The consequences of environmental big data information disclosure on hard-to-abate Chinese enterprises’ green innovation

Cheng Zhang\textsuperscript{a}, Bo Zhou\textsuperscript{b,*}, Qunwei Wang\textsuperscript{c}, Yaode Jian\textsuperscript{a}

\textsuperscript{a} School of Economics, Nanjing University of Finance and Economics, Nanjing 210023, PR China
\textsuperscript{b} School of Finance, Nanjing University of Finance and Economics, Nanjing 210023, PR China
\textsuperscript{c} College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, PR China

\textbf{Abstract}

Big data technology may improve the quality of environmental information disclosure and address the challenge of information asymmetry. This study adopts a difference-in-differences strategy to investigate the effectiveness of the big data-based environmental information disclosure system launched by the Chinese Ministry of Environmental Protection in 2013. The results reveal that, the system significantly stimulates green innovation in hard-to-abate enterprises with an entrepreneurial orientation. A mechanism exploration suggests that environmental big data information disclosure drives officials to implement environmental policies aimed at hard-to-abate enterprises and raises public concern regarding environmental quality; however, the mechanism of public attention is found to be ineffective in stimulating corporate green innovation in the sample period. This study identifies a useful channel for policymakers to develop green innovation by enhancing the current system of environmental big data information disclosure.

© 2024 The Authors. Published by Elsevier España, S.L.U. on behalf of Journal of Innovation & Knowledge. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

\textbf{Introduction}

Official and public governance are essential factors that drive corporate green innovation (\textit{Wu et al., 2023}; \textit{Xu et al., 2023}; \textit{Zhou & Ding, 2023}); however, it is difficult for governments and the public to collect accurate environmental information and evaluate local environmental quality. Environmental information asymmetry may cause biased judgment from governments and the public regarding environmental quality (\textit{Barwick et al., 2019}), diminishing government and public capabilities to sufficiently drive corporate green innovation. Therefore, addressing environmental information asymmetry is a notable issue for driving green innovation to achieve mutually beneficial outcomes between environmental governance and economic growth (\textit{Aghion et al., 2021}).

In theory, information disclosure is implemented to address the problem of information asymmetry faced by governments and the public (\textit{Greenstone et al., 2022}). In reality, numerous enterprises in many countries are already required to disclose environmental performance information (\textit{Döring et al., 2023}). However, this type of environmental information is disclosed annually or semiannually and can be easily greenwashed by enterprises (\textit{Delmas & Lessem, 2014}; \textit{Xia et al., 2023}; \textit{Zhou & Wang, 2023}). Therefore, inadequate information quality may reduce the credibility of environmental information disclosure (\textit{Shi et al., 2021}; \textit{Tu et al., 2019}).

To address the disadvantage in information quality, China implemented a new approach for environmental information disclosure to drive corporate green innovation, applying big data technology to collect and disclose city-level air pollution measurements. Air pollution, particularly at a city-level, is consistently the environmental issue that garners the highest level of concern and the highest frequency of policy implementation in China (\textit{Wang et al., 2021}; \textit{Zeng et al., 2023}). The Chinese Ministry of Environmental Protection (now the Ministry of Ecology and Environment) began an air monitoring program called the New Air Quality Standard Monitoring Program (NAQSMP) in 2012, establishing an air quality-based environmental big data information disclosure system for cities to publicly disclose detailed data on various types of real-time air monitoring to the community that was rolled out in three batches (\textit{Greenstone et al., 2022}). Through building this big data system, governments and public are able to understand real-time environmental information. Thus, this study argues that the NAQSMP – a system of environmental big data information disclosure – may adapt and alleviate the environmental information asymmetry between officials, public and enterprises, and motivate governments and public to drive corporate green innovation.

\textsuperscript{*} Corresponding author.
\textit{E-mail address: abc@163.com} (B. Zhou).

https://doi.org/10.1016/j.jik.2024.100474
2444-569X/© 2024 The Authors. Published by Elsevier España, S.L.U. on behalf of Journal of Innovation & Knowledge. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)}
The reality of China also poses some challenges to our investigation on the NAQSMP. One is that a large number of Chinese enterprises, especially hard-to-abate enterprises, still have relatively inadequate awareness and capacity of green innovation (Jiang & Bai, 2022). Thus, the effect of environmental big data information disclosure on corporate green innovation may merely exist in enterprises with specific characteristics; the specific characteristics should be developed through some theories, before the empirical investigation in this study. Another challenge is a wide variety of policy instruments issued by Chinese governments to govern corporate green innovation. For instance, because of environmental big data information disclosure, one hard-to-abate enterprise may receive a limit-to-growth policy, as well as a policy with a green growth idea. The differences in these policy instruments should be covered in this study.

Addressing the above challenges, this study took hard-to-abate enterprises of China as the sample, and had two objectives. First, this study seeks to measure the effect of the NAQSMP – a system of environmental big data information disclosure – on corporate green innovation. Second, this study aims to analyze how the system affect corporate green innovation, from the channels of official and public governances.

This study enriches the related literature in two aspects. First, this study fills the research gap of nexus between green innovation and environmental big data information disclosure. Existing literature mainly discusses the information asymmetry in some other fields, such as supply chain and financial intermediaries, and analyzes the corresponding information disclosure’s effect on enterprises (e.g. Niu et al., 2022; Yuan et al., 2021). Other literature is mainly on firm-level common environmental information disclosure (Tsang et al., 2022; Ren et al., 2023). Through setting China’s NAQSMP as an example, this study connects green innovation with environmental big data information disclosure.

Second, this study contributes to a growing body of literature on effect of environmental big data information disclosure, through investigating the corresponding mechanism. Being similar with Greenstone et al. (2022), a strand of literature has discussed the effect of China’s environmental big data information disclosure on pollution governance (e.g. Tan et al., 2021). However, little knowledge of the specific mechanisms has been found. This study aims to clarify what action took by governments and public, to respond to the environmental big data information disclosure and govern enterprises.

In what follows, this study first introduces institutional background of the studied environmental big data information disclosure in China, and reviews literature and develops hypotheses. Next, this study introduces methodology and data, presents empirical results and discussion, and summarizes the corresponding conclusion.

**Institutional background, literature review, and hypotheses development**

**Institutional background**

To meet the public need, the Chinese Ministry of Environmental Protection built the NAQSMP in 2012. Before this, along with the more and more serious air pollution, the demand for better air quality and better data on air pollution has increased. However, the only provided readings of the Air Pollution Index (API) cannot scientifically reflect the situation of air pollution. Accordingly, the Chinese government revised its air quality standards and established the NAQSMP in 2012. The NAQSMP replaced the API with a stricter standard of PM10. Three additional pollutants were added to the NAQSMP measurement, including PM2.5, O3 and CO. An automated national monitoring network was established to collect environmental information, and report it publicly.

As a system of environmental big data information disclosure, the NAQSMP is phased in different prefectural cities sequentially. The first phase of the NAQSMP requires cities in the major population and economic centers of China (totally 74 cities) to install automated monitoring stations by 2012. Averagely, there are approximately six stations to collect macro-level big data information on air pollution in each city. Concentrations of different air pollutants at monitoring stations are updated hourly and are available simultaneously and publicly on the Ministry of Environmental Protection website, provincial and municipal environment bureau websites, as well as on a large number of mobile apps and third-party websites. Then, the second phase requires all national model cities for environmental protection (totally 116 cities) to install stations by 2013; the third phase requires the remaining 177 cities to be completed by 2014. There are approximately four stations in each second-phase and third-phase city, and the completely same system of environmental big data information disclosure with the first-phase cities.

**Literature review and hypotheses development**

Based on the principal-agent theory (Ross, 1973), the relative lack of environmental participation by governments and public may be due to the asymmetry of environmental information with enterprises. Governments and public lack reliable sources of environmental information to understand the environmental situation, and then cannot properly exercise their right to govern enterprises.

For enterprises, if the system of environmental big data information disclosure is launched and the asymmetry is alleviated, the related information barrier is broken. The hard-to-abate enterprises anticipate that governments and public will soon be concerned about corporate environmental performance and even take action. To avoid to be published by governments and public, enterprises may choose to improve their environmental performances through some strategies, including developing green innovation. Some empirical researches have already verified the above argument (Foulon et al., 2002; Downar et al., 2021; Tomar, 2023). For instance, Using China’s 2008 mandatory corporate social responsibility (CSR) reporting policy as a quasi-natural experiment regarding environmental information disclosure, scholars found that after implementation of this policy, mandatory CSR reporting firms show substantially higher green innovation performance than non-CSR reporting firms (Ren et al., 2023).

Thus, the following hypothesis is proposed for empirical testing: **H1a**: The environmental big data information disclosure drives corporate green innovation.

However, not all hard-to-abate enterprises are good at developing green innovation. The motivation for technology innovation may influence impact of environmental information disclosure on corporate green innovation. Facing the pressure from environmental information disclosure, hard-to-abate enterprises can choose several strategies to address, including reducing production, investing in environmental protection, and green innovation (Cao et al., 2021; Zhang et al., 2022a, 2019; Zhou & Zhang, 2023). Compared to the former two strategies, green innovation is always with high risk. In this case, entrepreneurship may be a positive factor for green innovation. Entrepreneurship refers to new and established companies’ active search for new businesses and new products, and is not solely restricted to new investments (Harms et al., 2010). Both Bouncken et al. (2016) and Zhang et al. (2022b) found a positive relationship between entrepreneurship and innovation.

Thus, the following hypothesis is proposed for empirical testing: **H1b**: The environmental big data information disclosure drives corporate green innovation in enterprises with an entrepreneurial orientation.

Owing to environmental big data information disclosure, governments are able to reliably judge the environmental quality. Once the quality deteriorates, governments can strengthen the enforcement of environmental policies, which in turn promotes corporate
environmental protection. For example, treating the release of the Pollutant Information Transparency Index (PITI) in 2008 as a quasi-natural experiment regarding environmental information disclosure, scholars found that the disclosure significantly improved local governments’ environmental legislations, the size of environmental agencies and the frequency of environmental enforcements (Shi et al., 2021).

The encouraged environmental policies may be differential among enterprises. If officials note a deteriorating environmental quality through the disclosure system, they may firstly strengthen enforcements to larger-size enterprises and state-owned enterprises (SOEs). Larger-size enterprises may be with higher pollution emissions; regulating these enterprises, rather than smaller-size ones, is more likely to more quickly lead to improved environmental quality. The political connections between officials and SOEs may help a smooth environmental enforcement.

In addition to enforcements, governments may also issue the other two types of policies, including discharge fees and environmental subsidies, to larger-size enterprises and SOEs. For Chinese local governments, developing the local economy is an objective that goes hand in hand with local environmental governance. Local officials are more willing to provide economic help to larger-size enterprises because they generally have a higher productivity (Bai et al., 2020). The political connections of SOEs may provide access to governmental economic help (Fan et al., 2007; Zhou & Ding, 2023). Theoretically, compared to enforcements, discharge fees and environmental subsidies have more notable potential to help hard-to-abate enterprises to achieve win-win outcomes between environmental governance and economic activity (Meckling & Allan, 2020), and may thus be issued by local governments to larger-size enterprises and SOEs.

Environmental policies encouraged by environmental information disclosure may drive corporate green innovation. According to the Porter Hypothesis, both environmental enforcement and discharge fee are stress to firms, and thereby can promote corporate green innovation by incentivizing firms to change their production routines or increase efficiency, in order to release the stress (Porter & Van der Linde, 1995; Lian et al., 2022). As for environmental subsidy, it may enrich firms’ capital in the areas of green innovation. Green innovation depends on significant financial investments and incentives (Brav et al., 2018; Manso, 2011). Environmental subsidies will alleviate corporate financial constraint and improve capacity in risk-taking (Hu et al., 2017; Montmartin & Herrera, 2015; Stiglitz, 2015).

Thus, the following hypotheses are proposed for empirical testing:

H2a: The environmental big data information disclosure encourages environmental policies.

H2b: The environmental big data information disclosure drives environmental policies in larger-size enterprises and SOEs.

H3: The environmental big data information disclosure indirectly drives corporate green innovation through environmental policies.

Being similar to governments, public is also able to reliably judge the environmental quality through environmental big data information disclosure. Public has the demand for environmental quality (Deschenes et al., 2017), and once they note a deteriorating environmental quality through the disclosure, the willingness to pay attention to environmental issue may significantly rise (Tu et al., 2020; Wang & Zhou, 2021).

There are mixed views regarding whether public attention drives corporate green innovation. Some positive arguments suggest that public can be regarded as stakeholders; thus, public attention can be included in the environmental regulation system as informal regulation (Kathuria, 2007), and can affect corporate strategies including green innovation. Similar conclusions have been found in investigations in some countries (Da Motta, 2006; Cheng & Liu, 2018). Other arguments indicate that public attention does not have a powerful instrument like governmental enforcements, thus public cannot impact corporate strategy in green innovation. Some investigations in Mexico and China show that public attention is not, or not widely the determinant of corporate green technologies (Blackman & Kidegaard, 2010; Zhou & Ding, 2023).

Thus, the following hypotheses are proposed for empirical testing:

H4: The environmental big data information disclosure encourages public attention.

H5a: The environmental big data information disclosure indirectly drives corporate green innovation through public attention.

H5b: The environmental big data information disclosure cannot indirectly drive corporate green innovation through public attention.

Fig. 1 presents the structure of all hypotheses in this study.

Methodology

Main variables

This study uses a dummy variable to represent the launch of environmental big data information disclosure. Regarding the specific time points for cities to be included in the NAQMP, the first phase of the NAQMP requires cities in the main population and economic centers of the country to complete by 2012, the second phase requires all national model cities for environmental protection to complete by 2013, and the third phase requires the remaining cities to complete by 2014. Based on the latest administrative division codes in China, this study sequentially adjusts the list of cities implementing environmental big data information disclosure, and finally obtains two batch numbers for the sample of 104 cities in this paper: 53 in the first batch (2012) and 51 in the second batch (2013). Note that, according to the arrangement of the NAQMP, the first year of environmental big data information disclosure is the installation and debugging stage of monitoring stations, and the following year will be officially online. Therefore, this study sets the dummy variable of environmental big data information disclosure to 1 for the first batch, 0 for the second batch. For the first batch of a total of 53 cities since 2013, to 0 for the second batch of a total of 51 cities since 2014, to 1 for all cities since 2015, and to 0 for the remaining cases.

The natural logarithm of number of green patent authorizations is used to measure green innovation. Detailed information on
Table 1
Main variables in this study.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Variable names</th>
<th>Variable measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent</td>
<td>Info</td>
<td>Environmental big data information disclosure</td>
</tr>
<tr>
<td>Dependent</td>
<td>Ghino</td>
<td>Green innovation</td>
</tr>
<tr>
<td></td>
<td>Entrep</td>
<td>Entrepreneurial orientation</td>
</tr>
<tr>
<td></td>
<td>Enforcement</td>
<td>Environmental enforcement</td>
</tr>
<tr>
<td></td>
<td>Fee</td>
<td>Discharge fee</td>
</tr>
<tr>
<td></td>
<td>Subsidy</td>
<td>Environmental subsidy</td>
</tr>
<tr>
<td></td>
<td>Public</td>
<td>Public attention</td>
</tr>
</tbody>
</table>

Sources: The authors developed this table based on methodology in this study.

enterprises’ green patent authorizations is provided by the State Intellectual Property Office of China. This study selected the authorized patents related to alternative energy, energy saving, waste management, environment management, and renewable power generation to represent green patents (Zhang et al., 2022b).

As for environmental policies, this study includes and measures three types of policy instruments (Zhou & Ding, 2023). The first one is environmental enforcement, that is measured by the natural logarithm of number of enforcement records on an enterprise’ environmental issues. The second policy instrument is discharge fee, and is measured by the natural logarithm of amount of pollutant discharge fees paid by an enterprise. The third one is environmental subsidy, and is captured by the natural logarithm of amount of government environmental subsidies received by an enterprise.

This study uses the Baidu Index of public search ‘environmental pollution’ (huanjingwuran) to measure public attention. Baidu is the largest search engine service provider in China. By analyzing the big data information of its Baidu search, it obtains the statistical data of a keyword being searched, and then forms the Baidu Index of this keyword. Following the method adopted by El Ouadhiri et al. (2021) that using Google Index to measure public attention in the United States, this study obtains the indicators of public attention of each city in each year.

As for entrepreneurial orientation, this study calculated the sum of the four subcategories, including investment intensity in R&D, brand equity, organizational capital and human capital (He et al., 2020). To be specific, investment intensity in R&D is calculated as the ratio of an enterprise’s R&D expenditure to sales revenue. Investment intensity in brand equity is the ratio of an enterprise’s branding expenditure to sales revenue, and branding expenditure is calculated as 60% of advertising expenditure (He et al., 2020). Investment intensity in organizational capital is the ratio of 10% of the enterprise’s administration expense to sales revenue. According to De and Dutta (2007), in reality, a small proportion of administration expense is used to invest in organizational capital. This study selected 10% as this small proportion, as did He et al. (2020). Investment intensity in human capital is the ratio of 10% of the enterprise’s administration expense to sales revenue.

Table 1 summarizes the specific information regarding main variables.

**Empirical models**

This study applies a difference-in-differences (DID) strategy to test the hypotheses in Section 2.

The model to test H1a in this study is to regress the coefficients of environmental big data information disclosure (Info) on corporate green innovation (Ghino), and is set as Formula (1):

\[
Ghino_{it} = \beta_0 + \beta_1 \text{Info}_{it} + \beta_2 \text{Controls} + \beta_3 \text{i.City} + \rho_4 \text{i.Year} + \beta_5 \text{Firm} + \tau
\]

where \( \rho \) represent coefficients; \( \tau \) denotes a mean zero random error term; Controls is a series of controlled variables; and \( \text{i.City}, \text{i.Year}, \) and \( \text{i.Firm} \) denote city, year, and firm fixed effect, respectively. The H1a, the driving effect of environmental big data information disclosure on corporate green innovation can be proved, if \( \rho_1 \) is significantly positive.

To test H1b, the DID model is set as Formula (2):

\[
Ghino_{it} = \beta_0 + \beta_1 \text{Info}_{it} + \rho_2 \text{Info}_{it} \times \text{Entrep}_{it} + \rho_3 \text{Controls} + \beta_4 \text{i.City} + \rho_5 \text{i.Year} + \rho_6 \text{i.Firm} + \tau
\]

where Entrep represents entrepreneurial orientation. The H1b, the driving effect of environmental big data information disclosure on corporate green innovation in enterprises with an entrepreneurial orientation, can be proved, if \( \rho_2 \) is significantly positive.

The models to test H2a and H4 are similar to Formula (1); the only difference is that, when testing H2a and H4, this study replaced the dependent variable in Formula (1) into environmental policies and public attention, respectively. The model to test H2b is similar to Formula (2); the only difference is that, when testing H2b, this study replaced the variable Entrep in Formula (2) into the dummy variables representing larger-size enterprises and SOEs. To be specific, this study regards the enterprise with a sale size beyond the sample average as the larger-size one (Zhang & Zhou, 2023).

To test H3, this study establishes Formulas (3) and (4):

\[
\text{Enforcement}_{it} = \beta_0 + \beta_1 \text{Info}_{it} + \rho_2 \text{Controls} + \beta_3 \text{i.City} + \rho_4 \text{i.Year} + \beta_5 \text{i.Firm} + \tau
\]

\[
\text{Ghino}_{it} = \beta_0 + \beta_1 \text{Info}_{it} + \rho_2 \text{Enforcement}_{it} + \rho_4 \text{i.City} + \rho_5 \text{Controls} + \beta_4 \text{i.City} + \rho_6 \text{i.Year} + \rho_7 \text{i.Firm} + \tau
\]

where Enforcement represents one type of environmental policies in this study, that is environmental enforcement. The H3 with regard to environmental enforcement – the environmental big data information disclosure indirectly drives corporate green innovation through environmental enforcements – can be proved, if both \( \rho_2 \) in Formulas (3) and (4) are significantly positive. The H3 with regard to the other two types of environmental policies in this study can also be proved, if Enforcement in the above two formulas is replaced by Fee and Subsidy. In addition, the H5a – the environmental big data information disclosure indirectly drives corporate green innovation through public attention – can be proved, if Enforcement in the above two formulas is replaced by Public, and both \( \rho_1 \) in the two formulas are significantly positive. On the contrary, the H5b can be proved if both \( \rho_1 \) in the two formulas are insignificant.

The controlled variables (Controls) in Formulas (1) (2) (4) impact green innovation, and thus include asset size, gross liability ratio, cash flow, sales-expense ratio and capital intensity ratio of enterprises, and gross domestic product and industry ratio of city where the enterprise is located (Bu et al., 2020). The controlled variables
(Controls) in Formula (3), when the dependent variable is environmental policies, include age of enterprises, and pollutant index, environmental enforcement records, ratio of discharge fee and pollutant index, environmental expenditure, population size, ratio of secondary and tertiary industry output value, fiscal expenditure, and composite index of pollutants, of city where the enterprise is located (Zhou & Ding, 2023). The controlled variables (Controls) in Formula (3), when the dependent variable is public attention, include public attention on water pollution, gross domestic product per capita, size of population, amount of Internet users, land area, and composite index of pollutants of city where the enterprise is located.

The Pearson correlation coefficients of variables in all models were tested. All correlation coefficients remained below 0.5, indicating no significant multicollinearity problem for all adopted models.

Sample and data sources

In this study, the sample includes panel data of hard-to-abate listed companies on the Chinese Shanghai and Shenzhen Stock Exchange, located in the key cities of environmental protection in China, over the nine-year period of 2011–2019. The former Ministry of Environmental Protection defined 16 sectors as the pollution-intensive ones in China, and we regard them as hard-to-abate enterprises in China; they include the thermal power, steel, cement, electrolytic aluminum, coal, metallurgy, chemical, petrochemical, building materials, paper making, brewing, pharmaceutical, fermentation, textile, tanning and mining sectors. The specific sector to which the enterprise belongs can be found in the Wind database. The key cities of environmental protection in China include 4 province-level municipalities and 109 prefecture-level cities, and were selected by the Ministry of Environmental Protection in 2007 to illustrate the importance of ecological improvements to policymakers in these cities. This study also excluded the listed companies that were on the ST/ST stock exchange. To mitigate the influence of outliers, all firm-level continuous variables were winzorized at the 1st and 99th percentiles.

The enterprises covered in this study merely include listed companies, and the time span covered here starts with 2011, since in China only listed companies are required in September 2010 to reveal their annual environmental information. This provided data this study needed, such as the discharge fees and environmental subsidies of enterprises. The recent COVID-19 crisis may confound some of the metrics, thus the 2020 financial year is excluded from this study.

Green innovation data was gathered from the State Intellectual Property Office of China, as well as IPC Green Inventory developed by World Intellectual Property Organization. The data with regard to firm-level environmental enforcement records was gathered from enterprise environmental performance database developed by the Institute of Public & Environmental Affairs (IPE). The database contains information about enterprise environmental enforcement records in China, including the name of the enterprise, cause of the enforcement, year of record, and other variables. The data related to discharge fees and environmental subsidies was collected from annual reports of sampling enterprises. This study collected the data regarding public attention from the Baidu Index. The other data related to firm-level variables was gathered from the Wind and China Stock Market Accounting Research (CSMAR) databases. The other data related to city-level variables was collected from the China City Construction Statistical Yearbook, China Urban Construction Statistical Yearbook, and China Statistical Yearbook on Environment.

The descriptive statistics of main variables are presented in Table 2. The variable Info is a dummy variable, and has the mean of 0.799, the median of 1, and the standard deviation of 0.401, showing that most of sample cities established systems of environmental big data information disclosure at the early point of sample period. Note that the variable Ghino ranges from 0 to 6.771, with the mean of 0.209 and median of 0, highlighting that a large number of enterprises have no outcome of green innovation.

The three variables of environmental policies, namely Enforcement, Fee, and Subsidy, have the same median that is 0, indicating that more than half of samples receive no environmental policy. The mean, median, and standard deviation of public attention (Public) are 72.241, 65, and 41.191, respectively; the results reveal a similar level between samples. The gap between the top level and average of entrepreneurial orientation in sampling enterprises is notable, in light of the max (13.797) and mean (2.715) of Entrep.

Preliminary results

To preliminarily examine the effect of environmental big data information disclosure on corporate green innovation, that is H1a, this study firstly regressed Formula (1), Table 3 provides the corre-

---

**Table 2**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>6846</td>
<td>0</td>
<td>1</td>
<td>0.799</td>
<td>1</td>
<td>0.401</td>
</tr>
<tr>
<td>Ghino</td>
<td>6846</td>
<td>0</td>
<td>6.771</td>
<td>0.209</td>
<td>0</td>
<td>0.621</td>
</tr>
<tr>
<td>Enforcement</td>
<td>6751</td>
<td>0</td>
<td>1.386</td>
<td>0.052</td>
<td>0</td>
<td>0.195</td>
</tr>
<tr>
<td>Fee</td>
<td>6751</td>
<td>0</td>
<td>19.225</td>
<td>1.944</td>
<td>0</td>
<td>5.075</td>
</tr>
<tr>
<td>Subsidy</td>
<td>6751</td>
<td>0</td>
<td>20.572</td>
<td>3.944</td>
<td>0</td>
<td>6.284</td>
</tr>
<tr>
<td>Public</td>
<td>6751</td>
<td>0</td>
<td>164</td>
<td>72.241</td>
<td>65</td>
<td>41.191</td>
</tr>
<tr>
<td>Entrep</td>
<td>6846</td>
<td>0.002</td>
<td>13.797</td>
<td>2.715</td>
<td>2.525</td>
<td>2.600</td>
</tr>
</tbody>
</table>

Sources: The authors calculated this table with sample data.

**Empirical results and discussion**

**Descriptive statistics**

The descriptive statistics of main variables are presented in Table 2. The variable Info is a dummy variable, and has the mean of 0.799, the median of 1, and the standard deviation of 0.401, showing that most of sample cities established systems of environmental big data information disclosure at the early point of sample period. Note that the variable Ghino ranges from 0 to 6.771, with the mean of 0.209 and median of 0, highlighting that a large number of enterprises have no outcome of green innovation.

The three variables of environmental policies, namely Enforcement, Fee, and Subsidy, have the same median that is 0, indicating that more than half of samples receive no environmental policy. The mean, median, and standard deviation of public attention (Public) are 72.241, 65, and 41.191, respectively; the results reveal a similar level between samples. The gap between the top level and average of entrepreneurial orientation in sampling enterprises is notable, in light of the max (13.797) and mean (2.715) of Entrep.
sponding regression results. The estimation in column (1) controlled year and city fixed effects but no controlled variable or firm fixed effect; both the estimations in columns (2) and (3) controlled all controlled variables, however, the one in column (3) additionally controlled firm fixed effect.

None of coefficients of Info in the three columns were significant, revealing that there is no statistical relationship between environmental big data information disclosure and corporate green innovation. Accordingly, this study cannot confirm H1a.

Robustness checks

This study tested the following robustness issues. (a) The result in Table 3 is based on the DID method. The assumption underlying the unbiased estimation of a DID method presupposes the parallel trend. This study used the Event Study method to test the parallel trend assumption of result in column (3) of Table 3 (Beck et al., 2010). (b) Since China’s 13th Five Year Plan was issued in 2016, Chinese government announced a series of policies regarding environmental protection, while these policies might bring about unexpected impacts on corporate green innovation, as well as the results in Table 3. This study dropped the samples after year 2015, and re-regressed Formula (1). (c) The notable gap between green innovation of hard-to-abate enterprises in province-level municipalities and prefecture-level cities may impact the robustness. This study dropped the samples belonging to province-level municipalities, and re-regressed Formula (1). (d) Following Yang et al. (2022), this study reselected enterprises belonging to heavy industries as hard-to-abate enterprises. To be specific, the heavy industries include thermal power, steel, cement, electrolytic aluminum, coal, metallurgy, chemical, petrochemical, building materials, and mining sectors.

The result of parallel trend test was presented in Fig. 2. All of coefficients after the establishment of environmental big data information disclosure were insignificant, illustrating the parallel trend. Table 4 shows the results of the other three robustness checks. The estimation in column (1) dropped the samples after year 2015; the one in column (2) dropped the samples belonging to province-level municipalities; the one in column (3) reselected hard-to-abate enterprises. All coefficients of Public were insignificant. These checks proved the robustness of the estimations in Table 3.

Results in enterprises with entrepreneurial orientations

Based on the preliminary results, this study further examined H1b through regressing Formula (2), investigating the impact of environmental big data information disclosure on corporate green innovation in enterprises with entrepreneurial orientations. Table 5 reports the related results. The estimation in column (1) controlled all controlled variables, and year, city, and firm fixed effects; the one in column (2) dropped the samples after year 2015; the one in column (3) dropped the samples belonging to province-level municipalities.

All coefficients of Info x Entrep were significantly positive, revealing a positive impact of environmental big data information disclosure on corporate green innovation in enterprises with entrepreneurial orientations; accordingly, H1b was proved. In light of the coefficient of Info x Entrep in column (1), enterprises with entrepreneurial orientations responded to environmental big data information disclosure by scaling up green innovation, which resulted in a 5.0% increase in outcome of green innovation, that equals approximately 0.08 green patent authorizations per year.

Entrepreneurship helps enterprises anticipate the necessity of achieving coordinated development of environmental governance and economic activity under normalized environmental information disclosure; entrepreneurship also represents the advantages of enterprises’ experience and investment in research and development, human capital, and organizational management (Dess & Lumpkin, 2005). Therefore, enterprises with entrepreneurial orientations endeavor to develop green innovation, turning challenges into opportunities, and striving to realize green development under normalized environmental information disclosure.

Environmental big data information disclosure, environmental policies, and corporate green innovation

After identifying the positive impact of environmental big data information disclosure on corporate green innovation in some hard-to-abate enterprises, this study continued to investigate the role of
The impacts of environmental big data information disclosure on environmental policies.

<table>
<thead>
<tr>
<th></th>
<th>(1) Enforcement</th>
<th>(2) Fee</th>
<th>(3) Subsidy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>-0.032</td>
<td>-0.007</td>
<td>0.278</td>
</tr>
<tr>
<td></td>
<td>(-1.52)</td>
<td>(-0.01)</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Controlled variables and constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effect</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>City fixed effect</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>6751</td>
<td>6751</td>
<td>6751</td>
</tr>
<tr>
<td>A-R²</td>
<td>0.106</td>
<td>0.198</td>
<td>0.203</td>
</tr>
</tbody>
</table>

Notes: t values are shown in brackets; ***, **, and * indicate statistical significance at 1 %, 5 %, and 10 % levels, respectively. Sources: The authors developed this table using empirical results.

The impacts of environmental big data information disclosure on corporate green innovation through environmental policies.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>0.008</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Info×Larger</td>
<td>0.055***</td>
<td>(3.58)</td>
</tr>
<tr>
<td>Info×SOE</td>
<td>0.020**</td>
<td>(3.58)</td>
</tr>
<tr>
<td>Enforcement</td>
<td>0.057**</td>
<td>(1.98)</td>
</tr>
<tr>
<td>Fee</td>
<td>0.055**</td>
<td>(2.05)</td>
</tr>
<tr>
<td>Subsidy</td>
<td>0.001</td>
<td>(0.45)</td>
</tr>
</tbody>
</table>

Notes: t values are shown in brackets; ***, **, and * indicate statistical significance at 1 %, 5 %, and 10 % levels, respectively. Sources: The authors developed this table using empirical results.

Environmental policies regarding the impact, through regressing Formulas (3) and (4). The results of Formula (3), with regard to impacts of environmental big data information disclosure on environmental policies, are presented in Table 6. All coefficients of Info were insignificant, indicating that in the whole sample, there is no impact of environmental big data information disclosure on environmental policies; this study cannot confirm H2a.

This study further examined H2b, that is the positive impacts of environmental big data information disclosure on environmental policies in larger-size enterprises and SOEs. The estimation results are provided in Table 7. All of coefficients of Info×Larger and Info×SOE were significant and positive. The results highlight the positive impacts of environmental big data information disclosure on environmental policies in larger-size enterprises and SOEs; thus, H2b was proved.

In light of larger-size enterprises, coefficients of Info×Larger were 0.014, 0.778, and 1.551, when dependent variables were environmental enforcement, discharge fee, and environmental subsidy, respectively. The coefficients illustrate that environmental big data information disclosure leads to a 1.4 %, 77.8 %, and 155.1 % rise in the three types of environmental policies, respectively, indicating a notable encouragement to governmental actions induced by the disclosure, especially to discharge fees and environmental subsidies.

Based on the proved H2b, this study further examined H3, investigating whether environmental big data information disclosure indirectly drives corporate green innovation through environmental policies. The corresponding results are provided in Table 8. The coefficients of Enforcement and Fee were significantly positive, revealing that environmental big data information disclosure indirectly drives corporate green innovation through environmental enforcements and discharge fees. However, the coefficients of Subsidy were insignificant, showing that environmental subsidies cannot play a role between environmental big data information disclosure and corporate green innovation.

Environmental big data information disclosure provides accurate, real-time, and broad environmental information, so that local governments are able to understand local environmental quality and thus apply environmental policies to hard-to-abate enterprises. In this case, larger-size enterprises considerably impact local environmental quality, and can be interpreted as a rational choice of local environmental policies.

Political connections between SOEs and the government may make SOEs more susceptible to government attention, and this pressure may motivate SOEs to take a positive attitude toward environmental big data information disclosure and to be open to accepting environmental policies (Buijs & Verbeke, 2003; Zhang, 2017).

As for the two encouraged environmental policies, namely environmental enforcement and discharge fee, a series of studies have identified their positive impacts on green innovation (Yin et al., 2023; Zhang et al., 2020). The conclusion that environmental subsidy

Table 7

The impacts of environmental big data information disclosure on environmental policies in larger-size enterprises and SOEs.

<table>
<thead>
<tr>
<th></th>
<th>(1) Enforcement</th>
<th>(2) Fee</th>
<th>(3) Subsidy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Info</td>
<td>-0.039***</td>
<td>-0.380</td>
<td>-0.466</td>
</tr>
<tr>
<td></td>
<td>(-1.83)</td>
<td>(-0.77)</td>
<td>(-0.77)</td>
</tr>
<tr>
<td>Info×Larger</td>
<td>0.014**</td>
<td>0.778***</td>
<td>1.551***</td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(5.85)</td>
<td>(9.44)</td>
</tr>
<tr>
<td>Info×SOE</td>
<td>0.005***</td>
<td>0.177***</td>
<td>0.317***</td>
</tr>
<tr>
<td></td>
<td>(3.06)</td>
<td>(4.66)</td>
<td>(6.72)</td>
</tr>
<tr>
<td>Controlled variables and constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effect</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>City fixed effect</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>6751</td>
<td>6751</td>
<td>6751</td>
</tr>
<tr>
<td>A-R²</td>
<td>0.107</td>
<td>0.202</td>
<td>0.213</td>
</tr>
</tbody>
</table>

Notes: t values are shown in brackets; ***, **, and * indicate statistical significance at 1 %, 5 %, and 10 % levels, respectively. Sources: The authors developed this table using empirical results.
cannot drive corporate green innovation, is consistent with the practice. According to the rule with regard to environmental subsidies in China, the subsidies should be strictly provided and used in expenditure in purchasing environmental protection equipment, instead of investing in green innovation.

Environmental big data information disclosure, public attention, and corporate green innovation

After identifying the role of environmental policies, this study continued to investigate the role of public attention between environmental big data information disclosure and corporate green innovation. This study firstly examined H4, that is the positive impact of environmental big data information disclosure on public attention. The estimation result is presented in Table 9. The coefficient of *Info* in column (1) was 0.229 and significant, illustrating a 22.9% increase in public attention on environmental issues because of environmental big data information disclosure. In reality, the sample data tells that since the year environmental big data information disclosure began, there is a totally 34.6% increase in public attention on environmental issues. Thus, this study confirmed H4. Environmental big data information disclosure can explain approximately 66.2% (22.9%/34.6%) of the increase in public attention: it is essential during China’s rising public attention on environmental issues in recent years.

The result is in consistent with some previous studies. Greenstone et al. (2022) also studied the impact of environmental information disclosure in China, and found that the Baidu searches for “air purifiers” by public increased by about 20% as a result of the environmental information disclosure, which is a good reflection of the rising public attention. The magnitude of the increase is similar to the result of this study. Wang and Zhou (2021) also found that after the media publicized the environmental quality to public, public was enthusiastic to call the governmental hotline to express their increasing environmental concerns.

The NAQSMP, representing environmental big data information disclosure in this study, additionally measures PM2.5, O3 and CO, compared to the previous environmental information disclosure. The collected environmental big data information is reported publicly and real-timely. The information asymmetry between enterprises and public may be addressed, and public begin to consciously pay attention to environmental issues.

Although proving the existence of impact of environmental big data information disclosure on public attention, this study found no significant impact of environmental big data information disclosure on corporate green innovation through public attention, because the coefficient of *Public* in column (2) of Table 9 was insignificant, when the dependent variable was green innovation. Thus, H5b, instead of H5a, was proved by the empirical regression.

Note that in the estimation of column (2), to avoid the endogeneity problems between public attention and corporate green innovation, this study used an instrumental variable (IV) method. Based on the experience of Zhou and Ding (2023), the corresponding IV is the interaction of Dist and VC. The Dist represents the shortest geographical distance between a certain city and the most serious air-polluted region in China; while the VC means ventilation coefficient, reflecting the meteorological conditions that influence the speed of dispersion of pollutants in the air.

### Table 9

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Public</td>
<td>Green</td>
</tr>
<tr>
<td><strong>Info</strong></td>
<td>0.229***</td>
<td>−0.015</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(−0.29)</td>
</tr>
<tr>
<td><strong>Public</strong></td>
<td>0.633</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.19)</td>
<td></td>
</tr>
<tr>
<td>Controlled variables and constant</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year fixed effect</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>City fixed effect</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Firm fixed effect</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>936</td>
<td>6751</td>
</tr>
<tr>
<td>A-R²</td>
<td>0.999</td>
<td>0.749</td>
</tr>
</tbody>
</table>

Notes: t values are shown in brackets; ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels, respectively.

Sources: The authors developed this table using empirical results.

Conclusion

This study endeavors to untangle how environmental big data information disclosure drives corporate green innovation. Based on the developed hypotheses and a DID method, this study investigates Chinese hard-to-abate listed companies located in the key cities of environmental protection during the period 2011 to 2019.

The empirical results show that, environmental big data information disclosure drives corporate green innovation in enterprises with entrepreneurial orientations. To be specific, because of the disclosure, green innovation in enterprises with entrepreneurial orientations increases 5.0%, that equals approximately 0.08 green patent authorizations per year. This result plausibly explains the conclusions of Greenstone et al. (2022) and Tan et al. (2021). They find that China’s environmental big data information disclosure is effective to improve environmental performance: we confirm the information disclosure positively impacts corporate green innovation, and may further induces the improved environmental performance.

Governments play an essential role between environmental big data information disclosure and corporate green innovation, by implementing environmental policies. Environmental big data information disclosure encourages environmental enforcements, discharge fees, and environmental subsidies in larger-size enterprises and SOEs. Taking an example of larger-size enterprises, environmental big data information disclosure leads to a 1.4%, 77.8%, and 155.1% rise in the three types of environmental policies. Then, the encouraged environmental enforcements and discharge fees significantly drive corporate green innovation. However, although this study identified a significant and positive impact of environmental big data information disclosure on public attention, the encouraged public attention cannot drive corporate green innovation. Thus, this study fills the research gap of nexus between green innovation and environmental big data information disclosure, and contributes to a growing body of literature on effect of environmental big data information disclosure, through investigating the corresponding mechanism.

Findings from this study could shed light on a useful channel for policymakers to develop green innovation, that is establishing a system of environmental big data information disclosure. Policymakers, not only in China, could take advantage of rapidly developing big data technology to strengthen environmental disclosure system. To ensure and maximize the role of environmental big data information disclosure, this study recommends policymakers to educate firms on entrepreneurship, as merely enterprises with entrepreneurial orientations are able to achieve the positive impact of environmental big data information disclosure on green innovation. In light of the notable effect of the disclosure on public attention, policymakers and scholars need to analyze how to use public attention to drive corporate green innovation.

This study still has some limitations. The investigation regarding how environmental big data information disclosure impacts corporate resources to develop green technologies, has not been covered.
currently, and may further enrich the theory of green innovation. Besides, we only cover listed firms in this study. The results may be different in unlisted firms, and should be investigated on condition that we can collect solid data.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Social Science Foundation of China (grant number 22AZD095), the National Natural Science Foundation of China (grant numbers 71973062, 72303094), and the University Philosophy and Social Science Research Project of Jiangsu Province (grant number 2023SJV0278).

References


