The Hidden Information Content:

Evidence from the Tone of Independent Director Reports

Jiao Ji, Oleksandr Talavera[1] and Shuxing Yin

University of Sheffield Management School, Conduit Rd, Sheffield S10 1FL, UK
1 Swansea University School of Management, Bay Campus, Fabian Way, Swansea SA1 8EN, UK

Corresponding author email: jiao.ji@sheffield.ac.uk

ABSTRACT

The paper investigates the link between the information content of independent directors' reports (IDRs) and firm financial outcomes. By conducting sentiment analysis of 23,984 IDRs of the Chinese listed companies from 2004-2012, we find that the positive tone of IDRs is likely to signal an increase in future firm performance. We also posit that the tone of IDRs and its association with firm performance depends on director's incentives to monitor. Our results suggest that independent directors with greater career concerns (i.e., young directors or experts in accounting or finance) are more critical in evaluating firm fundamentals and express more negative tone in their reports. The relationship between the negative tone of IDRs and future firm performance is stronger for firms with greater monitoring needs. Overall, our evidence is consistent with the conjecture that career concerns motivate independent directors to disseminate information to external stakeholders.

JEL code： G30; M410

Key Words: Text Analysis, Tone, Independent Director Report, Corporate Governance

1. Introduction

Financial reporting and corporate disclosure are critical means for management to communicate firm performance and governance to outside stakeholders (Healy and Palepu, 2001; Loughran and McDonald, 2016). Prospects for corporate financial performance can be assessed through processing information from financial statements, earnings releases, and even conference call transcripts. However, quantitative information often provides investors with an incomplete picture of firm's current issues and future expectations. It can sometimes lead to misrepresentation or even manipulation.[1] Recent studies suggest to use textual analysis to investigate the tone of CEO disclosure since it reveals managers' optimism related to future earnings (Davis, Piger, and Sedor, 2012; Li, 2010a). However, such approach can be biased, due to the CEO's motivation to influence the perceptions of outside stakeholders (Huang, Teoh, and Zhang, 2014; Arslan-Ayaydin, Boudt, and Thewissen, 2016). Therefore, there is a great need to broaden the spectrum by exploring alternative channels of information dissemination originated from other stakeholders.

A possible channel of information dissemination stems from independent directors. To evaluate and authorise corporate decisions, independent directors possess a large amount of information about the company (e.g. Adams and Ferreira, 2007; Fama and Jensen, 1983). The tone of their disclosures could provide useful insights to outside stakeholders by indicating optimistic or pessimistic biases. This disclosure channel might not be perfect, as independent directors can be passive monitors (Hermalin and Weisbach, 1998; Tirole, 2001) or have restricted access to information (Adams and Ferreira, 2007) . In this study, we investigate whether the content of directors' disclosure contains useful information to outside stakeholders. We also examine

[1] Financial information might be misreported as it happened with Tesco in 2014. https://www.theguardian.com/business/2017/mar/28/tesco-agrees-fine-serious-fraud-office-accounting-scandal, accessed 12 December 2017.

the strength of this information channel depending on variations in directors' incentives to monitor.

Our analysis employs data of Chinese listed companies which provide us with a unique framework to explore this research focus. In 2004, the Chinese Securities Regulatory Commission (CSRC), the Chinese stock market regulator, mandated that reports of independent directors should be disclosed timely to the general public following board meetings. The reports comprise contents of the proposals discussed at board meetings and opinions of independent directors. This mandatory disclosure rule was created to improve corporate governance and market transparency, and to protect the interest of minority shareholders. The rule allows us to obtain a novel dataset of independent director's disclosure, which are not publicly available in the US or any other major markets.

Even though IDRs is a very important disclosure channel it does not always transmit important corporate information in Chinese corporate environment. About 98% of reports in our data are 'Agree' opinions, the binary outcome ('Agree or Dissent') per se cannot completely reveal the opinions of independent directors about firm performance. In this case, the tone of a report might reflect director's genuine evaluation of firm prospects. Moreover, the tone of such reports is less likely to be biased, since independent directors (mainly playing a monitoring role) do not have incentives to manipulate their tone to mislead outside stakeholders (Adams, Almeida, and Ferreira, 2009). Our sample contains approximate 24,300 independent director's reports from 2004 to 2012 which are processed by a machine learning technique utilizing Chinese word segmentation to capture the hidden content.

Our findings provide new evidence that the tone of board disclosures is correlated with future firm performance. Using fixed effects panel data analysis, we find that the positive tone of independent director's report indicates optimistic prospects for one-year future performance. Because we include numerous control variables that would normally be expected to impact the

use of tone in the reports, our measure of tone can be interpreted as unexpected tone given a firm's circumstances. Thus, the tone of the report has an explanatory power incremental to other factors in forecasting future performance. It also confirms that boards of directors are not passive and they understand the problems/opportunities related to business operations and *attempt* to convey information to outside stakeholder.

Which factors impact the language used in reports of independent directors? In recent years, numerous studies find that the tone of corporate disclosure is related to firm profitability, firm characteristics and managerial incentives (e.g., Davis et al., 2012; Arslan-Ayaydin et al., 2016; Huang, Teoh, and Zhang, 2014b). We first examine whether observable individual director-specific characteristics that are likely associated with cognitive optimism or pessimism (e.g., age, gender, educational and career experiences) explaining directors' tone in reports. Particularly, we focus on those characteristics that related to directors' career concern that has been proposed as the major motivation for independent directors (Fama, 1980). Previous studies on board of directors adopt firm level data share the same problem that board of directors are endogenously selected, any relation between board governance characteristics and firm performance could reflect the optimization of individual firms under different parameters rather than a causal relation resulting from the actions of directors (e.g., Hermalin andWeisbach, 1998). It is still not clear that how independent directors should reflect to career concern, and explain the variations in director behaviors due to lack of director-level data. Our study explores the unique director-level dataset of independent director reports to fulfil this research gap.

We find that young directors (Zajac and Westphal, 1996) and those directors with an accounting/finance background (Badolato, Ege, and Donelson, 2013) who concern more about career, are more likely to express a negative tone in their reports. The negative tone in reports shows the directors' tendencies to express a negative view on firm performance and decisions straightforwardly, which challenges the interest of management team and major shareholders.

Thus, our results suggest that higher reputation concern drive directors to play the monitoring role in firm decisions. This is consistent with the career concern models (Holmstrom, 1999) where agents who aspire to advance their careers work harder. In addition, we find that the total payment of independent director, gender and education level did not influence the variation of tone expression. Our results also show that firms with more board committees, greater board meetings, lower leverage, and controlled by private shareholders have an higher likelyhood of receiving a positive independent director report.

Prior research assumes that different managers make similar language choices in the same circumstances, in response to economic incentives (Davis, Ge, Matsumoto, and Zhang, 2015). The assumption appears reasonable regard for the potential litigation costs and reputation costs. During a firm crisis, directors concern more about their liabilities, and their relationships with management have a shorten expectation than those during good times, so directors are more likely to challenge the managirial decisions; however directors rarely dissent at board meetings in good times (Tirole, 2010). Thus we expect that positive /negative tone have different prediction power under good and bad firm circumstances. Particurlarly, in bad circirmstance, independent dirctors possibablely play an intensive monitoring role, and express a negative tone that is related to future performance. We investigate the validity of this conjecture by conducting a sub-sample analysis and find that in risky firms (those with high bankruptcy risk, liquidity risk, and overall bad-performance), the negative tone is significantly linked to firm future performance, whereas the positive tone is not associated with future performance.

We contribute to the literature in two important ways. First, we provide the first systematic evaluation of independent directors' disclosure. Prior research primarily considers the potential for tone to signal managers' private information about future performance. Our paper considers alternative channel of tone expressed by independent directors, and suggests that the tone of independent director has prediction power about firm performance. Moreover, previous studies

that employ the dataset of independent director reports in China only focus on the 'dissent' reports (Jiang, Wan, and Zhao, 2016; Tang, Du, and Hou, 2013). They find that firms with more severe agency problems are prone to experiencing independent directors saying 'no', and the dissent vote are partially determined by director' characteristics. However, the dissent reports only account for less than 2% in the dataset. These studies ignore the director's function in corporate governance in 'normal' situations. Using the reports with 'Agree' opinions, our research provides a systematic evidence with that the disclosure of independent director is informative to outside stakeholders about future performance, implying that it can reduce the information asymmetry between the management and outside stakeholders under the existing of controlling shareholders in most listed firms. It lends support to regulations mandating the disclosure of independent directors' opinions in China. It also supports the effectiveness of independent directors in offsetting the power of controlling shareholder in China, which is consistent with evidence from other economics (e.g., Dahya, Dimitrov, and McConnell, 2008).

Second, we contribute to literature on boards of directors by identifying observable reputation-related director-specific characteristics that explain the tone expression. Directors use directorships to signal internal and external markets for desiciosn agents that they are decision experts (Fama and Jensen, 1983). Prior empircal studies by finding that career opportunities for directors are related to their performance, and directors are rewarded with more career opportunities for "good" performance (e.g., Jiang, Wan, and Zhao, 2016). Our studies provide new evidence on reputation incentive by explaining how different reputation incentive captured by director-specific characteristics vary director's behavior- tone expression in their reports. Our findings suggest that only reputation related characteristics related to the tone expression, while other characteristics are not related to the tone of reports.

The rest of this paper is organized as follows. Section 2 introduces the institutional background of the independent director system in China, the implications of textual analysis using

a Naïve Bayesian learning algorithm, and hypotheses development. Section 3 presents the details of the method to classify the tone of independent director' reports, model design and sample statistics. Section 4 discusses the empirical results, and Section 5 presents the robustness check. Section 6 concludes the paper.

2. Institutional Background, Literature Review and Research Questions

2.1 Institutional background: the role of independent directors and their reports in China

Like many other emerging markets countries China's quality of corporate governance still needs substantial improvement. Its stock market, established around 1990, was initially aimed to provide state-owned enterprises (SOEs) capital and liquidity to develop and reform. Few decades later, listed companies, originated from SOEs still face strong government influence. Another problem is related to complicated legal structure and groups affiliations. (Allen, Qian, and Qian, 2005). These two aspects reduce the transparency of corporate operations and governance of private companies. To alleviate these issues and protect the interests of outside stakeholders, Chinese policy makers formally introduced boards independent directors in 2001.[2]

The role of the board of directors in China is comparable to that of those in the US. Boards of directors have a legitimate obligation to oversee a firm's strategic decisions and policies, and to select, review, compensate and terminate contracts of top management. The monitoring and advising functions are implemented after discussing and voting proposals at board meetings. A proposal has to obtain support by the majority of the board to be actioned. Also most

[2] The establishment of independent director system is based on "Guideline for the establishment of the independent director system in listed firms" by the China Securities Regulatory Commission (CSRC) in August 2001, then legal status is established in new "Corporate Law of China" 2005.

of these proposals (around 88%) are intiated during board meetings by a Chairman.[3] Independent directors are obligated to monitor these proposals to prevent firm decisions against the interest of minority shareholders.[4]

Although important to all stakeholders, the contents of the proposals and boards' opinions on these are not nomally released to general public due to data confidentiality. China initiated a unique requirement that mandates firms to disclose the contents of board meetings in a timely manner by *Reports of Independent Director*.[5] The reports contain the narratives of meeting topics (proposals), the opinions of the independent directors, and identities of the independent directors. This practice is aimed to improve transparency and provide outsiders with a timely notice of compromising circumstances. The novel practice allows building a dataset which connects the independent directors' attitude towards firm decisions

A few studies utilize data originating from reports to investigate dissenting votes of independent directors, which reflect a director's willingness to challenge the management and controlling shareholders on behalf of outside shareholders. Tang, Du, and Hou (2013) find that the stock market has a negative reaction to the disclosure of independent reports with dissenting votes. The probability of receiving a negative report is higher for firms with more agency problems. Jiang, Wan, and Zhao (2016) show that independent directors who have higher human capital concerns are more likely to have negative opinions in their reports, which results in

---

[3] Source: a 2004 survey of 204 firms by the Research Center at the Shanghai Stock Exchange.

[4] The regulation of independent director system in China was meant to be a "legal transplant" from U.S. corporate governance law and practice (Clarke, 2006). For example, independent directors can be nominated by board and supervisory board members or shareholders who has at least one percentage share; appointment decisions are made by shareholders' meeting. They constitute at least one third of board, only serve maximally two terms (3 years per term). They are not employees in the listed firms, or have kinships with corporate employees. They do not have over 1 % share of listed firm, or they do not have kinships with nature people in 10 largest shareholders. They are not employees or kin to employees in firms which have over 5 % share of the listed company or in firms which are 5 largest shareholders.

[5] "Code of information disclosure for listed firms: Annual reports" in 2004.

obtaining more directorships and lowering the likelihood of regulatory sanctions. As a result, stakeholders take actions to improve corporate governance following dissent by independent directors - decreasing inter-corporate borrowing, reducing bank loan, and dismissing the incumbent CEO/Chairman (Jiang *et al*., 2016; Tang *et al*., 2013). Overall, these results suggest that independent directors' dissension is useful to protect the interests of outside stakeholders.

However, the negative opinions are found in only 328 reports, which accounts for 1.15% of 28,634 reports from 2004 to 2012.[6] Ji, Talavera, and Yin (2016) use the full dataset to obtain the meeting frequency for discussion of different topics and find that the number of nomination and growth strategies meetings could alter the relationship between CEO turnover/compensation and firm performance. The results indicate that 'agree' reports contain more information than the 'agree/dissent opinions'. Therefore, this study takes a step forward and examines whether the narratives in independent director reports contain influential information content.

2.2 Sentiment analysis in Accounting and Finance, tone of independent director reports and firm performance

With substantial increase in computing power over last few decades, researchers in accounting and finance start applings the textual analysis methods to large amount of unstructured qualitative information contained in corporate disclosures, including news articles, earnings conference calls, Securities and Exchange Commission (SEC) filings, and text from social media (see e.g., Li, 2010b; Loughran and Mcdonald, 2016). Empirical studies (e.g., Das and Chen, 2007; Li, 2010a; Tetlock, 2007) have actively inspected the impact of qualitative information on equity valuations and have shown that the sentiment of these disclosures is correlated

[6] The low dissension rate is not unique in China, which is around 2% in Israel firms (Schwartz-Ziv and Weisbach, 2013). In fact, director dissensions are hardly observed, the cases in China are precious to researchers who interested in board function.

with future earnings, stock returns, and even the likelihood of future fraudulent activities by management.

The views of managers has been in the center of corporate disclosure research, since managers are important corporate insiders who have relatively precise and complete information about corporate operations and future challenges to the firm. Li (2010b) argues that the earnings quality extracted from managerial disclosure can be '*incremental or even superior to existing empirical measures*', since there is a high information symmetry between managers and outsiders. Empirical studies have shown that the Linguistic patterns and textual analysis of management's views have predictive power of future firm profitability, such as the CEO online letter to Shareholder (Segars and Kohut, 2001), CEO forward-looking statements in annual reports (Li, 2010a), and management views on firms' competitive environment (Li, Lundholm, and Michael, 2013). Therefore, it has been documented that manager's disclosures have significant connections with firm profitability and earning quality.

To our best knowledge, there is no evidence of the informativeness of directors' disclosure. This may be because directors in developed economies can only express their opinion about firm performance and governance to the public via annual reports. This type of disclosure is stylized and modified, so it may not show directors' communication patterns. The communications by boards of directors offer researchers a great setting in which to recognise private information sets, namely seeing corporate operations from director' eyes. Independent director reports (IDRs) are an innovative source of corporate disclosures, which can contain value information content regarding firm operations which cannot be explained by financial data.

Independent directors are often criticized for passive monitoring due to the lack of independence and motivation to mitigate agency costs (Tirole, 2010). However, directors have demand to protect/enhance their reputation by signaling their commitment to practice good cor-

porate governance so they can obtain more seats in listed firms. It has been suggested by theoretical and empirical literature that reputation could be a strong incentive for independent directors to oversee firm operations (Fama and Jensen, 1983; Fich and Shivdasani, 2007; Levit and Malenko, 2016). Moreover, independent directors necessarily use independent reports to avoid legitimacy risk stemming from illegal behavior by the firms or the executives which would damage shareholders' interests.

We expect the narratives in the reports include information about firm operations, and explore whether these reports convey indications about future firm performance. Independent directors who have negative (positive) information about firm operations or future strategies, should issue IDRs with a negative (positive) tone. We hypothesize that the tone of reports is positively associated with future firm performance. The empirical examinations are a joint test of (a) whether independent directors provide information about future firm performance in their reports, (b) whether they have different attitudes about future regardless of whether the majority of their opinions are in "agreement" and (c) whether the sentiment analysis method-machine learning technique captures the information content in reports.

If the sentiment analysis method is able to measure the tone of independent directors, then evidence that the tone of IDRs predicts future performance is coherent with the hypotheses that independent directors shows attitudes in their reports and IDRs contain information content. However, if the IDRs' tone, based on our measure, is not related to future performance, we cannot reject hypothesis (a) and (b) because the result can be because of the low power of the method applied.

2.3 Determinants of IDR's Tone

The content of reports is determined not only by independent directors' characteristics, but is also influenced by corporate governance and firm performance which is the foundation of

directors' analysis.

2.3.1 Independent directors' incentives, experience and other characteristics.

Director's age and number of directorships are employed to proxy for director's reputation incentive for monitoring. Age could be a factor influencing directors' reputation concerns and behaviors (Zajac and Westphal, 1996). For example, young directors have longer career paths, so they have more motivations to build a good reputation. Jiang *et al*. (2016) document that younger directors are more likely to issue dissension reports to reveal governance problems. Directorships is the number of seats of the director in listed companies, which has been widely used in empirical literature as a proxy for the human capital in developed countries such as the U.S. (Yermack, 2004). Financial/accounting experts have a lower probability of obtaining directorship seats, because these directors are often have lower social status compared to other directors (Badolato, Ege, and Donelson, 2013). The frauds in listed firm damage the reputation of accounting and financial experts, therefore these experts are more likely to disclose negative components of firm operations.

Other director-level factors such as monetary incentive, gender diversity and other features of directors are included in our study. We use director's cash payments to measure monetary incentive. Although equity based compensation has been used for independent directors in the U.S. (Ryan and Wiggins, 2004), it has not been broadly adopted in Chinese listed firms. Adams and Ferreira (2009) suggest that female directors are more likely to allocate more effort to monitor, in terms of better meeting attendance and joining more committees. Female presence on board could "push" male directors to work harder as well. Hence gender diversity has a positive influence on board efficiency and firm outcomes. Gender diversity is measured by the percentage of women in independent directors. Other characteristics such as education level and work experience have been controlled as well.

2.3.2 Corporate Governance Indicators and Firm Features.

Empirical studies have explored a number of board characteristics which influence board monitoring and advising activities, which affect corporate governance and firm outcomes. Liu *et al*. (2015) find comprehensive and robust evidence to support that board independence has a positive effect on firm performance in China. Particularly, they document that independent directors have a positive impact on reducing insider self-dealing. The choice of optimal board size reflects the tradeoff between costs of monitoring and benefits considering the work efficiency as a group. Several studies examine the relationship between board size and firm performance and support that larger board size has mixed effect on firm value depending on the development stage of the firms (Eisenberg, Sundgren, and Wells, 1998; Yermack, 1996). Remuneration, audit and nomination committees could meet more often than the entire board, and these committees could function more efficiently because of higher independence, smaller size, and greater degree of expertise. Empirical research find that various committees provide means to intensify monitoring and have positive effects on reducing related corporate governance issues (Conyon and Peck, 1998; Klein, 2002; Shivdasani and Yermack, 1999). Although the CSRC encourages listed firms to set up committees to improve their corporate governance system, it does not specifically mandate the number and type of committees which should be established. Hence, firms decide which committees are established, so firm with more committees have higher corporate governance quality. Board meeting frequency is used to quantify the level of monitoring activities (Vafeas, 1999).

Firm size captures many properties of a company's business environment. Larger firms provide directors greater visibility, status (Adams and Ferreira, 2009), reimbursement (Ryan and Wiggins, 2004), and the possibility of gaining extra directorships (Yermack, 2004). One major role of independent director in China is to prevent tunnelling behaviour (related party transactions) and earnings manipulation in listed firms. Account receivables, as the main accrual

anomaly, has been used as an instrument to manipulate earnings in China (Peng, Wei, and Yang, 2011). Abnormal accounts receivable could be negatively related to the tone of reports. Ownership structure affects firm fundamentals, such as firm performance, corporate disclosure transparency and CEO turnover in China (e.g. Kato and Long, 2006b). We use several variables to control the ownership effects: a state-owned enterprise indicator, share owned by largest shareholder, and the ratio of shares owned by largest shareholder to the second largest one. We also include standard firm performance measure – ROA and Tobin's Q, and capital structure measure – leverage ratio.

3. Methodology

3.1 Implementation of Naïve Bayesian Algorithm for Sentiment Analysis

Sentiment analysis is the computational research regarding the tone or opinions of textual information using natural language processing, which has been widely employed in analyzing customers' reviews and social media users' behaviours. There are two standard methods for the sentiment analysis: dictionary-based approach and machine learning. The dictionary-based approach uses a predefined dictionary of positive and negative words to match the words, phrases or sentences into groups (Also called 'bag-of-words' model).[7] In the accounting and finance literature, four different word lists have been extensively used by researchers: Henry (2008), Harvard's GI, Diction, and Loughran and Mcdonald (2011). However, the method does not consider the unexpected effect or inaccurate result due to the same word having a different tone in various industries or topics. For example, consider a sentence from the Chinese independent director reports translated into English "The asset has had good profitability and huge market

[7] Recent financial studies show that negative tone impacts firm returns by using dictionary method (Loughran and Mcdonald, 2011; Ahmad *et al*., 2015).

potential, the attrition only occurs short-term, we still expect the company will follow the future development to provide a greater return for investors." Based on National Taiwan University Semantic Dictionary (NTUSD), the sentence has a lot of positive words (such as "profitability" and "greater"). However, we can see that the overall tone tends to be negative.

We adopt the machine learning method, the Naïve Bayes methods, that apply algorithms as a classification problem to mitigate the disadvantage of the dictionary-based approach.[8] We use a partition of the complete corpus of textual data to train a classifier based on linguistic features, then use the classifier to score the remaining corpus. The words in the training set are tokenized as "positive", "negative", or some other sentiment, for instance, "calm", "tense", "excited" and "upset" depend on the circumstance. The statistical inference picks up sentiment classification rules from the trained set and applies these rules to the entire textual data. The machine learning approach has the advantage of processing the particular textual data by constructing customised classifiers, which can be trained efficiently under supervised learning. The Naı¨ve Bayes machine learning approach has been widely used to analyze disclosures in the U.S. market such as annual report fillings (Li, 2010a; Purda and Skillicorn, 2015), analyst reports (Huang, Zang, and Zheng, 2014), and newspaper articles concerning U.S. merger announcements (Buehlmaier and Zechner, 2016).

All Chinese words in the reports have been segmented before we apply the machine learning method. Unlike English corpus, different combinations of Chinese characters often have different meaning. The Character Based Generative model is used because it provides the highest accuracy rate (94%) for Chinese segmentation (Wang, Zong, and Su, 2012).

[8] The Naïve Bayes methods has been applied in accounting and finance research to investigate the sentiment of sorts of financial text, such as stock message postings (Antweiler and Frank, 2004), 10-K filing (Li, 2010a; Purda and Skillicorn, 2015), and analyst reports (Huang, Zang, and Zheng, 2014).

Using segmented words as terms, we create a term-document matrix that describes the frequency of terms that occur in a given document, where in rows correspond to the occurrence of terms in the document and columns correspond to the terms. Then we can reduce sentences to a list of words (d) by frequency in the sentences. The aim is to classify the sentence into a specific category (c) from a set of all predefined categories (positive, negative and neutral). Let $\{w_1, \dots, w_t\}$ be a predefined set of sentences with t features. Let $n_i(d)$ be the occurrence of w_i in document d, we have the document vector $d = (n_1(d), \dots, n_t(d))$. So the best category can be described as $c^* = argmax_c P(c|d)$. Using Bayes' theorem, the conditional probability is

$$P(c|d) = \frac{P(d|c)P(c)}{P(d)}$$

Where $P(c)$ is the prior probability of a category occurs, $P(d|c)$ is the prior probability that a given document set is classified by a category. $P(d)$ is the prior probability that a given document set occurs. We assume all documents are independent, thereby the problem is equivalent to:

$$P(c|d) = \frac{P(c) * P(w_1|c) * \dots * P(w_t|c)}{P(d)}$$

Since we have three categories, $P(d)$ has no effect in c^*. It can be eliminated. The equation can be rewritten as follows:

$$P(c|d) = P(c) * P(w_1|c) * \dots * P(w_t|c)$$

And the document categorization algorithm is described as

$$c^* = argmax_c(c) * P(w_1|c) * \dots * P(w_t|c)$$

The assumption is independence for each document, in that the probability of each word appearing in a document is unaffected by the presence or absence of each other word in the document. Although the conditional independence assumption does not fully hold in reality, the

Naïve Bayesian algorithm has little effect on the results and still deliver accurate categorization (Lewis, N'edellec, and Rouveirol, 1998).

We use SnowNLP (a Chinese sentiment package) in Python to conduct the sentiment calculation. The package includes a pre-defined corpus from various sources (e.g. financial newspapers and social media). This package has been used previously in the computational linguistics area (e.g., (Chen, Wan, and Xu, 2016; Ouyang, Li, and Li, 2016). Researcher subjectivity is circumvented by the utilizing this computation, when compared to manually training a corpus using independent director reports.

3.2 Sample

3.2.1 Preparation of data

Our sample contains data from non-financial main board firms listed on Shanghai and Shenzhen stock exchanges over the period of 2004 to 2012 from The China Stock Market and Accounting Research (CSMAR) Database. The database has been used for research published in world leading journals. We select the start period 2004 instead of 2002 because the amount of independent director reports increases significantly after a two-year transition period.[9] Reports with disagree opinions (around 400) are excluded, as these already indicate the "strongest" negative tone. It yields 23984 reports from 2004 to 2012.

Table 1 presents variable definitions and descriptive statistics of the main variables. The tone for each report (*Tone_each*) is a continuous variable, ranging from 0 to 1, representing negative view to positive view with 0.5 imply a neutral. The average tone is 0.458 which is

[9] The Information Disclosure Standards (CSRC, 2005) further clarified the disclosure requirements, improving the quality and quantity of the reports. After 2005, the number of independent directors' reports increased accordingly. In the same time, the "Code of information disclosure for listed firms: Annual reports" (2004) enhance the timely disclosure of IDRs.

slightly negative (median 0.447) about future proposals as shown in Panel A of Table 1. The slightly negative sentiment of director's disclosures is consistent with sentiment of manager's disclosures. For example, Li (2010a) documents there is a slightly negative sentiment of the CEO forward-looking statement from management discussion and analysis of 10-K and 10-Q fillings in the US.

For firm *i* with *k* reports in year *t*, we define its annual tone (*Tone_year*) as the average tone of all the reports for a firm. $Tone\_year_{i,t}$ is created as the independent variable merging with annual financial data to test whether the tone of IDRs predicts future performance. The annual average tone (*Tone_year*) is 0.463, similar to the tone of individual reports (*Tone_each*). We then define the dummy variable *positive tone* that equals one when the annual average tone for a firm is greater than 0.7, and zero otherwise. Correspondingly, *negative tone* is a dummy variable, taking the value of one when the annual average tone of a firm is less than 0.3, and zero otherwise. The number of reports with negative tone (27.6%) is more than that of ones with positive tone (20.1%).

To mitigate the influence of extreme values, all firm level data were winsorized at the top and bottom 1%. The final dataset includes 1,437 firms and 11,249 firm year observations.

Insert Table 1 about here

3.2.2 Descriptive Statistics for main variables

Panel B of Table 1 presents independent director's characteristics. The average age of independent director is 51.8 years, which is significant larger than the average age 49.2 of entire

board by *t*-test (in Panel C). Annual payment is 48, 000 RMB (around 4,800 GBP). Females account for 13.8% in independent directors. Almost one third of independent directors have working experiences in accounting and finance area (or law area).

Panel C and Panel D of Table 1 demonstrate features of entire boards and firms respectively. Generally, a board of a Chinese listed firm has 9 directors with one female, one third of who are independent directors. It has three committees and holds 9 meetings per year. 70% of listed firms are controlled or strong influenced by state, state agents or SOEs.

3.3 Model Design

3.3.1 Model (1) the tone of IDRs and future firm performance

To examine whether the tone of IDRs predict future performance, we apply panel data analysis with firm and year fixed effects using the following model:

$$FirmPerformance_{i,t} = Tone\_year_{i,t-1} + Controls_{i.t-1} \qquad (1)$$

Where $FirmPerformance_{i,t}$ is measured as return on assets (ROA), since it is the most common measure of performance for Chinese listed companies (e.g., Liu, Miletkov, Wei, and Yang, 2015).[10] We also use return on sales (ROS) as a performance measure, the untabulated regression results are similar to the ones using ROA. The independent variable is $Tone\_year_{i,t-1}$ capturing the annual average tone of all the IDRs in firm *i* at year *t-1*. Independent and executive

[10] We do not employ return on equity (ROE) as a performance measure since it is often manipulated to satisfy a SEO requirement. Tobin's Q, though widely used to proxy firm performance in the existing literature, is not considered a proper performance measure for Chinese listed firms. Most Chinese listed firms originated from state-owned enterprises (SOEs) with majority shares not tradable in the secondary market. The non-tradable shareholders, mainly governments or state-owned legal persons, typically acquire their shares of stocks at prices significantly lower than the initial public offering prices. Since there are big pricing gaps between tradable and non-tradable shares, Tobin's Q would not correctly reflect firm financial performances or firm values.

board members are groups that make business decisions, we control the board characteristics for capturing the factors that could influence their work efficiency: *age_board, #female, board size, % independence, number of committees* and *meeting frequency*. Firm status (*Log(total assets), ROA and Tobin's Q*), ownership structure (*SOE, Largest Share and Largest1to2*) and debtor's interests (*Leverage*) are controlled in $FirmControls_{i,t}$. The definitions of key variables are exhibited in Table 1.

Fixed effects estimators are applied to mitigate the heteroscedasticity and endogeneity issues of omitted variables and measuring errors by control unobserved firm invariance influenced. To partially mitigate the simultaneity endogeneity issue, we use lagged values of independent and control variables to facilitate causality. We also estimate this model using IV-GMM method for further identification in the robustness check section.

3.3.2 Model (2): Determinants of IDRs' Tone

Which factors affect the tone of independent director reports? An empirical examination linking the features of independent directors and firm characteristics takes the following form:

$$Tone\_each_{i,k,t} = \beta\ IndependentFeatures_{i,t} + \gamma\ BoardContorls_{i,t} + \delta\ FirmControls_{i,t} \quad (2)$$

Where $Tone\_each_{i,k,t}$ is the tone of report *k* of firm *i* at year t. $\text{IndependentFeatures}_{i,t}$ is a vector of variables reflecting the features of independent directors of firm *i* at time *t.* It mainly includes proxies for reputation concerns (*Age_outside and Directorships_outside*), the monitory incentive (*Pay_outside*), education and working experience (*Education_outside and Financial/ Law expert*). The empirical specification leads to omitted variable concerns, so our estimations includes firm and time fixed effects for controlling firm and time heterogeneity.

4. Empirical Results

4.1 The Relationship between Firm Performance and the Tone of IDRs

In this section, we assess the implications of IDR's tone created by Chinese word segmentation and Naïve Bayesian machine learning algorithm for a firms' future performance. We control the numeric financial information variables which have been explored in the Chinese literature and have been documented to related to firm performance (e.g. Chen, Firth, and Xu, 2009; Liu *et al*., 2015; Ma, Naughton, and Tian, 2010). Thus we provide evidence that the tone contains information contents beyond financial information in annual reports.

Table 2 reports the results of the panel data regressions with firm and year fixed effects from 2004 to 2014 for testing whether the tone of IDRs can predict future performance - Model (1). The dependent variable is firm's profitability – ROA. Lagged value of explanatory variables are employed. Column (1) and (2) of Table 2 includes the IDR's tone and all the control variables, and Column (3) only contain control variables in the regression. The coefficients on lagged *tone_year* in Column (1) and Column (2) of Table 2 are significantly positive, which indicates that the IDRs tone has at least one year of prediction power. It is similar to the prediction power of the tone of CEO statement's on firm future performance in the US (Li, 2010a).

Insert Table 2 about here.

One major concern for listed SOEs is that the particular ownership structure may alter business operation and manage/director motivations. Specifically, top management and directors in SOEs are more likely to have political connections and political aspirations. Hence, we introduce the interaction term between *tone_year* and *SOE* to capture the difference between

SOEs and non-SOEs in the relationship of the IDR's tone with firm performance. It is displayed in Column (2) of table 2. The coefficient on this interaction is negatively significant, which implies that the tone in non-SOEs has a higher predictive power than that for SOEs.

Control variables have consistent signs and similar magnitudes of their coefficients for the models with and without *tone_year* variables. The coefficients on *female_board* and *% independence* are significant and positive, suggesting that more female directors and independent directors on the board are beneficial for firm performance, in line with literature (e.g. Liu *et al*., 2015). Listed SOEs or firms with larger assets are more likely to have lower profitability. Moreover, the share owned by the largest shareholder is positively associated with firm performance.

We then substitute the *tone_year* with the *positive_tone* and the *negative_tone* indicators to estimate the relationship between the IDRs tone and future performance. Table 3 illustrates the regression results of firm performance on the *positive_tone* (Column - 1) and the negative tone (Column - 2) respectively. The *positive_tone* is positively related with firm performance, whereas the negative tone has a negative relationship with firm performance. The scale of coefficient on the negative tone shown in Column (2) is larger than the one of the positive tone shown in Column (1), which indicates that the negative tone contains more information content for predicting future performance. This may be because the key role of independent director's is monitoring required by the CSRC, so they tend to oversee whether firm decisions are made according to legal requirements. To protect themselves using IDRs, they will care more about the negative impacts on firm performance. To compare whether negative and positive tone have the same predictive power, regressions of standardised variables are also employed to further robustness check, and the results remain similarly. Generally, the results from Table 2 and Table 3 support that there is a positive relationship between the IDR's tone and one year ahead future firm performance.

Insert Table 3 about here.

To examine whether the tone have prediction power more than one year ahead, Table 4 reports regression results including one year lagged and two year lagged positive tone (Column - 1) and negative tone (Column – 2) of IDRs. They are regressed on the determinants of the tone and year-firm fixed effects. The lagged one year positive / negative tone has significant relationship with current firm performance. However, the coefficients on two year lagged tone are not significant and obviously smaller, although they have the same sign as the one year lagged value. We estimate the relationship between multi-year lagged tones and current firm performance, by regressing performance on two, three and four year lagged tone. The coefficients on three and four year lagged tone are not significant and the magnitude of the coefficients is even smaller (nearly zero). Therefore, IDR's tone is only related to one year ahead future performance. This is reasonable based on the nature of independent director reports. Independent director's monitoring focuses on the related party transactions (over half of meeting proposals) which only impact assets/debts and earnings in short term.

Insert Table 4 about here.

Overall, we find that IDRs' tone is positively correlated to future firm performance after controlling for other factors that might impact future firm performance. It shows that the tone includes information about firm performance, also supports that our measure for the IDR's tone is accurate.

4.2 Determinants of IDRs' Tone

In this section, we investigate which factors influence the tone of independent director's reports in each report. Table 5 shows regression results with random effects based on model (2)-determinants of IDR's tone. The dependent variable is $Tone\_each_{i,k,t}$ in Column (1) and (2), the sentiment for report *k* of firm *i* at year *t*, regressed on its hypothesized determinants as discussed in section 2.3. The dependent variable is *negative_tone* shown in Column (3), whereas *positive_tone* is shown in Column (4). The results in Column (3) and (4) are coefficients using the logit regression. Column (1), (3) and (4) of Table 5 include all the explanatory variables, and Column (2) contains board controls and firm controls. We also estimate the models with firm fixed effects, and the untabulated results remain similar. Results with random effects are reported for two reasons: the result of the Hausman test cannot reject the null hypothesis that the preferred model is random effects; in China, a huge proportion of listed firms (about 70%) are SOEs, and this time invariant fundamental difference may have an impact on the tone, which cannot be shown in the firm fixed effects models.

Insert Table 5 about here.

The tone of IDRs is influenced by independent director's characteristics. The Adjusted R-square in Column (1) is larger than the one in Column (2), indicating that modeling with independent director's characteristics has a higher explanatory power. The coefficient on independent directors age is negatively significant as shown in Column (1) and (4), while it is shown to be positive in Column (3). Thus the age of independent directors is positively associated with the IDR's tone. Younger independent directors are more (less) likely to issue reports with neg-

ative (positive) tone. It is consistent with our hypothesis that younger directors have more reputation concerns, and therefore monitor more intensively. The coefficient on financial expert in Column (1) is significantly negative, while positive in Column (3), suggesting that independent directors with financial or accounting experience are more conservative when they issue reports, and they have a higher probability to issue negative reports. It supports the policy of adding financial experts to a board to oversee firm risk.

In addition, Table 4 shows that several board and firm features impact the IDR's tone. The coefficients on number of committees are significant in all the columns, being positive in Column (1), (2) and (4) while being negative in Column (3), which implies that board with more committees have more reports with a positive tone. An explanation is that boards with more committees work more efficiently, so they have more information and confidence in the firm, so the IDR's tone trends to be more positive. Another reason could be that the establishment of committees is not compulsory, so firms with more committees may have a higher quality of corporate governance, and correspondingly directors express more positive tone. Board activities, measured using meeting frequency, is positive associated with the tone (and the likelihood to obtain a positive report). However, the it is not related to the probability of receiving a report with a negative tone. Ji *et al*. (2016) shows that using meeting frequency cannot precisely capture board monitoring efforts, Chen, Firth, Gao, and Rui (2006) also find that meeting frequency tends to have a positive relationship with frauds of firms and managers. We do not claim a causality relationship between the tone and meeting frequency using our model, because these two variables are simultaneously determinated by corporate governance factors. Moreover, SOEs tend to have reports with less positive tone. The leverage ratio has a negative impact on the IDR's tone, consistent with our hypothesis that firms with higher leverage ratio have higher bankrupt risk, so independent director will be more cautious to monitor firm operation and express less positive tone.

5. The IDR's Tone and the Need to Monitor

Is the positive association between the tone of reports and firm performance stronger for firms with higher risk and greater monitoring needs? If the answer to this question is "Yes", it not only delivers more evidence on the relationship between information content of director's disclosures and firm performance, but this would also provide important evidence that independent directors are not "passive" in their roles - in fact they proactive address the issues in the firm. We hypothesis that the negative tone for poorly performing or high-risk firms contains "information" that is significantly related to firm performance. To investigate this question, we divide firms into two groups depending on the level of monitoring needs. We expect that firms with higher liquidity risk (measured by Cash/short-term debt), greater bankruptcy risk (measured by total debt/ total assets) and poor performance are more likely to attract directors' attention than other firms, so the negative tone in such firm reports is more strongly linked with future performance.

Table 6 reports regression results for each sub-sample. All regressions include the sets of control variables introduced in Section 4.1 with firm and year fixed effects. Panel A of Table 6 shows the results for 'poor', 'median' and 'good' performance firms in Column (1), (2) and (3) respectively. For each firm-year, we first calculate the difference between ROA and the average industry return. Then we rank firm performance by the sum of this difference over the number operating periods. The poorly performing firms are defined as the lower third, medium performing firms are the middle third, and good performing firms comprise the higher third. Panel B and C document results for firms with low and high liquidity - Column (1) and (2), and firms with low and high bankruptcy risk – Column (3) and Column (4). The industry average for liquidity and leverage ratios are used to divide firms into related sub-samples. The regression coefficients on the negative and positive tone are also reported.

Insert Table 6 about here.

The results suggest that the negative relationship between negative tone and future firm performance is stronger for those firms that are likely to have greater monitoring needs from independent directors. These are the companies with poor performance, above average leverage and above average liquidity risk. However, the coefficients on the negative tone are insignificant in the sub-samples of companies with low monitoring needs. Interestingly the coefficients on positive tone is not significant in all sub-samples except for good performing firms. This is because director reports are required to disclose the proposals and their opinions, in line with their responsibility of monitoring.

Overall, the results in Table 6 provide strong evidence that directors recognize the issues in listed firms, and use their reports to express their concerns.

6. Robustness check

Some empirical literature on board diversity (e.g. Carter *et al*., 2003) suggests that the diversity of directors could enhance firm value by offering more monitoring and resources. Hence, we examine whether the diversity of independent directors affects the tone of IDRs by including age, remuneration, education and directorship diversity in the determinates of the IDR's tone (Model 2) [11]. The main results are similar, and the coefficients on these variables are not significant.

Although we do not claim the causality effect between the tone of director report can improve the future firm performance, endogeneity issues are still a big concern. We use the IV-

[11] Diversity is measured by standard method, the ratio of the standard deviation to the mean.

GMM method for further identification (Baum, Schaffer, and Stillman, 2007; Wooldridge, 2002). Following the literature about Chinese boards (e.g. Liu, Uchida, and Yang, 2012), industry level of corporate governance factors could affect firm corporate governance, but this would not directly affect firm performance. Thus average industrial board independence (one year lagged), the IDR's tone (one year and two year lagged value), and independent director's age (one year lagged) are used as instruments. Table 7 reports the regression results with the IV-GMM estimator. The regressions include the set of control variables introduced in Section 4.1, errors are firm and year clustered. The *p*-value of the Hansen J statistic for the over-identification test is 24.02% (larger than 10%), so we cannot reject the null the joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the equation. The *p*-value of the LM statistic for under-identifying restrictions is larger than 10.5%, implying that the instruments are correlated with the tone. The C-D statistic is 224.052, which is larger than the 5% critical value of the Stock and Yogo test statistic, which indicates that our instruments are not weakly identified. Therefore, our IV-GMM estimator is valid and the relationship between the tone and future performance is unchanged.

Insert Table 7 about here.

---

7. Conclusion

This study analyzes the information content of independent director reports. These reports disclose independent director's evaluation of important business decisions based on firm funda-

mentals. Whether these reports are informative is an open question. Although the content (narratives in the reports) is based on independent directors' overall judgement of the firms current and future operations, these reports might not be informative. Independent directors might be reluctant to reveal their disagreement or linger concerns about the firm.

We employ Chinese word segmentation and a Naïve Bayesian machine learning algorithm to measure the tone of independent director reports from a dataset of more than 24, 000 records from 2004 to 2012. We find that the tone of IDRs is positively related to future firm performance. It has explanatory power incremental to other variables. The IDR's tone is associated with independent director's age, work expertise, the number of board committees, the board meeting frequency, state owned ownership and the leverage ratio. We also find that the negative tone of IDRs is negatively associated with firm performance for firms with greater monitoring needs. Our findings remain significant following a variety of robustness test.

Our results suggest that boards of directors have private information about firm profitability. The directors have the capability and willingness to disclose their opinions if they have a proper channel. Director disclosures could be used by outside stakeholders to monitor firm operations. These disclosures could also be a mechanism for inspecting boards of directors' diligence in their work. We provide new evidence to support the regulations of launching independent director system in China coupled with the disclosure requirement of independent director about firm's proposals. Previous literature only focuses on dissenting opinions in board disclosures (Jiang *et al*., 2016; Ma and Khanna, 2016; Tang *et al*., 2013).

Researchers increasingly focus on board behavior and interactions among themselves and with management. While most of these studies rely on interviews and survey data, empirical study with a large sample in response to director's opinions or attitudes is rare[12] due to data

[12] Studies that inspect board actions normally use extreme and clear specifications for board actions, for example,

availability. Our paper is the first to analyze independent director disclosures using a statistical learning methodology with a large sample set. Empirical analyses in this studies are joint assessments of the economic hypotheses and machine learning method. Our results show that this statistical learning algorithm can be used to analyze Chinese corporate disclosure, and this approach could be useful for future research.

anti-takeovers (e.g. McWilliams and Sen, 1997) and CEO dismissals (e.g. Weisbach, 1988). Schwartz-Ziv and Weisbach (2013) observe board meetings of 11 Israeli firms and conduct analysis using the private data on meeting minutes over period 2007-2009.

Reference

Adams, R., Almeida, H., & Ferreira, D. 2009. Understanding the relationship between founder-CEOs and firm performance. *Journal of Empirical Finance*, 16(1): 136–150.

Adams, R. B., & Ferreira, D. 2007. A theory of friendly boards. *Journal of Finance*, 62(1): 217–250.

Adams, R. B., & Ferreira, D. 2009. Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, 94(2): 291–309.

Allen, F., Qian, J., & Qian, M. 2005. Law, finance, and economic growth in China. *Journal of Financial Economics*, 77(1): 57–116.

Antweiler, W., & Frank, M. Z. 2004. Is all that talk just noise? The information content of Internet stock message boards. *Journal of Finance*, 59(3): 1259–1294.

Arslan-Ayaydin, Ö., Boudt, K., & Thewissen, J. 2016. Managers set the tone: Equity incentives and the tone of earnings press releases. *Journal of Banking and Finance*, 72: S132–S147.

Badolato, P., Ege, M., & Donelson, D. C. 2013. Audit committee financial expertise an dearnings management: the role of status. *Journal of Accounting and Economics*, 58(2–3): 71.

Baum, C. F., Schaffer, M. E., & Stillman, S. 2007. Enhanced routines for instrumental variables/GMM estimation and testing. *Stata Journal*, 7(4): 465–506.

Bertrand, M., & Schoar, A. 2003. Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics*, 118(4): 1169–1208.

Buehlmaier, M. M. M., & Zechner, J. 2016. Financial Media, Price Discovery, and Merger Arbitrage. *Working Paper*, (Buehlmaier, Matthias M. M. and Zechner, Josef, Financial Media, Price Discovery, and Merger Arbitrage (September 25, 2017). CFS WP No. 551. Available at SSRN: https://ssrn.com/abstract=2858999).

Carter, D. A., Simkins, B. J., Simpson, W. G., Borokhovich, K., Crutchley, C., Elson, C., Fry, M., Hersch, P., Li, J., & Longhofer, S. 2003. Corporate Governance, Board Diversity, and Firm Value. *The Financial Review*, 38: 33–53.

Chen, G., Firth, M., Gao, D. N., & Rui, O. M. 2006. Ownership structure, corporate governance, and fraud: Evidence from China. *Journal of Corporate Finance*, 12(3): 424–448.

Chen, G., Firth, M., & Xu, L. 2009. Does the type of ownership control matter? Evidence from China's listed companies. *Journal of Banking and Finance*, 33(1): 171–181.

Chen, G., Wan, Y., & Xu, X. 2016. An analysis of the sales and consumer preferences of e-cigarettes based on text mining of online reviews. *2016 3rd International Conference on Systems and Informatics, ICSAI 2016* (pp. 1045–1049).

Clarke, D. C. 2006. The independent director in Chinese corporate governance. *Delaware Journal of Corporate Law*, 31(1): 126–228.

Conyon, M. J., & Peck, S. I. 1998. Board control, remuneration committees, and top management compensation. *Academy of Management Journal*, 41(2): 146–157.

Dahya, J., Dimitrov, O., & McConnell, J. J. 2008. Dominant shareholders, corporate boards, and corporate value: A cross-country analysis. *Journal of Financial Economics*, 87(1): 73–100.

Das, S. R., & Chen, M. Y. 2007. Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*, 53(9): 1375–1388.

Davis, A. K., Ge, W., Matsumoto, D., & Zhang, J. L. 2015. The effect of manager-specific optimism on the tone of earnings conference calls. *Review of Accounting Studies*, 20(2): 639–673.

Davis, A. K., Piger, J. M., & Sedor, L. M. 2012. Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. *Contemporary Accounting*

Research, 29(3): 845–868.
Eisenberg, T., Sundgren, S., & Wells, M. T. 1998. Larger Board Size and Decreasing Firm Value in Small Firms. *Journal of Financial Economics*, 48(1): 35–54.
Fama, E. F. 1980. Agency problems and the theory of the firm. *The Journal of Political Economy*, 88(2): 288–307.
Fama, E. F., & Jensen, M. C. 1983. Separation of ownership and control. *The Journal of Law and Economics*, 26(2): 301–325.
Fich, E. M., & Shivdasani, A. 2007. Financial fraud, director reputation, and shareholder wealth. *Journal of Financial Economics*, 86(2): 306–336.
Healy, P. M., & Palepu, K. G. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*.
Henry, E. 2008. Are Investors Influenced By How Earnings Press Releases Are Written? *Journal of Business Communication*, 45(4): 363–407.
Hermalin, B. E., & Weisbach, M. S. 1998. Boards of directors as an endogenously determined institution: A survey of the economic literature. *American Economic Review*, 88(1): 96–118.
Huang, A. H., Zang, A. Y., & Zheng, R. 2014a. Evidence on the information content of text in analyst reports. *Accounting Review*, 89(6): 2151–2180.
Huang, X., Teoh, S. H., & Zhang, Y. 2014b. Tone Management. *The Accounting Review*, 89(3): 1083–1113.
Ji, J., Talavera, O., & Yin, S. 2016. *CEO Dismissal, Compensation and Topics of Board Meetings: The Case of China*.
Jiang, W., Wan, H., & Zhao, S. 2016. Reputation Concerns of Independent Directors: Evidence from Individual Director Voting. *Review of Financial Studies*, 29(3): 655–696.
Kato, T., & Long, C. 2006. CEO turnover, firm performance, and enterprise reform in China: Evidence from micro data. *Journal of Comparative Economics*, 34(4): 796–817.
Klein, A. 2002. Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics*, 33(3): 375–400.
Levit, D., & Malenko, N. 2016. The Labor Market for Directors, Reputational Concerns and Externalities in Corporate Governance. *The Journal of Finance*, LXXI(2): 1–52.
Lewis, D. D., N'edellec, C., & Rouveirol, C. 1998. Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval. *Machine Learning: ECML-98*.
Li, F. 2010a. The information content of forward- looking statements in corporate filings-A Na¨ıve bayesian machine learning approach. *Journal of Accounting Research*, 48(5): 1049–1102.
Li, F. 2010b. Textual Analysis of Corporate Disclosures : A Survey of the Literature. *Journal of Accounting Literature*.
Li, F. 2010c. The Information Content of Forward-Looking Statements in Corporate Filings - A Naïve Bayesian Machine Learning Approach. *Journal of Accounting Research*, 48(5): 1049–1102.
Li, F., Lundholm, R., & Michael, M. 2013. A measure of competition based on 10-k filings. *Journal of Accounting Research*, 51(2): 399–436.
Liebman, B. L., & Milhaupt, C. J. 2008. Reputational sanctions in China's securities market. *Columbia Law Review*.
Liu, C., Uchida, K., & Yang, Y. 2012. Corporate governance and firm value during the global financial crisis: Evidence from China. *International Review of Financial Analysis*, 21: 70–80.
Liu, Y., Miletkov, M., Wei, Z., & Yang, T. 2015. Board independence and firm performance

in China. *Journal of Corporate Finance*, 30: 223–244.
Loughran, T., & Mcdonald, B. 2011. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66(1): 35–65.
Loughran, T., & Mcdonald, B. 2016. Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54(4): 1187–1230.
Ma, J., & Khanna, T. 2016. Independent directors' dissent on boards: Evidence from listed companies in China. *Strategic Management Journal*, 37(8): 1547–1557.
Ma, S., Naughton, T., & Tian, G. 2010. Ownership and ownership concentration: Which is important in determining the performance of China's listed firms? *Accounting and Finance*, 50(4): 871–897.
Masulis, R. W., & Mobbs, S. 2014. Independent director incentives: Where do talented directors spend their limited time and energy?? *Journal of Financial Economics*, 111(2): 406–429.
McWilliams, V. B., & Sen, N. 1997. Board Monitoring and Antitakeover Amendments. *Journal of Financial and Quantitative Analysis*, 32(4): 491–505.
Ouyang, S., Li, C., & Li, X. 2016. A peek into the future: Predicting the popularity of online videos. *IEEE Access*, 4: 3026–3033.
Peng, W. Q., Wei, K. C. J., & Yang, Z. 2011. Tunneling or propping: Evidence from connected transactions in China. *Journal of Corporate Finance*, 17(2): 306–325.
Purda, L., & Skillicorn, D. 2015. Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection. *Contemporary Accounting Research*, 32(3): 1193–1223.
Ryan, H. E., & Wiggins, R. A. 2004. Who is in whose pocket? Director compensation, board independence, and barriers to effective monitoring. *Journal of Financial Economics*, 73(3): 497–524.
Schwartz-Ziv, M., & Weisbach, M. S. 2013. What do boards really do? Evidence from minutes of board meetings. *Journal of Financial Economics*, 108(2): 349–366.
Segars, A. H., & Kohut, G. F. 2001. Strategic communication through the World Wide Web: An empirical model of effectiveness in the CEO's letter to shareholders. *Journal of Management Studies*, 38(4): 535–556.
Shivdasani, A., & Yermack, D. 1999. CEO Involvement in the Selection of New Board Members: An Empirical Analysis. *The Journal of Finance*, 54(5): 1829–1853.
Tang, X., Du, J., & Hou, Q. 2013. The effectiveness of the mandatory disclosure of independent directors' opinions: Empirical evidence from China. *Journal of Accounting and Public Policy*, 32(3): 89–125.
Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62(3): 1139–1168.
Tirole, J. 2001. Corporate governance. *Econometrica*, 69(1): 1–35.
Tirole, J. 2010. *The theory of corporate finance*. Princeton University Press.
Vafeas, N. 1999. Board meeting frequency and firm performance. *Journal of Financial Economics*, 53(1): 113–142.
Wang, K., Zong, C., & Su, K.-Y. 2012. Integrating Generative and Discriminative Character-Based Models for Chinese Word Segmentation. *ACM Transactions on Asian Language Information Processing*, 11(2): 1–41.
Warther, V. a. 1998. Board effectiveness and board dissent: A model of the board's relationship to management and shareholders. *Journal of Corporate Finance*, 4: 53–70.
Weisbach, M. S. 1988. Outside directors and CEO turnover. *Journal of Financial Economics*, 20: 431–460.
Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. *MIT Press*.

MIT Press.
Yermack, D. 1996. Higher market valuation of companies with a small board of directors. *Journal of Financial Economics*, 40(2): 185–211.
Yermack, D. 2004. Remuneration , Retention , and Reputation. *The Journal of Finance*, LIX(5): 2281–2308.
Zajac, E. J., & Westphal, J. D. 1996. Director Reputation, CEO-Board Power, and the Dynamics of Board Interlocks. *Administrative Science Quarterly*, 41(3): 507–529.

Table 1
Variable definitions and descriptive statistics

Panel A The Tone of independent director' reports

	Definition	Obs.	Mean	Median	S.D.	p5	p95
Tone_each	The tone of report[s], a continuous variable, ranged from 0 to 1 from a negative view to a positive view; 0.5 means neutral	23984	0.458	0.475	0.447	0	1
Tone_year	The average of tones of reports in a firm	11249	0.463	0.500	0.303	0	1
Positive tone	Indicator variable takes one when tone_year is larger than 0.7	11249	0.201	0	0.401	0	1
Negative tone	Dummy variable, equals to one when tone_year is smaller than 0.3	11249	0.276	0	0.447	0	1

Panel B Independent directors' characteristics

	Definition	No. of firms	Mean	Median	S.D.	p5	p95
Age	The average age of independent directors in a firm	1437	51.755	51.333	6.258	42.000	63.000
Age_outside	Log value of Age	1437	3.939	3.938	0.121	3.738	4.143
Pay (1000 RMB)	The average monetary compensation of independent directors in a firm	1437	48	45	28.32	15	100
Pay_outside	Log value of Pay	1437	10.664	10.714	0.548	9.798	11.513
% female_outside	Ratio of women in independent directors	1437	0.138	0	0.199	0	0.5
Directorships_outside	The number of directorships	1437	1.893	1.714	0.786	1	3
Education_outside	The education level:5- Ph.D., 4-Master, 3-Bachelor, 2-College and 1-High School and lower	1437	4.025	4	0.656	3	5
Financial expert	Dummy variable, equals to 1 when a firm has at least one financial expert	1437	0.283	0	0.451	0	1
Law expert	Indicator variable, takes 1 when a firm has at least one law expert	1437	0.360	0	0.480	0	1

Panel C Board features

	Definition	No. of firms	Mean	Median	S.D.	p5	p95
Age_board	The average age of director	1437	49.166	49.222	3.885	42.800	55.444

Table 1
continued

# female	The number of female directors	1437	1.037	1	1.069	0	3
Board size	The number of directors	1437	9.350	9	1.914	7	13
% independence	The share of independent director on board	1437	0.355	0.333	0.046	0.308	0.444
Number of committees	The number of committees	1437	3.349	4	1.369	0	4
Meeting frequency	The number of board meetings	1437	8.737	8	3.718	4	16

Panel D Firm performance and characteristics

		No. of firms	Mean	Median	S.D.	p5	p95
Log (total assets)	Log value of total assets	1437	21.641	21.536	1.117	20.024	23.694
SOE	Dummy variable equals to one when firm is controlled by state agent or state enterprises	1437	0.72	1	0.449	0	1
ROA (%)	The ratio of net income to total assets	1437	2.66	2.81	5.8	-7.5	10.72
Leverage	Debt over total assets	1437	0.524	0.527	0.236	0.185	0.819
Largest Share	The percentage of share held by largest shareholder	1437	37.744	35.897	15.03	16.148	63.74
Largest1to2	The ratio of share owned by largest over second largest	1437	24.820	6.220	62.220	1.130	101
Tobin's Q	Firm market value divided by firm value	1437	1.487	1.122	1.221	0.383	3.854

This table shows all the descriptive statics and definitions of main variables for the sample of non-financial main board listed firm in Shanghai and Shenzhen Stock exchange over the period 2004 to 2012.

Table 2

Regression results of firm performance on the tone of independent director reports

	(1)	(2)	(3)
Tone_year $_{t-1}$	0.574***	1.602***	
	(0.198)	(0.379)	
Tone_year $_{t-1}$ × SOE $_{t-1}$		-1.446***	
		(0.402)	
Age_board $_{t-1}$	0.027	0.029	0.028
	(0.039)	(0.039)	(0.039)
Female_board $_{t-1}$	0.257***	0.260***	0.263***
	(0.099)	(0.098)	(0.099)
Financial expert $_{t-1}$	0.085	0.093	0.085
	(0.208)	(0.208)	(0.208)
Law expert $_{t-1}$	-0.300	-0.287	-0.298
	(0.210)	(0.209)	(0.209)
Board size $_{t-1}$	-0.040	-0.042	-0.038
	(0.080)	(0.080)	(0.080)
% Independence $_{t-1}$	5.594**	5.626**	5.652**
	(2.269)	(2.268)	(2.268)
Number of committees $_{t-1}$	-0.015	-0.020	-0.017
	(0.075)	(0.075)	(0.075)
Meeting frequency $_{t-1}$	0.029	0.028	0.031
	(0.024)	(0.023)	(0.024)
Log (total assets) $_{t-1}$	-1.970***	-1.960***	-1.969***
	(0.226)	(0.225)	(0.226)
SOE $_{t-1}$	-0.938*	-0.522	-0.948*
	(0.539)	(0.512)	(0.542)
Largest Share $_{t-1}$	0.106***	0.106***	0.106***
	(0.014)	(0.014)	(0.014)
Leverage $_{t-1}$	-0.975	-0.911	-0.997
	(0.914)	(0.911)	(0.917)
Firm and Year Dummies	Yes	Yes	Yes
Adjusted R-squared	9.1%	9.3%	9.0%
Observations	8866	8866	8866

This table reports the results of panel data regressions with firm and year fixed effects from 2004 to 2014 to test whether the tone of IDRs could predict future performance (Model 2). The dependent variable is firm's probability - ROA. Column (1) and (2) of Table includes the IDR's tone and all the control variables, and Column (3) only involves control variables.
* p<0.10, ** p<0.05, *** p<0.01

Table 3

Regression results of firm performance on the positive and negative tone

	(1)	(2)
Positive tone $_{t-1}$	0.291**	
	(0.122)	
Negative tone $_{t-1}$		-0.334**
		(0.138)
Age_board $_{t-1}$	0.028	0.028
	(0.039)	(0.039)
Female_board $_{t-1}$	0.263***	0.256***
	(0.099)	(0.099)
Financial expert $_{t-1}$	0.084	0.087
	(0.208)	(0.208)
Law expert $_{t-1}$	-0.298	-0.299
	(0.210)	(0.210)
Board size $_{t-1}$	-0.039	-0.041
	(0.080)	(0.080)
% Independence $_{t-1}$	5.616**	5.600**
	(2.266)	(2.269)
Number of committees $_{t-1}$	-0.016	-0.016
	(0.075)	(0.075)
Meeting frequency $_{t-1}$	0.029	0.029
	(0.024)	(0.024)
Log (total assets) $_{t-1}$	-1.972***	-1.964***
	(0.226)	(0.226)
SOE $_{t-1}$	-0.946*	-0.940*
	(0.541)	(0.540)
Largest Share $_{t-1}$	0.106***	0.107***
	(0.014)	(0.014)
Leverage $_{t-1}$	-0.995	-0.988
	(0.916)	(0.914)
Firm and Year dummies	Yes	Yes
Adjusted R-squared	9.0%	9.1%
Observations	8866	8866

Table 4 documents the regression results of firm performance on positive_tone (Column - 1) and negative tone (Column - 2) respectively. Other variables are the same with ones used in Table 3. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 4

The IDR's Tone and over one year firm performance

	(1)	(2)
$Positive\ tone_{t-1}$	0.278**	
	(0.139)	
$Positive\ tone_{t-2}$	0.033	
	(0.150)	
$Negative\ tone_{t-1}$		-0.282**
		(0.140)
$Negative\ tone_{t-2}$		-0.067
		(0.122)
Control variables	Included	Included
Firm and year dummies	Yes	Yes
Adjusted R-squared	9.0%	9.0%
N	7645	7645

Table 5 reports regression results including one year lagged and two year lagged positive tone (Column - 1) and negative tone (Column – 2) of IDRs. They are regressed on the determinants of the tone and year- firm fixed effects.
* p<0.10, ** p<0.05, *** p<0.01

Table 5 Determinants of tone of independent director reports: director and firm level regressions

	(1)Tone_each	(2) Tone_each	(3) negative	(4) positive
Age_outside	0.080**		-0.382**	0.405**
	(0.033)		(0.166)	(0.171)
Pay_outside	0.007		-0.042	0.009
	(0.008)		(0.038)	(0.038)
%Female_outside	-0.009		0.063	-0.056
	(0.018)		(0.090)	(0.092)
Directorships_outside	0.001		-0.003	0.003
	(0.002)		(0.011)	(0.012)
Education_outside	-0.001		0.010	-0.022
	(0.005)		(0.027)	(0.028)
Financial expert	-0.018**		0.103**	-0.049
	(0.008)		(0.042)	(0.042)
Law expert	-0.001		-0.011	-0.005
	(0.009)		(0.044)	(0.045)
Board size	0.003	0.002	-0.015	0.009
	(0.002)	(0.002)	(0.011)	(0.011)
% Independence	0.114	0.084	-0.713*	0.474
	(0.076)	(0.073)	(0.377)	(0.383)
Number of committees	0.009***	0.008***	-0.047***	0.048***
	(0.003)	(0.003)	(0.016)	(0.017)
Meeting frequency	0.002**	0.002**	-0.001	0.015***
	(0.001)	(0.001)	(0.005)	(0.005)
Log (total assets)	0.001	-0.000	0.005	0.009
	(0.005)	(0.004)	(0.023)	(0.024)
Account receivables	0.050	0.064	-0.236	0.266
	(0.043)	(0.041)	(0.212)	(0.218)
ROA	0.113	0.098	-0.329	0.722*
	(0.074)	(0.070)	(0.367)	(0.375)
Tobin's Q	0.003	0.002	-0.014	0.010
	(0.003)	(0.003)	(0.016)	(0.016)
SOE	-0.025***	-0.028***	0.089*	-0.172***
	(0.009)	(0.009)	(0.046)	(0.046)
Largest Share	-0.000	-0.000	0.001	-0.000
	(0.000)	(0.000)	(0.001)	(0.001)
Largest1to2	-0.000	-0.000*	0.001	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)
Leverage	-0.091***	-0.079***	0.397***	-0.375***
	(0.024)	(0.023)	(0.117)	(0.119)
Firm, industry and year dummies	Yes	Yes	Yes	Yes
Adjusted R-squared	5.42%	5.26%		
(Pseudo) R-squared			3.93%	3.65%
Observations	22244	23855	22238	22241

The table reports regression results with random effects based on model (1). The dependent variable is $Tone\_each_{i,k,t}$ in Column (1) and (2), the sentiment for report k in firm i at year t, regressed on its hypothesized determinants discussed in 2.3.1. The dependent variable is *negative_tone* in Column (3), whereas *positive_tone* in Column (4). The results in Column (3) and (4) are coefficients using logit estimator. Column (1), (3) and (4) of Table 2 include all explanatory variables, and Column (2) contains board controls and firm controls.* p<0.10, ** p<0.05, *** p<0.01

Table 6 The IDR's Tone and The Necessities for Monitoring

Panel A

The predicting power of the IDR's Tone in firms with Poor, Middle or Good Performance

	(1) Poor performance ROA < Lower one third	(2) Middle Lower 1/3 <ROA < Higher 1/3	(3) Good performance ROA > Higher one third
Negative tone $_{t-1}$	-0.661**		0.008
	(0.263)		(0.160)
Positive tone $_{t-1}$	-0.087		0.466***
	(0.343)		(0.173)
Tone_year $_{t-1}$		0.549***	
		(0.211)	
Control variables	Included	Included	Included
Firm and year dummies	Yes	Yes	Yes
Adjusted R-squared	11.7%	9.6%	11.5%
N	2843	3004	3019

Panel B

The predicting power of the IDR's Tone in firms with liquidity and bankruptcy risk levels

	ROA			
	(1)	(2)	(3)	(4)
	Liquidity > Mean	Liquidity < Mean	Leverage < Mean	Leverage > Mean
Negative tone $_{t-1}$	-0.234	-0.452**	-0.214	-0.372**
	(0.161)	(0.212)	(0.187)	(0.180)
Positive tone $_{t-1}$	0.076	-0.023	0.219	-0.106
	(0.154)	(0.233)	(0.184)	(0.193)
Control variables	Includes	Includes	Includes	Includes
Firm and Year Dummies	Yes	Yes	Yes	Yes
R-squared	6.5%	10.3%	5.8%	10.1%
N	4330	4367	4209	4488

Table 6 reports regression results for each sub-sample. All regressions include the sets of control variables introduced in Section 4.2 with firm and year fixed effects. Panel A of Table 6 shows results for poor, median and good performance firms, respectively in Column (1), (2) and (3). For each firm-year, we firstly calculate the difference between ROA and the average industry return. Then we rank firm performance by the sum of the difference over the number operating periods. The poor performed firms are defined as the lower third, median performed ones are the middle third, good performed firms are the higher third. Panel B and C document results for firms with low and high liquidity - Column (1) and (2), and firms with low and high bankruptcy risk – Column (3) and Column (4). The industry average of liquidity and leverage ratio are used to divide firms into related sub-samples.

* p<0.10, ** p<0.05, *** p<0.01

Table 7

Robustness check for the relationship between firm performance and tone of reports

	Firm performance (ROA)
Tone_year $_{t-1}$	3.203*
	(1.894)
Age_board $_{t-1}$	0.004
	(0.023)
Female_board $_{t-1}$	-0.009
	(0.072)
Financial expert $_{t-1}$	0.050
	(0.169)
Law expert $_{t-1}$	0.121
	(0.199)
Board size $_{t-1}$	-0.019
	(0.038)
% Independence $_{t-1}$	-3.394
	(2.171)
Number of committees $_{t-1}$	0.051
	(0.058)
Meeting frequency $_{t-1}$	-0.027
	(0.019)
Log (total assets) $_{t-1}$	0.712***
	(0.166)
SOE $_{t-1}$	-0.579***
	(0.176)
Largest Share $_{t-1}$	0.019**
	(0.008)
Leverage $_{t-1}$	-6.176***
	(0.631)
N	7644
P-value of LM statistic	0.105
Cragg-Donald F statistic	224.052
P-value of Hansen J statistic	0.223

Table 7 reports the results of regressing of firm future performance on the tone with IV-GMM estimator. industrial average board independence (one year lagged), the IDR's tone (one year and two lagged value), and independent director's age (one year lagged) are used as instruments. The regressions include the sets of control variables introduced in Section 4.2, errors are firm and year clustered.