Local media sentiment towards pollution and its effect on corporate green innovation

Local Media Sentiment towards Pollution and its Effect on Corporate Green Innovation*

Yu He
School of Economics and Management, China University of Geosciences (Beijing), Beijing, 100083, China
Shanglin Lu†
China School of Banking and Finance, University of International Business and Economics, Beijing, 100029, China
Ran Wei
Research Center for Rural Economy, Ministry of Agriculture and Rural Affairs, Beijing, 100810, China
Shixuan Wang
Department of Economics, University of Reading, Reading, RG6 6EL, UK

Abstract

This study examines the effect of local media sentiment on corporate green innovation based on a textual analysis of China's provincial official party newspapers from 2007 to 2018. The results show that the negative sentiment in official media positively influences firms' green innovation, measured by the number of green patents and green patent citations. This positive effect is more pronounced when firms have weaker internal or external governance structures, when the regional punitive measures are less stringent, or after the incentive measures are implemented, suggesting that official media plays a governance role in corporate green innovation. Further analysis shows that the negative sentiment from market-oriented media impedes green innovation and does not affect the relationship between official media sentiment and green innovation. Taken together, our findings reveal the real effects of local media negative sentiment on technological progress and pollution controls through its pressure on firms to engage in green innovation.

Keywords: Media Sentiment, Green Innovation, Environmental Pollution, Corporate Governance
JEL: G34, Q55

* Acknowledgement: We are grateful to the Editor, the anonymous Associate Editor, and four anonymous referees for their useful comments to improve the presentation of our findings. Shanglin Lu is supported by the National Natural Science Foundation of China (Grant No. 72303032) and the Fundamental Research Funds for the Central Universities in UIBE (22QD07). Declaration of competing interest: none.
† Correspondence to: Shanglin Lu, China School of Banking and Finance, University of International Business and Economics, Beijing, 100029, China. E-mail: lushanglin@uibe.edu.cn

1. Introduction

Green innovation encompasses green processes and products that are related to energy-saving, pollution-prevention, waste recycling, green product designs, or corporate environmental management (Chen et al., 2006). Both anecdotal evidence and empirical studies suggest that promoting green innovation is an urgent issue. According to the survey of the Lancet Commission on Pollution and Health, global environmental pollution was responsible for an estimated 9 million premature deaths in 2015 (Das and Horton, 2018), around 15 times more than all wars and other forms of violence.[1] Besides the health effects, the estimated welfare losses due to pollution are more than US $4.6 trillion per year, accounting for 6.2% of global economic output (Landrigan et al., 2017). Existing environmental policies (i.e., environmental regulation) may cause trade-offs between environmental protection and economic development (Porter, 1991; Shen et al., 2017).

Given the sheer magnitude of environmental pollution and the unique manner in which green innovation helps to develop clean technology, an expanding body of literature is exploring the drivers of green innovation (Hojnik and Ruzzier, 2016; Cuerva et al., 2014; Cai and Zhou, 2014; Lin et al., 2014; Demirel and Kesidou, 2011; Bai and Lyu, 2023). However, few studies provide empirical evidence on how the informal regulations affect firms' green innovation. We address this gap by focusing on an important form of informal regulation: media news on environmental pollution. In particular, we investigate whether and how local media sentiment, defined as the provincial official party newspaper tone on environmental pollution in the area where the firms are domiciled, affects firms' decisions to engage in green innovation.

We develop two competing hypotheses on the green innovation effect of negative sentiment based on the literature examining the effect of media on corporate behaviors. Our first hypothesis postulates that negative sentiment in relation to pollution impedes green

[1] The news title is "Global pollution is the world's biggest killer and a threat to survival of mankind, study finds", which was written by Stanglin and published in 2017.

innovation through the information channel. Media news can reduce information asymmetry between firms and the public by collecting, selecting, certifying, and repackaging information (Miller, 2006; Dyck et al., 2008). Meanwhile, the media will express positive or negative opinions which provide important information for lenders or shareholders who are seeking to evaluate firm risk (Luo, 2012; Lu et al., 2014). In particular, firms in regions with positive news have a stronger financing ability because of a lower environmental risk. On the contrary, negative news increases firms' financing constraints, particularly financing for innovation. In other words, lenders or shareholders are generally risk averse and are less likely to benefit from a firm's innovation success (e.g., Stiglitz, 1985; Hall and Lerner, 2010; Chang et al., 2019). Negative news impedes lenders' or shareholders' risk-taking, thus increasing firms' financial constraints. As such, we expect negative sentiment to discourage firms from engaging in green innovation. We label this mechanism the information channel.

By contrast, our second hypothesis argues that negative sentiment on pollution enhances green innovation through the governance channel. Prior studies show that media exerts a corporate governance role and has the ability to expose fraud and other harmful behaviors (e.g., Miller, 2006; Dyck et al., 2008; Dyck et al., 2010). The media governance role functions through managers' reputations and administrative intervention (e.g., Liu and McConnell, 2013; Pollock et al., 2008; Cahan et al., 2015; Li and Shen, 2010; Snyder and Strömberg, 2010; Zhou et al., 2016a, 2016b). Specifically, media news influences managers' reputations and managers will respond to the reports and improve the governance to prevent the negative news from affecting their future job opportunities and salary. Furthermore, the negative news can attract attention from environmental regulators. The administrative intervention will then cause firms to shut down or to reduce production (Pan et al., 2019). In such negative sentiment conditions, firms have to foster green innovation in order to go green. On the contrary, positive news puts very little pressure on firms and managers then decrease their level of green innovation because innovations have a high likelihood of failure and cannot generate immediate gains for shareholders (e.g., Holmstrom, 1989; Hall and Lerner, 2010; Chang et al., 2019). Taken together, we expect a negative sentiment in relation to pollution to foster firms' green innovation by performing an external governance function. We label this mechanism the governance channel.

We choose to investigate this issue in China because China presents an ideal laboratory setting to study the effect of negative sentiment on firms' green innovation. Compared to developed countries (i.e., U.S., U.K.), China is under more pressure to improve its environmental performance. According to the Environmental Performance Index 2018, China is ranked 120th out of 180 countries. China has also placed green innovation at the high level of its national strategy. From the perspective of statistics, we can also see that the quantity of green patents from China's listed firms has grown rapidly. The number of green patents granted has increased from 101 in 2000 to 10,223 in 2017. Moreover, the market features of China's newspapers make it possible to construct a province-level media environment which firms are faced with. China's local official newspapers are owned by the provincial party committee of the CPC and they have a stronger power to affect firm behavior (Djankov et al., 2003; Qin et al., 2018). Compared to the national newspapers (e.g., the Wall Street Journal, The New York Times) used in studies based on U.S. samples (Tetlock, 2007), local newspapers exhibit greater variation.

Using a sample of firms listed on China's two mainland stock exchanges (i.e., Shanghai and Shenzhen) from 2007 to 2018, we empirically investigate the effect of provincial official media negative sentiment on firms' green innovation. Within the different media channels, we restrict the analysis to the role of the local state-controlled newspaper. This is because local media are more likely to discover information from local sources than from national newspapers (Gurun and Butler, 2012) and official newspapers carry greater authority (Qin et al., 2018). Different from other media (e.g., Twitter, Microblog), the printed newspapers' space is limited and valuable, which leads to an increase in news credibility and visibility (Tewksbury and Althaus, 2000; Bucher and Schumacher, 2006; Bajo and Raimondo, 2017). Meanwhile, there exists a strong link between social media and news media (Gan et al., 2020). The provincial newspaper news can convey the firms' external information environment. Following prior studies (e.g., Qi et al., 2018), we measure green innovation using the number of green patents and green patent citations.

Our main results show that a negative sentiment in relation to pollution enhances firms' green innovation. Specifically, we measure the level of negative sentiment in a province-level region during each year using the tone of the news relating to pollution. We document a positive relationship between negative sentiment and firms' green innovation. We conduct a battery of tests to ensure that our main findings are robust to alternative measures and subsamples. These findings are consistent with the governance hypothesis, in that negative sentiment on pollution performs an external governance function which then places greater pressure on firms to foster green innovation.

We perform several tests to mitigate endogeneity concerns arising from omitted variables and reverse causality. Among these tests, we control for industry-by-year fixed effects, province-level variables and firm-level news, which may potentially be correlated with both negative sentiment and firms' green innovation. We follow Chang et al. (2019) and incorporate firms' past innovation investments/outputs into the baseline to explicitly account for reverse causality. Furthermore, we employ the standard deviation of elevation in the province, where the enterprise is located, as the instrumental variable for local media sentiment. Our findings remain the same. Collectively, our tests of endogeneity point to a causal effect of media negative sentiment on firms' green innovation.

Next, we explore the cross-firm heterogeneity of our results to better understand the channel through which negative sentiment affects firms' green innovation. In our governance argument, negative sentiment can affect firms' green innovation by playing the role of external governance. In this case, the effect of negative sentiment on green innovation should be more pronounced for firms with weaker governance mechanisms. Following prior studies (e.g., Chang et al., 2019), we link governance structure to firms' external and internal governance mechanisms and find a stronger negative sentiment-green innovation relationship for firms with a weaker governance structure, suggesting that official media can complement the corporate governance mechanism. Meanwhile, this positive effect is more pronounced when the punitive measures are weak, or after the incentive measures are implemented. Moreover, the financial constraints do not affect the negative sentiment-green innovation relationship, running counter to the information channel. This is because, under the information channel, firms subject to stronger financing constraints should reduce the level of green innovation that reacts to the negative sentiment.

Finally, we explore how the market-oriented media sentiment affects the official media sentiment-green innovation relationship. The results show market-oriented media negative sentiment has a negative impact on green innovation, both in terms of green innovation quantity and quality. This means that the market-oriented media mainly plays the role of information, and a negative sentiment will increase corporate external restrictions and reduce corporate green innovation. Additionally, the inclusion of market-oriented media sentiment does not affect the positive relationship between official media negative sentiment and firm green innovation. Moreover, we examine the real effect of official media negative sentiment on different types of green innovation. Specifically, we document that official media negative sentiment has significant effects on green inventory patents, green utility patents and green patent claims. These results further indicate that official media negative sentiment can impact both firms' green innovation quantity and innovation quality.

By identifying a causal effect of the negative sentiment on firms' green innovation, our study contributes to the existing literature in at least three important ways. First, our paper reveals the effects of China's official media news on firms' green innovation, adding a new way of national ecological governance besides environmental regulation. The existing literature is limited to exploring the role of environmental regulation in promoting firms' environmental responsibility, ignoring the influence of informal institutions. We extend this literature by showing that the official party newspaper can be used as an informal supervision to promote enterprises' green innovation.

Second, our study also contributes to the ongoing debate on the effects of the media. Existing studies focus on the economic consequences of financial media or social network media reports on corporate decision-making (e.g., Miller, 2006; Chahine et al., 2015; Liu and McConnell, 2013), ignoring the possible influence of party newspapers. Provincial party newspapers have clear political and public opinion guidance, outstanding performance in information content, credibility, and supervision roles. This study not only analyzes the influence of party newspaper media on enterprise green innovation but also makes a comparative study of the different influences of market media.

Third, our paper combines the Latent Dirichlet Allocation model (LDA) and the dictionary method to accurately extract regional news on the topic of corporate environmental pollution from more than 1.4 million provincial party newspaper news. This combination greatly reduces the cost and subjectivity of information extraction and provides a new idea for accurate and rapid extraction of text information.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data, variables, and summary statistics. Section 4 presents the main empirical results regarding the effect of negative sentiment on firms' green innovation. Section 5 reports the additional analysis. Section 6 concludes this paper.

2. A Brief Review of the Relevant Literature

Our paper contributes to three strands of literature: recent studies that focus on the effects of media sentiment on corporate policies, research on the determinants of corporate green innovation, and research on the application of textual analysis in finance.

Media is not a neutral mouthpiece, and it can express positive or negative opinions. (You and Wu, 2012). Prior research has revealed that media sentiment is informative and has significant economic consequences (Maghyereh and Abdoh, 2022; Polyzos, 2022). Kalamara et al. (2020) show that sentiment from newspaper text can improve forecasts of macroeconomic variables. A stream of literature has shown that media sentiment produces an impact on market price and trading volume (Tetlock, 2007), stock performance (Fang et al., 2009; Feng et al., 2022), and IPO underpricing (Bajo and Raimondo, 2017). Rognone et al. (2020) uncover a positive Bitcoin reaction to both positive and negative news. Sapkota (2022) highlights that news sentiment has significant impact on Bitcoin volatility. Yu et al. (2023) find that ESG news sentiment affects stock crash risk through reducing ESG incidents, information asymmetry, and agency costs. Except for the asset pricing perspective, recent studies have focused on corporate behaviors. Thus far, multiple views exist on the role of the media in influencing corporate policies. On the one hand, media news can reduce information asymmetry and can help others to assess firm risk (Miller, 2006). Specifically, negative news reported by the media can detect corporate fraud (Dyck et al., 2010), and can expose food safety accidents (Zhou et al., 2016a) and pollution accidents (Jia et al., 2016; Zeng et al., 2018; Pan et al., 2019). Meanwhile, the media are likely to report favorable news in relation to firms with better CSR performances (Cahan et al., 2015). Kothari et al. (2009) show that a positive corporate image portrayed by the media can decrease the cost of capital. Thus, the media has an information function and positive news can provide firms with economic benefits. In the context of this paper, positive (negative) news will decrease (increase) firms' external restrictions, such as financial constraints, which are a key determinant for green innovation.

On the other hand, Dyck et al. (2008) and Zhou et al. (2016b) argue that the media performs a governance role through managers' reputations and administrative intervention. As a result, negative news can urge firms to engage in innovation to prevent a loss of reputation or to avoid regulatory intervention. In line with this governance view, Liu and McConnell (2013) find that a manager will abandon value-reducing acquisition attempts when the acquisition in question receives a higher level of media attention with a negative tone. The results imply that the media plays a role in guiding managers' decisions and has a positive role in governance through managers' reputational capital. Li and Shen (2010) argue that, in addition to the reputation mechanism, administrative intervention is even more effective in China's case because firm development is strongly dependent upon the government. In order to avoid attracting the attention of regulators, firms reported on through negative news will correct irregularities and reduce the likelihood of a repeat offense (Zhou et al., 2016b; Yang and Zhao, 2012). Taken together, the evidence on how media sentiment affects corporate policies is mixed. Our study adds to this debate by documenting that the media plays more of a governance role, than an information role. In turn, negative sentiment can foster firms' green innovation.

Extant studies have identified various factors affecting firms' green innovation. Several studies consider developed economies when examining the drivers of green innovation. For example, Demirel and Kesidou (2011) show that environmental regulations and cost saving are the main drivers of green innovation using a dataset of 289 UK firms. Horbach et al. (2012) find similar results using German firms. Although their findings provide useful insights, these findings may not be relevant for developing countries. Liao et al. (2018) use listed firms in China to show that, among the external drivers, it is only regulation and fiscal and taxation measures which are the main driving factors for environmental innovation. Cai and Zhou (2014) reveal that external pressures (i.e., environmental regulations, customers' green demands, competitors) affect green innovation through internal drivers, including technological and organization capabilities. Bai and Lyu (2023) show that executive's environmental protection background positively affects corporate green innovation, and media attention enhances this positive relation. This study has not yet examined the direct impact of media news on green innovation. To our knowledge, we are the first to show that media sentiment is an important driver for firms' green innovation.

With the development of natural language processing technology, researchers are not limited to standardized information such as numbers but use text analysis technology to extract non-standardized information such as text. Textual data contains higher dimensional information, because each word in the text can represent a dimension of information (Gentzkow et al., 2019). Further, textual data can portray subjective indicators. By counting the positive and negative words in texts, it is possible to measure media sentiment (Tetlock, 2007) and management sentiment (Jiang et al., 2019). However, it has proven difficult to extract the textual topic due to its high-dimensional character. Some studies use dictionary methods to extract keywords to capture topics, such as economic policy uncertainty (Baker et al., 2016), political risk (Hassan et al., 2019), and climate change exposure (Sautner et al., 2023). The topics identified by dictionary method are more intuitive, but the result relies on the choice of dictionary. Latent Dirichlet Allocation model (LDA) developed by Blei et al. (2003) classifies topics based on semantic analysis. A benefit of this method is that its classification process is more objective, because it uses Bayesian algorithm for continuous learning. Our paper firstly uses dictionary method to screen the news, and then employs LDA model to conduct more detailed classification. The combination of the two methods not only improves the speed of text processing, but also overcomes the subjectivity and incompleteness of the dictionary.

Collectively, unlike prior studies that have mainly focused on the investor reaction or accounting effects of media sentiment, we are among the first to show that media sentiment in relation to pollution can have real effects on a firm's green innovation output. By doing so, we add to the literature that examines the real effects of the media (e.g., Dyck et al., 2008; Liu and McConnell, 2013; Li and Shen, 2010) by revealing an important micro-level channel (i.e., corporate green innovation) through which the media could affect pollution-controlling and economic growth. Meanwhile, we also provide a new determinant for green innovation.

3. Variables and Data

3.1. Sample and data selection

Our initial sample consists of all A-share firms listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange in China. Our sample starts from 2007 and ends in 2018, determined by the availability of media text data and control variables[2]. We obtained news articles published by China's provincial official party newspapers from the full-text database of Chinese newspapers, and we have engaged in a (semi-)manual process to standardize various formats of text, reconcile them together, and carefully scrutinize the cleaned data.[3] To measure the firms' green innovation outputs, we rely on the World Intellectual Property Organization's (WIPO) IPC Green Inventory to determine whether the patent is a green innovation output.[4] Green patents and citations are obtained from the Incopat Database and China National Intellectual Property Administration.[5] As suggested by Hall et al. (2001) and Chang et al. (2019), the patent application year is closer to the time of invention than the granted date. We use the application year of the granted green patent to merge with other firm variables. The data

[2] Our data sample ends in 2018 due to the availability of the full-text dataset of Chinese newspapers purchased at the time of this research.

[3] It is worth noting that we end up with 1.4 million provincial Party newspaper news that used in this research.

[4] IPC Green Inventory consists of seven main categories, including Alternative Energy Production, Transportation, Energy Conservation, Waste Management, Agriculture/Forestry, Administration Regulatory or Design Aspects and Nuclear Power Generation. See more details at https://www.wipo.int/classifications/ipc/en/green_inventory/ (January 6, 2023).

[5] The Incopat Database is operated by BEIJING INCOPAT CO., LTD., which is a global IP data service provider. The database collects the IPC and other basic information of each patent of all domestic and foreign companies. We refer to the website https://www.incopat.com (January 6, 2023) for more information about the database.

on firm locations are taken from the RESSET database. Financial information is obtained from the China Stock Market & Accounting Research (CSMAR) database.

In line with common practice (e.g., Chang et al., 2015), we exclude firms in the financial industry. We further exclude observations with missing values for the variables used in the empirical analyses. In order to mitigate the effect of outliers, we winsorize all variables (except the dummy) at the top and bottom 1%. These restrictions result in a final sample that consists of 18,708 firm-year observations from 3,083 firms for 2007-2018.

Appendix A shows the sample distribution by year and by industry. Panel A of Table A.1 shows that the number of observations in our sample increases from 317 in 2007 to 2,720 in 2018. Panel B reports the sample distribution by the one-digit code of the CSRC 2012 industry classification. We find that our sample observations are mainly from the manufacturing industry (74.05%) and information, software, and information technology (6.75%).

3.2. Measuring local media sentiment

The media does not always play a neutral role when conveying information and it will express positive or negative opinions, which is called media sentiment (Tetlock, 2007). Before we measure the media negative sentiment in relation to pollution, we need to identify the media pollution-news published by China's official provincial party newspapers.[6] We use the dictionary method and topic model to select which news stories report on the provincial pollution situation. The approach of extracting pollution-news is presented graphically in Figure 1. It contains the following five steps:

[6] China's official provincial party newspapers occupy the commanding heights of public opinion in the provinces (Chang et al., 2020).

Step (i). We obtain news articles published by China's provincial official party newspapers from the full-text database of important Chinese newspaper. We then remove the news with incomplete basic information (e.g., time of publication, the newspaper name of publication). We finally collected approximately 1.4 million news articles from the provincial official newspapers published in China.

Step (ii). We construct two important dictionaries. The first list is the pollution dictionary, based on the Sogou input method and manual screening. The dictionary contains 190 words about environmental pollution, such as air pollution, water pollution, and ecological system. The second list consists of both positive and negative sentimental words. The list of positive (negative) words is a collection of 4,566 (4,370) words based on not only positive (negative) emotions but also the positive (negative) evaluation of Zhiwang. These words are useful for our sentiment analysis in identifying the positive or negative tone of the news.

Step (iii). We use jieba, a word segmentation package commonly used in the analysis of Chinese text data, to break the textual content into words. In the process of jieba separation, we add the pollution dictionary and sentiment dictionary to the jieba list. By adding these two lists to jieba, we can better split out words related to pollution or sentiment. To speed up the textual analysis, we remove stop words such as "is", "that", and "of", which have little economic meanings. At this point, we have the news text consisting of the words.

Step (iv). We use the dictionary method to identify news related to environmental pollution. If the key words in the pollution dictionary are included in the news stories, we consider the news to be on pollution. Otherwise, it is non-pollution news. Based on the dictionary method analysis, we extracted 400,000 news articles related to environmental pollution.

Step (v). We use a topic model to further screen the news on pollution. The topic model is a Bayesian generative model developed by Blei et al. (2003). It can reduce the level of subjective error in dictionary analyses. We use Latent Dirichlet Allocation (LDA), which has been used in much of the existing literature (e.g., Huang et al., 2018; Bao and Datta, 2014; Ellingsen et al., 2020), to filter the pollution news. We set 30 topics to analyze the news articles and then compile the 20 words with the highest probability of being used for each topic[7]. By reading the news in every topic, we choose topic 29 to represent firms' environmental pollution.[8] There are more than 20,000 news articles on pollution in all sample provinces.

[Insert Figure 1 here]

We use the dictionary method to construct provincial media negative sentiment measures. The use of a sentiment measure is very common in financial text research. For example, Tetlock (2007) use the frequency of negative words and positive words to gauge the tone of firm-specific news. Loughran and McDonald (2011) develop a dictionary with positive words and negative words to identify the tone of the financial content. Following prior studies (e.g., Liu

[7] We used the "Gensim" package to run the LDA model. When doing the textual analysis, we need to set three parameters: the topic smoothing (α), the term smoothing (β), and the number of topics. We used the coherence score to decide the number of topics (Kaplan and Vakili, 2015). The optimal number is 30. The parameters α and β are automatically set as "1/number of topics".

[8] The highest-probability words in topic 29 are environmental protection (环保), pollution (污染), discharge (排放), environment (环境), enterprise (企业), environmental protection (环境保护), works (工作), govern (治理), project (项目), key point (重点), implement (实施), country (国家), reduce (减少), cycle (循环), clean (清洁), purpose (目标), energy (能源), energy conservation and emissions reduction (节能减排), pollutant (污染物), and use (利用).

and McConnell, 2013; Jiang et al., 2019), we use the dictionary method to calculate the frequency of positive (negative) words in every news article on pollution. The sentiment dictionary consists of positive word list and negative word list. The positive (negative) list contains positive (negative) emotion words and evaluation words of Zhiwang, with a total of 4566 (4370) words. The specific procedure for determining the negative sentiment variable is as follows:

$$MSENT = \frac{Negative\ Words - Positive\ Words}{Positive\ Words + Negative\ Words} \tag{1}$$

where $Positive\ Words$ is the number of positive words in the pollution news article, and $Negative\ Words$ is the number of negative words in the pollution news article.

3.3. Measuring green innovation

The number of patents is an effective indicator of the level of innovation output (Chang et al., 2019). Our first measure of green innovation output, named GPT, is the natural logarithm of one plus the number of patents applied for during the current year, and that which are eventually granted. There are two main reasons for choosing this variable. First, the time interval from application to the granting of a green patent is 1.5 years on average, and the longest time for some patents can be up to 12 years.[9] Meanwhile, on average, a 2.9 (0.7) years lag exists between the date of the invention (utility) green patent application and the date that it is granted. Therefore, if we use the number of patents applied for and granted during the current year, it may underestimate the green innovation output. Second, we do not use the proportion of green patents on all patents because the ratio tends to represent the attention paid to green innovation, which cannot accurately reflect an enterprise's true environmental governance ability.

[9] According to the detailed date of the green patent, we calculate the average time interval between the application year and the granted year.

Our second measure of green innovation output, named $GPTC$, is the natural logarithm of one plus the number of patent citations, which is adjusted by the fixed effects of industry and application year. Patent citation represents the innovation technology value. The variable has been used to represent the innovation output in much of the existing literature (e.g., Hall et al., 2001; Chang et al., 2015, 2019). The forward citation is closely related to the type of green patent. We adjust the patent citation by subtracting the citation mean in the same industry and year. We also use the raw citation and another adjusted citation in the robustness checks.

3.4. Control variables

To isolate the effect of negative sentiment on the green innovation output, we control for an array of firm characteristic variables that prior studies have documented as being important determinants of a firm's innovation output (e.g., Qi et al., 2018; Hao and Liang, 2019; Chang et al., 2019). Chang et al. (2019) argue that larger and older firms have more patents and citations. We therefore use the natural logarithm of the firms' total assets ($SIZE$) to control for firm size and use the natural logarithm of one plus the years since the firm was listed in the exchange to represent firm age (AGE). Capital structure and the cash flow can affect the financing ability for innovation. We use the ratio of the book value of debts to total assets to represent the leverage of the firm (LEV) and use the net cash flow from operations divided by total assets to account for cash flow (CFO). To control for a firm's profitability and growth opportunities, we include return on assets (ROA) and the book-to-market ratio (BM). The percentage of shares held by the largest ten shareholders (TOP) is included to account for the firms' equity structure. To control for innovation input, we included the R&D expenses ($R\&D$), defined as the R&D expenses divided by total assets. In line with the study by Hirshleifer et al. (2012) and Chang et al. (2019), we assign a zero R&D value for observations with missing R&D information. We include industry fixed effects and year fixed effects to control for the unobserved industry characteristics and time variation.

3.5. Descriptive statistics

Panel A of Table 1 presents the descriptive statistics for the variables used in the analysis. The sample consists of 18,708 firm-year observations. The mean value of GPT is 0.417 with a standard deviation of 0.854. This shows that the number of green patents applied for by the firms and eventually granted is 0.517 each year on average. The other green innovation output variable, $GPTC$, has a median value -0.018, which shows that many of the patent citations values are lower than the industry average for green patents. Turning to the negative sentiment measure, we find that the mean of $MSENT$ is -0.657. This suggests that the provincial official party newspapers of China tend to use a positive tone when reporting pollution-related news. The other control variables also exhibit a level of variability, which captures the basic characteristics of the companies.

We present the correlation matrix of the variables in Panel B of Table 1. The correlation between the negative sentiment and firm green innovation is positive and statistically significant at the 1% level. These results are consistent with previous studies which document that older and larger firms are relatively more likely to generate more green patents. To further validate the relationship between negative sentiment in relation to pollution news and corporate green innovation, we next conduct multivariate tests to rule out any confounding factors and to mitigate the omitted variable bias.

[Insert Table 1 here]

4. Key Findings

4.1. The baseline model

To investigate the relationship between provincial media negative sentiment on pollution and green innovation output, the baseline regression specification is written as follows:

$$GreenInnovation_{i,t} = \beta_0 + \beta_1 MSENT_{j,t-1} + \lambda Controls_{i,t-1} + \delta \mathrm{Indus}_i + \theta Year_t + \varepsilon_{i,t} \quad (2)$$

where $GreenInnovation_{i,t}$ represents our green innovation output measures (GPT and $GPTC$) for firm i in year t. The key independent variable is negative sentiment on pollution ($MSENT$) in province j, where a firm's headquarters are located during year $t-1$. $Controls$ is a series of control variables, including firm size ($SIZE$), leverage (LEV), return on assets (ROA), equity structure (TOP), cash flow (CFO), age (AGE), book-to-market value (BM) and R&D expenses ($R\&D$). All control variables are measured in year $t-1$. We also control for the industry fixed effects and year fixed effects in the regressions. The standard errors of the estimated coefficients are corrected for heteroscedasticity[10].

Table 2 presents the baseline results. We use Eq. (2) to estimate the effect of negative sentiment on firm green innovation. Columns (1) and (2) present the results for the green patent count and the citation count, respectively. In both columns, the coefficients of $MSENT$ are positive and statistically significant at the 1% level (coefficient = 0.393, *t*-statistics = 4.1 and coefficient = 0.167, *t*-statistics = 3.7 in columns (1) and (2)). The results suggest that firms located in a province where the media reports on pollution news with a negative tone tend to increase their level of green innovation output. Economically, a one-standard-deviation increase in $MSENT$ increases the number of green patents by 7.38% on average.[11] The coefficients of the control variables are consistent with the prior literature (e.g., Chang et al., 2019). Larger

---

[10] The results remain when the standard errors are clustered at industry or firm.

[11] Specifically, $d[\ln(1+y)]/dx = [1/(y+1)]dy/dx, dy = (y+1)dx \times d[\ln(1+y)]/dx, d[\ln(1+y)]/dx = 0.393$. A-one-standard deviation increase in $MSENT$, so $dx = 0.064$. The change in green patent number (GPT) from its mean value ($\ln(1+Patent) = 0.417, y = patent = 0.517$) is then equal to $0.393 \times (1+0.517) \times 0.064 = 0.038$, which amounts to 7.38% of the mean value of the green patents. Because the patent citation is adjusted by industry and year, we do not show the economic significance of the coefficients for the independent variable $GPTC$.

firms and those with a higher level of R&D expenses are more prone to increase their level of green innovation. Firms with lower book-to-market ratios, or a shorter listed age are more innovative in terms of the quantity of green patents applied for, as this may be considered a way for a firm to achieve a rapid breakthrough.

Taken together, our baseline results in Table 2 confirm a positive relationship between negative sentiment towards pollution and firm green innovation output. The results are consistent with the governance view that if the provincial official media reports the polluting-news in a more negative tone, the media pressure may force firms to be more environmentally friendly to prevent punishment by the regulators. That said, media pressure stemming from negative reporting sentiment is a form of external governance, which could enhance firm green innovation. In turn, the results do not support the information view that negative sentiment reduces a firm's financing ability, thus reducing the number of green patents.

[Insert Table 2 here]

4.2. Robustness tests

We perform a number of additional tests to ensure that our baseline results are robust to alternative variable definitions and subsamples. Following Byun and Oh (2018), we create three alternative variables to measure media sentiment towards pollution. We first construct a new negative sentiment variable that combines both the negative and positive media attention for a firm. Specifically, we classify the polluting-news into two categories according to news tone, i.e., $(Negative\ Words - Positive\ Words)/(Negative\ Words + Positive\ Words)$. If the tone of a news article is higher than the median of all news tones used, the news is judged as being negative news, otherwise it is positive. We calculate the new media negative sentiment

($MSENT1$) as the difference between the proportion of negative news and the proportion of positive news, i.e., $MSENT1 = [(the\ number\ of\ negative\ news - the\ number\ of\ positive\ news)/total\ news] \times 100$. The alternative media negative sentiment variable is $MSENT2$, defined as $(Negative\ Words - Positive\ Words)/Total\ \ Words$. Finally, according to Mohammad and Turney (2013), and Polyzos et al. (2023), we employ a new sentiment dictionary, named NRC lexicon in Chinese, to develop the negative sentiment variable ($MSENT3$). The calculation is the same as Eq. (1). Panel A, Panel B, and Panel C of Table 3 report the respective regression results of replacing $MSENT$ with $MSENT1$, $MSENT2$, and $MSENT3$ in Eq. (2). The coefficients of $MSENT1$, $MSENT2$, and $MSENT3$ are positive and statistically significant for both measures of green innovation. As such, our results are robust to alternative measures of media sentiment towards environmental pollution.

We also use two alternative green innovation output variables to alleviate the measure errors stemming from patent citations. The first is denoted as $GPTC\_RAW$ and is measured as the natural logarithm of one plus the raw citations. In line with the study by Chang et al. (2019), we use another method to adjust the patent citations by scaling the raw citation counts by the average citation counts applied for in the same year and in the same industry, which is denoted as the second dependent variable $GPTC\_ADJ$. The results using $GPTC\_RAW$ and $GPTC\_ADJ$ as independent variables are presented in columns (1) and (2) of Panel D in Table 3, respectively. In both columns, the coefficients of negative sentiment are positive and statistically significant, suggesting that the positive relationship between negative sentiment and firm green innovation is robust to alternative measures of green innovation.

Finally, we construct a subsample consisting of firm-year observations before 2017 and re-estimate the regression in Eq. (2). As the patents require around two years to be granted, the number and citations of patents applied for in the years 2017 and 2018 are underestimated. In order to reduce the level of error, we only retain the sample from the period 2007 to 2016. We use the subsample to re-estimate the relationship between negative sentiment and green innovation and report the results in Panel E of Table 3. The coefficients of negative sentiment remain positive and statistically significant at the 1% level for both measures of green innovation. Taken together, our battery of tests show that the positive innovation effect of negative sentiment is robust to alternative measure of regression variables, as well as to the choice of subsample.

[Insert Table 3 here]

4.3. Tests for endogeneity

Although we document a robust relationship between negative sentiment and green innovation, the results are potentially subject to two types of endogeneity. The first is omitted variables that may affect negative sentiment and firms' green patents. The second is reverse causality in that the environmental performance of a firm may affect the media sentiment towards any pollution news. We perform several tests to mitigate endogeneity concerns. We tabulate the results of the endogeneity tests in Table 4 where we only report the coefficients of negative sentiment and newly added variables for the sake of brevity.

4.3.1 Tests on omitted variables

Our first strategy is to control for fixed effects. In the baseline regression, we control for year and industry fixed effects. However, there may be some industry effects that change by year. We control for industry-by-year fixed effects and re-estimate the relationship between negative sentiment and firm green innovation outputs. Panel A.1 of Table 4 suggests that our main results continue to hold.

Second, we control for three province-level variables that influence media reports and firm green innovation. The first variable is provincial pollution measured by the mean of PM2.5 in the last year. Media plays a monitoring role and reports a greater quantity of news on pollution if the province has serious environmental problems (Jia et al., 2016). On the other hand, firms in highly polluted provinces have two choices in relation to increasing their green innovation output. The cost of environmental violation is relatively low in highly polluted provinces, and this saves a considerable amount of environmental governance costs, meaning there is sufficient capital to invest in innovation activities. The other choice is that, in an external environment with serious environmental pollution, firms drift along and still create a vast amount of productive investment, while neglecting their own green innovation. The second variable is provincial environmental regulation, measured by the natural logarithm of the number of laws on pollution prevention existing in the last year. Firms in provinces with high levels of environmental regulation may carry out green innovation activities to reduce the environmental cost in the long run, or they may reduce their innovation because of the high compliance cost. We also control for provincial economic development, which can enhance firms' green innovation. We therefore control for the three province-level variables and re-estimate the relationship between negative sentiment towards pollution and firm green innovation. The results in Panel A.2 of Table 4 show that the coefficients on negative sentiment remain positive and statistically significant, suggesting that our results are not driven by the level of provincial pollution, environmental regulation and economic development.

Furthermore, we control for firm-level media coverage and re-estimate Eq. (2). Compared with province-level media sentiment, firm-level media coverage creates more pressure for firms to behave in a certain way. Li and Shen (2010) find that the greater the media exposure for firms, the higher probability of the firms correcting their behavior in response to the violation. Yu et al. (2011) suggest that media attention encourages firms to conduct earnings management. If certain companies are exposed to a scandal, the provincial party's official newspaper will also report the related news. Therefore, the firm-level media coverage may be an important omitted variable affecting both the province-level media sentiment and corporate green innovation. To address this concern, we collect firm-level news from the CASMAR database and construct two variable proxies for firm-level media coverage. The first is, $TOTAL\_NEWS$ which is the firm-level total news, defined as the natural logarithm of one plus the number of total news articles on the firm. $POLLUTE\_NEWS$ is defined as the natural logarithm of one plus the number of news articles on a firm's polluting activities. We incorporate $TOTAL\_NEWS$ and $POLLUTE\_NEWS$ in Eq. (2), respectively. The results in Panels A.3 and A.4 of Table 4 show that the positive relationship between the negative sentiment towards pollution and firm green innovation is robust to the inclusion of firm-level media coverage in the baseline regression.

Finally, we control for several linguistic features of the polluting-news, including the length, specificity, and intensity of the news reported by the official newspaper. $MLENGTH$ is the average news article length, defined as the natural logarithm of the number of words in the pollution-related news. $MSPECIFICITY$ is the average news specificity, defined as the number of specific words divided by the total words. Following Chang et al. (2020), we define the specific words as being the particular entity names, including names of locations, listed firms, and money units. $MINTEN$ is the average news intensity, calculated as the number of pollution-related words scaled by the total words in the news. We include them in Eq. (2) and the associate results are reported in Panel A.5 of Table 4. The coefficients of negative media sentiment are positive and significant, suggesting that our results are not primarily driven by the linguistic features of the polluting-news.

4.3.2 Tests on reverse causality

The reverse causality may not be a significant issue because the negative sentiment on pollution is a provincial variable, which cannot be affected by firm's decision on green innovation activities. However, if most of the firms in a province are paying little attention to green innovation investments, the provincial media will report more news on pollution to encourage the firms to increase their number of green patents, because the region's green growth targets are difficult to achieve. In line with Chang et al. (2019), we incorporate a firm's past innovation investments and past green innovation outputs into Eq. (2) to explicitly account for reverse causality. Specifically, we measure past innovation investments ($PAST\_R\&D$) as the rolling average R&D/Assets in the past three years (from year $t-2$ to $t-4$). The past green innovation output is defined as the rolling average of GPT and $GPTC$ in the past three years (from year $t-2$ to $t-4$) for both green innovation measures. The results presented in panels B.1 and B.2 of Table 4 reveal that the coefficients of negative sentiment remain positive and statistically significant, and this is consistent with the baseline regression.

4.3.3 The instrumental variable approach

To further alleviate the endogeneity concerns, we employ an instrumental variable approach. Specifically, we employ the standard deviation of elevation in the province, where the firm is located, as an instrumental variable for local media negative sentiment. From a relevance perspective, media negative sentiment in one province should be negatively correlated with the standard deviation of elevation. The reason is that the elevation affects the cost of information collection (Liu and Ma., 2020), which makes it difficult for the media to find environmental pollution problems in time. Moreover, this instrumental variable is likely to meet exclusion criteria because the elevation is a geographical factor and should not directly affect firm innovation outputs.

We construct the province-level standard deviation of elevation and implement the instrumental variable analysis using a two-stage procedure (2SLS). The first-stage regression (un-tabulated) shows that the standard deviation of elevation is negatively and significantly correlated with the province media negative sentiment. The weak instrument test generates a p-value of less than 0.01, which demonstrates that the standard deviation of elevation meets the relevance criteria and is not a weak instrumental variable. The results of the second-stage regression are shown in Panel C of Table 4. The coefficients on the fitted media negative sentiment of pollution are positive and statistically significant, consistent with the baseline results.

[Insert Table 4 here]

5. Further Analysis

Our baseline results suggest that a negative sentiment has significantly positive effects on firms' green innovation output, consistent with the governance view that the media plays an external monitoring role in firm governance structure. In this section, we conduct a number of tests to better understand the channel through which official media negative sentiment towards pollution enhances firm green innovation performance. We first examine the governance channel. Generally, it is the governance channel which plays the key role in shaping the negative sentiment-innovation relationship, and the media can be seen as a supplement to the existing structure of firm governance. For firms with weaker governance mechanisms, we expect the media sentiment effect on green innovation to be stronger. Meanwhile, we test how institutional background affects the relationship between negative sentiment and green innovation. We also conduct several tests to examine the possibility of the information channel, in that negative news impedes firm green innovation by increasing the level of financing constraint. Finally, we explore how different types of media affect firm green innovation. The above-mentioned negative sentiment is extracted from provincial official newspapers. We then measure the sentiment from market-oriented media to compare the difference between the two types of media. In addition, we analyze the effects of official media negative sentiment on different types of green patent, which helps us to understand firms' green innovation strategies.

5.1. Cross-sectional heterogeneity

We use a variable proxy for governance structure one year prior to applying for a green patent to partition our sample in several ways. We first examine how a firm's internal governance structure affects our results. We measure a firm's internal governance based on whether a firm's CEO serves as the chairman of the board (CEO). The internal governance is better if the CEO does not serve as the chairman of the board. We re-estimate Eq. (2) for the two subsamples with stronger and weaker governance structures. The results are presented in Panel A of Table 5. To be concise, we only tabulate the coefficients of negative sentiment for different subsamples. The results show that the effect of negative sentiment on pollution is more pronounced for firms with weaker internal governance mechanisms. The coefficients of negative sentiment are significantly smaller for firms with stronger internal governance structures.

We also examine how the relationship changes for firms with different external governance structures. We use the variable of whether the firm hires Big Four auditors to divide the sample into stronger governance firms (hires Big Four auditors) and weaker governance firms (does not hire Big Four auditors). The results are presented in Panel B of Table 5. Consistent with our expectation, the coefficients of negative sentiment are positive and statistically significant mainly for firms with weaker external governance structures. In addition, among firms with stronger governance, the coefficients of the negative sentiment are not significantly different from zero.

We further partition the sample according to provincial economic development, which is a province-level external governance variable. The Pollution Haven Hypothesis and Pollution Source Nearby Transfer are hot issues in environmental economics. The implication is that the varying intensity of environmental regulation among regions leads to the transfer of some polluting firms to regions with weaker environmental regulation, especially to certain economically backward regions. In order to attract capital and to accelerate regional economic development, environmental requirements will then be relaxed, and the cross-land transfer of enterprises will be promoted in reverse (Becker and Henderson, 2000). Firms in more developed regions may face a higher level of external governance pressure. Firms are then divided into two groups according to regional GDP. If the regional GDP is above (below) the median of the entire sample, the province-level external governance is strong (weak). We then re-estimate Eq. (2) for the two groups. The results are presented in Panel C of Table 5. We find that the coefficients of negative sentiment on pollution are significant for firms in less economically developed provinces. The effects are insignificant or significantly smaller for firms in more developed provinces. Collectively, our cross-sectional results show that the positive effect of negative sentiment on green innovation is more pronounced when firms have weaker governance structures. The findings reveal that the external governance of negative sentiment is complementary to other forms of governance structures.

We also explore how institutional background influences the relationship between negative sentiment and corporate green innovation. Concerning the great harm caused by environmental pollution, environmental governance has attracted a great deal of attention from economists and policy makers (e.g., Jin and Shen, 2018). China's approach to environmental governance is dominated by environmental regulation. The regulator imposes administrative penalties on environmental pollution to increase the cost of pollution, and then supervises enterprises to reduce pollution. The regulator also implements incentive policies to encourage enterprises to carry out environmental improvements. Media news, as a tool for the government to guide public opinion, may complement environmental regulation policies. We use the proportion of firms penalized for environmental pollution at the provincial level to measure punitive measures. If the proportion is above (below) the median of the entire sample, the regulation is strong (weak). We use the regional green support level to measure incentive measures. We construct a composite variable from the perspective of local financial institution support and fiscal support[12]. If the variable is above (below) the median of the entire sample, the incentive is strong (weak). We re-estimate Eq. (2) for the two groups according to punitive measures and environmental incentives, respectively. The results are presented in Panel D of Table 5. We find that the coefficients of negative sentiment on pollution are significant for firms in areas where punitive measures are weak, suggesting that punitive measures and negative sentiment complement each other, with negative sentiment playing a greater watchdog role when environmental punitive measures are weak, consistent with governance channel. Meanwhile, under incentives, negative sentiment has a stronger impact on corporate green innovation.

The baseline results, which show that a negative sentiment has positive effects on firms' green innovation, run contrary to the information channel. To further eliminate concerns, we also explore whether the information channel can explain the negative sentiment-innovation relationship. Media news could reduce the information asymmetry between firms and the public by collecting, selecting, certifying, and repackaging information (Miller, 2006; Dyck et al., 2008). Specifically, prior studies suggest that firms suffer from external financing constraints when they are engaged in innovative activities. Negative sentiment towards pollution conveys negative information to the public and impedes lenders' or shareholders' risk taking, thus increasing firms' financial constraints. The increased financing constraints caused by negative

[12] The local financial institution support includes green credit, green securities, and green insurance. The fiscal support includes investment in pollution control and government spending on environmental protection.

sentiment could impede firms' green innovation. As such, we expect that the positive effect on green innovation should be less for firms with greater financing constraints. To test this conjecture, we partition our sample using financing constraint proxies and separately estimate the effect of negative sentiment on green innovation for more and for less financially constrained firms. We measure financial constraints using two variables i.e., FC_SIZE and $FC\_DIVIDEND$. Specifically, FC_SIZE is a dummy variable with a value of one if a firm's size is above the sample median, and zero otherwise. $FC\_DIVIDEND$ indicates whether a firm pays dividends and takes the value of one if the firm pays dividends, and zero otherwise. A firm is defined as being less financially constrained if the firm's size is greater than the sample median and if the firm pays dividends. Accordingly, we construct our subsamples of more and less financially constrained firms based on the different values of financial constraint proxies (i.e., FC_SIZE and $FC\_DIVIDEND$) and re-estimate the regression in Eq. (2) for these subsamples. The associated results are reported in Panel E of Table 5. The coefficients on negative sentiment are both positive and statistically significant for the two groups. More importantly, the difference in coefficients for the two groups is not significant. This finding runs counter to the information channel, suggesting that our main results are not driven by negative sentiment increasing firms' financial constraints and impeding firms' green innovation. That said, the evidence in Panel E of Table 5 does not support the alternative explanation that the negative sentiment effect on green innovation works through the information channel.

[Insert Table 5 here]

5.2. Different types of media

To recap, we find that negative sentiment from a provincial official newspaper, namely stated-controlled media, has positive effects on firm green innovation outputs. We point out the role of state-controlled media in the way it affects corporate behavior, which is different from the prior studies which are limited to a particular type of media, this being market-oriented media. In this section, we further consider the effects of market-oriented media on firm green innovation and explore whether these two types of media play different roles in the process of firm green innovation.

Compared with state-controlled media, market-oriented media is less censored and has a higher degree of freedom to report news that satisfies the demand of readers (Houstan et al., 2011; Dyck et al., 2010). This independence, however, is a double-edged sword. Market-oriented media tends to exaggerate the news stories or cast rumors to attract readers' attention (Mullainathan and Shleifer, 2002). The negative news stories on pollution, in particular, are eye-catching and they have become the most newsworthy (Jia et al., 2016). Conversely, state-controlled media is a form of government regulation. Empirically, You et al. (2018) show that negative news from the market-oriented media has a stronger power to influence the likelihood of a forced executive turnover, than those from state-controlled media. By contrast, negative news, especially rumors from commercial newspapers, exacerbates managerial short-termism. This then impedes corporate innovation (Yang et al., 2017) or leads to firms conducting fake green M&As (Pan et al., 2019). Therefore, the empirical evidence is mixed and the effects of news from market-oriented media cannot be ignored. We first examine how the market-oriented media sentiment affects firm green innovation and then explore whether it affects the relationship between the state-controlled media sentiment and firm green innovation.

We measure market-oriented media sentiment using the top three Chinese market-oriented business newspapers, namely *21st Century Business Herald*, *First Financial Daily*, and *China Business Journal*. We use the dictionary method to judge whether the news reported by the market-oriented media contains particular province (city) names and pollution-related words. If the province or city name is mentioned when a news article discusses environmental pollution, we classify the news article as belonging to that particular province. Through the above approach, we convert the nation-level media news into province-level media news. We then use the dictionary method to count the number of positive and negative words in the news. We calculate the market-oriented media negative sentiment ($MSENT\_MARKET$) using the same method as $MSENT$ (i.e., Eq. (1)). We replace $MSENT$ with $MSENT\_MARKET$ in Eq. (2) to investigate the impact of market-oriented media sentiment on firm green innovation. The results reported in columns (1) and (2) of Table 6 show that the coefficients of $MSENT\_MARKET$ are negative and statistically significant for both green innovation measures, revealing that negative news from market-oriented media exacerbates managerial short-termism, which then impedes corporate green innovation. That said, market-oriented media newspapers may serve more of an information function as opposed to the governance function of provincial party newspapers. The negative sentiment from market-oriented media newspapers may trigger penalty mechanisms in asset markets, which increase the difficulties of financing for green innovation. Moreover, we include both market-oriented media negative sentiment ($MSENT\_MARKET$) and state-controlled media negative sentiment ($MSENT$) in Eq. (2), and the results are reported in columns (3) and (4) of Table 6. The coefficients of $MSENT$ ($MSENT\_MARKET$) remain positive (negative) significantly. This finding supports our argument that negative news from state-controlled media plays a governance role by affecting firm green innovation.

[Insert Table 6 here]

5.3. Different types of green innovation

Our analysis has shown that official media negative sentiment towards pollution affects firms' green innovation output significantly (i.e., patents and citations). In this section, we test how negative sentiment affects different types of green innovations to further understand firms' innovation strategies. We classify all green patents into invention patents and utility patents and re-estimate Eq. (2) for these two types of patents. The two independent variables are the natural logarithm of one plus the number of invention patents and the natural logarithm of one plus the number of utility patents, respectively. Columns (1) and (2) of Table 7 report the results and show that negative news on pollution leads to an increase in the number of invention patents and utility patents. We also use green patent claims, which is an important variable, to measure the quality of patents (Hao and Liang, 2019) as an additional analysis. Column (3) of Table 7 presents the results and shows that negative sentiment has a significant effect on green patent quality. The results imply that negative sentiment plays a monitoring role in urging firms to increase their green patent quantity and quality.

[Insert Table 7 here]

6. Conclusion

We study the relationship between local official media negative sentiment on pollution and firms' green innovation. In particular, we propose two competing views on the effect of negative sentiment on firms' green innovation. The information channel posits that media acts as an information intermediary. Thus, negative media sentiment increases a firm's level of financial constraints, thereby impeding green innovation. Whereas the governance channel maintains that negative news performs a governance function through managers' reputations and administrative intervention. Thus, a negative sentiment towards pollution has positive effects on green innovation.

Using a sample of Chinese listed firms from 2007 to 2018, we find that a negative sentiment from local official media has positive effects on firms' green innovation, which is consistent with the governance view. Our results are robust to a variety of tests on alternative variable measures, subsamples and endogeneity issues. Further analysis shows that the negative sentiment towards pollution has a stronger effect on firms with weaker governance mechanisms.

Moreover, we examine how market-oriented media sentiment affects the relationship between official media negative sentiment and green innovation. The results show that the negative sentiment from market-oriented media impedes green innovation and does not affect the relationship between state-controlled media negative sentiment and green innovation. We also explore the effects of official media negative sentiment on different types of green innovation and the results show that negative sentiment can enhance green inventory patents, green utility patents and patent claims.

Our findings document a significant effect of negative sentiment on firms' green innovation through the governance channel, thus revealing a new determinant of green innovation. This paper also highlights that provincial official media can report more negative news on pollution in order to foster firms' green innovation outputs. On the contrary, the market-oriented media has a strong information function by exposing pollution incidents. We suggest the local government encourages the market-oriented media to cooperate with the government and financial institutions to strengthen the exchange of information, thereby effectively identifying polluting enterprises and improving the efficiency of resource allocation. The government should guard against speculation in the market-oriented media to avoid managerial short-termism on the part of firms.

Finally, we stress three limitations of our study. For one thing, since we extract the news sentiment from China's provincial official party newspaper, our results should be generalized with caution, especially to social media and commercial newspapers. For another thing, although our findings suggest negative sentiment can promote corporate green innovation, we do not conclude that the negative sentiment can produce positive social and economic outcomes. Thus, in the future research, we can investigate whether the beneficial effect of negative sentiment on green innovation indeed improve firms' environmental performance (e.g., higher environmental score or lower level of emissions) and financial performance. Last but not least, future research can update data on the media sentiment measure to monitor the changes in the function of provincial official newspapers.

References

Bao, Y., Datta, A., 2014, Simultaneously discovering and quantifying risk types from textual risk disclosures. *Management Science* 60, 1371-1391.

Baker, S. R., Bloom, N., Davis, S. J., 2016, Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131, 1593-1636.

Bucher, H., Schumacher, P., 2006, The relevance of attention for selecting news content. An eye-tracking study on attention patterns in the reception of print and online media. *Communications 31*, 347-368.

Bajo, E., Raimondo, C., 2017, Media sentiment and IPO underpricing. *Journal of Corporate Finance* 46, 139-153.

Bai, X. Y., Lyu, C., 2023, Executive's environmental protection background and corporate green innovation: evidence from China. *Sustainability* 15, 4154.

Becker, R., Henderson, V., 2000, Effects of air quality regulations on polluting industries. *Journal of Political Economy* 108, 379-421.

Blei, D., Ng, A., Jordan, M., 2003, Latent dirichlet allocation. *Journal of Machine Learning and Research* 3, 993-1022.

Byun, S. K., Oh, J. M., 2018, Local corporate social responsibility, media coverage, and shareholder value. *Journal of Banking and Finance* 87, 68-86.

Chahine, S., Mansi, S., Mazboudi, M., 2015, Media news and earnings management prior to equity offerings. *Journal of Corporate Finance* 35, 177-195

Cai, W. G., Zhou, X. L., 2014, On the drivers of eco-innovation: empirical evidence from China. *Journal of Cleaner Production* 79, 239-248.

Cahan, S. F., Chen, C., Chen, L., Nguyen, N. H., 2015, Corporate social responsibility and media coverage. *Journal of Banking and Finance 59*, 409 - 422.

Chang, X., Chen, Y., Wang, S. Q., Zhang, K., Zhang, W., 2019, Credit default swaps and corporate innovation. *Journal of Financial Economics* 134, 474-500.

Chang, X., Fu, K., Low, A., Zhang, W., 2015, Non-executive employee stock options and corporate innovation. *Journal of Financial Economics* 115, 168-188.

Chen, Y., Lai, S., Wen, C., 2006, The influence of green innovation performance on corporate advantage in Taiwan. *Journal of Business Ethics* 67, 331–339.

Das, P., Horton, R., 2018, Pollution, health, and the planet: time for decisive action. *Lancet* 391, 407-408.

Demirel, P., Kesidou, E., 2011, Stimulating different types of eco-innovation in the UK: Government policies and firm motivations. *Ecological Economics* 70, 1546-1557.

Dyck, I. J., Volchkova, N., Zingales, L., 2008, The corporate governance role of the media: Evidence from Russia. *Journal of Finance* 63, 1093-1135.

Dyck, A., Morse, A., Zingales, L., 2010, Who blows the whistle on corporate fraud? *The Journal of Finance* 65, 2213-2253.

Djankov, S., McLiesh, C., Nenova, T., Shleifer, A., 2003, Who owns the media? *Journal of Law and Economics* 46, 341-381.

Ellingsen, J., Larsen. V. H., Thorsrud. L. A., 2020, News media versus FRED-MD for macroeconomic forecasting. *Applied Econometrics* 37, 63-81.

Fang, L., Peress, J., 2009, Media coverage and the cross-section of stock returns. *Journal of Finance* 64, 2023-2052.

Feng, L. B., Fu, T., Shi, Y. L., 2022, How does news sentiment affect the states of Japanese stock return volatility? *International Review of Financial Analysis* 84, 102267.

Gan, B. Q., Alexeev. V., Bird, R., Yeung, D., 2020, Sensitivity to sentiment: News vs social media. *International Review of Financial Analysis* 67, 101390.

Gentzkow, M., Kelly, B., Taddy, M., 2019, Text as data. *Journal of Economic Literature* 57, 535-574.

Gurun, U. G., Butler, A. W., 2012, Don't believe the hype: Local media slant, local advertising, and firm value. *Journal of Finance* 67, 561-598.

Hassan, T. A., Hollander, S., Van Lent, L., Tahoun, A., 2019, Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics* 134, 2135-2202.

Horbach, J., Rammer, C., Rennings, K., 2012, Determinants of eco-innovations by type of environmental impact: The role of regulatory push/pull, technology push and market pull. *Ecological Economics* 78, 112-122.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2001, The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). *National Bureau of Economic Research*.

Hall, B. H., Lerner, J., 2010, The financing of R&D and innovation. Handb. Econ. Innov. 1, 609–639.

Hao, X., Liang, Q., 2019, Does hedging foreign exchange risk promote corporate innovation. *Journal of World Economy* 9, 151-172.

Huang, A. H., Lehavy, R., Zang, A. Y., Zheng, R., 2018, Analyst information discovery and interpretation roles: A topic modeling approach. *Management Science* 64, 2833-2855.

Hojnik, J., Ruzzier, M., 2016, What drives eco-innovation? A review of an emerging literature. *Environmental Innovation and Societal Transitions* 19, 31-41.

Holmström, B., 1989, Agency costs and innovation (No. 214). IUI Working Paper.

Hirshleifer, D., Low, A., Teoh, S. H., 2012, Are overconfident CEOs better innovators? *Journal of Finance* 67, 1457-1498.

Houston, J. F., Lin, C., Ma, Y., 2011, Media ownership, concentration and corruption in bank lending. *Journal of Financial Economics* 100, 326-350.

Jia, M., Tong, L., Viswanath, P. V., Zhang, Z., 2016, Word power: The impact of negative media coverage on disciplining corporate pollution. *Journal of Business Ethics* 138, 437-458.

Jiang, F., J. Lee, Martin, X., Zhou, G., 2019, Manager sentiment and stock returns. *Journal of Financial Economics* 132, 126-149.

Jin, G., Shen, K. R., 2018, Beggar-thy-neighbor or neighborly?— Environmental regulation enforcement interactions and urban productivity growth. *Management World* 34, 43-55.

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. *Accounting Review* 84, 1639-1670.

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., Kapadia, S., 2020, Making text count: economic forecasting using newspaper text. Staff Working Paper No. 865.

Kaplan, S., K. Vakili, 2015, The double-edged sword of recombination in breakthrough innovation, *Strategic Management Journal* 36, 1435-1457.

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

Li, P., Shen, Y., 2010, The corporate governance role of media: Empirical evidence from China. *Economic Research Journal* 4, 14-27.

Lin, H., Zeng, S. X., Ma, H. Y., Qi, G. Y., Tam, V. W., 2014, Can political capital drive corporate green innovation? Lessons from China. *Journal of Cleaner Production* 64, 63-72.

Liu, B., McConnell, J. J., 2013, The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics* 110, 1-17.

Liu, C., Ma, Q., 2020, The impact of network infrastructure construction on total factor productivity growth: a quasi-natural experiment based on the “Broadband China” pilot policy. *Chinese Journal of Population Science* 3, 75-88+127-128.

Liao, Z., Xu, C. K., Cheng, H., Dong, J., 2018, What drives environmental innovation? A content analysis of listed companies in China. *Journal of Cleaner Production* 198, 1567-1573.

Landrigan, P. J., Fuller, R., Acosta, N. J. R., et al., 2017, The Lancet Commission on pollution and health. *Lancet*, published online Oct 19.

Luo, J., 2012, The Effects and difference of media on the cost of equity and debt. *Review of Investment Studies* 9, 95-112.

Lu, W., Guan, F., Zhang, P., Deng, Y., 2014, Media coverage, information disclosure environment and equity cost. *Accounting Research* 12, 66-96.

Miller, G. S., 2006, The press as a watchdog for accounting fraud. *Journal of Accounting Research* 44, 1001-1033.

Mohammad, S. M., Turney, P. D., 2013, Crowdsourcing a word-emotion association lexicon, *Computational Intelligence* 29, 436-465.

Mullainathan, S., Shleifer, A., 2002, Media Bias. NBER Working Paper.

Polyzos, E., Fotiadis, A., Huan, T. C., 2023, From heroes to scoundrels: exploring the effects of online campaigns celebrating frontline workers on COVID-19 outcomes, *Technology in Society* 72, 102198.

Porter, M. E., 1991, America’s Green Strategy. *Scientific American* 264, 168.

Pollock, T. G., Rindova, V. P., Maggitti, P. G., 2008, Market watch: Information and availability cascades among the media and investors in the US IPO market. *Academy of Management Journal* 51, 335-358.

Pan, A., Liu, X., Qiu, J., Shen, Y., 2019, Can green M&A of heavy polluting enterprises achieve substantial transformation under the pressure of media? *China Industrial Economics* 2, 174-192.

Qin, B, Strömberg, D., Yanhui, W., 2018, Media bias in China. *American Economic Review* 108, 2442-76.

Qi, S., Lin, S., Cui, J., 2018, Do environmental rights trading schemes induce green innovation? Evidence from listed firms in China. *Economic Research Journal* 12, 129-143.

Rognone, L., Hyde, S., Zhang, S. S., 2020, News sentiment in the cryptocurrency market: an empirical comparison with Forex. *International Review of Financial Analysis* 69, 101462.

Sautner, Z., Van Lent, L., Vilkov, G., Zhang, R., 2023, Firm-level climate change exposure. *Journal of Finance* 78, 1449-1498.

Stiglitz, J. E., 1985, Credit markets and the control of capital. *Journal of Money, Credit and Banking* 17, 133-152.

Snyder, Jr. J. M., Strömberg, D., 2010, Press coverage and political accountability. *Journal of Political Economy* 118, 355-408.

Shen, K., Jin, G., Fang, X., 2017, Does environmental regulation cause pollution to transfer nearby? *Economic Research Journal* 5, 44-59.

Sapkota, N., 2022, News-based sentiment and bitcoin volatility. *International Review of Financial Analysis* 82, 102183.

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

Tewksbury, D., Althaus, S. L., 2000, Differences in knowledge acquisition among readers of the paper and online versions of a national newspaper. *Journalism and Mass Communication Quarterly* 77, 457-479.

Yang, D., Zhao, C., 2012, Media monitoring, media governance and managers' compensation. *Economic Research Journal* 6, 116-126.

Yu, Z., Tian, G., Qi, B., Zhang, H., 2011, Corporate governance mechanism of media: Based on the perspective of earnings management. *Management World* 9, 127-140.

Yu, H. X., Liang, C. Y., Liu, Z. H., Wang, H., 2023, News-based ESG sentiment and stock price crash risk. *International Review of Financial Analysis* 88, 102646.

Yang, D. G., Chen, H. W., Liu, Q. L., 2017, Media pressure and corporate innovation. *Economic Research Journal* 52, 125-139.

You, J. X., Zhang, B. H., Zhang, L., 2017. Who captures the power of the pen? *Review of Financial Studies* 31, 43-96.

You, J., Wu, J., 2012, Spiral of silence: Media sentiment and the asset mispricing. *Economic Research Journal* 7, 141-152.

Zeng, H., Li, S., Zhou, Z., Xiao, X., 2018, Water disclosure, media coverage, and firm risk. *Accounting Research* 4, 89-96.

Zhou, K., Yang, H., Wu, Y., 2016a, Research on collaborative food safety governance involving media, capital market and government. *Economic Research Journal* 51, 58-72.

Zhou, K., Ying, Q., Zhong, C., 2016b, Can media coverage improve corporate governance? Evidence from fraud by listed firms in China. *Journal of Financial Research* 6, 193-206.

Table 1. Summary Statistics

The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. GPT is the natural logarithm of one plus the number of patents applied for, during the current year, which are eventually granted. $GPTC$ is the natural logarithm of one plus the patent citations, which is adjusted using the industry-year fixed effect. $MSENT$ is the province-level negative sentiment of polluting-news articles published by China's official party newspapers in the province where a firm is located, calculated as $(Negative\ Words - Positive\ Words)/(Negative\ Words + Positive\ Words)$. $SIZE$ is the natural logarithm of a firm's total assets. LEV is the book value of debts scaled by total assets. ROA is the return of total assets. TOP is the percentage of shares held by the largest ten shareholders. CFO is net cash flow from operations divided by total assets. AGE is the natural logarithm of one plus the number of years since a firm went public. BM is the book value scaled by market value. $R\&D$ is R&D expenses scaled by total assets. Panel A reports the summary statistics and Panel B presents the correlation matrix for the main variables, where numbers in bold indicate statistically significant at the 1% level.

Panel A: Summary statistics

Variables	N	Mean	SD	Q1	Median	Q75
GPT	18708	0.417	0.854	0.000	0.000	0.693
$GPTC$	18708	0.000	0.365	-0.143	-0.018	0.000
$MSENT$	18708	-0.657	0.064	-0.696	-0.656	-0.613
$SIZE$	18708	22.039	1.328	21.089	21.823	22.736
LEV	18708	0.408	0.206	0.242	0.399	0.563
ROA	18708	0.042	0.076	0.018	0.042	0.070
TOP	18708	59.651	15.462	48.844	61.093	71.880
CFO	18708	0.043	0.073	0.005	0.042	0.083
AGE	18708	1.960	0.857	1.322	2.079	2.708
BM	18708	0.342	0.159	0.224	0.319	0.442
$R\&D$	18708	0.019	0.023	0.003	0.016	0.027

Panel B: Correlation matrix

	GPT	$GPTC$	$MSENT$	$SIZE$	LEV	ROA	TOP	CFO	AGE	BM
$GPTC$	**0.369**									
$MSENT$	**0.043**	**0.027**								
$SIZE$	**0.264**	**0.137**	0.007							
LEV	**0.116**	**0.058**	**-0.032**	**0.541**						
ROA	0.017	0.017	-0.019	**-0.019**	**-0.306**					
TOP	**0.028**	**0.023**	0.016	**0.071**	**-0.166**	**0.202**				
CFO	**0.021**	-0.004	0.004	**0.081**	**-0.131**	**0.374**	**0.094**			
AGE	**0.040**	0.003	-0.012	**0.474**	**0.442**	**-0.190**	**-0.445**	**0.032**		
BM	**-0.019**	-0.008	-0.001	**-0.071**	**-0.522**	**0.099**	**0.236**	-0.006	**-0.328**	
$R\&D$	**0.134**	**0.054**	**0.092**	**-0.194**	**-0.212**	**0.069**	0.017	**0.053**	**-0.165**	-0.018

Table 2. Effect of Media Sentiment on Green Innovation

This table reports the regression results for the relation between negative sentiment and corporate green innovation output. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. GPT is the natural logarithm of one plus the number of patents applied for, during the current year, which are eventually granted. $GPTC$ is the natural logarithm of one plus the patent citations, which is adjusted using the industry-year fixed effect. $MSENT$ is the province-level negative sentiment of polluting-news articles published by China's official party newspapers in the province where a firm is located, calculated as $(Negative\ Words - Positive\ Words)/\ (Negative\ Words + Positive\ Words)$. $SIZE$ is the natural logarithm of a firm's total assets. LEV is the book value of debts scaled by total assets. ROA is the return of total assets. TOP is the percentage of shares held by the largest ten shareholders. CFO is net cash flow from operations divided by total assets. AGE is the natural logarithm of one plus the number of years since a firm went public. BM is the book value scaled by market value. $R\&D$ is R&D expenses scaled by total assets. All explanatory variables are measured at $t-1$ in the regressions. Industry fixed effect and year fixed effect are included. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Dependent variables	(1) GPT	(2) $GPTC$
$MSENT$	0.393^{***}	0.167^{***}
	(0.095)	(0.045)
$SIZE$	0.256^{***}	0.067^{***}
	(0.010)	(0.005)
LEV	0.010	-0.047^{**}
	(0.049)	(0.021)
ROA	0.086	0.027
	(0.069)	(0.027)
TOP	-0.003^{***}	-0.001^{***}
	(0.000)	(0.000)
CFO	-0.240^{***}	-0.146^{***}
	(0.082)	(0.036)
AGE	-0.120^{***}	-0.043^{***}
	(0.011)	(0.005)
BM	-0.184^{***}	-0.093^{***}
	(0.058)	(0.025)
$R\&D$	5.811^{***}	1.241^{***}
	(0.918)	(0.267)
$CONS$	-4.738^{***}	-1.141^{***}
	(0.204)	(0.097)
Industry fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
N	18708	18708
Adjusted R-squared	0.155	0.034

Table 3. Robustness Checks

This table presents the results of robustness tests. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. In Panels A, B, and C, we use three alternative variables to measure negative sentiment. $MSENT1$ is defined as difference between the proportion of negative polluting-news and the proportion of positive polluting-news. $MSENT2$ is defined as $(Negative\ Words - Positive\ Words)/Total\ Words$. $MSENT3$ is calculated as $(Negative\ Words - Positive\ Words)/(Negative\ Words + Positive\ Words)$, in which we use NRC Lexicon to extract the number of positive and negative words. Column (1) presents the results using GPT as the dependent variable. Column (2) presents the results using $GPTC$ as the dependent variable. In Panel D, we use another two alternative variables to measure green innovation. Column (1) presents the results using $GPTC\_RAW$, which is calculated as the natural logarithm of one plus the raw citations. Column (2) presents the results using $GPTC\_ADJ$, which is defined as the patent citations scaling by the average citations applied for in the same year and in the same industry. Panel E presents the results using sub-sample. We delete the sample in the year 2017 and 2018. Column (1) presents the results using GPT as the dependent variable. Column (2) presents the results using $GPTC$ as the dependent variable. All regressions include industry and year fixed effects. All regressions include the control variables as those in Table 2 and their coefficients are not tabulated. Detailed variable definitions are in the legend of Table 2. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Alternative measure of media sentiment, $MSENT1$ (N = 18708)		
	(1)	(2)
	GPT	$GPTC$
$MSENT1$	0.1071***	0.0368***
	(0.0202)	(0.0097)
Panel B: Alternative measure of media sentiment, $MSENT2$ (N = 18708)		
	(1)	(2)
	GPT	$GPTC$
$MSENT2$	7.551***	2.382***
	(1.378)	(0.641)
Panel C: Alternative measure of media sentiment, $MSENT3$ (N = 18708)		
	(1)	(2)
	GPT	$GPTC$
$MSENT3$	0.521***	0.296***
	(0.134)	(0.063)
Panel D: Alternative measures of green innovation ($N_{(1)}$ = 18708, $N_{(2)}$ = 13669)		
	(1)	(2)
	$GPTC\_RAW$	$GPTC\_ADJ$
$MSENT$	0.171***	2.397*
	(0.045)	(1.267)
Panel E: Using the sub-sample before 2017 (N = 13559)		
	(1)	(2)
	GPT	$GPTC$
$MSENT$	0.584***	0.271***
	(0.114)	(0.069)

Table 4. Tests on Endogeneity

This table presents the results on addressing endogeneity concerns. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. Panel A presents the regression results on concerning about omitted variables. Panel A.1 presents the regression results with industry-by-year fixed effects. Panel A.2 presents the regression results controlling for the province-level characteristics, which include regional pollution and regional law for pollution. The regional pollution variable, PM is calculated as the annual mean of PM2.5 in the air of all cities in the province last year. The regional law variable, LAW is the natural logarithm of the number of regional laws on pollution last year. GDP is defined as the natural logarithm of gross domestic product. Panel A.3 presents the regression results controlling for firm-level total media news that is the natural logarithm of the number of media news on the firm last year. Panel A.4 presents the regression results controlling for the firm-level media news on pollution that is the natural logarithm of firm's media news on pollution last year. Panel A.5 presents the regression results controlling for linguistic features of the pollution-related news. The news length, $MLENGTH$, is defined as the natural logarithm of the number of words in the news. The news intensity, $MINTEN$, is calculated as the number of pollution-related words scaled by the total words in the news. The news specificity, $MSPECIFICITY$, is defined as the number of specific words divided by the total words in the news. Panel B presents the regression results on concerning about reverse causality. $PAST\_R\&D$ is the average of R&D/Total Asset in the past three years ($t-2\ to\ t-4$). $PAST\_GPT$ ($PAST\_GPTC$) is the average of GPT ($GPTC$) in the past three years ($t-2\ to\ t-4$). Panel C presents the results of the second stage of 2SLS regression. We use the standard deviation of regional elevation as an instrument variable. All regressions include industry and year fixed effects. All regressions include the control variables as those in Table 2 and their coefficients are not tabulated. Detailed variable definitions are in the legend of Table 2. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Controlling for omitted variables		
	GPT	$GPTC$
Panel A.1: *Controlling for industry-by-year fixed effects* (*N* = 18708)		
$MSENT$	0.391***	0.169***
	(0.096)	(0.170)
Panel A.2: *Controlling for province-level characteristics* (*N* = 15725)		
$MSENT$	0.272**	0.183***
	(0.114)	(0.064)
PM	-0.008	-0.003
	(0.019)	(0.009)
LAW	-0.014	0.005
	(0.015)	(0.007)
GDP	0.066***	0.011
	(0.013)	(0.007)
Panel A.3: *Controlling for the firm-level total news* (*N* = 18708)		
$MSENT$	0.381***	0.166***
	(0.095)	(0.045)
$TOTAL\_NEWS$	0.033***	0.002
	(0.006)	(0.004)
Panel A.4: *Controlling for the firm-level news on pollution* (*N* = 18708)		
$MSENT$	0.343***	0.148***
	(0.094)	(0.044)
$POLLUTE\_NEWS$	0.138***	0.054***
	(0.009)	(0.006)
Panel A.5: *Controlling for the linguistic features of news* (*N* = 18708)		
$MSENT$	0.271***	0.163***
	(0.097)	(0.044)

$MLENGTH$	-0.061* (0.036)	0.008 (0.017)
$MINTEN$	-2.859*** (0.962)	0.106 (0.422)
$MSPECIFICITY$	4.541*** (0.775)	0.173 (0.320)
Panel B: Tests on reverse causality		
	GPT	$GPTC$
PanelB.1: Controlling for past innovation investments (N = 18708)		
$MSENT$	0.360*** (0.094)	0.165*** (0.045)
$PAST\_R\&D$	6.161*** (0.865)	0.380 (0.302)
PanelB.2: Controlling for past innovation outputs (N = 18708)		
$MSENT$	0.155** (0.073)	0.127*** (0.043)
$PAST\_GPT$	0.931*** (0.014)	
$PAST\_GPTC$		0.332*** (0.026)
Panel C: Using the standard deviation of regional elevation as an instrument variable (the second stage of 2SLS regression, N = 18708)		
	GPT	$GPTC$
$InstrumentedMSENT$	1.553*** (0.315)	0.372** (0.145)

Table 5. Cross-sectional Differences in the Effects of Media Sentiment on Green Innovation

This table partitions firms into subsamples and re-estimates the regressions in Table 2 for the subsamples. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. In Panel A, we divide the sample into two groups according to whether a firm's CEO serves as the chairman of the board. In Panel B, we split the sample into two groups based on auditor quality ($BIG4$), which equals one if a company's auditor is one of the international top 4 audit firms, and zero otherwise. In Panel C, the sample is split based on the median of provincial development. Provincial development is measured using GDP. Firms have stronger (weaker) governance when their CEO does not (does) serve as the chairman of the board, when their auditor is (is not) one of the international top 4 audit firms, or when their provincial GDP is above (below) the sample median. In Panel D.1, we divide the sample into two groups according to the sample median of provincial punitive measures, which are defined as the proportion of firms penalized for environmental pollution. In Panel D.2, the sample is split based provincial incentive measures. In Panel E.1, we divide the sample into two groups according to the sample median level of firm size. A firm is defined as less financially constraint if firm size above sample median ($FC\_SIZE = 1$, *LFC*), and otherwise is more financially constraint ($FC\_SIZE = 0$, *MFC*). In Panel E.2, we split the sample into two groups based on dividend indicator. A firm is defined as less financially constraint ($FC\_DIVIDEND = 1$, *LFC*), and otherwise is more financially constraint ($FC\_DIVIDEND = 0$, *MFC*). All regressions include industry and year fixed effects. All regressions include the control variables as those in Table 2 and their coefficients are not tabulated. Detailed variable definitions are in the legend of Table 2. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Dependent variables	(1)	(2)	(3)	(4)
	GPT		$GPTC$	
Panel A: *Partitioning the sample based on firms' internal governance mechanisms* (CEO, N_{weak}=5342, N_{strong}=13164)				
	Weak	Strong	Weak	Strong
$MSENT$	0.770***	0.224**	0.282***	0.120**
	(0.184)	(0.112)	(0.090)	(0.052)
Panel B: *Partitioning the sample based on firms' external governance mechanisms* ($BIG4$, N_{weak}=17628, N_{strong}=1080)				
	Weak	Strong	Weak	Strong
$MSENT$	0.306***	0.350	0.147***	0.224
	(0.092)	(0.555)	(0.044)	(0.279)
Panel C: *Partitioning the sample based on province-level external governance mechanisms* (GDP, N_{weak}=9814, N_{strong}=8765)				
	Weak	Strong	Weak	Strong
$MSENT$	0.965***	0.090	0.303***	0.128**
	(0.161)	(0.122)	(0.080)	(0.052)
Panel D: *Partitioning the sample based on institutional background*				
Panel D.1: *Partitioning the sample based on punitive measures* ($Punish$, N_{high}=9328, N_{low}=9380)				
	Low	High	Low	High
$MSENT$	0.775***	0.023	0.318***	0.025
	(0.132)	(0.143)	(0.084)	(0.031)
Panel D.2: *Partitioning the sample based on incentive measures* ($Incent$, N_{low}=9057, N_{high}=8854)				
	Low	High	Low	High
$MSENT$	0.182	0.613***	0.101*	0.285***
	(0.124)	(0.172)	(0.060)	(0.077)
Panel E: *Partitioning the sample based on firms' financial constraints*				
Panel E.1: *Partitioning the sample based on firms' size* (FC_SIZE, N_{MFC}=9354, N_{LFC}=9354)				
	MFC	*LFC*	*MFC*	*LFC*

$MSENT$	0.482***	0.384**	0.138***	0.223***
	(0.107)	(0.158)	(0.048)	(0.074)

Panel E.2: *Partitioning the sample based on firms' dividend indicator* ($FC\_DIVIDEND$, N_{MFC}=3754, N_{LFC}=14477)

	MFC	*LFC*	*MFC*	*LFC*
$MSENT$	0.356*	0.404***	0.312***	0.144***
	(0.183)	(0.112)	(0.086)	(0.053)

Table 6. Different Types of Media

This table reports the regression results for the negative sentiment-innovation relation for different types of media. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. Columns (1) and (2) report the results for the relation between market-oriented media sentiment ($MSENT\_MARKET$) and firm green innovation using GPT and $GPTC$ as dependent variables, respectively. Market-oriented media sentiment, $MSENT\_MARKET$, is calculated as $(Negative\ words - Positive\ words)/(Negative\ words + Positive\ words)$ using Chinese Market-oriented newspapers, including *21th Century Business Herald*, *First Financial Daily*, and *China Business Journal*. Columns (3) and (4) report the results where the regression equation include market-oriented media sentiment ($MSENT\_MARKET$) and stated-controlled media sentiment ($MSENT$) together, using GPT and $GPTC$ as dependent variables, respectively. All regressions include industry and year fixed effects. All regressions include the control variables as those in Table 2. Detailed variable definitions are in the legend of Table 2. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Dependent variables	(1) GPT	(2) $GPTC$	(3) GPT	(4) $GPTC$
$MSENT$			0.380***	0.163***
			(0.095)	(0.045)
$MSENT\_MARKET$	-1.396***	-0.506***	-1.327***	-0.477***
	(0.409)	(0.180)	(0.409)	(0.180)
$SIZE$	0.256***	0.067***	0.256***	0.067***
	(0.010)	(0.005)	(0.010)	(0.005)
LEV	0.013	-0.047**	0.017	-0.045**
	(0.049)	(0.021)	(0.049)	(0.021)
ROA	0.084	0.026	0.086	0.027
	(0.069)	(0.027)	(0.069)	(0.027)
TOP	-0.003***	-0.001***	-0.003***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
CFO	-0.234***	-0.144***	-0.233***	-0.144***
	(0.082)	(0.036)	(0.082)	(0.036)
AGE	-0.121***	-0.043***	-0.120***	-0.043***
	(0.011)	(0.005)	(0.011)	(0.005)
MB	-0.181***	-0.092***	-0.179***	-0.091***
	(0.058)	(0.025)	(0.058)	(0.025)
$R\&D$	5.835***	1.252***	5.793***	1.234***
	(0.918)	(0.268)	(0.916)	(0.267)
$CONS$	-5.993***	-1.614***	-5.672***	-1.477***
	(0.360)	(0.174)	(0.364)	(0.171)
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
N	18708	18708	18708	18708
Adjusted R-squared	0.155	0.034	0.156	0.035

Table 7. Different Types of Green Innovation

This table reports the regression results for the relation between negative sentiment and different types of corporate green innovation output. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. Column (1) presents the results using green invention patent (GPT_INV) as dependent variable. GPT_INV is defined as the natural logarithm of one plus the number of green invention patents. Column (2) presents the results using green utility patent (GPT_UTI) as dependent variable. GPT_UTI is defined as the natural logarithm of one plus the number of green utility patents. Column (2) presents the results using green utility patent (GPT_UTI) as dependent variable. GPT_CLAIM is defined as the natural logarithm of one plus the number of green patent claims. All regressions include industry and year fixed effects. Detailed variable definitions are in the legend of Table 2. The standard errors corrected for heteroskedasticity are displayed in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Dependent variables	(1) GPT_INV	(2) GPT_UTI	(3) GPT_CLAIM
$MSENT$	0.275***	0.282***	0.316***
	(0.078)	(0.070)	(0.073)
$SIZE$	0.221***	0.158***	0.119***
	(0.009)	(0.008)	(0.006)
LEV	-0.104**	0.054	0.109***
	(0.040)	(0.036)	(0.037)
ROA	0.034	0.026	0.187***
	(0.056)	(0.050)	(0.052)
TOP	-0.002***	-0.001***	-0.003***
	(0.000)	(0.000)	(0.000)
CFO	-0.181***	-0.166***	-0.284***
	(0.067)	(0.059)	(0.064)
AGE	-0.082***	-0.087***	-0.081***
	(0.009)	(0.008)	(0.008)
MB	-0.217***	-0.072*	-0.017
	(0.048)	(0.040)	(0.043)
$R\&D$	4.831***	2.835***	4.152***
	(0.757)	(0.462)	(0.695)
$CONS$	-4.124***	-2.947***	-2.217***
	(0.180)	(0.156)	(0.124)
Industry fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
N	18708	18708	18708
Adjusted R-squared	0.155	0.123	0.120

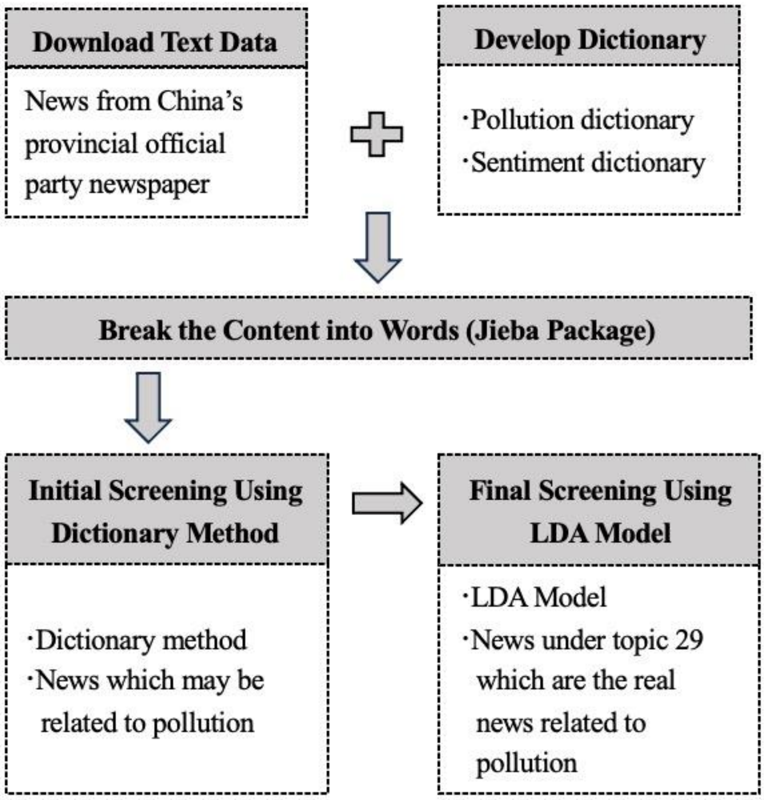

Figure 1. The Approach of Extracting Pollution-news

Appendix A

Table A1. Sample Distribution over Time and across Industries

This table presents the sample distribution by year and across industries. The sample consists of firms listed on the Shanghai or Shenzhen Stock Exchanges between 2007 and 2018. Panel A reports the distribution of sample observations by year. Panel B presents the distribution of sample observations across industries as classified according to one-digit China Securities Regulatory Commission Industry Code (2012).

Panel A: Distribution of firms by year

Year		Number of observations	Percentage
2007		317	1.69%
2008		480	2.57%
2009		774	4.14%
2010		1156	6.18%
2011		1476	7.89%
2012		1714	9.16%
2013		1734	9.27%
2014		1805	9.65%
2015		1965	10.50%
2016		2138	11.43%
2017		2429	12.98%
2018		2720	14.54%
Total		18708	100.00%
Panel B: Distribution of observations by industries			
A	Agriculture, forestry, and fishing	216	1.15%
B	Mining	364	1.95%
C	Manufacturing	13854	74.05%
D	Electricity, heat, gas, water	399	2.13%
E	Construction	551	2.95%
F	Wholesale and retail	514	2.75%
G	Transportation, storage, and postal	374	2.00%
H	Accommodation and catering	24	0.13%
I	Information, software, and information technology	1263	6.75%
K	Real estate	287	1.53%
L	Leasing and business services	147	0.79%
M	Scientific research and technical services	192	1.03%
N	Water conservancy, environment, and public facilities management	144	0.77%
O	Residential services, repairs and other services	11	0.06%
P	Education	10	0.05%
Q	Health and social work	31	0.17%
R	Culture, sports, and entertainment	155	0.83%
S	Conglomerates	172	0.92%