stock *i* of day *t*.

effect of reported in-house meetings on crash risk is concentrated in firms with higher informed trading.

<Insert Table 8 about here>

The second channel we explore is ownership structure. Since the majority of in-house meetings are initiated by institutional investors such as sell-side analysts or investment fund managers, institutional investors who are present are first to benefit from attending the in-house meeting. However, retail investors can also benefit after the in-house meeting reports become publicly available, and the benefit is positively associated with the percentage of retail investors. In other words, the effect of in-house meetings in decreasing information asymmetry (and crash risk) is expected to be more prominent for firms with a higher percentage of retail investors (lower institutional ownership). To test this conjecture, we collected data on institutional ownership and partitioned the sample based on the median of institutional ownership for each year. We then run regressions for two sub-samples and compare the effect of in-house meeting reports on future stock crash risk between the two sub-samples. The results are reported in Table 9. In columns (1)-(2) ((3)-(4)), the crash risk measure is *DUVOL* (*NCSKEW*). The coefficients on L_INH are significantly negative (insignificant) in the lower institutional ownership sub-sample (higher institutional ownership sub-sample), consistent with our prediction that the effect of in-house meetings on crash risk is concentrated in firms with lower institutional ownership.

<Insert Table 9 about here>

6. Conclusion

Since July 2012, all SZSE-listed firms have been mandated to electronically publish a standard summary report for each private in-house meeting through the SZSE web portal. This paper analyzed the association between disclosure of in-house meetings and stock crash risk. First, we developed a model to explicate the impact of the disclosure regulation for in-house meetings on future crash risk. Our model predicts a negative association between the number of in-house meetings and stock price crash risk. It also predicts that the textual characteristics of in-house meeting reports (average readability and sentiment) have a significant impact on crash risk. Second, based on the analysis of data collected from Chinese listed firms between 2012 and 2017, we find results consistent with the model's predictions.

Our research contributes to the literature in multiple ways. First, we add to the disclosure literature. Prior research on in-house meetings largely focuses on the potential benefits received by a select group of investors (i.e., in-house meetings attendees). Our study explores the broad capital market consequences of the disclosure of in-house meeting reports in terms of impact on reducing the future stock crash risk. Second, we extend the literature on stock crash risk by identifying the release of in-house meeting reports as one determinant of future crash risk. Finally, we contribute to the research on textual analysis in accounting and finance by focusing on the characteristics of in-house meeting reports: this subset is publicly accessible thanks to the adoption of the recent disclosure regulation in SZSE.

Reference:

Bushee, B., Gerakos, J., & Lee, L.F (2018). Corporate Jets and Private Meetings with Investors. *Journal of Accounting and Economics*, 65, 358-379..

Bowen, R., Dutta, S., Tang, S., & Zhu, P. (2018). Inside the "black-box" of private in-house meeting. *Review of Accounting Studies*, 19, 487-527.

Callen, J., Khan, M., & Lu, H. (2013). Accounting quality, stock price delay and future stock return. *Contemporary Accounting Research*, 30, 269-295.

Campbell, J., Chen, H., Dhaliwal, D. S., Lu, H., & Steele, L. (2014). The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies*, 19, 396–455.

Chen, J., Hong, H., & Stein, J. C. (2001). Forecasting crashes: trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics*, 61, 345–381.

Dasgupta, S., J. Gan, and N. Gao (2010): Transparency, price informativeness, and stock return synchronicity: Theory and evidence. *Journal of Financial and Quantitative Analysis*, 45, 1189-1220.

DeFond, M.L., Hung, M., & Li, S. (2015). Does mandatory IFRS adoption affect crash risk? The Accounting Review, 90, 265-299.

De Franco, G., Hope, O.K., Vyas, D., & Zhou, Y. (2015). Analyst report readability. Contemporary Accounting Research, 32, 76-104.

Davis, A. K., Ge, W., Matsumoto, D., & Zhang, J. (2015). The effect of manager-specific optimism on the tone of earnings conference calls." *Review of Accounting Studies*, 20, 639-673.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7, 197-226.

Durnev, A. & Mangen, C. (2011). The real effects of disclosure tone: evidence form restatement. Available at http://ssrn.com/abstract=1650003

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, 68, 1267-1300.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication* 45, 363-407.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics*, 94, 67–86.

Jin, L., & Myers, C.S. (2006). R-square around the world: New theory and new tests. *Journal of Financial Economics*, 79, 257–292.

Kim, J.-B., Li, Y., & Zhang, L. (2011a). Corporate tax avoidance and stock price crash risk: firm-level analysis. *Journal of Financial Economics*, 100, 639–662.

Kim, J.-B., Li, Y., & Zhang, L. (2011b). CFOs versus CEOs: equity incentives and crashes. *Journal of Financial Economics*, 101, 713-770.

Kim, J.-B., Wang, Z., & Zhang, L. (2016a). CEO overconfidence and stock price crash risk. *Contemporary Accounting Research*,33, 1720-1749.

Kim, J.-B., Li, L., Lu, L.Y., & Yu, Y. (2016b). Financial statement comparability and expected stock price crash risk. *Journal of Accounting and Economics*, 61, 294-312.

Kothari, S. P., Li, X; & Short, J.E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis. *The Accounting Review* 84, 1639-1670.

Kothari, S., Shu, S., & Wysocki, P. (2009). Do manager withhold bad news? *Journal of Accounting Research*, 47, 241-267.

Lehavy, R., Li, F., & Merkley, K. (2011) The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, 86, 1087-1115.

Lee, Y.J. (2012). The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research*, 29, 1137-1170.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economic*, 45, 221-247.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66, 35-65.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, 69, 1643-1671.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54, 1187-1230.

Miller B. (2010). The effect of reporting complexity on small and large investor trading. *The Accounting Review*, 85, 2107-2143.

Rees, L., & Twedt, B. (20120. Reading between the lines: an empirical examination of qualitative attributes of financial analysts' reports. Journal of Accounting and Public Policy, 31, 1-21.

Rogers, J.L., Van Buskirk,., A & Zechman, S.L. (2011). Disclosure tone and shareholder litigation. The Accounting Review, 86,

Tetlock P.C. (2007). Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance* 62, 1139-1168.

Tetlock, P.C., Saar-Tsechansky, M., & MacSkassy, S.(2008). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance* 63, 1437-1467.
Solomon, D., & Soltes, E. (2015). What are we meeting for? The consequence of private meetings with investors. Journal of Law and Economic, 58, 325-355.

You, H., & Zhang, X. (2009). Financial reporting complexity and investor under-reaction to 10-K information. *Review of Accounting Studies*, 14, 559-586.

Vega, C., 2006. Stock price reaction to public and private information. *Journal of Financial Economics*, 82(1), pp.103-133.

Durnev, A., Morck, R., Yeung, B. and Zarowin, P., 2003. Does greater firm-specific return variation mean more or less informed stock pricing?. *Journal of Accounting Research*, *41*(5), pp.797-836.

Table 1: Descriptive Data on in-house meetings

Year	No. of in-house meetings	No. of hosting firms	No. of institutional investors	No. of management participants
2012	2711	667	4.04	2.17
2013	6400	963	4.51	2.14
2014	7324	1114	5.33	2.25
2015	6257	1145	6.60	2.33
2016	6920	1238	7.11	2.25
2017	6364	1228	7.92	2.25

Table 2: Descriptive statistics

	Mean	Std. Dev.	Min	P25	Median	P75	Max	N
$NCSKEW_{t+1}$	-0.185	1.121	-5.109	-0.871	-0.184	0.527	4.695	8532
$DUVOL_{t+1}$	-0.104	0.894	-2.193	-0.707	-0.135	0.469	2.089	8532
L_INH_t	0.480	0.861	0.000	0.000	0.000	0.693	4.419	8532
$NCSKEW_t$	-0.294	1.095	-5.109	-0.944	-0.284	0.384	4.694	8532
$DTURN_t$	0.358	16.793	-117.542	-7.700	0.570	8.828	46.901	8532
$SIGMA_t$	0.069	0.029	0.025	0.049	0.061	0.083	0.160	8532
RET_t	0.006	0.010	-0.015	-0.001	0.005	0.011	0.039	8532
$SIZE_t$	15.543	1.053	11.943	14.969	15.544	16.146	18.157	8532
MB_t	4.433	4.970	0.013	1.978	3.095	4.970	34.086	8532
LEV_t	0.394	0.225	0.005	0.218	0.381	0.555	0.996	8532
ROA_t	0.258	0.387	-1.041	0.060	0.216	0.434	1.578	8532
$ACCM_t$	0.168	0.266	0.011	0.051	0.093	0.168	1.796	8532

This table presents the descriptive statistics of variables use in the analysis. The sample contains 8532 firm-year observations from 2012 to 2017. P25 (P75) is the 25^{th} (75th) percentile of the variable's distribution. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. L_INH is defined as the natural log of (number of investor in-house meetings + 1) for a firm in the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years.

Table 3: Correlation table

	$NCSKEW_{t+1}$	$DUVOL_{t+1}$	L_INH_t	$NCSKEW_t$	$DTURN_t$	$SIGMA_t$	RET_t	$SIZE_t$	MB_t	LEV_t	ROA
$DUVOL_{t+1}$	0.928***										
L_INH_t	0.007	0.008									
$NCSKEW_t$	-0.009	-0.008	0.005								
$DTURN_t$	-0.109***	-0.139***	0.011	-0.049***							
$SIGMA_t$	-0.036**	-0.035**	0.088***	-0.134***	0.443***						
RET_t	-0.011	-0.044***	0.071***	-0.293***	0.588***	0.606***					
$SIZE_t$	-0.004	0.006	0.146***	-0.009	-0.018	0.190***	-0.007				
MB_t	0.059***	0.046***	0.019	-0.050***	0.086***	0.291***	0.261***	0.059***			
LEV_t	-0.041***	-0.046***	-0.090***	-0.016	0.018	-0.055***	-0.035**	-0.053***	0.116***		
ROA_t	0.011	0.011	0.115***	0.005	-0.003	-0.138***	-0.028**	0.244***	-0.032**	-0.355***	
$ACCM_t$	0.008	0.003	-0.061***	0.006	-0.002	0.016	-0.007	0.001	0.103***	-0.054***	0.055***

This table presents the Pearson correlation of variables use in the analysis. The sample contains 8532 firm-year observations from 2012 to 2017. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. L_INH is defined as the natural log of (number of investor in-house meetings + 1) for a firm in the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively.

Table 4: Difference-in-Difference test

	(1) $DUVOL_{t+1}$	(2) $NCSKEW_{t+1}$	(3) $DUVOL_{t+1}$	(4) $NCSKEW_{t+1}$
POST*SZSE	-0.091***	-0.112***	-0.092*	-0.136**
	[-2.993]	[-2.931]	[-1.804]	[-2.180]
SZSE	0.053**	0.067**	0.053**	0.068**
	[2.281]	[2.196]	[2.536]	[2.469]
$DTURN_t$	-0.001*	-0.001	-0.002*	-0.002
	[-1.879]	[-1.560]	[-1.874]	[-1.519]
$SIGMA_t$	0.079	-0.821	-0.558	-1.312
	[0.167]	[-1.338]	[-0.842]	[-1.513]
$NCSKEW_t$	0.039***	0.051***	0.029**	0.052***
	[4.489]	[4.548]	[2.363]	[3.250]
RET_t	10.349***	14.970***	8.829***	11.944***
	[7.940]	[8.804]	[4.593]	[4.864]
$SIZE_t$	-0.011	-0.022*	0.029***	0.007
	[-1.215]	[-1.956]	[2.608]	[0.500]
MB_t	0.004***	0.008***	-0.000	0.002
	[2.787]	[3.728]	[-0.076]	[0.582]
LEV_t	-0.048	-0.033	0.072	0.112*
	[-1.300]	[-0.693]	[1.477]	[1.802]
ROA_t	0.034	0.057**	0.058**	0.097***
	[1.635]	[2.165]	[2.234]	[2.840]
$ACCM_t$	-0.024	0.004	0.034	0.085
	[-0.801]	[0.097]	[0.736]	[1.403]
Observations	12,643	12,643	6,319	6,319
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.071	0.063	0.0432	0.0473

This table presents the results of the regressions examining the impact of mandatory disclosure regulation in July 2012 on crash risk. The full sample contains 12643 firm-year observations from 2010 to 2015. Column (3) and (4) present results based on PSM matched sample, which contains 6319 firm-year observations. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. POST is a time dummy which equals to 1 for observations after 2011 and 0 otherwise. SZSE is a dummy which equals to 1 for firms listed on SZSE and 0 for firms listed on SHSE. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Table 5: Number of in-house meetings and future stock crash risk**.**

	Full sample		SZSE	
	(1)	(2)	(3)	(4)
	$DUVOL_{t+1}$	$NCSKEW_{t+1}$	$DUVOL_{t+1}$	$NCSKEW_{t+1}$
L_INH_t	-0.025**	-0.034**	-0.031**	-0.041**
	[-2.260]	[-2.504]	[-2.125]	[-2.325]
$DTURN_t$	-0.001	-0.001	-0.002	-0.002
	[-1.378]	[-1.159]	[-1.297]	[-1.394]
$SIGMA_t$	0.204	-0.620	-0.049	-1.237
	[0.332]	[-0.795]	[-0.057]	[-1.144]
$NCSKEW_t$	0.050***	0.058***	0.028*	0.026
	[4.817]	[4.365]	[1.881]	[1.373]
RET_t	12.163***	17.139***	8.306***	12.775***
	[7.493]	[8.206]	[3.734]	[4.527]
$SIZE_t$	-0.022**	-0.023*	0.036**	0.052**
	[-2.040]	[-1.757]	[1.965]	[2.290]
MB_t	0.006***	0.010***	0.004	0.007*
	[2.857]	[3.748]	[1.098]	[1.800]
LEV_t	-0.093**	-0.110*	-0.030	-0.036
	[-1.971]	[-1.818]	[-0.443]	[-0.408]
ROA_t	0.060**	0.067*	0.088**	0.067
	[2.128]	[1.923]	[2.226]	[1.334]
$ACCM_t$	-0.042	-0.022	-0.021	0.023
	[-1.208]	[-0.513]	[-0.424]	[0.365]
Observations	8,532	8,532	4,354	4,354
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.094	0.078	0.087	0.066

This table presents the results of the regressions examining the impact of in-house meeting intensity on crash risk. The full sample contains 8532 firm-year observations from 2012 to 2017, while the SZSE sample contains 4354 firm-year observations for firms listed on SZSE. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. L_INH is defined as the natural log of (number of investor in-house meetings + 1) for a firm in the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Table 6: In-house meeting readability and future stock crash risk.

	In-house meeting sample	
	(1)	(2)
	$DUVOL_{t+1}$	$NCSKEW_{t+1}$
Readability_t	-0.111**	-0.131*
	[-2.014]	[-1.875]
$DTURN_t$	-0.001	-0.003
	[-0.915]	[-1.303]
$SIGMA_t$	1.126	0.267
	[1.031]	[0.195]
$NCSKEW_t$	0.021	0.012
	[1.085]	[0.475]
RET_t	3.575	6.845*
	[1.231]	[1.852]
$SIZE_t$	0.028	0.047
	[1.026]	[1.404]
MB_t	0.010*	0.012*
	[1.844]	[1.790]
LEV_t	-0.128	-0.137
	[-1.358]	[-1.135]
ROA_t	0.005	0.002
	[0.081]	[0.033]
$ACCM_t$	-0.038	0.026
	[-0.641]	[0.347]
Observations	2,381	2,381
Industry fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Adj. R-squared	0.074	0.045

This table presents the regression results examining the relation between average Readability of in-house meeting reports and crash risk. The sample contains 2381 firm-year observations from 2012 to 2017 for firms listed on SZSE with in-house meeting reports. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. Readability is calculated as negative value of natural logarithm of average length of sentences in in-house meeting reports for a firm over the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Table 7: In-house meeting sentiment and future stock crash risk.

	Full sample		SZSE	
	(1)	(2)	(3)	(4)
	$DUVOL_{t+1}$	$NCSKEW_{t+1}$	$DUVOL_{t+1}$	$NCSKEW_{t+1}$
$Negative_t$	-0.130**	-0.143**	-0.125**	-0.133*
	[-2.488]	[-2.131]	[-2.187]	[-1.801]
$DTURN_t$	-0.001	-0.001	-0.002	-0.002
	[-1.385]	[-1.162]	[-1.326]	[-1.420]
$SIGMA_t$	0.193	-0.646	-0.030	-1.218
	[0.315]	[-0.827]	[-0.035]	[-1.125]
$NCSKEW_t$	0.049***	0.057***	0.027*	0.026
	[4.768]	[4.311]	[1.857]	[1.359]
RET_t	12.008***	16.942***	8.053***	12.488***
	[7.399]	[8.107]	[3.619]	[4.422]
$SIZE_t$	-0.023**	-0.025*	0.030*	0.043*
	[-2.198]	[-1.934]	[1.656]	[1.911]
MB_t	0.006***	0.010***	0.004	0.007*
	[2.850]	[3.748]	[1.133]	[1.856]
LEV_t	-0.093**	-0.110*	-0.032	-0.038
	[-1.968]	[-1.811]	[-0.468]	[-0.438]
ROA_t	0.057**	0.063*	0.082**	0.058
	[2.035]	[1.790]	[2.070]	[1.140]
$ACCM_t$	-0.040	-0.018	-0.015	0.032
	[-1.140]	[-0.428]	[-0.300]	[0.503]
Observations	8,532	8,532	4,354	4,354
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.094	0.078	0.087	0.066

This table presents the regression results examining the relation between average sentiment of in-house meeting reports and crash risk. The full sample contains 8532 firm-year observations from 2012 to 2017, while the SZSE sample contains 4353 firm-year observations for firms listed on SZSE. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. POST is a time dummy which equals to 1 for observations after 2011 and 0 otherwise. Negative is the average probability of negative sentences of all in-house meeting content based on "Baidu deep learning NLP model" for a firm over the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Table 8: in-house meetings, informed trading, and future stock crash risk.

	Informed trading intensity based on SYNCH			
	Low	High	Low	High
	$DUVOL_{t+1}$	$DUVOL_{t+1}$	$NCSKEW_{t+1}$	$NCSKEW_{t+1}$
L_INH_t	-0.019	-0.045**	-0.021	-0.062***
	[-0.936]	[-2.279]	[-0.833]	[-2.589]
$DTURN_t$	-0.002	-0.000	-0.002	-0.002
	[-1.021]	[-0.150]	[-0.693]	[-0.826]
$SIGMA_t$	-0.215	0.801	-1.393	-0.506
	[-0.180]	[0.678]	[-0.910]	[-0.345]
$NCSKEW_t$	-0.001	0.046**	-0.005	0.043*
	[-0.059]	[2.391]	[-0.189]	[1.711]
RET_t	16.483***	1.511	20.539***	4.650
	[3.902]	[0.558]	[3.803]	[1.330]
$SIZE_t$	-0.015	0.085***	-0.012	0.115***
	[-0.488]	[3.589]	[-0.338]	[3.931]
MB_t	0.008	0.001	0.015**	0.002
	[1.415]	[0.176]	[2.023]	[0.414]
LEV_t	-0.141	0.097	-0.149	0.109
	[-1.323]	[1.072]	[-1.068]	[0.974]
ROA_t	0.082	0.093*	0.079	0.070
	[1.154]	[1.954]	[0.847]	[1.183]
$ACCM_t$	-0.083	0.010	-0.057	0.062
	[-1.106]	[0.139]	[-0.597]	[0.716]
Observations	2,201	2,151	2,201	2,151
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.106	0.093	0.083	0.070

This table presents the results of the regressions examining the impact of in-house meeting intensity on crash risk for firms with high/low informed trading intensity based on SYNCH. The sample is restricted to 4352 firm-year observations from 2012 to 2017 for firms listed on SZSE. SYNCH is defined as stock return synchronicity after log transformation, calculated based on Jin and Myers (2006). NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. L_INH is defined as the natural log of (number of investor in-house meetings + 1) for a firm in the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Table 9: in-house meetings, institutional ownership, and future stock crash risk.

	Institutional Ownership			
	High	Low	High	Low
	$DUVOL_{t+1}$	$DUVOL_{t+1}$	$NCSKEW_{t+1}$	$NCSKEW_{t+1}$
L_INH_t	0.012	-0.063***	0.012	-0.085***
	[0.660]	[-2.861]	[0.508]	[-3.120]
$DTURN_t$	0.000	-0.003*	-0.000	-0.004*
	[0.141]	[-1.760]	[-0.163]	[-1.825]
$SIGMA_t$	2.605**	-1.415	2.263	-2.900*
	[1.997]	[-1.221]	[1.393]	[-1.941]
$NCSKEW_t$	-0.002	0.044**	0.001	0.034
	[-0.088]	[2.193]	[0.021]	[1.258]
RET_t	3.963	10.779***	7.624*	15.397***
	[1.224]	[3.555]	[1.864]	[3.892]
$SIZE_t$	0.089***	0.026	0.117***	0.037
	[3.325]	[1.002]	[3.457]	[1.163]
MB_t	0.012**	-0.000	0.014**	0.004
	[2.127]	[-0.096]	[2.142]	[0.886]
LEV_t	-0.089	0.002	-0.051	-0.045
	[-0.921]	[0.015]	[-0.413]	[-0.343]
ROA_t	-0.027	0.163***	-0.052	0.141*
	[-0.466]	[2.939]	[-0.730]	[1.937]
$ACCM_t$	-0.121*	0.056	-0.101	0.125
	[-1.690]	[0.814]	[-1.187]	[1.375]
Observations	2,204	2,150	2,204	2,150
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Adj. R-squared	0.0802	0.113	0.0652	0.0831

This table presents the results of the regressions examining the impact of in-house meeting intensity on crash risk for firms with high/low institutional ownership. The sample is restricted to 4354 firm-year observations from 2012 to 2017 for firms listed on SZSE. NCSKEW is computed by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. DUVOL is calculated as the natural log ratio of volatility in the "down" sample to volatility in the "up" sample. L_INH is defined as the natural log of (number of investor in-house meetings + 1) for a firm in the year. DTURN the average monthly stock turnover minus the average monthly share turnover of the prior year. SICMA is the standard deviation of firm-specific weekly returns over the year. RET is the arithmetic average of firm-specific weekly returns over the year. SIZE is the natural log of (Firm's market capitalisation). MB is measured as market capitalisation scaled by the book value of equity. LEV is calculated using total long term debt scaled by total assets for a firm in the year. ROA is defined as income before extraordinary items scaled by total assets. ACCM is the sum of accrual quality value based on the modified Jones model (Dechow et al., 1995) of the previous three years. *, ** and *** denote significance levels at 10%, 5% and 1% respectively. T-statistics robust to heteroscedasticity and clustered by firm are reported in the square brackets.

Appendix 1

Example of in-house meeting report

Security code: 002594	Corporate Name: BYD
Type of investors relation activities	■In-house Investor meeting □analyst meeting □Media interview □ Performance announcement meeting □Public news meeting □Road show □Site visit □Others (Conference call)
Time	23/01/2015 10:00
Meeting participants	China Life Insurance Assets Management Co, security investment department head Xiao Rao and his colleagues; China Merchants Securities analyst Liushen Wang
Management attended	Chuanfu Wang, Qian Li and Haijin Wang
Content of the meeting	The meeting focused on the BYD section Q 1: what is the current progress of the company's new energy vehicles? Answer: the company will divided the new energy vehicles into two major markets: public transportation and private cars. in the field of public transportation, the company has pure electric vehicles, including pure electric bus K9, pure electric taxi e6 cars areas, restricted to charging infrastructure and range anxiety problems, the company mainly focuses on plug-in hybrid cars. On 20 January this year, the company has just announced the new plug-in hybrid SUVs-Tang, the market reaction is very good. Qin model, which was put on sale at the end of 2013, received a very strong demand, accumulated a large number of orders in hand. In the next one or two months, the company will release new capacity, and then the capacity bottleneck will be solved. Q2: sales and market share of the company's new energy vehicles? Answer: according to the number of auto industry association, 2014, BYD new energy automobile owns market share of around 30%; for Plug-in hybrid cars, BYD's market share is over 80%, the Qin model, since the end of 2013, has quickly become China's new energy automobile sales champion models. If the battery capacity is sufficient, sales of BYD new energy vehicles can be further improved Q3: what is the future plan of the company's new energy vehicles? Answer: in the future, the company will continuously launch more new energy vehicle models, including the newly launched plug-in SUV model tang, the small plug-in SUV model song yuan, and the pure electric version of sedan model Qin

	EV. With the enrichment of models and the launch of new battery capacity, it is expected that the sales of new energy vehicles will have a rapid growth Q4: what are the characteristics and advantages of the company's automobile batteries? Answer: BYD auto uses the lithium iron phosphate batteries, the cycle life and safety ranks top among car battery that has been globally commercialized. In the operation of the Shenzhen 850 e6 vehicles and 780 K9 models, accumulative total mileage has reached more than 200 million kilometers, the security and stability have been fully validated. company of new energy vehicles in Europe, America and South Asia and other regions also validate the stability of the leading technology and quality Q5: does the company have enough funds to expand capacity? Answer: after the completion of the placement in Hong Kong stock market, the company has about 8 billion yuan in cash, plus operating cash flow, which is expected to be enough to support the capital expenditure needed for capacity expansion in the future. In addition, the company has a large number of bank credit lines, which is relatively convenient to replenish funds

Appendix 2: The theoretical model for crash risk

General Setting. We start by introducing the general setting in Jin and Myers (2006). We work with a representative firm with an operating asset. Let K_t denote the intrinsic value of the firm, as the present value of future operating cash flows, let C_t denote future cash flow and let r denote the risk-free rate assumed to be constant. Conditioning on the information set F_t, the firm's value can be written as:

$$K_t = \sum_{k=1}^{\infty} \frac{\mathrm{E}[C_{t+k} \mid \mathrm{F}_t]}{(1+r)^k}.$$

The firm is operated by a manager, understood as the insider of the company, on behalf of the investors of the firm who own all the shares of the company, which are also referred to as outsiders of the company. The manager can intercept cash flow, which cannot be verified by the investors. The manager will take as much as possible as long as they can satisfy the investors' demand of dividend. The manager prefers to stay in the firm indefinitely to capture future cash flows from the firm. However, he will have to abandon the firm under certain conditions. We will describe this abandonment option in detail in the next section. If the inside managers do not pay enough dividend, the investors will initiate a collective action to fire the manager and take over the firm. This, however, requires a fixed cost proportional to the value of the firm, and the outsiders can only get a net value of αK_t after the takeover, where $0 < \alpha < 1$ represents the effectiveness of protection of investors' property rights.

The firm's cash flow generating process is:

$$C_t = K_0 X_t,$$

where K_0 is the initial investment, and X_t represents a random shock to the cash flow process. Following Jin and Myers (2006), we assume that X_t can be decomposed into three independent shocks:

$$X_t = f_t + \theta_{1,t} + \theta_{2,t},$$

where f_t is the market or macroeconomic shock which is publicly available, $\theta_{1,t}$ and $\theta_{2,t}$ capture firm-specific cash flow shocks. Insiders observe both firm-specific shocks, but outsiders only observe $\theta_{1,t}$, thus $\theta_{2,t}$ is understood as the private information about the firm. Following the notion that $\theta_{2,t}$ represents the investors' unbiased forecast error about the private information, the unconditional mean of $\theta_{2,t}$ is set to zero. To ease notation, we will use the filtration $\mathcal{F}_t$ to denote all the cash flow information available to the insider, and $\mathcal{G}_t \subset \mathcal{F}_t$ to denote the cash flow information available to the outsiders.

Operation Cycle. To study the effect of abandonment, we need to discuss the consequence if a manager decides to abandon the company at time t. Jin and Myers (2006) suggest three different possibilities, but do not provide a detailed picture of the implication and aftermath of an abandonment to the firm's value and the return process. In our model we simplify the assumption of the abandonment event, which allows us to study the relationship between abandonment and crash risk explicitly. To begin with, we depict the actions of the manager and the investors for each period t, which are understood as the normal operation cycle of the firm.

Definition 1: The **Operation Cycle** of the firm at time t is the following list of actions by the manager and the investors at each t.

1. The manager and investors observe f_t, $\theta_{1,t}$ while the manager observes $\theta_{2,t}$ in addition. The investors only have a prior knowledge about $\theta_{2,t}$ which equals its unconditional distribution.

2. The manager and the investors calculate the required dividend Y_t based on the new information.

3. The manager calculates the ex-dividend current value of the firm K_t and the ex-dividend intrinsic value of the firm to the investors $\mathrm{E}[K_t \mid \mathcal{G}_t]$.

   (a) If $K_t > p\mathrm{E}[K_t \mid \mathcal{G}_t]$, the manager will not abandon the firm and will pay the required dividend Y_t.

   (b) If $K_t \leq p\mathrm{E}[K_t \mid \mathcal{G}_t]$, the manager will abandon the firm by taking all the current cash flow with him and refusing to pay the dividend.

4. The investors decide based on the received dividend as follows:

   (a) If Y_t meets the requirement of the investors calculated in 1(a), the investors are happy to let the manager run the firm for another period.

   (b) If $Y_t = 0$, the investors pay for a collective action to take over the firm. Upon take-over, the investors observe the private cash flow $\theta_{2,t}$ and realize the current value of the firm which is αK_t. The investors hire another manager and the operation cycle repeats from step 1.

Our settings in Definition 1 deviate from Jin and Myers (2006) in terms of the abandonment condition. In Jin and Myers (2006), the abandonment condition for the manager is specified as an option, which is economically sensible but renders it difficult for us to consider the probability of abandonment explicitly. Thus, we consider a simplified setting as described above. We assume that the manager decides to abandon the firm whenever he realizes that if he continues to run the firm and pay the required dividend, the current ex-dividend value of the firm is only p fraction of the intrinsic value to the investors, where $0 < p < 1$ represents the tolerance of the company for concealing bad news, or the inverse of the cost for the managers to hide bad performance that is known to both insiders and outsiders. Intuitively, a smaller p means that it is cheaper or easier for the manager to conceal bad news without being discovered. The quantity p can be linked to the stringency of the company's accounting policy, and the wealth of the manager himself to absorb bad performances.[12]

With the operating cycle in mind, we introduce the binary random variable $Q_t \in \{0,1\}$ which is one if the manager defaults at time t and zero otherwise. The cash flow generating processes are assumed to be regime-switching AR(1) processes as follows:

$$
\begin{aligned}
f_{t+1} &= m + \varphi f_t + \varepsilon_{t+1}, \\
\theta_{1,t+1} &= \mu + \varphi(\theta_{1,t} + Q_t \theta_{2,t}) + \xi_{1,t+1}, \\
\theta_{2,t+1} &= (1 - Q_t)\varphi\theta_{2,t} + \xi_{2,t+1}.
\end{aligned}
$$

In the above model, ε_t , $\xi_{1,t}$ and $\xi_{2,t}$ are assumed to be i.i.d. Gaussian white noise processes with variances σ_f^2, σ_1^2 and σ_2^2 respectively. The intercept m and μ controls for the fixed cash inflow for the company per period, and $0 < \phi < 1$ controls for the persistence of the cash flows. To ensure that the investors always receive a positive

12 In our model we assume that p is fixed for all managers. However, we can allow for heterogeneous p to account for heterogeneity in managers' personal wealth. The results of our model remain qualitatively unchanged.

dividend, we will restrict $f_t + \theta_{1,t}$ to be positive. We can write $X_t = X_0 + \phi X_{t-1} + \lambda_t$ where $X_0 = m + \mu$ and $\lambda_t = \varepsilon_t + \xi_{1,t} + \xi_{2,t}$.

This design has two important properties. Firstly, this structure has a clear economic interpretation. If the manager chooses not to default with $Q_t = 0$, then f_t, $\theta_{1,t}$ and $\theta_{2,t}$ are just normal AR(1) processes which reduces to the case in Jin and Myers (2006). However, when the manager defaults at time t ($Q_t = 1$), the private cash flow shock $\theta_{2,t}$ is subsumed into $\theta_{1,t}$, which is due to the fact that $\theta_{2,t}$ is observed upon abandonment. The private cash flow shock is thus reset to zero.

Secondly, the total cash flow shock process $\{X_t\}_{t\geq 1}$ is independent of $\{Q_t\}_{t\geq 1}$, since Q_t cancels out in the summation. This property ensures that the actual value of the firm K_t is independent of abandonment. This is also intuitive as abandonment is caused by information asymmetry between insiders and outsiders, rather than a change in the productivity of the firm. Assuming that the investors can always find a good substitute of the current manager that has comparable capability of running the firm, abandonment should not have an impact on the actual value of the firm. This also prevents the value of the firm from diverging to zero or infinity.

We introduce an important concept that we will use extensively in constructing and solving our model:

Definition 2: Let t_0 be the previous period in which a manager abandons the firm. The **steady state** is the period t where $t - t_0 \to \infty$.

As the cash flow processes are all AR(1) processes, convergence to steady state is very fast for practically relevant parameter values of φ as argued in Jin and Myers (2006). Therefore we will also follow this approach and solve our model in the steady state, which greatly simplifies our model. We make the following assumption to ensure that a steady state can be obtained in our model.

Assumption 1: (Rare Abandonment) In the steady state, the unconditional one-step ahead probability of abandonment satisfies $\mathrm{E}[\Pr[Q_t = 1 \mid \mathrm{F}_{t-1}]] \approx 0$.

This assumption is reasonable because if the unconditional probability of one-step ahead abandonment is large, the model may never reach the steady state. As we will show later, an abandonment corresponds to a crash in stock price, and since crash in stock price is defined as a very rare event in our paper, this result is also justified empirically. We also make the following assumption about the magnitude of private information:

Assumption 2: (Gradual Abandonment) We assume that $\Pr[Q_t = 1 \mid \mathrm{F}_{t-1}, \theta_{2,t-1} = 0] \approx 0$.

This assumption suggests that it is extremely unlikely for the manager to abandon the firm based on a bad private cash flow for just one period. This is also realistic since from the intuition in Jin and Myers (2006), the abandonment only occurs when the accumulated hidden private cash flow is large and negative, and managers are usually able to absorb bad performance for one period. Thus, the abandonment is extremely unlikely to occur if the hidden private cash flow is neutral just one period before. Also, since upon abandonment the private information shock is reset to zero, this assumption also implies that it is unlikely to observe sequential abandonments, which is also sensible.

Finally, we note that the abandonment here should also not be understood literally as the manager's departure from the firm. The key ingredient in this model is that hidden negative cash flow is accumulated to the extent that the manager cannot withhold it. Therefore, the abandonment event should be interpreted as the sudden release of accumulated negative cash flow rather than whether the manager actually abandons the firm.

In-house Meeting Reports. A major innovation of our model is that we account explicitly for the disclosure of private information in the model and study its impact on crash risk. Following SZSE's mandatory disclosure policy in July 2012, we interpret this policy as an external shock to the operation cycle at time t^*. For $t \le t^*$, the investors do not receive any information about the private cash flow, and cannot verify its existence. Therefore, the investors' prediction about the private cash flow at time $t \le t^*$ in the steady state is specified as the unconditional distribution of $\theta_{2,t}$, which is known to the investors:

$$\theta_{2,t} \mid G_t \sim \mathrm{N}\left(0, \frac{\sigma_2^2}{1-\varphi^2}\right). \tag{A.1}$$

For $t \ge t^*$, the manager is required to provide reports on the in-house meetings to all investors. These reports contain information about the private cash flow of the firm $\theta_{2,t}$, and the investors will consequently trade based on this extra information and adjust their estimated value of the firm accordingly. We therefore append the operation cycle in Definition 1 as follows:

Definition 3: (Augmented Operation Cycle) For every $t \ge t^*$, the following step is appended to the operation cycle of the firm after step 4:

5. The investors receive a set of signals $\{s_{n,t}\}_{n=1:N_t}$, where N_t is the number of in-house meeting reports and $s_{n,t} \sim i.i.d.\mathrm{N}(\theta_{2,t}, 1/\tau_t)$.

The variable τ_t is understood as the precision of each individual report at time t, which can be interpreted as the readability of the in-house meeting reports. Each $s_{n,t}$ can be interpreted as the sentiment of individual reports, and the sample mean $\overline{s}_t = \sum_{n=1}^{N_t} s_{n,t} / N_t$ is a measure of average sentiment of the reports. We will call the quantity $S_t = N_t \tau_t$ the **quality** of the in-house meeting reports, which we assume to be known to both the investors and the manager. Clearly, a larger N_t and τ_t reduces the variance of the sample mean $\overline{s}_t$, which yields a more precise signal about the private cash flow.

We would like to stress that the signals $\overline{s}_t$ does not affect the manager's decision of abandonment at time t, as the investors only receive these signals after the abandonment decision. This is justified by that the manager will decide the abandonment choice as soon as the private cash flow shock realizes after step 1. However, the investors will have to arrange the in-house meetings and wait for the manager to publish the reports to learn about the most recent private cash flow shock. As a result, $\overline{s}_t$ will be received with delay and used in the calculate for the required dividend at time $t+1$.

Finally, we give a summary of the information sets that will be used extensively in the rest of the model to avoid confusion.

Definition 4: The information sets $\mathcal{G}_t^- \subseteq \mathcal{G}_t \subseteq \mathcal{G}_t^+ \subseteq \mathcal{F}_t$ are defined as follows:

1. $\mathcal{G}_t^-$ contains all information about publicly available cash flow information (f_t and $\theta_{1,t}$).

2. $\mathcal{G}_t$ contains $\mathcal{G}_t^-$ and all available signals up to time $t-1$.

3. $\mathcal{G}_t^+$ contains $\mathcal{G}_t$ and the signals received at time t.

4. $\mathcal{F}_t$ contains $\mathcal{G}_t^+$ and the private cash flow $\theta_{2,t}$.

Prior to time t^*, $\mathcal{G}_t^-$, $\mathcal{G}_t$ and $\mathcal{G}_t^+$ are identical, and the difference only matters when $t > t^*$, as we need to clearly specify what information is used in the decision of abandonment for the manger. $\mathcal{F}_t$ should be understood as the information set available to the insider, which automatically updates when the manager is required to publish signals.

Abandonment and Crash. We firstly show that abandonment leads to unexpected crash in stock price when $t \le t^*$, that is, no signals are available to the investors to learn about $\theta_{2,t}$. From Definition 1, the manager will abandon the firm when $K_t \le p\mathrm{E}[K_t \mid \mathcal{G}_t]$. Since K_t is independent of Q_t, the current value of the firm at any t is:

$$K_t = \sum_{k=1}^{\infty} \frac{\mathrm{E}[C_{t+k} \mid \mathcal{F}_t]}{(1+r)^k} = \frac{(1+r)K_0 X_0}{r(1+r-\varphi)} + \frac{\varphi K_0}{1+r-\varphi}(f_t + \theta_{1,t} + \theta_{2,t}).$$

Note that the above result can be found in Equation (A.3) of Jin and Myers (2006). Without any signals, the value of the firm to the investors is:

$$\mathrm{E}[K_t \mid \mathcal{G}_t] = \frac{(1+r)K_0 X_0}{r(1+r-\varphi)} + \frac{\varphi K_0}{1+r-\varphi}(f_t + \theta_{1,t}).$$

This allows us to define the abandonment condition formally after some simple algebra:

Definition 5: The Abandonment Condition for $t \le t^*$ is defined as the following event:

$$\theta_{2,t} + (1-p)(f_t + \theta_{1,t}) \le B(p),$$

where $B(p) \equiv (p-1)\dfrac{(1+r)X_0}{r\varphi} < 0$ is a constant. The manager will abandon the firm at time $t \le t^*$ if the abandonment condition above is met.

Since $B(p) < 0$, we see that the manager will only abandon the firm if the hidden cash flow shock is sufficiently negative. The amount of concealable negative cash flow increases in magnitude as p decreases, since a small p means that it is easier or cheaper for the manager to hide bad news, which leads to a larger negative hidden cash flow without abandonment of the manager. Also we see that the manager is able to conceal a larger negative cash flow as f_t and $\theta_{1,t}$ increase, as the amount of concealment is proportional to the value of the firm.

The process $D_t \equiv \theta_{2,t} + (1-p)(f_t + \theta_{1,t})$ plays an important row in predicting the event of abandonment. From the Gaussian assumption of ε_t, $\xi_{1,t}$ and $\xi_{2,t}$, it is not hard to see that D_t has the following conditional distribution in the steady state:

$$\begin{aligned} D_t \mid \mathrm{F}_{t-1} &\sim \mathrm{N}\,((1-p)X_0 + \varphi D_{t-1}, (1-p)^2(\sigma_f^2 + \sigma_1^2) + \sigma_2^2), \\ D_t \mid \mathrm{G}_{t-1} &\sim \mathrm{N}\,((1-p)X_0 + \varphi(1-p)(f_{t-1} + \theta_{1,t-1}), \varphi^2\sigma_2^2/(1-\varphi^2) + (1-p)^2(\sigma_f^2 + \sigma_1^2) + \sigma_2^2). \end{aligned} \tag{A.2}$$

Define the manager and the investors' one-step ahead prediction of abandonment probability as $\pi_t \equiv \Pr(Q_t = 1 \mid \mathrm{F}_{t-1})$ and $\varpi_t = \mathrm{E}[Q_t = 1 \mid \mathrm{G}_{t-1}]$, from (A.2) it follows directly that[13]:

$$\begin{aligned} \pi_t &= \Phi\Big(\frac{X_0(p-1)(1+r+r\varphi)/(r\varphi) - \varphi D_{t-1}}{\sqrt{(1-p)^2(\sigma_f^2+\sigma_1^2)+\sigma_2^2}}\Big), \\ \varpi_t &= \Phi\Big(\frac{X_0(p-1)(1+r+r\varphi)/(r\varphi) - \varphi(1-p)(f_{t-1}+\theta_{1,t-1})}{\sqrt{\varphi^2\sigma_2^2/(1-\varphi^2)+(1-p)^2(\sigma_f^2+\sigma_1^2)+\sigma_2^2}}\Big). \end{aligned} \tag{A.3}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. The manager's one-step ahead abandonment probability π_t depends inevitably on the path of $\theta_{2,t}$, and is only known by

[13] Note that technically, investors also learn $\theta_{2,t}$ from that the manager does not abandon the firm at time $t-1$, or $Q_{t-1} = 0$. However, since we define crash as a rare event, knowing $Q_{t-1} = 0$ does not add any substantial information about $\theta_{2,t}$. We will therefore ignore this condition, which greatly simplify our analysis.

the manager. Clearly, this probability will be large only when D_{t-1} is largely negative, that is, when $\theta_{2,t-1}$ more than completely offsets $(1-p)(f_{t-1}+\theta_{1,t-1})$. Since we assume $f_t+\theta_{1,t}$ to be positive and $\theta_{2,t}$ to be mean zero, for most values of $\theta_{2,t}$ the probability will be close to zero until large negative shocks accumulate in $\theta_{2,t}$.

The investors' one-step ahead abandonment probability ϖ_t has the following property:

Proposition 1: Under Assumption 2 with $t \le t^*$, it holds that $\varpi_t \approx 0$.

Proof. Using (A.3) it is straightforward to show that:

$$\begin{aligned}\varpi_t &< \Phi\Big(\frac{X_0(p-1)(1+r+r\varphi)/(r\varphi)-\varphi(1-p)(f_{t-1}+\theta_{1,t-1})}{\sqrt{(1-p)^2(\sigma_f^2+\sigma_1^2)+\sigma_2^2}}\Big) \\ &= \Pr[Q_t=1 \mid \mathrm{F}_{t-1}, \theta_{2,t-1}=0] \approx 0,\end{aligned}$$

where the approximation is due to Assumption 2. This completes the proof. ■

We see that when signals are absent, the investors predict an almost zero abandonment probability for the next period. Therefore, from the perspective of the manager, the hidden negative cash flow is 'undetected' by the investors. In fact, as long as ϖ_t is significantly different from zero, the investors will require a premium on the return that compensates for the potential loss on abandonment. The manager is therefore incentivised to conceal private cash flow so that the investors can only have a neutral, or innocent prior, which necessarily leads to the result in Proposition 1.

We are now in the position to solve the equilibrium dividend policy and examine the properties of the realized return of investor in the absence of signals. Let V_t^{ex} denote the ex-dividend market value of the firm. Since by Proposition 1 the investors are almost always expected to receive a dividend, the equilibrium dividend policy is identical to that in Jin and Myers (2006) as the solution to the following set of equations:

$$\begin{aligned} &\mathrm{E}[V_t^{ex} \mid \mathcal{G}_t] = \frac{\mathrm{E}[Y_{t+1} \mid \mathcal{G}_t] + \mathrm{E}[V_{t+1}^{ex} \mid \mathcal{G}_t]}{1+r}, \\ &\mathrm{E}[V_t^{ex} \mid \mathcal{G}_t] = \mathrm{E}[\alpha K_t \mid \mathcal{G}_t]. \end{aligned}$$

Thus in this equilibrium ignoring bubbles and empty promises, the optimal dividend is $Y_t^* = \alpha \mathrm{E}[C_t \mid \mathcal{G}_t]$, which is a fixed proportion of the cash flow expected by the investors. As discussed before, the manager will pay this dividend unless the abandonment condition is triggered. Conditioning on that the manager does not default at $t+1$, the realized one-period return is:

$$\begin{aligned} r_{t+1}^{nd} &= \frac{\mathrm{E}[V_{t+1}^{ex} \mid \mathcal{G}_{t+1}] + Y_{t+1}^*}{\mathrm{E}[V_t^{ex} \mid \mathcal{G}_t]} - 1 \\ &= r + \frac{(1+r)(\varepsilon_{t+1} + \xi_{1,t+1})}{X_0(1+r)/r + \varphi(f_t + \theta_{1,t})}, \end{aligned}$$

where the *nd* denotes the condition that $Q_{t+1} = 0$, or 'non-default'. Under the Gaussian assumption of ε_t and ξ_t, we see that distribution of r_{t+1}^{nd} conditioning on $\mathcal{F}_t$ is normally distributed with mean r.

However, if the manager abandons the firm at time $t+1$, the investors will not receive any dividend. Instead, the investors will have to pay for the collective action to fire the investor and the realized value of the company becomes αK_{t+1}. We therefore see that:

$$r_{t+1}^{d} = \frac{\alpha K_{t+1}}{\mathrm{E}[V_t^{ex} \mid \mathcal{G}_t]} - 1,$$

where the superscript d represents the condition $Q_{t+1} = 1$, or 'default'. Perhaps the first thing to notice is that we strictly have $r_{t+1}^{d} < r_{t+1}^{nd}$ since $p < 1$ and that no dividend is received when the manager defaults. Assuming $K_{t+1} = p\mathrm{E}[K_{t+1} \mid \mathrm{G}_{t+1}]$ so that the manager abandons the firm without overshooting the abandonment boundary $B(p)$, we can simplify the above to:

$$r_{t+1}^{d} = \varphi p - 1 + \frac{pX_0(1+r-\varphi)/r + \varphi p(\varepsilon_{t+1} + \xi_{1,t+1})}{X_0(1+r)/r + \varphi(f_t + \theta_{1,t})}.$$

Therefore, we see that r_{t+1}^{d} is also conditionally normal with mean:

$$\mathrm{E}[r_{t+1}^{d} \mid \mathrm{F}_t] = \varphi p - 1 + \frac{pX_0(1+r-\varphi)/r}{X_0(1+r)/r + \varphi(f_t + \theta_{1,t})} \le p - 1 + \frac{\varphi pr}{1+r} < r,$$

where the inequality can be easily proven by noting that the expectation is bounded above with $f_t = \theta_{1,t} = 0$. Also, if we take $f_t + \theta_{1+t} \approx X_0/(1-\varphi)$ which is the unconditional value, then we can show that $\mathrm{E}[r_{t+1}^{d} \mid \mathrm{F}_t] \approx p - 1 < 0$. This is clear evidence supporting that the abandonment of a manager leads necessarily to a crash in stock price.

We proceed to show that this crash can be translated into the skewness of the return distribution. To this end, we firstly introduce some notations: $\mu_t \equiv \mathrm{E}[r_t \mid \mathrm{F}_{t-1}]$, $h_t \equiv \mathrm{V}[r_t \mid \mathrm{F}_{t-1}]$, and $\Delta_t = \mu_t^{nd} - \mu_t^{d}$. As the actual return r_{t+1} is a mixture of two normally distributed returns, we have:

$$\begin{aligned}
\mu_{t+1} &= \pi_{t+1}\mu_{t+1}^{d} + (1-\pi_{t+1})\mu_{t+1}^{nd}, \\
h_{t+1} &= \pi_{t+1}h_{t+1}^{d} + (1-\pi_{t+1})h_{t+1}^{nd} + \pi_{t+1}(1-\pi_{t+1})\Delta_{t+1}^{2}.
\end{aligned}$$

We finally arrive at the following result:

Proposition 2: The conditional skewness of realized return $skew(r_{t+1} \mid \mathrm{F}_t) < 0$ if $0 < \pi_{t+1} \le \frac{1}{2} - \frac{3(h_{t+1}^{nd} - h_{t+1}^{d})}{2\Delta_{t+1}^{2}}$.

Proof. The condition that $skew(r_{t+1} \mid \mathrm{F}_t) < 0$ is equivalent to the condition that $\mathrm{E}[(r_{t+1} - \mu_{t+1})^3 \mid \mathrm{F}_t] < 0$. We now study this expectation:

$$\mathrm{E}[(r_{t+1} - \mu_{t+1})^3 \mid \mathrm{F}_t] = \pi_{t+1}\mathrm{E}[(r_{t+1}^d - \mu_{t+1})^3 \mid \mathrm{F}_t] + (1-\pi_{t+1})\mathrm{E}[(r_{t+1}^{nd} - \mu_{t+1})^3 \mid \mathrm{F}_t]. \quad \text{(A.4)}$$

The following quantities will be useful:

$$\begin{aligned} \mu_{t+1}^d - \mu_{t+1} &= -(1-\pi_{t+1})\Delta_{t+1}, \\ \mu_{t+1}^{nd} - \mu_{t+1} &= \pi_{t+1}\Delta_{t+1}. \end{aligned}$$

We now rewrite: $\mathrm{E}[((r_{t+1}^d - \mu_{t+1}^d) + (\mu_{t+1}^d - \mu_{t+1}))^3 \mid \mathrm{F}_t]$ and solve for this expectation explicitly:

$$\begin{aligned} &\mathrm{E}[((r_{t+1}^d - \mu_{t+1}^d) + (\mu_{t+1}^d - \mu_{t+1}))^3 \mid \mathrm{F}_t] = \mathrm{E}[((r_{t+1}^d - \mu_{t+1}^d) - (1-\pi_{t+1})\Delta_{t+1})^3 \mid \mathrm{F}_t] \\ &= -3h_{t+1}^d(1-\pi_{t+1})\Delta_{t+1} - (1-\pi_{t+1})^3\Delta_{t+1}^3. \end{aligned} \quad \text{(A.5)}$$

Note that we have used the fact that the conditional skewness for both r_{t+1}^d and r_{t+1}^{nd} is zero. Similarly:

$$\mathrm{E}[((r_{t+1}^{nd} - \mu_{t+1}^{nd}) + (\mu_{t+1}^{nd} - \mu_{t+1}))^3 \mid \mathrm{F}_t] = 3h_{t+1}^{nd}\pi_{t+1}\Delta_{t+1} + \pi_{t+1}^3\Delta_{t+1}^3. \quad \text{(A.6)}$$

Substituting equation (A.5) and (A.6) into (A.4) and rearranging yield:

$$\mathrm{E}[(r_{t+1} - \mu_{t+1})^3 \mid \mathrm{F}_t] = \Delta_{t+1}\pi_{t+1}(1-\pi_{t+1})(3(h_{t+1}^{nd} - h_{t+1}^d) + \Delta_{t+1}^2(2\pi_{t+1} - 1)).$$

Since $\Delta_{t+1} > 0$, we see that for the above moment to be negative, it must hold that:

$$0 < \pi_{t+1} \le \frac{1}{2} - \frac{3(h_{t+1}^{nd} - h_{t+1}^d)}{2\Delta_{t+1}^2}. \quad \text{(A.7)}$$

This completes the proof. ■

There will be a bound on π_{t+1} for the skewness to be negative because the conditional skewness of both r_{t+1}^{nd} and r_{t+1}^d is zero due to normality. Intuitively, when π_{t+1} is small, the majority mass of r_{t+1} will be centered around r, but this distribution has an excess left tail due to a possible crash. As in our steady state the crash probability is close to zero, unconditionally there will be only a small number of jumps in a sufficiently long sample, which indicates that the unconditional return distribution over a certain period of time is also likely to have an excess left tail with negative skewness.

To conclude, we have shown that in the steady state, if investors do not receive any signals, unexpected abandonment of the manager leads to a crash in stock price, which is also translated into a negative conditional skewness of realized return.

Disclosure of In-house Meeting Reports. We now discuss the scenario in which the in-house meeting reports $\{s_{n,t}\}_{t=1:N_t}$ become available to the investors at time $t > t^*$ due to the mandatory disclosure policy. Prior to the abandonment decision at time t, signals $\bar{s}_{t-1}$ from the previous period are available to the investors, which contains useful information about $\theta_{2,t-1}$ that can be used to learn about $\theta_{2,t}$. In the steady state, the investors' posterior belief about $\theta_{2,t}$ conditioning on $\mathcal{G}_t$ and $\mathcal{G}^+_{t-1}$ is:

$$\theta_{2,t} \mid \mathcal{G}_t \overset{d}{=} \theta_{2,t} \mid \mathcal{G}^+_{t-1} \sim \mathrm{N}\,(\varphi\hat{\theta}_{2,t-1}, \varphi^2\hat{\sigma}^2_{2,t-1} + \sigma_2^2), \tag{A.8}$$

where $\overset{d}{=}$ denotes equal in distribution, $\hat{\sigma}^2_{2,t} \equiv \left(S_t + \frac{1-\varphi^2}{\sigma_2^2}\right)^{-1}$ and $\hat{\theta}_{2,t} \equiv \hat{\sigma}^2_{2,t} S_t \bar{s}_t$ are the posterior mean and variance of $\theta_{2,t}$ given $\bar{s}_t$, respectively. This result is intuitive as $\mathcal{G}_t$ does not contain any additional information about $\theta_{2,t}$ than $\mathcal{G}^+_{t-1}$, and the investors can only rely on $\bar{s}_{t-1}$ to learn about $\theta_{2,t}$.

In comparison to (A.1), we can see that the availability of $\bar{s}_{t-1}$ updates the investors' belief about the value of the firm due to a more precise estimate of $\theta_{2,t}$. Specifically,

$$\mathrm{E}[K_t \mid \mathcal{G}_t] = \frac{(1+r)K_0 X_0}{r(1+r-\varphi)} + \frac{\varphi K_0}{1+r-\varphi}(f_t + \theta_{1,t} + \varphi\hat{\theta}_{2,t-1}).$$

As the manager is aware of the signals, he also needs to take the investors' new estimate of firm value $\mathrm{E}[K_t \mid \mathcal{G}_t]$ into consideration when making decisions about his abandonment, as it now partially reveals the hidden private information. Consequently, the abandonment condition is updated:

Definition 6. The abandonment condition for $t > t^*$ is defined as the following event:

$$(\theta_{2,t} - \varphi\hat{\theta}_{2,t-1}) + (1-p)(f_t + \theta_{1,t}) \leq B(p), \tag{A.9}$$

where $B(p) \equiv (p-1)\dfrac{(1+r)X_0}{r\varphi} < 0$ is a constant. The manager will abandon the firm at time t if the abandonment condition above is met.

We therefore see that the abandonment condition now depends on the discrepancy between the actual private cash flow $\theta_{2,t}$ and the investors' prediction based on the information available at $\mathcal{G}_t$. This change of abandonment condition leads to a change in the one-step ahead estimate of crash probability, which can be derived directly from (A.2), (A.8) and (A.9):

Proposition 3: The manager and investors' one-step ahead estimate of crash probability for $t > t^*$, denoted by $\pi^+ \equiv \Pr(Q_t = 1 \mid \mathcal{F}_{t-1})$ and $\varpi^+ \equiv \Pr(Q_t = 1 \mid \mathcal{G}^+_{t-1})$, can be expressed as follows:

$$\pi_t^+ = \Phi\Big(\frac{X_0(p-1)(1+r+r\varphi)/(r\varphi) + \varphi\hat{\theta}_{2,t-1} - \varphi D_{t-1}}{\sqrt{(1-p)^2(\sigma_f^2 + \sigma_1^2) + \sigma_2^2}}\Big),$$

$$\varpi_t^+ = \Phi\Big(\frac{X_0(p-1)(1+r+r\varphi)/(r\varphi) - \varphi(1-p)(f_{t-1} + \theta_{1,t-1})}{\sqrt{\varphi^2\hat{\sigma}_{2,t-1}^2 + (1-p)^2(\sigma_f^2 + \sigma_1^2) + \sigma_2^2}}\Big).$$

It is immediate that we also have $\varpi_t^+ \approx 0$ in this case by the same logic as the proof in Proposition 1. In the presence of the signals, the investors believe that $\mathrm{E}[\theta_{2,t} \mid \mathrm{G}_{t-1}^+] - \varphi\hat{\theta}_{2,t-1} = 0$, so that their expected private cash flow one period ahead is an accurate forecast of the actual private hidden cash flow. Therefore the result reduces to Assumption 2, which states that the abandonment should not happen based on merely one bad private cash flow shock.

We argue that the most practically relevant case is when $\theta_{2,t-1} \ll \hat{\theta}_{2,t-1} < 0$. In this case, the hidden cash flow is largely negative, which results in a large crash probability at time t. However, this information is not fully revealed by the signals. In this case, our model predicts the following results:

Proposition 4. For $t > t^+$, conditioning on that $\theta_{2,t-1} \ll \hat{\theta}_{2,t-1} < 0$, π_t^+ satisfies $\pi_t^+ < \pi_t$, $\frac{\partial \pi_t^+}{\partial \overline{s}_{t-1}} > 0$, $\frac{\partial \pi_t^+}{\partial S_{t-1}} < 0$ and that:

$$\lim_{S_{t-1}\to\infty} \pi_t^+ \approx 0.$$

Proof. The first claim can be derived directly by comparing the expression of π_t and π_t^+.

For the second and the third claims, we note that $\hat{\theta}_{2,t-1} = \hat{\sigma}_{2,t-1}^2 S_{t-1}\overline{s}_{t-1}$, therefore $\frac{\partial \hat{\theta}_{2,t-1}}{\partial \overline{s}_{t-1}} > 0$ and $\frac{\partial \hat{\theta}_{2,t-1}}{\partial S_{t-1}} < 0$ hold, which imply the two claims. The last claim follows by noting that $\hat{\theta}_{2,t} \xrightarrow{p} \theta_{2,t}$ as $S_t \to \infty$, and π_t^+ reduces to the case of Assumption 2 in the limit. This completes the proof. ■

The above proposition has several crucial implications that support our empirical analysis. In the scenario with $\theta_{2,t-1} \ll \hat{\theta}_{2,t-1} < 0$, we firstly see that the availability of signals will always reduce the actual crash probability comparing to the case with no signals. This results from the following mechanisms: (1) the investors use $\bar{s}_{t-1}$ to obtain a more precise estimate of $\theta_{2,t-1}$, and consequently a more precise forecast of the firm value $\mathrm{E}[K_t \mid \mathcal{G}_{t-1}^{+}]$. As the manager's abandonment depends on the discrepancy between the actual and the perceived value of the firm, the presence of the signals reduces such discrepancy thus lowering the crash probability; (2) Interpreting $\theta_{2,t-1} - \hat{\theta}_{2,t-1}$ as the amount of concealed private cash flow at time $t-1$, we see that a negative $\hat{\theta}_{2,t-1}$ results in less negative private cash flow being concealed by the manager. Since the tolerance of concealing bad news is fixed, this also leads to a lower crash probability.

Secondly, the actual crash probability is monotonically increasing in $\bar{s}_{t-1}$, the average sentiment of in-house meetings at time $t-1$. This is intuitive as the estimate $\hat{\theta}_{2,t-1}$ is already biased upwards towards zero, therefore an increase in $\bar{s}_{t-1}$ would actually imply that the investors are less aware of the actually large hidden cash flow. This allows the manager to conceal more negative cash flow and inflate the actual crash probability.

Thirdly, Proposition 4 predicts that the actual crash probability decreases monotonically as S_{t-1} increases, that is, the quality of the signals received at time $t-1$ in terms of N_{t-1} and τ_{t-1}. This provides theoretical justifications to our claim that larger quantity and more readable in-house meeting reports should reduce crash probability for the next period. Finally, we show that the crash probability can be reduced to zero if the investors receive signals of the utmost quality. In this case, the firm is close to completely transparent with no information asymmetry between the insider and the outsiders, and the abandonment condition will never be triggered. Our results thus emphasis the importance of timely and accurate disclosure of private information as an effective tool to prevent a price crash in the future due to information asymmetry.

We conclude our model by showing that the crash risk as measured by skewness comoves with π_t^+, so that our results in Proposition 4 is applicable to another measure of crash risk. As $\varpi_t^+ \approx 0$ holds in the case with in-house meeting reports, the optimal dividend policy is solved by the following system of equations:

$$\begin{aligned} \mathrm{E}[V_t^{ex} \mid \mathcal{G}_t^+] &= \frac{\mathrm{E}[Y_{t+1} \mid \mathcal{G}_t^+] + \mathrm{E}[V_{t+1}^{ex} \mid \mathcal{G}_t^+]}{1+r}, \\ \mathrm{E}[V_t^{ex} \mid \mathcal{G}_t^+] &= \mathrm{E}[\alpha K_t \mid \mathcal{G}_t^+]. \end{aligned}$$

This yields the solution $Y_t^* = \alpha \mathrm{E}[C_t \mid \mathcal{G}_t]$, which is the same dividend policy as the no signal case with an augmented information set. Note that Y_t^* here is adapted to $\mathcal{G}_t$ instead of $\mathcal{G}_t^+$ because the manager needs to decide the dividend payment for period t before the signals for time t are produced. The realized returns r_{t+1}^{nd} and r_{t+1}^{d} under learning are also similar to their counterparts in the no signal case:

$$\begin{aligned} r_{t+1}^{nd} &= r + \frac{(1+r)(\varepsilon_{t+1} + \xi_{1,t+1})}{X_0(1+r)/r + \varphi(f_t + \theta_{1,t} + \hat{\theta}_{2,t})}, \\ r_{t+1}^{d} &= \varphi p - 1 + \frac{pX_0(1+r-\varphi)/r + \varphi p(\varepsilon_{t+1} + \xi_{1,t+1})}{X_0(1+r)/r + \varphi(f_t + \theta_{1,t} + \hat{\theta}_{2,t})}. \end{aligned}$$

It is evident that the properties of the returns in the learning setting remains largely unchanged compared to the perfect concealment case. However, different from the perfect concealment case, in the learning case the physical crash probability π_{t+1} is now a function of the signals received at time t. We conclude our model on discussing the implication of

S_t of signals and skewness of the return in the learning scenario. Deriving the exact marginal effect is difficult due to the fact that both r_{t+1}^{nd} and r_{t+1}^{d} are function of $\hat{\theta}_{2,t}$, which is by itself a function of S_t. Therefore, we provide some indirect evidence supporting the argument that an increase in the quality of the signals lead to a reduction in crash risk as measured by skewness. Firstly, we have that:

Proposition 5. Under the assumption that $skew(r_{t+1} \mid \mathrm{F}_t) < 0$, it holds that $\frac{\partial skew(r_{t+1} \mid \mathrm{F}_t)}{\partial \pi_{t+1}^{+}} < 0$ for $\pi_{t+1}^{+} \downarrow 0$.

Proof. We firstly state without proof that $skew(r_{t+1} \mid \mathrm{F}_t)$ is a continuous function of π_{t+1}^{+}. Since condition (A.7) is still applicable to the case with signals, we see that in the region $(0, \frac{1}{2} - \frac{3(h_{t+1}^{nd} - h_{t+1}^{d})}{2\Delta_{t+1}^{2}})$, $skew(r_{t+1} \mid \mathrm{F}_t) < 0$. Since r_{t+1}^{nd} is normally distributed conditionally, it holds that $skew(r_{t+1} \mid \mathrm{F}_t) = 0$ as $\pi_{t+1} = 0$. Therefore it must hold that $\frac{\partial skew(r_{t+1} \mid \mathrm{F}_t)}{\partial \pi_{t+1}^{+}} < 0$ for $\pi_{t+1}^{+} \approx 0$ as the continuity of $skew(r_{t+1} \mid \mathrm{F}_t)$ implies that $skew(r_{t+1} \mid \mathrm{F}_t)$ must be monotonically decreasing in π_{t+1}^{+} around the origin. This completes the proof. ■

This result shows that, in the case of $skew(r_{t+1} \mid \mathrm{F}_t) < 0$ and $\pi_{t+1}^{+} \approx 0$, we would expect the skewness to revert back to zero from below as $\pi_{t+1}^{+} \downarrow 0$. Now we argue that the unconditional marginal effects of S_t on π_{t+1}^{+} will be mostly negative, because the largest impact of S_t comes from the scenario $\theta_{2,t-1} \ll \hat{\theta}_{2,t-1} < 0$ depicted in Proposition 4, whereas in the other direction this marginal effect will be close to zero thus negligible due to Assumptions 1 and 2. As a result, we would expect in the steady state, future unconditional skewness of return will decrease when signals of better quality are given in the current period. Combining this with Propositions 4 and 5, we can conclude that in our model, an increase in the quality of signals S_t reduces the discrepancy between the market value and current value of the firm, which leads to a smaller crash probability and a less negatively skewed return distribution. For the same logic, we can conclude that an increase in the average sentiment of signals $\overline{s}_t$ has the opposite effect as an increase in S_t.

Appendix 3: Variable definitions

Variable	Definition
$NCSKEW$	Negative skewness of firm-specific weekly returns, calculated by taking the negative of the third moment of firm-specific weekly return for each sample year and dividing it by the standard deviation of firm-specific weekly return raised to the third power. See equation (2) for details
$DUVOL$	The natural log ratio of volatility in the "down" sample to volatility in the "up" sample. For each stock i over a fiscal year period, we separate the days with firm-specific daily returns above (below) the mean of the period and put them in an "up" ("down") sample. We then calculate the samples' standard deviation and compute the natural log of the ratio of the variance in the "down" sample to the variance of the "up" sample. See equation (3) for details
L_INH	Natural log of (number of firm-specific investor in-house meetings + 1)
$DTURN$	Detrended average monthly stock turnover value, calculated as the average monthly stock turnover minus the average monthly share turnover of the prior year
$SIGMA$	The standard deviation of firm-specific weekly returns over the year
RET	The arithmetic average of firm-specific weekly returns over the year
$SIZE$	Natural log of (Firm's market capitalisation)
MB	market capitalisation scaled by the book value of equity
LEV	Total long term debt scaled by total assets
ROA	Income before extraordinary items scaled by total assets
$ACCM$	Measure of accounting quality, which is the sum of accrual quality value of the previous three years. Accrual quality measure is based on the modified Jones model (Dechow et al., 1995).
$SYNCH$	Stock return synchronicity after log transformation, calculated based on Jin and Myers (2006). See equation (7) and (8) for detail.
$IAPI$	Price impact of information asymmetry, calculated following Amihud (2002). See equation (9) for detail.
$Negative$	The average probability of negative sentences of all in-house meeting content based on "Baidu deep learning NLP model" for a firm over the year
$Readability$	Negative value of natural logarithm of average length of sentences in in-house meeting reports for a firm over the year