Received: 16 April 2023 | Revised: 11 December 2023 | Accepted: 20 December 2023

DOI: 10.1002/bse.3677

RESEARCH ARTICLE

Business Strategy and the Environment WILEY

Corporate digitalization and green innovation: Evidence from textual analysis of firm annual reports and corporate green patent data in China

Liting Fang[1] | Zhaohua Li[2]

[1]School of Economics and Management, Fuzhou University, Fuzhou, China

[2]Department of Financial and Business Systems, Faculty of Agribusiness and Commerce, Lincoln University, Lincoln, New Zealand

Correspondence
Zhaohua Li, Department of Financial and Business Systems, Faculty of Agribusiness and Commerce, Lincoln University, PO Box 85084, Lincoln, New Zealand.
Email: zhaohua.li@lincoln.ac.nz

Funding information
This work was financially supported from the National Natural Science Foundation of China (Grant No. 72003034), National Social Science Foundation of China (Grant No. 22BTJ026), Natural Science Foundation of Fujian Province in China (Grant No. 2021J01113).

Abstract

We investigate the effect of corporate digitalization capabilities on green innovation among Chinese-listed firms. Using a panel dataset of 2908 companies from 2011 to 2020, we use textual analysis and entropy weighting on corporate annual reports to construct a yearly corporate digitalization index. Our findings show that corporate digitalization promotes green innovation, as evidenced by patent applications and grants. This relationship is stronger for firms with fewer financial constraints and in provinces with strong intellectual property protection. We also find the national digital policy of the "Internet Plus" strategy has a stronger positive effect on corporate green innovation for corporations with a higher degree of digitalization. Our results are robust to various alternative measures, econometric models, and different samples.

KEYWORDS

corporate digitalization, financing constraints, green innovation, green patent applications and grants, local intellectual property protection

1 | INTRODUCTION

Sustainability has become a critical issue in today's world, with growing concern about climate change, resource depletion, and the negative impact of economic activities on the environment. The search for effective solutions becomes ever more urgent. In this context, green innovation stands out as an effective way to reduce environmental risk, pollution and other negative impacts of resource use compared with relevant alternatives (Castellacci & Lie, 2017). However, the journey toward green innovation is complex, given the length and uncertainty of innovation. Therefore, recognizing and understanding the factors that drive green innovation are therefore essential.

As we navigate the challenges of green innovation, the digital era presents an opportunity. In the digital economy, information is presented in bits (Goldfarb & Tucker, 2019). The process of converting information into a digital format is a technical process of capturing, digitizing, and processing data. This is often called *digitization*. Scanning a printed report to create a PDF file or converting handwritten patient records into electronic health records in a hospital, are examples of

Abbreviations: CNRDS, China Research Data Service Platform. The source of corporate patent data used in the study; CSMAR, China Stock Market and Accounting Research database. One of the data sources for the study; Digi, Natural logarithm of the actual summed number of digitalization words. This term is used to measure the total digitalization disclosure in the study; Digi_MDA, Ratio of total digitalization disclosures to the number of corporate Management Discussion and Analysis (MD&A) disclosures. It represents a specific measure of digitalization in the MD&A section of annual reports; DV, Digitalization-related vocabularies. Used in the construction of the firm-year level digitalization index; Intangible, Natural logarithmic form of digital intangible measure. Derived from keywords associated with digital technology in the Intangible Asset section of annual reports; MD&A, Management Discussion and Analysis. A section of a company's annual report where management discusses various aspects of the company, including its financials and future outlook.

digitization. Another example of digitization is companies changing from using paper-based invoicing to generating and sending digital invoices. When raw data are turned into structured data for a specific purpose, the result is not very useful until firms work out how to use it and take action (Veldkamp, 2023). This leads to *digitalization*, which refers to the use of digital technology to create value with a broader impact on society (Gobble, 2018). Digitalization can thus serve as driver to enhance green innovation by enabling more efficient, environmentally friendly business processes. At the firm level, digitalization can involve automating jobs that were performed manually to improve operational efficiency or using digital technologies such as cloud computing or artificial intelligence to manage inventory and production, forecast customer demand, and make predictions on price movements. In short, digitization is a process and digitalization is the transformation of the socioeconomic environment through processes of digital artifact adoption, application, and utilization (Gradillas & Thomas, 2023).

Building on this background and rooted in the resource-based view, firms establish competitiveness from resources and abilities that are valuable, rare, and not easily copied, both tangible and intangible (Barney, 1991). The natural resource-based theory expands this view by incorporating environmental factors in response to escalating ecological concern (Hart, 1995; Hart & Dowell, 2011). Bridging these two, the dynamic capability theory highlights the necessity for firms to adapt and evolve with technological advances to sustain competitiveness. In this business context, digitalization capabilities, as focused on by Annarelli et al. (2021), are identified as dynamic capabilities. They provide opportunities for increased functionality, reliability, efficiency, and optimization to create customer value. This paper proposes that leveraging digitalization capabilities enables firms to enhance green innovation. We examine the underlying mechanisms for this relationship by exploring how factors at national, provincial, and firm level influence that relationship.

Authors have examined the role of corporate digitalization in driving green innovation. Wei and Sun (2021) show digital technologies can support the development and implementation of environmentally friendly manufacturing practices. Li and Shen (2021) find corporate digitalization affects green innovation through internal controls and institutional ownership. Peng et al. (2022) find there is a U-shape relationship between digitalization and green transformation. However, the existing studies have limitations. First, there is conceptual uncertainty and a need for a clear, robust understanding of digitalization. Gradillas and Thomas (2023) find there are 26 different distinct definitions attributed to varying degrees of digitalization. A simplistic measure of corporate digitalization used in the literature is not adequate to capture the full spectrum of digitalization. This situation calls for a more comprehensive measure and a more detailed understanding of digitalization capabilities. Furthermore, prior studies tend to rely on data from manufacturing firms (Liu et al., 2023; Wei & Sun, 2021), overlooking the fact that digitalization applies to other sectors. For instance, logistic companies apply digitalization to optimize route planning, reducing fuel consumption and emissions. In healthcare, it can lead to paperless operations and efficient resource management. In the finance sector, digitalization enables paperless transactions and remote services, reducing the footprint of physical branches. Therefore, a comprehensive measure to capture applications and their impact on green innovation is needed for all companies and across a longer sample periods.

Lastly, green innovation is more complex and novel than general innovation (Barbieri et al., 2020). This complexity requires us to examine the driving forces of green innovation at various levels, including national, local, and the firm. Although previous studies often examined the underlying mechanism in a firm such as green credit (Ning et al., 2023), internal control and institutional ownership (Li & Shen, 2021) and top management environmental attention (Sun & He, 2023), they tend to overlook the coordinated effort required across all levels to foster an environment for sustainable, innovative growth.

This paper aims to address these gaps. We use a textual analysis technique and an entropy weighting method to construct a yearly corporate digitalization index for Chinese listed firms from 2011 to 2020 to measure their digitalization capabilities. We investigate the effect of digitalization capabilities on green innovation, measured by green patent applications and grants. After documenting a positive effect, we examine national digital policy by exploring the extent to which national digital policy affects corporate green innovation, depending on the degree of corporate digitalization. We then move to the provincial level to see how the effect of corporate digitalization on green innovation is differs by the degree of local intellectual property protection. Finally, we examine a firm internal factor, that is, financial constraints.

Our paper contributes to the literature in several ways. First, we use a comprehensive digitalization-related vocabulary (DV) covering four aspects of digitalization: digital technology, platform, smart manufacturing, and modern information integration. Then we use entropy weighting in text mining to construct the annual measure of corporate digitalization capability of each firm. We not only develop a deeper understanding of digitalization over the full spectrum but also identify areas that mostly drive green innovation. Lastly, our analyses provide policy implications for a national digital economic policy, provincial intellectual property rights protection, and firm financial constraints that can be useful for policy makers. The findings highlight that sustainability is a complex issue requiring collaborative efforts from multiple stakeholders. Our results show which internal and external institutional factors promote sustainable practices.

The reminder of this paper is organized as follows. Section 2 reviews the relevant literature and theories and develops the hypotheses. Section 3 describes the data and methodology for the empirical analysis and section 4 presents the main empirical results. Section 5 details supplementary analyses and robustness test results. Section 6 presents the discussion and concludes the paper.

2 | THEORY AND HYPOTHESES DEVELOPMENT

2.1 | Corporate digitalization and green innovation

In the face of global challenges marked by rapid resource depletion and fast population growth, sustainable practices have become increasingly important. Though cities cover only about 3% of the world's land surface, they account for over 75% of global natural resource consumption. Human activities in these urban areas contribute to multiple environmental challenges that are expected to intensify because the global urban population is projected to increase by 70%–80% by 2050 (Perrotti et al., 2021). Society stands at the crossroads of environmental and economic sustainability. The sustainability challenge includes aspects of environmental, social and economic dimensions.

According to the resource-based theory, firms are entities that use resources and capabilities to achieve desired outcomes. To establish corporate competitiveness, firms must build internal structures and processes to develop resources and abilities that are valuable, rare, and not easily copied (Barney, 1991). Such resources can be physical and financial assets or intangible assets like human capital or patents and organization structure. Building on this, the natural resource-based theory then incorporates the importance of environmental considerations. Numerous studies have reported the escalation of environmental deterioration and emphasize the need for business to shift toward more adaptive and innovative eco-solutions (Ghobakhloo et al., 2021; Hart, 1995; Hart & Dowell, 2011). In this context, Barbieri et al. (2020) examine the nature and impact of green technological change. Their study finds that green patents are more complex and novel than general innovation. This complexity arises from the multifaceted challenges of addressing environmental issues that demand integrated solutions and the collaboration of multiple technological components. Their novelty is often because of a departure from traditional technological pathways in a search for sustainable solutions.

The nature of valuable, rare resources can evolve with technological advances that emphasize that a firm must possess dynamic abilities to address rapidly changing environments, especially as technological advances reshape the landscape of valuable and rare resources (Teece et al., 1986). As businesses have navigated this evolving landscape, corporate digital transformation emerges as a beacon of hope. Digital transformation is rooted in the term digitalization, which is defined as the use of digital technologies to innovate the business model, thereby creating new revenue streams and value-producing opportunities (Kohtamäki et al., 2019; Parida et al., 2019). Digitalization capabilities, as defined by Annarelli et al. (2021), are organizational dynamic capabilities that enable firms to pervasively combine digital assets and business resources and leverage digital networks to innovate products, services, and processes for organizational learning and customer value creation and manage innovation to ensure a sustained competitive advantage.

Digital capabilities have several dimensions. One dimension is the extent to which well-developed digital options (i.e., greater digitized process reach and richness and greater knowledge reach and richness) contribute to higher levels of digital technology capability (Sambamurthy et al., 2003). Another dimension is digital platform capability. "Digital platform" is digital information technology that enables businesses to share, edit, update, and homogenize data on a large scale (Cenamor et al., 2019). From this, digital platform capability is the technical competence of businesses' digital platforms to facilitate information sharing, electronic connectivity, and online partner collaboration (Wang et al., 2022). The evolution of advanced technologies has also given rise to smart manufacturing capabilities, the integration of advanced manufacturing technologies and data analytics to improve production processes' efficiency and sustainability (Xia et al., 2019). Lastly, in this age of information abundance, modern information capabilities have become crucial, allowing firms to manage, analyze, and use vast amounts of data to drive decision-making and planning (Rai et al., 2006). As shown by Annarelli et al. (2021), digitalization capabilities are more than just static assets. They embody the essence of dynamic capabilities that enable firms to nimbly adapt, innovate, and reconfigure their resources to address the unique complexity and novelty green innovation requires.

Studies have shown several business benefits of digitalization. Ramirez Lopez et al. (2019) find the Internet of Things, and its application, can lower energy consumption by 20% in the healthcare sector. Not only direct energy consumption cost, but also cost saving in human capital productivity can be achieved. Cette et al. (2022) find the use of digital technologies improves a firm's labor productivity by about 23% and its total factor productivity by about 17%. The employment of information and communications technology specialists and the use of big data lower the labor costs by 2.5%. The savings in energy consumption and human costs resulting from digitalization release resources for green innovation development. Digitalization, with its inherent dynamism, fosters agility and rapid innovation in firms (Sambamurthy et al., 2003).

As digital transformation progresses, can companies use the tools to develop capabilities to address the pressing environmental challenges? This is where the idea of green innovation enters. Green innovation, defined as new or improved product, process, technology or practice innovation to avoid or mitigate environmental damage, stands at the forefront of such practices (Rennings, 2000). Improved information collection and processing can help reduce the uncertainties of green innovation. An important public good is information on the performance of a technology, customer preferences and market structure. For instance, when people have incomplete information about the scale of energy savings, a firm is unlikely to invest in the technology (Popp et al., 2010). With the help of digital technology, companies can better understand and meet the demand for green products and services when they have access to more and accurate information on customer preferences (Lenka et al., 2017). Improved information on regulations and policies, such as tax credits, subsidies and other incentives for green innovation, can help companies make more informed decisions about where to invest and how to structure their investments (Xia et al., 2022). Improved information collection and processing can also help companies understand the risks of new technologies becoming obsolete or the risks of regulatory changes. By leveraging agility and innovation stemming from digitalization capabilities, firms can more effectively monitor, manage, and innovate their processes, leading to the development and implementation of green innovations.

Together, it becomes evident that firms equipped with digitalization capabilities, when employed dynamically, are poised to increase green innovation. Such innovation not only addresses current the environmental challenges but also offers a sustainable competitive advantage in the marketplace. We form our first hypothesis as follows:

H1. Corporate digitalization has a positive effect on corporate green innovation.

2.2 | National digitalization policy and green innovation

In 2015, the Chinese government announced its "Internet Plus" strategy that aims to promote digitalization and drive economic growth by incorporating emerging digital technologies into traditional industries.[1] The policy emphasizes the development of big data, cloud computing, and the Internet of Things (IoT), and the improvement of e-commerce, online finance, and digital media. In addition, the government has invested heavily in the development of the country's infrastructure to support the growth of digital technology and the digital economy.

This "Internet Plus" strategy is the foundation framework to enhance the digital infrastructure across the nation. The strategy includes several core areas: the integration of the retail sector, service virtualization, the evolution of mobility, the digitization of social life, and the digitalization of industrial IoT and supply chain management, that collectively contribute to digital urbanization (Bu et al., 2021). Cities are building infrastructure that uses sensors and IoT to monitor structural health, manage waste, and ensure the efficient operation of urban systems. This has led to increased corporate investment in green technologies and processes like electric vehicles and renewable energy. Corporations are motivated to innovate green products and services to meet the urban sustainability standards and ongoing market demand. The massive-scale implementation of this national strategy compels the traditional manufacturing and service industries to reinvent their production methods, business models, and profit mechanisms. For instance, "industry 4.0," also known as the integration of digital technology into production, improves energy efficiency and decreases waste through intelligent, automated operations (Jamwal et al., 2021).

However, it is essential to consider the varying degrees of digitalization among corporations. Although some may have comprehensive digital integration, others may be at an early stage. Access to an integrated society where information systems across various domains are interconnected through financial technology (FinTech), mobile internet, e-commerce platforms, and supply chains, is beneficial. Corporations with an advanced degree of digitalization can better use resources. The widespread use of FinTech helps alleviate financial constraints by opening up new funding opportunities for corporations that, in turn, facilitate new green products and services (Ding et al., 2022). In summary, the large-scale national rollout of the "Internet Plus" strategy has laid a strong foundation from which corporates can leverage. Therefore, corporations with higher levels of digitalization have better green innovation than corporations with a lower level of digitalization. We propose our second hypothesis as follows:

H2. The national digital policy of "Internet Plus" strategy has a stronger positive effect on corporate green innovation for corporations with a higher degree of digitalization.

2.3 | Local intellectual property protection (IPR), corporate digitalization and green innovation

We now turn our attention to a local factor that may affect the relationship between digitalization and green innovation. Specifically, if H1 is supported, it is helpful to examine whether the effectiveness of the digitalization is influenced by regional protection of the green innovation. The literature shows that good legal institutions have a positive influence on economic outcomes in financial markets (La Porta et al., 2013). China's effort to improve IPR protection is highlighted by the establishment of three specialized IP courts and the creation of 46 IPR protection centers aimed at offering quick, collaborative services with local government to address difficulties in evidence collection and IPR enforcement costs (Huang, 2017). Good intellectual property rights protection increases corporate innovation (Fang et al., 2017) because IPR protects the competitive advantage of innovating firms, promotes technology transfer and provides a mechanism for inventors to monetize their innovation outputs. A good legal framework thus addresses innovation's externality problem.

There is evident regional disparity in the level of IPR protection across China, with more developed districts offering higher protection levels than less developed ones. Each province has a distinct enforcement effect (Yao & Rao, 2009). In provinces with strong IPR protection, corporations are more likely to invest in digitalization for green innovation because of the reduced risk of imitation and higher chance of receiving a return.

In provinces where IPR is weak, the double externality problem in environmental innovation, i.e., negative externalities from environmental pollution and positive externalities from technology, becomes pronounced (Nordhaus, 2021). Negative externalities arise when the cost of environmental pollution is not reflected in the market price of goods and services, resulting in a failure to internalize polluters' costs. This leads to overconsumption of natural resources and increased pollution. Positive externalities arise when the benefits of technology are not captured by the inventor, leading to inadequate realization of environmental innovation benefits. This can contribute to underinvestment in green innovation, because companies may not see the financial benefits of investing in new technologies or processes. If local government fails to effectively combat intellectual property infringement, there is an increased risk of green innovation being

[1]The official Chinese government document outlining the Internet Plus strategy is written in chapter 26 of the "Outline of the Thirteenth Five-Year Plan for National Economic and Social Development of the People's Republic of China (2016-2020)."

stolen by competitors, which discourages innovative activity. Based on these arguments, we present the third hypothesis as follows:

H3. The effect of corporate digitalization on corporate green innovation is smaller in provinces with poor intellectual property rights protection.

2.4 | Financial constraint, corporate digitalization, and green innovation

After considering the influence of regional IPR protection, we shift our focus to an internal firm factor, the role of financial constraint. Innovation is characterized by uncertainty, a high failure rate, and time-consumption. Financial constraint has been widely acknowledged as a significant barrier to firms' innovation activity (Hall & Lerner, 2010). In addition, financial constraint may disproportionately affect smaller firms or those in emerging markets, where access to capital is often more limited (Beck et al., 2008).

If H1 is supported, then we will proceed to examine whether the efficacy of digitalization in promoting green innovation is influenced by varying degrees of corporate financial constraint. Green innovation has stricter requirements in terms of finance, technology, and staff quality compared with typical innovation activity. This elevated standard requires substantial financial assistance. Firms with financial constraints may be unable to fully capitalize on their digitalization efforts, potentially limiting their capacity to invest in green innovation (Hottenrott & Peters, 2012). This challenge is further attenuated by the capital-intensive nature of digitalization. The successful integration of digital technologies, such as IoT, AI, and big data analytics, into environmental sustainability practices requires both initial investment and ongoing financial support for maintenance and upgrades.

The necessity for financial resources is further highlighted by Chwiłkowska-Kubala et al. (2023), who demonstrate that the availability of financial resources can influence the readiness for digitalization. Furthermore, Acharya and Xu (2017) find publicly listed firms in external finance dependent industries invest more in R&D and have a better patent portfolio than private counterparts, highlighting the positive role of financing in innovation.

Financial constraints extend their impact beyond just limiting a firm's ability to spend on R&D. They also affect a firm's capacity to attract and retain talented employees and commercialize innovative products or services, among other things. These limitations can negatively affect a firm's ability to engage in collaborative partnerships that are often important in driving green innovation. Therefore, we propose the fourth hypothesis as follows:

H4. The effect of corporate digitalization on green innovation is smaller for firms with financial constraints than for their counterparts without financial constraints.

The conceptual framework of digitalization's effect on corporate green innovation and the underlying mechanisms are illustrated in Figure 1.

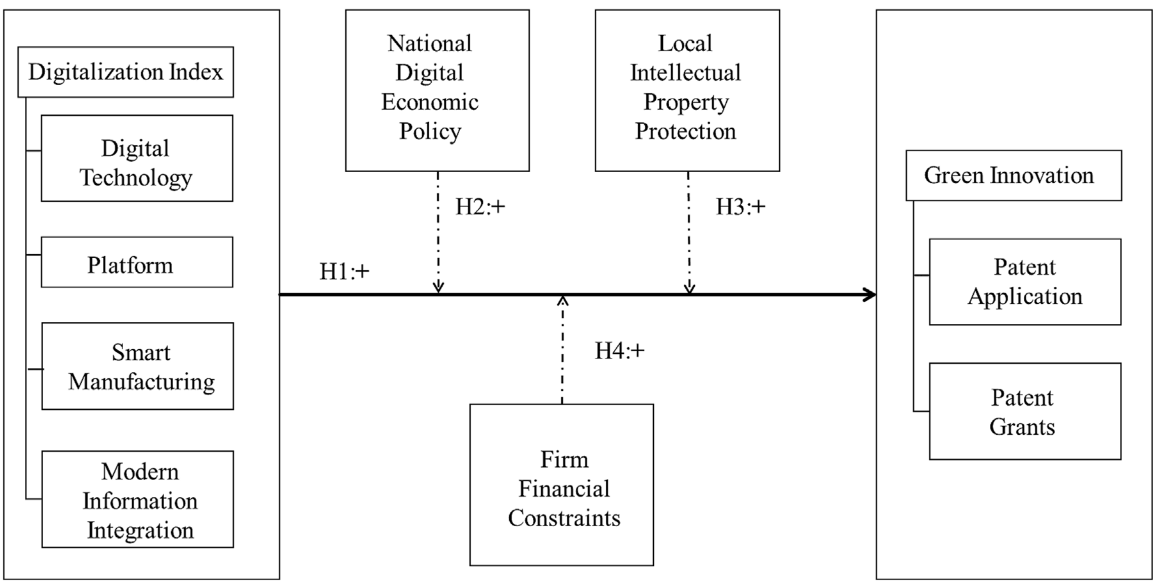


FIGURE 1 Conceptual model of the study. (a) Corporate digitalization in China (2011–2014). (b) Corporate digitalization in China (2015–2020).

3 | DATA AND METHODOLOGY

3.1 | Sample

We select our sample from all A-share listed Chinese companies from 2011 to 2020.[2] Following previous studies that used accounting and firm characteristic data (Chen et al., 2023; Liu et al., 2023), we screen the raw data according to the following criteria: (1) exclude listed companies flagged with ST or PT (which are Special Treatment firms and Particular Transfer firms) and firms in the financial sector[3]; (2) exclude companies with asset-liability ratios exceeding [0,1]; (3) exclude firms with missing key variables such as total operating revenue, R&D investment and Tobin's q value; (4) exclude listed companies with fewer than 3 years of listing and/or 3 years of continuous operation. To reduce the impact of outliers on the empirical results, continuous variables are winsorized at the 99% and 1% percentile. Robust standard errors were used to correct the standard errors of all regression estimates. The process produces a database that is both robust and representative for the study's objectives. It is about to reduce the impact of extreme values or sources of bias, ensuring data completeness. The final sample is 2908 listed companies with a total of 19,907 unbalanced firm-year observations.

The corporate patent data used in this study are from the China Research Data Service Platform (CNRDS) database. Macro data are from the China Statistical Yearbook of the National Bureau of Statistics for each year; the rest of the data are from the China Stock Market and Accounting Research (CSMAR) database.

3.2 | Measures

3.2.1 | Corporate digitalization

Constructing our firm-year level digitalization index involves three steps: (1) use a comprehensive digitalization-related vocabulary as the reference for textual analysis; (2) match the digitalization-related vocabularies (DV) from a firm's annual reports; and (3) construct the index using entropy weighting. We use the key words used by Zhao (2021) and Zhuo and Chen (2023). There are four categories: digital technology capability, platform capability, smart manufacturing capability, and modern information integration capability. Then we match the key words with a firm's annual report to count the keyword frequency. The full word list and count data are presented in Table A1. Finally, entropy weighting is used to calculate the digitalization index for each firm per year. The calculation method is as follows.

Since we have four categories of digitalization words in the list, the initial information matrix for firm i is

$$x_{i,j} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,4} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,4} \end{bmatrix} \tag{1}$$

where x_{ij} represents the value (i.e., counts of the keywords in the word list of Table A1) of the jth subcategories in firm i with $i=1,2,\cdots,n$, and $j=1, 2, 3, 4$. The subcategories are digital technology, platform, smart manufacturing, and modern information integration. The first step is standardization of the measured values based on Equation (2). The data need to be standardized to eliminate the influence of different dimensions.

$$y_{ij} = \frac{x_{ij} - \min\{x_{1j},\cdots,x_{nj}\}}{\max\{x_{1j},\cdots,x_{nj}\} - \min\{x_{1j},\cdots,x_{nj}\}} \tag{2}$$

where $\max\{x_{1j},\cdots,x_{nj}\}$ and $\min\{x_{1j},\cdots,x_{nj}\}$ are the maximum and minimum values of the jth dimension in *firm i*, respectively.

Step 2 Calculate the weight of the jth indicator in the firm i:

$$p_{ij} = \frac{y_{ij}}{\sum_{i=1}^{n} y_{ij}}, i=1,2,\cdots,n, j=1,2,3,4 \tag{3}$$

Step 3 Calculate the j entropy value of the indicators:

$$e_j = -k\sum_{i=1}^{n} p_{ij} \cdot ln(p_{ij}),\ j=1,2,3,4 \tag{4}$$

where $k = 1/\ln(n) > 0, e_j \geq 0$.

Step 4 Calculate the difference coefficient:

$$d_j = 1 - e_j, \quad j=1,2,3,4 \tag{5}$$

Step 5 Calculate the weights of each indicator

$$w_{ij} = \frac{d_j}{\sum_{j=1}^{4} d_j}, \quad j=1,2,3,4 \tag{6}$$

Step 6 Calculate the comprehensive index for each firm:

$$F_i = \sum_{j=1}^{4} w_{ij} y_{ij}, \quad i=1,2,\cdots,n \tag{7}$$

The final composite number is the digitalization index that measures the degree of digitalization of each firm. It is our measure to proxy for digitalization capability. We then repeat all

[2]A-shares are stocks of Chinese firms listed on the Shanghai and Shenzhen exchanges, traded in Renminbi, mainly for local investors. B-shares, listed on the same exchanges, are quoted in foreign currencies and were created primarily for foreign investors but are now accessible to locals. Meanwhile, H-shares represent mainland Chinese companies listed on the Hong Kong Stock Exchange, traded in Hong Kong Dollars, and available to foreign investors. All represent Chinese firms but differ in listing location, currency, and investor accessibility.

[3]The ST (Special Treatment) and PT (Particular Transfer) designations in the Chinese stock market are set by the China Securities Regulatory Commission (CSRC). The designations serve as a warning to investors about the company's financial stability. PT designation indicates a higher level of financial distress compared to ST.

steps every year to get yearly data. The frequency count of the *j*th subcategories represents the subcomponents of digitalization.

3.2.2 | Green innovation

Based on existing research (Barbieri et al., 2020; Guo et al., 2023; Liu et al., 2023), we use the number of green patents (the sum of those independently and jointly obtained) applications and patent grants by listed companies as the proxy for corporate green innovations (*Green_Inno*). CNRDS provides a patent classification of each patent. It can be matched to the Green Patent list released by the State Patent Statistics Bureau and the World Intellectual Property Organization (WIPO) to identify green patents.[4] Then we further construct the green *invention* patent applications and grants ($Green\_Invention^{app}$, $Green\_Invention^{grant}$) and green *utility model* patent applications and grants ($Green\_Utility^{grant}$, $Green\_Invention^{grant}$) according to patent type. Though green patent applications reflect the volume of green innovation efforts, the granted patents signify the quality, because they have passed rigorous examination. By considering both applications and grants, we offer a comprehensive representation of green innovation in Chinese firms. To avoid losing firm-year observations with zero patents, we add one to the actual patent counts before taking the natural logarithm.

3.2.3 | Intellectual property rights (IPR) protection

Local IPR is calculated as the proportion of patent litigation filings to patent applications in a given province. Provinces are then classified based on the annual median ratio: those above the median have strong IPR protection and vice versa. Firms located in provinces with high (low) IPR are considered to operate under a strong (weak) IPR protection environment.

3.2.4 | Financial constraints

Financial constraint is a widely studied concept in finance literature and refers to the limitations or restrictions that a firm faces in accessing financing. It can arise from various factors such as limited availability of capital, high cost of debt and equity financing, poor credit rating, and limited access to financial markets. The literature has used many methods to measure financial constraints such as the KZ index (Kaplan & Zingales, 1997) and WW index (Whited & Wu, 2006)., Each has its own strengths and weakness (Hadlock & Pierce, 2010). Many models to estimate financial constraints rely on endogenous financial choices, therefore, Hadlock and Pierce (2010) recommend the use of only firm size and firm age, two exogenous firm characteristics, to identify constrained firms because financial constraints fall sharply as young and small firms begin to mature. Following Hadlock and Pierce (2010), a financial constraint measure is constructed using Equation (8):

$$SA_{i,t} = 0.043 \times Size_{i,t}^2 - 0.04 \times Age_{i,t} - 0.737 \times Size_{i,t} \tag{8}$$

To distinguish the firm size variable in Equation (8) from our main Equation (9), we use the natural logarithm of book assets to proxy for *size*. *Age* is the number of years the firm has been listed.

The SA index is designed to capture the degree of financial constraint faced by firms, with lower (or more negative) values indicating fewer constraints and higher (or less negative) values implying more constraints. A firm is coded annually as a low financially constrained firm when its SA index is equal to or less than the median SA index and coded as a high SA firm otherwise.

3.2.5 | Other variables

The following firm and corporate governance variables are selected as control variables because prior literature shows they affect corporate innovation. We also include firm size (*Firm Size*), state ownership (*SOE*), leverage (*Lev*), cash ratio (*Cash*), profitability (ROA), Tobin's q value (*Tobin's Q*), equity concentration (*Top10*), R&D investment (*R&D*) and local economic development (*GDP*). The definitions of these variables are given in Table A2.

3.3 | Models

To test Hypothesis 1, we estimate the relationship between green innovation and *digitalization* with the following model:

$$\text{Green\_Innovation}_{i,t} = \alpha_0 + \alpha_1 \text{Digitalization}_{i,t-1} + \gamma' X_{i,t} + \mu_t + \delta_f + \varepsilon_{i,t} \tag{9}$$

where *Green_Innovation* is the innovation output represented by *Green_Inno*, *Green_Invention*, and *Green_Utility*, respectively, for firm *i* in year *t*, and superscripts ^{app} and ^{grant} are used to indicate whether it is a patent application or patent grant variable; α_1 is the estimated coefficient of the key explanatory variable of *digitalization*; and *X* is a vector of firm level characteristics in Equation (9). We include year fixed effects (μ_t), industry fixed effects (δ_f) and $\varepsilon_{i,t}$ is the error term.

Traditional linear regression models, like OLS, estimate the mean of the response variable (*digitalization*) as a linear function of *green_innovations*, but digital capability sometimes leads to higher green innovation output and, in other cases, lower output. In that case, digital development may be associated with different outcomes at the low and high ends of *green innovations*. Quantile regression provides a good model to evaluate how digitalization is associated with the

[4]The "International Patent classification (IPC) Green Inventory" has been developed by the IPC Committee of Experts to facilitate searches for patent information related to Environmentally Sound Technologies (ESTs). These ESTs are referenced by the United Nations Framework Convention on Climate Change (UNFCCC). For the full list, please access it from here: https://www.wipo.int/classifications/ipc/green-inventory/home

distribution of *green_innovations*. To test this hypothesis, we run quantile regressions of the following form:

$$Q\tau(\text{Green\_Innovation}_t|\text{Digitalization}_t) \\ = \alpha_\tau + \beta_\tau \text{Digitalization}_t + \gamma'_\tau X_t + \varepsilon_t \quad (10)$$

where $Q\tau(\text{Green\_Innovation}_t|\text{Digitalization}_t)$ is the best linear predictor of the quantile τ of variables Green_Inno, Green_Invention, and Green_Utility, conditional on values of digitalization. We estimate Equation (10) at the 10%, 25%, 50%, 75% and 90% quantiles to show the "shape" of β. The results are presented in section 4.2 and Table 4.

For Hypothesis 2, we construct the following models.

$$\text{Green\_Innovation}_{i,t} = \theta_0 + \theta_1 \text{Digitalization}_{i,t-1} + \theta_2 \text{Policy}_{i,t-1} \\ + \theta_3 \text{Digitalization}_{i,t-1} * \text{Policy}_{i,t-1} + \gamma' X_{i,t} + \mu_t + \delta_f \\ + \varepsilon_{i,t} \quad (11)$$

where the digital economic policy (*Policy*) equals one for 2015 and beyond, and zero otherwise. We use the interaction term (*digitalization*policy*) to examine how green innovation changes for high digitalization firms after the implementation of a digital policy, compared with low digitalization firms. We hope to identify a positive spillover of digital economic policy to green innovation via corporate digitalization development; that is, we expect θ_3 to be positive. Implementation of a digital economic policy can also help to establish a causal relationship between digitalization and green innovation. In particular, θ_3 captures the changes in green innovation outcomes between high and low digitalization firms, thereby accounting for unobserved time-invariant factors that may affect both digitalization and green innovation. To test Hypotheses 3 and 4, we re-run model (9) based on subsamples, partitioned by local the intellectual property protection variable and firm financial constraints variable defined earlier.

4 | ANALYSES AND RESULTS

4.1 | Descriptive statistics

Table 1 presents the descriptive statistics of the study's main variables. The average firm submits 7.48 green patent applications each year, comprising 3.9 invention patents and 3.4 utility model patents. Average firms get 4.213 green patents granted, on average, including 1.141 inventions and 2.95 utility model patents. Green invention patents have more applications than utility model patents but lower approval rates. This shows there are more stringent criteria for invention patents.

The mean and standard deviation of digitalization are 0.03 and 0.043, respectively, with a maximum of 0.257, indicating that the digitalization level of different firms varies greatly. Figure 2 shows the average digitalization index across different provinces for our sub sample periods: 2011–2014, and 2015–2020. A darker color indicates a higher level of digitalization. Overall, it can be seen that there is a greater degree of digitalization in the second sample period and a wide range of differences across regions. The average corporate leverage ratio is close to 40%, the average return on assets is about 4%, and the top 10 shareholding is about 59.5%, on average. There are large differences in the corporate size, equity concentration and R&D investment in the sample companies, which may influence the behavior of their innovation activity.

4.2 | The results for Hypothesis 1

Table 2 reports the regression results of digitalization on total corporate green innovation and on two types of green innovation outputs (inventions and utility model). In columns (1) to (6), the estimated coefficients of the independent variables are all significantly positive at the 1% level with the different specifications of year and industry fixed effects and control variables. This indicates that a higher degree of corporate digitalization can lead to an effective increase in corporate green innovation, for both invention and utility model patents. The magnitude of the coefficients for green patent grants is lower than that for applications. For the economic significance, column (1), for example, shows one unit increase in the firm digitalization capability index leads to the number of green patent applications in the next year increasing by about 0.2774. The baseline result supports our main hypothesis.

For the control variables in Table 2, *firm size*, *state ownership*, *R&D investment*, and local economic conditions positively affect green innovation output, particularly for patent applications, indicating that larger firms and SOEs are more motivated to invent green technology because of their advantages. Firms with high R&D investment have access to sufficient funding sources for innovative projects and, therefore, increase their green innovation output. A robust local economy encourages corporate innovation in green technology. High ownership concentration is detrimental to green innovation.

Table 3 reports the digitalization capabilities in greater detail by dividing into the results into four sub-indices: *digital technology capability*, *platform capability*, *smart manufacturing capability*, and *modern information integration capability*. The results in Table 3, columns (1)–(4), show that the coefficients of all sub-indices are significant at the 1% level for green patent applications. Among them, platform capability has biggest impact on green innovation and modern information integration capability the least impact. By building a digital platform that brings together stakeholders in the green economy, firms can create a network effect that can accelerate the development and adoption of green innovation (Singh et al., 2022). In columns (5)–(8), the magnitude of the coefficients of all sub-indices is smaller in for green patent grants and only two are statistically significant: digital technology capability and platform capability. This shows those two are important for innovations in both quantity and quality. The other two do not directly contribute to innovations meeting the stringent criteria set by patent offices for granting patents.

The quantile regression results are shown in Table 4. Overall, digitalization consistently, significantly drives green innovation output

TABLE 1 Descriptive statistics of the variables.

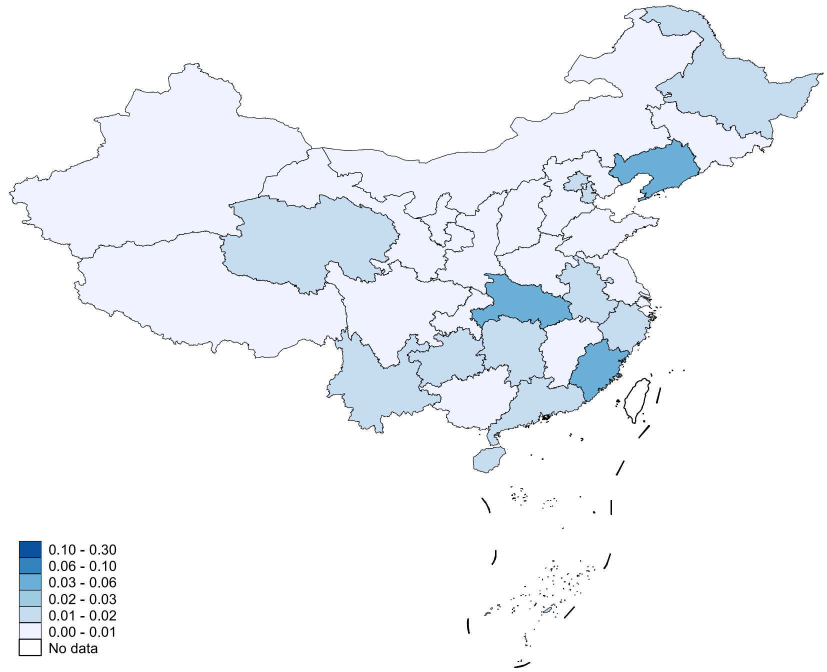
Variable	Obs	Mean	Standard deviation	Min	Median	Max
Green innovation variables						
Number of green innovation patent applications (raw)	19,907	7.480	20.880	0	1	215
Number of green invention patent applications (raw)	19,907	3.913	12.322	0	0	131
Number of green utility model patent applications (raw)	19,907	3.420	8.939	0	0	88
$Green\_Inno^{app}$	19,907	1.078	1.248	0	0.693	5.375
$Green\_Invention^{app}$	19,907	0.727	1.042	0	0	4.883
$Green\_Utility^{app}$	19,907	0.738	1.015	0	0	4.489
Number of green innovation patents granted (raw)	19,907	4.213	11.354	0	0	109
Number of green invention patents granted (raw)	19,907	1.141	3.858	0	0	35
Number of green utility model patents granted (raw)	19,907	2.955	7.726	0	0	72
$Green\_Inno^{grant}$	19,907	0.822	1.070	0	0	4.700
$Green\_Invention^{grant}$	19,907	0.337	0.690	0	0	3.584
$Green\_Utility^{grant}$	19,907	0.677	0.972	0	0	4.290
Digitalization variables						
Digitalization index	19,907	0.030	0.043	0	0.013	0.257
Digital technology	19,907	0.693	1.013	0	0.000	4.111
Platform	19,907	0.524	0.829	0	0.000	3.434
Smart manufacturing	19,907	1.024	0.941	0	1.099	3.689
Modern information integration	19,907	0.707	0.790	0	0.693	3.219
Control variables						
Firm size	19,907	21.487	1.414	18.655	21.323	25.868
SOE	19,907	0.312	0.463	0	0	1
Leverage	19,907	0.399	0.197	0.028	0.388	0.883
Cash	19,907	0.188	0.132	0.017	0.152	0.812
ROA	19,907	0.041	0.064	−0.435	0.040	0.214
Tobin's Q	19,907	1.056	0.323	0.594	0.977	2.483
Top10	19,907	59.498	14.732	22.338	60.655	91.037
R&D	19,907	17.858	1.460	13.274	17.827	22.331
GDP	19,907	11.181	0.451	9.842	11.223	12.013

across all quantiles from Q10 to Q90, for both quantity and quality. The magnitude of the coefficients increases with quantile. This shows that the impact of digitalization on green innovation is more pronounced at higher quantile levels of green innovation outcomes. In column (2), which focusses on the impact of digitalization on green invention patent applications, the coefficient at the 90th quantile (Q90) is 3.8502. This indicates that, among the top 10% of green invention patent applications (those at the higher end of the distribution), digitalization has a notably strong, positive effect. Conversely, at the 10th quantile (Q10), the coefficient is 1.9094. This suggests that, at the lower end of the distribution (the bottom 10% of cases), although digitalization still positively influences green invention patent applications, its impact is less pronounced than at the upper end. Comparing green invention innovation with green utility model innovation in both applications and grants, we find larger coefficients in columns (2) and (5) than in columns (3) and (6). This shows that digital development has a greater effect on green invention patents than on utility patents. As invention patents are more innovative in nature than utility model patents, this is a positive, encouraging finding. The magnitude of coefficients across all quantiles is smaller than those for green patent applications, consistent with the observation from Table 2.

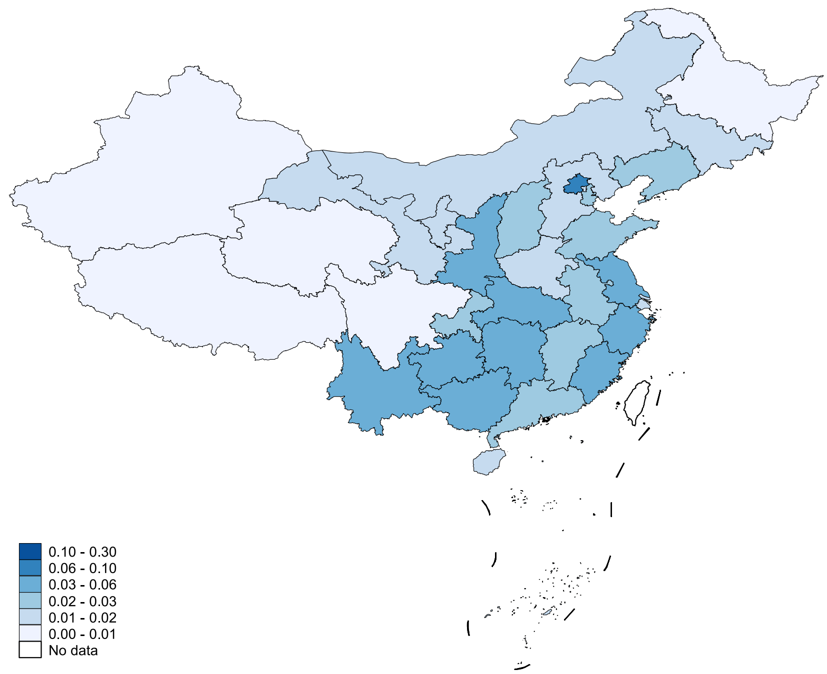
4.3 | Results for Hypothesis 2

The results in Table 5, columns (1)–(3), show that digital economic policy (θ_2) has a positive impact on total green innovation output and green invention patents, but not on green utility model patent applications. More importantly, the θ_3 is positive and significant in all three columns. This further shows how a nation's digital strategy and corporate digitalization capability can collectively enhance corporate green innovative outputs. For the results in columns (4)–(6), the coefficients of θ_2 are not significant and only one coefficient of θ_3 is significant in column 4.

FIGURE 2 Corporate digitalization increase in China 2011–2020.



Panel (a) Corporate digitalization in China (2011-2014)



Panel (b) Corporate digitalization in China (2015-2020)

4.4 | Results for Hypothesis 3

We now examine whether the effects of digitalization on green innovation vary with local intellectual property (IPR) protection. The results in Table 6, Panel A, columns (1) and (2), show that digitalization significantly increases corporate green innovation output in both groups, but the increase is greater for firms in high IPR provinces. The same conclusion can be drawn for the invention patents (columns 3–4) and utility model patents (columns 5–6). We conduct the SUR (Seemingly Unrelated Regression) test to formally examine the differences in the significance of the coefficients of the *digitalization* for the two groups.[5] The results in Panel B are like those in Panel A. The SUR

[5]When comparing regression coefficients across groups, the SUR method can be used to test for differences in the coefficients between groups. By jointly estimating the regression equations for each group, the SUR method takes into account the correlations between the error terms and provides more efficient estimates than estimating the equations separately.

TABLE 2 The overall impact of digitalization on corporate green innovation—Regression results.

Variable	(1) Green_Inno^{app}	(2) Green_Invention^{app}	(3) Green_Utility^{app}	(4) Green_Inno^{grant}	(5) Green_Invention^{grant}	(6) Green_Utility^{grant}
Digitalization$_{t-1}$	2.7742***	2.7054***	1.0330***	1.0348***	0.7056***	0.8314***
	(0.2234)	(0.1993)	(0.1856)	(0.2033)	(0.0098)	(0.1385)
Firm Size$_{t-1}$	0.1102***	0.1145***	0.0778***	0.7062***	0.2734	0.6811***
	(0.0108)	(0.0094)	(0.0089)	(0.0261)	(1.4112)	(0.0072)
SOE$_{t-1}$	0.1109***	0.1357***	0.0316*	1.0173***	0.2825	−0.0042
	(0.0205)	(0.0180)	(0.0171)	(0.3061)	(0.2436)	(0.0178)
Lev$_{t-1}$	0.6759***	0.3759***	0.6640***	0.5874***	0.0004	0.0233***
	(0.0594)	(0.0505)	(0.0491)	(0.1121)	(0.2372)	(0.0051)
Cash$_{t-1}$	0.3267***	0.3947***	0.1739***	0.2351	0.0452	0.0156*
	(0.0659)	(0.0565)	(0.0524)	(0.1875)	(0.2398)	(0.0079)
ROA$_{t-1}$	0.0587	−0.1620	0.0581	0.5561***	0.0603	0.0214
	(0.1415)	(0.1186)	(0.1134)	(0.0267)	(0.2649)	(0.0217)
Tobin's Q$_{t-1}$	−0.0403	0.0652**	−0.1312***	−0.8202***	−0.0126	0.0127
	(0.0318)	(0.0269)	(0.0259)	(0.2352)	(0.2465)	(0.0136)
Top10$_{t-1}$	−0.0026***	−0.0017***	−0.0011**	0.0625	−0.0445	0.0032
	(0.0006)	(0.0005)	(0.0005)	(0.3151)	(0.3253)	(0.0156)
R&D$_{t-1}$	0.3310***	0.2722***	0.2278***	0.4778***	0.0602	0.0224
	(0.0085)	(0.0073)	(0.0069)	(0.0865)	(0.2609)	(0.0243)
GDP$_{t-1}$	0.0700***	0.0874***	0.0579***	0.0294	0.0986	0.0123
	(0.0212)	(0.0183)	(0.0176)	(0.2926)	(0.2751)	(0.0147)
Constant	−8.1442***	−7.8085***	−5.7227***	−6.9672***	−5.7226***	3.4073***
	(0.2962)	(0.2636)	(0.2497)	(0.2751)	(0.3095)	(0.1552)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999
Adj-R^2	0.3414	0.3164	0.3120	0.3182	0.2903	0.2885
F	572.0395	450.2893	386.7911	237.3659	207.1572	241.7935

Note: The heteroscedastic robust standard deviation is reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

test shows that the impact of digitalization on green innovation under different IPR regimes is statistically significant. Hypothesis 3 is supported. Therefore, a sound local IPR protection environment can make the relationship between digitalization and green innovation stronger.

4.5 | Results for Hypothesis 4

The results in Table 7, Panel A, columns (1) and (2), show that the digitalization of firms increases green patent applications significantly for firms with both high and low financial constraints, although the increase is bigger for those with low financial constraints. The same result holds for columns (3)–(4), for green invention patent applications, and for columns (5)–(6), for green utility patent applications. The results in Panel B are similar to those in Panel A. The SUR test result shows that the impact of digitalization on green innovation for firms with different financial constraints is statistically significant. Our hypothesis 4 is supported.

5 | SUPPLEMENTARY ANALYSES AND ROBUSTNESS TESTS

5.1 | The instrumental variable approach

Our results may be biased by omitted variables and unobservable factors that influence firms' green innovation behavior. In this section, we use an instrumental variable approach (two-stage estimation) to address this concern. An ideal instrument should be correlated with

TABLE 3 Sub-indices of digitalization impact on corporate green innovation—Regression results.

Variable	(1) $Green\_Inno^{app}$	(2) $Green\_Inno^{app}$	(3) $Green\_Inno^{app}$	(4) $Green\_Inno^{app}$	(5) $Green\_Inno^{grant}$	(6) $Green\_Inno^{grant}$	(7) $Green\_Inno^{grant}$	(8) $Green\_Inno^{grant}$
Digital technology$_{t-1}$	0.1630***				0.1047***			
	(0.0099)				(0.0213)			
Platform$_{t-1}$		0.8387***				0.0743***		
		(0.0577)				(0.0165)		
Smart manufacturing$_{t-1}$			0.1781***				0.0679	
			(0.0089)				(0.0492)	
Modern information integration$_{t-1}$				0.0880***				0.0521
				(0.0105)				(0.0508)
Firm Size$_{t-1}$	0.1135***	0.1083***	0.1248***	0.1127***	0.2681***	0.2343***	2.0964***	0.0345***
	(0.0107)	(0.0109)	(0.0107)	(0.0108)	(0.0732)	(0.0791)	(0.3021)	(0.0051)
SOE$_{t-1}$	0.1106***	0.0971***	0.1070***	0.0960***	0.5603***	0.3971***	−0.0043	−0.0066***
	(0.0204)	(0.0207)	(0.0204)	(0.0206)	(0.0321)	(0.0222)	(0.0071)	(0.0001)
Leverage$_{t-1}$	0.6648***	0.6866***	0.6181***	0.6807***	−0.0151	0.0013	0.0003	0.0005
	(0.0592)	(0.0597)	(0.0591)	(0.0596)	(0.0096)	(0.0051)	(0.0023)	(0.0004)
Cash$_{t-1}$	0.2988***	0.3742***	0.3346***	0.3401***	0.0254***	0.0085***	0.0004	−0.0005
	(0.0657)	(0.0663)	(0.0655)	(0.0661)	(0.0071)	(0.0011)	(0.0012)	(0.0006)
ROA$_{t-1}$	0.0817	0.0480	0.0057	−0.0050	0.0433*	0.0324***	−0.0056	−0.0005
	(0.1411)	(0.1431)	(0.1411)	(0.1427)	(0.0221)	(0.0102)	(0.0052)	(0.0013)
Tobin's Q$_{t-1}$	−0.0569*	−0.0273	−0.0021	−0.0235	−0.0521	0.0233	0.0107	−0.0023
	(0.0317)	(0.0321)	(0.0316)	(0.0320)	(0.0474)	(0.0225)	(0.0109)	(0.0017)
Top10$_{t-1}$	−0.0024***	−0.0026***	−0.0028***	−0.0028***	−0.4633	1.5471**	0.0337	0.1053***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(1.1801)	(0.6212)	(0.2552)	(0.0341)
R&D$_{t-1}$	0.3212***	0.3407***	0.3142***	0.3357***	0.3687***	0.2408***	−0.0124	0.0177***
	(0.0085)	(0.0085)	(0.0084)	(0.0084)	(0.0761)	(0.0412)	(0.0168)	(0.0021)
GDP$_{t-1}$	0.0622***	0.0826***	0.0669***	0.0765***	0.6689***	0.4505***	0.0224	0.0109***
	(0.0211)	(0.0213)	(0.0210)	(0.0212)	(0.0761)	(0.0413)	(0.0169)	(0.0021)
Constant	−7.9670***	−8.3546***	−8.2266***	−8.3441***	−7.1442***	−6.3465***	−7.8075***	−8.3545***
	(0.2951)	(0.2979)	(0.2937)	(0.2964)	(0.2763)	(0.4052)	(0.2635)	(0.2972)

(Continues)

TABLE 3 (Continued)

Variable	(1) Green_Inno^{app}	(2) Green_Inno^{app}	(3) Green_Inno^{app}	(4) Green_Inno^{app}	(5) Green_Inno^{grant}	(6) Green_Inno^{grant}	(7) Green_Inno^{grant}	(8) Green_Inno^{grant}
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999	16,999	16,999
Adj-R^2	0.3464	0.3355	0.3507	0.3381	0.3067	0.2872	0.2982	0.2843
F	591.4432	544.3428	603.4507	553.9082	275.3228	306.7145	283.5023	235.3127

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

digitalization but uncorrelated with *green innovation*. We use a geographic instrument variable for digitalization, i.e., the average digitalization value of other firms in the same year, industry, and location (*local_digital*). We lag the instrument by 1 year. Table 8 presents the results of the IV estimations. The first stage regression result, which has *digitalization* as the dependent variable and *local_digital* as the independent variable, is shown in column (1). The coefficient of *local_digital* is positive and significant at the 1% level. The p-values for the Kleibergen-Paap rk LM statistics are all less than 0.01, strongly rejecting the non-identifiable hypothesis. The Kleibergen-Paap rk Wald F statistic is significantly larger than the critical value (i.e., 8.96) of 15% threshold of the Stock-Yogo weak instrumental variable test. This indicates that the model does not have the problem of weak instrumental variables. In the second stage, we use the fitted value of digitalization calculated from the first stage as the independent variable. The coefficients of the fitted digitalization in columns (2)–(7) are positive and significant. Our main result remains unchanged because digitalization still significantly contributes to the green innovation outputs after the use of instrumental variables.

5.2 | Fixed effect model and repeated sampling

To address possible omitted regional factors in the regression findings, province fixed effects are added to control for factors that do not vary over local resources. The regression results are shown in Table 9, Panels A and B, and columns (1) to (3). The coefficients on digitalization are significantly positive for the columns with little difference from the previous results.

Repeated random sampling using the bootstrapping method alleviates sample selection bias and yields more conservative statistical inferences. We bootstrap the sample 1000 times for the results in Table 9, Panels A and B, and columns (4) to (6). The repeated random sampling regression results are not much different from the baseline regression results, further strengthening the robustness of our findings.

5.3 | Alternative measures for firm digitalization

The literature records different ways to construct measures of firms' digitalization. Errors from different measures can affect results. In this section, we present four other measures to check the robustness of our main results. For the first three measures, we use the traditional word frequency method, which is a technique used in natural language processing (NLP) and computational linguistics to analyze the frequency of words or terms in a set of texts. It involves counting the number of times each word or term appears in a text and then ranking them according to their frequency. For our first measure, we use the words in Appendix A.1 and match them to corporate annual reports. The word frequency data are recorded and aggregated to measure the total digitalization disclosure. To avoid losing firm-year observations with zero digitalization words, we add one to the actual

TABLE 4 Quantile regression results of digitalization on green innovation.

Quantile	(1) Green_Inno^{app}	(2) Green_Invention^{app}	(3) Green Utility^{app}	(4) Green_Inno^{grant}	(5) Green_Invention^{grant}	(6) Green_Utility^{grant}
Q10	2.4627***	1.9094***	0.5497**	1.2374***	0.6590***	0.5369**
	(0.3391)	(0.3189)	(0.2719)	(0.1982)	(0.1731)	(0.2249)
Q25	2.5610***	2.1161***	0.6695***	1.5201***	0.7695***	0.6693**
	(0.2677)	(0.2592)	(0.2219)	(0.2520)	(0.2107)	(0.2471)
Q50	2.7287***	2.5193***	0.9201***	1.7391***	0.8205***	0.9156***
	(0.2257)	(0.2439)	(0.2130)	(0.2334)	(0.2013)	(0.2312)
Q75	2.9532***	3.1667***	1.3262***	2.1772***	1.0261**	1.1202***
	(0.3583)	(0.4606)	(0.4201)	(0.5460)	(0.4221)	(0.2302)
Q90	3.1583***	3.8502***	1.7172**	2.7520***	1.8172**	1.6125***
	(0.5518)	(0.7708)	(0.6806)	(0.6507)	(0.7508)	(0.3742)
Controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Observations	16,999	16,999	16,999	16,999	16,999	16,999

Note: The heteroscedastic robust standard deviations are reported in parentheses. Q10, Q25, Q50, Q75, and Q100 represent the 10th, 25th, 50th, 75th, and 100th quantiles of the distribution of green innovation variables conditional on values of *digitalization*.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

summed number before taking the natural logarithm (*Digi*). For our second measure, we repeat the above steps, but search only for words appearing in the Management Discussion and Analysis (MD&A) section of the annual report to measure MD&A disclosures, then take the ratio of the total digitalization disclosures to the number of corporate MD&A disclosures (*Digi_MDA*). The third measure uses the keywords in Qi et al. (2020) associated with digital technology, such as "software," "information system," "firewall," "cloud service," "terminal," and "domain name." These keywords are extracted from the Intangible Asset section of the annual report to create a digital intangible measure that is then converted into a natural logarithmic form as intangible (*Intangible*).

Our final measure is the Digital Financial Inclusion Index of China, developed by Guo et al. (2020). This index is based on personal account data obtained from the Ant Group. It offers insights into the adoption of FinTech at the regional level, such as Alipay account usage, third-party digital payments through Alipay, and investment, credit, and insurance services provided by Ant Group's financial technology platform. As this index is based on personal transaction accounts that are highly representative, it is often used as an indicator of regional digital financial development (Yang & Zhang, 2022). The financial inclusion index of the province where the business is located (*Dig Fin Index*) is used as a measure of a firm's digitalization level.

We use each of our measures in turn to replace *Digitalization* and report the results in Table 10. The estimated coefficients in both Panels A and B are significantly positive for the first two measures (columns (1) and (2)), but not for the latter two measures (columns (3) and (4)). The insignificant result for *Intangible* could be because this measure obtains information from only the intangible asset section and no other parts of firms' annual reports. Digitalization involves a process of adopting digital technologies to improve business processes and gain a competitive advantage. A narrow focus on one section is not helpful in generating green innovation. The lack of significance for the Digital Financial Inclusion Index could be because the index uses personal data to gauge how digital financial products are accessible to households and not firms' activities.

5.4 | Alternative econometric models

Because of the right-skewed distribution of corporate patents (Jaffe & Trajtenberg, 2002), we use Tobit regression instead of OLS with the results reported in Table 11, columns (1) to (3). Additionally, the number of patents can only be a non-negative number, so we use Poisson regression to replace OLS regression and present the results in Table 11, columns (4) to (6). Table 11 shows that the estimated coefficients of the independent variables are all significant at the 1% level, indicating that our main results are robust to alternative models.

5.5 | Different samples

There are firms in our sample that have never applied for green patents. We eliminate such firms from our sample and repeat the regressions, a step that also helps mitigate potential sample selection. The results are presented in Table 12, Panels A and B, columns (1) to (3).

TABLE 5 The impact of digital economic policy on green innovation analysis.

Variable	(1) $Green\_Inno^{app}$	(2) $Green\_Invention^{app}$	(3) $Green\_Utility^{app}$	(4) $Green\_Inno^{grant}$	(5) $Green\_Invention^{grant}$	(6) $Green\_Utility^{grant}$
$Digitalization_{t\text{-}1}$	0.1390***	0.0573**	0.0669***	0.0542***	0.0197***	0.0146***
	(0.0102)	(0.0223)	(0.0192)	(0.0112)	(0.0043)	(0.0025)
$Policy_{t\text{-}1}$	0.0819**	0.0649**	0.0263	0.0312	0.0365	0.0248
	(0.0322)	(0.0286)	(0.0355)	(0.0257)	(0.0264)	(0.0257)
Digitalization * $Policy_{t\text{-}1}$	0.1418***	0.0862***	0.0558*	0.0631*	0.0163	0.0465
	(0.0167)	(0.0286)	(0.0315)	(0.0317)	(0.0262)	(0.0503)
Firm $Size_{t\text{-}1}$	0.1672***	0.1407***	0.1222***	0.2475***	0.3343***	0.8465***
	(0.0189)	(0.0167)	(0.0149)	(0.0581)	(0.0672)	(0.1341)
$SOE_{t\text{-}1}$	0.1785***	0.1805***	0.1052***	0.3845***	0.5196***	0.0072
	(0.0403)	(0.0364)	(0.0326)	(0.0950)	(0.0202)	(0.0097)
$Leverage_{t\text{-}1}$	0.3183***	0.1947***	0.3417***	0.0032	−0.0085	0.0002
	(0.0852)	(0.0742)	(0.0693)	(0.0054)	(0.0059)	(0.0011)
$Cash_{t\text{-}1}$	0.1932***	0.2204***	0.1761***	0.0083*	0.0098***	0.0011
	(0.0722)	(0.0621)	(0.0592)	(0.0042)	(0.0021)	(0.0012)
$ROA_{t\text{-}1}$	0.4176***	0.2727**	0.2896***	0.0243**	0.0476***	0.0023
	(0.1382)	(0.1192)	(0.1055)	(0.0102)	(0.0112)	(0.0024)
Tobin's $Q_{t\text{-}1}$	0.0943***	0.1128***	0.0243	−0.0046	−0.0082	0.0027
	(0.0307)	(0.0266)	(0.0251)	(0.0231)	(0.0254)	(0.0032)
$Top10_{t\text{-}1}$	−0.0002	0.0004	0.0000	−0.1682	0.8785	0.0448
	(0.0009)	(0.0008)	(0.0008)	(0.6850)	(0.6463)	(0.0927)
$R\&D_{t\text{-}1}$	0.1962***	0.1628***	0.1363***	0.0068**	0.0021	0.0006
	(0.0130)	(0.0114)	(0.0101)	(0.0025)	(0.0025)	(0.0007)
$GDP_{t\text{-}1}$	0.0460	0.0423	0.0423	0.5239***	0.0158**	0.0064***
	(0.0365)	(0.0318)	(0.0296)	(0.0412)	(0.0063)	(0.0021)
Constant	−7.0585***	−6.0118***	−5.3467***	−6.7863***	−7.1207***	−5.2053***
	(0.5099)	(0.4424)	(0.4250)	(0.2237)	(0.5441)	(0.2541)
Year FE	NO	NO	NO	NO	NO	NO
Industry FE	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

The estimated coefficients are significantly positive for green patent applications and grants, except for green utility patent applications.

Beijing, Shanghai, and Guangzhou are three of the largest, most economically developed cities in China. They have a higher GDP per capita, and a higher standard of living than many other cities in China. These cities also have a more diverse and developed economy with a greater emphasis on services and high-tech industries, whereas many other cities in China have a more manufacturing-based economy. To avoid the possibility that our results could be driven by such an effect, we exclude firms located in Beijing, Shanghai, and Guangzhou. The results for the excluded sample presented in Table 12, Panels A and B, columns (4) to (6) show that digitalization coefficients are significantly positive for all columns, except for green utility patent grants. The results from the different samples are consistent with our main results.

6 | DISCUSSION

Rapid global resource depletion and population growth call for urgency in sustainability. In this context, green innovation is important because it involves modifying production processes or using improved technology and products. These changes aim to minimize harm to the environment thus promoting long-run, sustainable growth in economic and social contexts (Huang & Li, 2017). Building on a natural resource-based view and the dynamic capability theory, we

TABLE 6 Heterogeneity analysis of intellectual property (IPR) protection's impact on green innovation.

Panel A. Green patent applications

	(1)	(2)	(3)	(4)	(5)	(6)
	$Green\_Inno^{app}$		$Green\_Invention^{app}$		$Green\_Utility^{app}$	
Variable	High IPR	Low IPR	High IPR	Low IPR	High IPR	Low IPR
$Digitalization_{t-1}$	3.7469***	1.7423***	3.5618***	1.7948***	1.8599***	0.1724
	(0.3055)	(0.3234)	(0.2759)	(0.2860)	(0.2505)	(0.2701)
Firm $Size_{t-1}$	0.0857***	0.1288***	0.0896***	0.1347***	0.0579***	0.0940***
	(0.0140)	(0.0167)	(0.0119)	(0.0146)	(0.0117)	(0.0137)
SOE_{t-1}	0.0232	0.1727***	0.0639***	0.1710***	−0.0406*	0.0851***
	(0.0276)	(0.0312)	(0.0244)	(0.0270)	(0.0231)	(0.0262)
$Leverage_{t-1}$	0.7597***	0.6187***	0.4517***	0.3283***	0.7243***	0.6232***
	(0.0778)	(0.0916)	(0.0653)	(0.0787)	(0.0650)	(0.0747)
$Cash_{t-1}$	0.4319***	0.2109**	0.4739***	0.3071***	0.2665***	0.0661
	(0.0871)	(0.1003)	(0.0746)	(0.0859)	(0.0696)	(0.0791)
ROA_{t-1}	0.3257*	−0.1972	0.0986	−0.4021**	0.2917**	−0.1929
	(0.1794)	(0.2288)	(0.1506)	(0.1899)	(0.1423)	(0.1861)
Tobin's Q_{t-1}	−0.0684	−0.0041	0.0375	0.0997**	−0.1324***	−0.1300***
	(0.0424)	(0.0478)	(0.0359)	(0.0404)	(0.0350)	(0.0382)
$Top10_{t-1}$	−0.0066***	0.0016*	−0.0051***	0.0018**	−0.0041***	0.0021***
	(0.0008)	(0.0009)	(0.0007)	(0.0008)	(0.0007)	(0.0007)
$R\&D_{t-1}$	0.3391***	0.3182***	0.2726***	0.2699***	0.2371***	0.2119***
	(0.0110)	(0.0131)	(0.0095)	(0.0113)	(0.0090)	(0.0107)
GDP_{t-1}	−0.0841**	0.1743***	−0.0706**	0.1777***	−0.0534*	0.1335***
	(0.0355)	(0.0287)	(0.0301)	(0.0245)	(0.0298)	(0.0240)
Constant	−5.8533***	−9.7335***	−5.3657***	−9.4044***	−4.0950***	−6.7936***
	(0.4565)	(0.4261)	(0.3928)	(0.3778)	(0.3858)	(0.3582)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	9570	7429	9570	7429	9570	7429
Adj-R^2	0.3211	0.3808	0.2890	0.3630	0.2828	0.3618
F	325.2441	266.6894	239.5575	223.9490	221.5534	180.1763
SUR test for *digitalization* coefficients	−2.005*** $Chi^2 = 20.39$		−1.767*** $Chi^2 = 19.86$		−1.688*** $Chi^2 = 21.07$	

Panel B. Green patents granted

	(1)	(2)	(3)	(4)	(5)	(6)
	$Green\_Inno^{grant}$		$Green\_Invention^{grant}$		$Green\_Utility^{grant}$	
Variable	High IPR	Low IPR	High IPR	Low IPR	High IPR	Low IPR
$Digitalization_{t-1}$	1.6341***	1.0324	1.5512***	1.0873***	1.0572*	0.7742
	(0.1232)	(1.2142)	(0.2078)	(0.1756)	(0.5301)	(0.5702)
Firm $Size_{t-1}$	0.0158***	0.0092	0.2587***	0.5758***	−0.7864***	0.0093
	(0.0032)	(0.0228)	(0.0421)	(0.1673)	(0.2301)	(0.0301)
SOE_{t-1}	0.0521	0.0487***	−2.8012**	1.9277***	−0.0185	−0.0157***
	(0.0438)	(0.0101)	(1.2031)	(0.5142)	(0.0297)	(0.0032)
Lev_{t-1}	0.0025***	−0.0016***	0.0254	−0.0153	0.0045	−0.0033**
	(0.0002)	(0.0001)	(0.0243)	(0.0241)	(0.0034)	(0.0012)
$Cash_{t-1}$	0.0486***	0.0051	−0.0253	1.0232	0.0213	0.0947**

(Continues)

TABLE 6 (Continued)

Panel B. Green patents granted						
	(1)	(2)	(3)	(4)	(5)	(6)
	Green_Inno^{grant}		Green_Invention^{grant}		Green_Utility^{grant}	
Variable	High IPR	Low IPR	High IPR	Low IPR	High IPR	Low IPR
	(0.0156)	(0.0192)	(0.6812)	(0.6401)	(0.0980)	(0.0381)
$ROA_{t\text{-}1}$	0.0037*	0.0055*	−0.1386*	0.0956	−0.0041	0.0098*
	(0.0019)	(0.0027)	(0.0697)	(0.0824)	(0.0116)	(0.0050)
Tobin's $Q_{t\text{-}1}$	0.0078***	0.0087***	0.0591	0.1556**	−0.0316***	−0.0032
	(0.0021)	(0.0022)	(0.0715)	(0.0681)	(0.0101)	(0.0051)
$Top10_{t\text{-}1}$	0.0014***	0.0032***	0.0452**	0.0867***	0.0012	−0.0004
	(0.0001)	(0.0005)	(0.0201)	(0.0174)	(0.0031)	(0.0012)
$R\&D_{t\text{-}1}$	0.0023***	0.0026***	0.0261**	0.0133**	0.0012	−0.0001
	(0.0004)	(0.0006)	(0.0122)	(0.0061)	(0.0022)	(0.0001)
$GDP_{t\text{-}1}$	0.0124***	0.0234***	0.5814***	0.4506***	0.0122*	0.0074***
	(0.0011)	(0.0031)	(0.0432)	(0.0421)	(0.0061)	(0.0020)
Constant	−4.2807***	−4.6554***	−2.5363	−5.1812***	−2.7156	−4.5443***
	(1.2829)	(0.2561)	(2.0673)	(0.3157)	(2.3921)	(0.3458)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	9570	7429	9570	7429	9570	7429
Adj-R^2	0.3320	0.3256	0.2982	0.3126	0.2895	0.3036
F	212.3126	207.6894	303.3752	324.7528	272.6321	309.5107
SUR test for *digitalization* coefficients	−0.0495** Chi^2 = 12.06		−0.0456** Chi^2 = 11.83		−0.0512*** Chi^2 = 14.04	

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

investigate one important factor, corporate digitalization, on how it affects green innovation and how this effect is influenced by national, regional and internal factors. Using Chinese A-share listed companies from 2010 to 2020, we find that corporate digitalization has a positive effect on firms' green innovation. As corporate digitalization capabilities advance, the beneficial effect grows stronger. The positive effect is more pronounced in firms with fewer financial constraints and in regions with strong intellectual property protection. The national digital policy of "Internet Plus" has a stronger positive effect on corporate green innovation for corporations with a higher degree of digitalization. These results have important implications for existing knowledge and for practice and public policy.

6.1 | Contributions to the literature

First, this paper makes an important contribution to our understanding of corporate digitalization that provides dynamic capabilities (Annarelli et al., 2021) consisting of sensing, seizing, and transforming (Teece, 2007; Teece et al., 1986). Our work not only aligns with established theories but extends them by contextualizing dynamic capabilities in the field of digitalization.

In our paper, we look at four dimensions of digitalization: digital technology capability, platform capability, smart manufacturing capability, and modern information integration capability. Digital technology capability is a firm's ability to use digital technologies to create value. This may include having the necessary hardware, software infrastructure and expertise and skills needed to manage and use the technology effectively (Sambamurthy et al., 2003). It reflects the sensing aspect of dynamic capabilities, where firms identify and assess technological opportunities. Our result enriches the sensing dimension by highlighting how digital technologies contribute to a firm's ability to perceive and interpret technological trends and opportunities.

Platform capability shows a firm's ability to build and manage a digital platform to facilitate interactions or transactions between customers, suppliers and partners, such as creating or maintaining an online marketplace (Wang et al., 2022). This reflects how firms use their resources to seize opportunities. Our work illustrates how seizing opportunities in the digital economy goes beyond resource allocation and covers the creation and management of digital ecosystems.

TABLE 7 Heterogeneity analysis of financial constraints' impact on green innovation.

Panel A: Green patent applications						
	(1)	(2)	(3)	(4)	(5)	(6)
	Green_Inno^{app}		Green_Invention^{app}		Green_Utility^{app}	
Variable	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint
Digitalization$_{t-1}$	2.0720***	3.5494***	1.9392***	3.5559***	0.7854***	1.3424***
	(0.2929)	(0.3480)	(0.2564)	(0.3121)	(0.2482)	(0.2837)
Firm Size$_{t-1}$	0.1192***	0.0754***	0.1328***	0.0657***	0.0891***	0.0484***
	(0.0168)	(0.0143)	(0.0148)	(0.0121)	(0.0137)	(0.0118)
SOE$_{t-1}$	0.2385***	0.0758***	0.2473***	0.0979***	0.1291***	0.0050
	(0.0343)	(0.0259)	(0.0307)	(0.0225)	(0.0294)	(0.0213)
Leverage$_{t-1}$	0.7974***	0.5972***	0.3741***	0.3961***	0.7621***	0.6104***
	(0.0905)	(0.0784)	(0.0772)	(0.0665)	(0.0748)	(0.0646)
Cash$_{t-1}$	0.5615***	0.1141	0.5716***	0.2531***	0.3275***	0.0285
	(0.0912)	(0.0961)	(0.0790)	(0.0809)	(0.0724)	(0.0769)
ROA$_{t-1}$	0.1163	0.2320	−0.1574	0.0478	0.0273	0.3013**
	(0.2107)	(0.1883)	(0.1750)	(0.1585)	(0.1675)	(0.1521)
Tobin's Q$_{t-1}$	−0.0281	−0.0823*	0.0765*	0.0164	−0.1163***	−0.1699***
	(0.0475)	(0.0429)	(0.0403)	(0.0360)	(0.0391)	(0.0347)
Top10$_{t-1}$	−0.0039***	−0.0044***	−0.0021**	−0.0038***	−0.0022***	−0.0027***
	(0.0009)	(0.0008)	(0.0008)	(0.0007)	(0.0008)	(0.0006)
R&D$_{t-1}$	0.3474***	0.3022***	0.2916***	0.2446***	0.2372***	0.2053***
	(0.0143)	(0.0103)	(0.0127)	(0.0087)	(0.0118)	(0.0084)
GDP$_{t-1}$	0.1074***	0.0196	0.1193***	0.0443*	0.0975***	−0.0005
	(0.0305)	(0.0296)	(0.0266)	(0.0254)	(0.0258)	(0.0240)
Constant	−9.0170***	−6.1812***	−8.8735***	−5.6658***	−6.5454***	−3.9283***
	(0.4155)	(0.4158)	(0.3723)	(0.3563)	(0.3559)	(0.3417)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	8392	8606	8392	8606	8392	8606
Adj-R^2	0.4121	0.2702	0.3939	0.2408	0.3822	0.2404
F	384.6292	207.0379	300.1994	167.9679	256.5930	141.5111
SUR test for *digitalization* coefficients	1.477*** Chi2 = 10.59		1.617*** Chi2 = 16.08		0.557 Chi2 = 2.19	
Panel B: Green patents granted						
	(1)	(2)	(3)	(4)	(5)	(6)
	Green_Inno^{grant}		Green_Invention^{grant}		Green_Utility^{grant}	
Variable	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint
Digitalization$_{t-1}$	0.7923***	1.5349***	1.0139***	1.0468*	0.3765	0.6743***
	(0.1280)	(0.2571)	(0.2303)	(0.5221)	(0.3802)	(0.1724)
Firm Size$_{t-1}$	0.0082*	0.0256***	0.1092	0.3464***	0.2788***	1.2673***
	(0.0041)	(0.0041)	(0.0716)	(0.0821)	(0.0312)	(0.2460)
SOE$_{t-1}$	0.0013	0.0065***	0.4846***	0.4187***	0.0069***	0.0058***
	(0.0014)	(0.0012)	(0.0251)	(0.0254)	(0.0012)	(0.0014)
Lev$_{t-1}$	0.0657***	−0.0254	1.9298***	−0.5562	0.0618*	0.0947

(Continues)

TABLE 7 (Continued)

Panel B: Green patents granted						
	(1)	(2)	(3)	(4)	(5)	(6)
	$Green\_Inno^{grant}$		$Green\_Invention^{grant}$		$Green\_Utility^{grant}$	
Variable	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint	High financial constraint	Low financial constraint
	(0.0190)	(0.0169)	(0.6060)	(0.7976)	(0.0310)	(0.1082)
$Cash_{t-1}$	0.0008***	−0.0005	0.0025	0.0107***	−0.0005	0.0013*
	(0.0001)	(0.0009)	(0.0026)	(0.0031)	(0.0006)	(0.0007)
ROA_{t-1}	0.0882***	0.0144	0.8361***	0.4958	0.0463***	−0.0577
	(0.0091)	(0.0245)	(0.2760)	(1.3952)	(0.0101)	(0.1893)
Tobin's Q_{t-1}	−0.0702***	0.0123	0.6187***	2.0635	0.0215***	−0.0153
	(0.0041)	(0.0221)	(0.1212)	(1.3102)	(0.0061)	(0.1771)
$Top10_{t-1}$	0.0759***	0.0034	0.1613	1.3351	0.0103	−0.0403
	(0.0073)	(0.0221)	(0.2125)	(1.3110)	(0.0125)	(0.1776)
$R\&D_{t-1}$	0.0846***	0.0018	0.9913***	0.3596	0.0126	−0.0876
	(0.0141)	(0.0251)	(0.0312)	(1.4415)	(0.0241)	(0.1954)
GDP_{t-1}	0.0976***	0.0018	0.4886*	1.2407	0.0085	0.0538
	(0.0093)	(0.0224)	(0.2450)	(1.3521)	(0.0157)	(0.1832)
Constant	−5.0951***	−6.3656***	−7.1192***	−5.0116***	−6.8532***	−8.0582***
	(0.3653)	(0.3922)	(0.2972)	(0.4323)	(0.4561)	(0.5192)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	8392	8606	8392	8606	8392	8606
Adj-R^2	0.3251	0.2806	0.3031	0.2913	0.3057	0.3202
F	264.6367	297.9812	309.8076	276.0356	258.3502	221.5037
SUR test for *digitalization* coefficients	0.0415*** Chi2 = 11.07		0.0542*** Chi2 = 14.91		0.0502*** Chi2 = 12.24	

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

Smart manufacturing capability is the capacity to use digital technologies and data to optimize and automate a firm's manufacturing processes (Xia et al., 2019). Examples include using sensors and other devices to monitor and control production and use data analytics to improve efficiency and quality. It is a reflection of the transforming aspect of dynamic capabilities. We contribute to this capability by demonstrating how digitalization transforms corporations, not just through automation but also leading to green innovation.

Modern information integration capability refers to a firm's ability to integrate and manage data from different internal and external sources and customer interactions. Data analytics from that integration generate insights and inform decision-making (Rai et al., 2006). It is a multifaceted dynamic capability, not just about internal change but also about integrating external data for comprehensive business intelligence.

Our second contribution is in the measurement of digitalization. The literature on measuring corporate digitalization is broadly divided into three categories: ratio calculation, survey or case study and textual analysis. Qi et al. (2020) use the ratio calculation approach with the proportion of a company's digital intangible assets to intangible assets as an indicator of digitalization. Annarelli et al. (2021) conducted a very comprehensive review of digitalization, examining 249 relevant papers across multiple disciplines. They find most empirical papers rely on surveys or case studies to measure digitalization. Surveys typically gather data through respondents' self-reported outcomes, which aren't subject to public scrutiny or verification. A growing number of recent papers have used textual analysis to measure digitalization in China (Li & Shen, 2021; Peng et al., 2022; Wu et al., 2023). This method is based on the frequency of digital economy-related words in firm annual reports or the MD&A section to assess their digitalization capability. Corporate annual reports and MD&A sections represent formal communication channels where firms disclose their strategic priorities, investment, and key information to public investors and stakeholders. When a firm emphasizes its digital initiatives, it is more than an expression. It signals strategic importance and management attention to the

TABLE 8 Instrumental variable results of digitalization on green innovation.

Variable	(1) First: $Digitalization_{t\text{-}1}$	(2) Second: $Green\_Inno^{app}$	(3) $Green\_Invention^{app}$	(4) $Green\_Utility^{app}$	(5) $Green\_Inno^{grant}$	(6) $Green\_Invention^{grant}$	(7) $Green\_Utility^{grant}$
Fitted ($Digitalization_{t\text{-}1}$)		1.1695***	0.4531***	1.1017***	1.0695***	1.0034*	0.6107***
		(0.2027)	(0.0824)	(0.0679)	(0.2071)	(0.5024)	(0.0686)
$Local\_digital_{t\text{-}1}$	1327.2350***						
	(18.7062)						
Control variables	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999	16,999
Adj-R^2	0.3714	0.3569	0.3307	0.3263	0.3467	0.3203	0.3174
F statistic	1869.25***	——	——	——	——	——	——
Kleibergen-Paap rk LM statistic	654.023***	——	——	——	——	——	——
Kleibergen-Paap rk Wald *F* statistic	1869.25***	——	——	——	——	——	——

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

TABLE 9 Province fixed effect and bootstrapped sampling—Robustness test.

Panel A: Green patent applications

	(1)	(2)	(3)	(4)	(5)	(6)
	With province fixed effect			Bootstrapped sampling		
Variable	$Green\_Inno^{app}$	$Green\_Invention^{app}$	$Green\_Utility^{app}$	$Green\_Inno^{app}$	$Green\_Invention^{app}$	$Green\_Utility^{app}$
$Digitalization_{t-1}$	2.6119***	2.5958***	0.8947***	2.7742***	2.7054***	1.0330***
	(0.2234)	(0.2000)	(0.1850)	(0.2203)	(0.2022)	(0.1825)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−3.7158	−6.2809***	−1.5366	−8.1442***	−7.8085***	−5.7227***
	(2.3910)	(1.9829)	(1.9672)	(0.2996)	(0.2545)	(0.2395)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	NO	NO	NO
N	16,999	16,999	16,999	16,999	16,999	16,999
Adj-R^2	0.3509	0.3258	0.3202	0.3414	0.3164	0.3120
F	555.5340	437.7915	378.5114	——	——	——

Panel B: Green patents granted

	(1)	(2)	(3)	(4)	(5)	(6)
	With province fixed effect			Bootstrapped sampling		
Variable	$Green\_Inno^{grant}$	$Green\_Invention^{grant}$	$Green\_Utility^{grant}$	$Green\_Inno^{grant}$	$Green\_Invention^{grant}$	$Green\_Utility^{grant}$
$Digitalization_{t-1}$	1.0312***	0.7906***	0.7845***	0.8427***	0.6153***	0.8031
	(0.1435)	(0.1012)	(0.1047)	(0.2032)	(0.1357)	(0.5253)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−7.3185***	−7.1052***	−11.4527***	−8.7354***	−7.4023***	−9.7435***
	(0.2975)	(0.2952)	(0.5708)	(0.5260)	(0.3756)	(0.4582)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	NO	NO	NO
N	16,999	16,999	16,999	16,999	16,999	16,999
Adj-R^2	0.3125	0.3109	0.3132	0.3402	0.3312	0.2863
F	336.7013	407.6402	267.2013	——	——	——

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

topic. Drawing on these studies, we advance this method by using a digitalization-related vocabulary across four categories to ensure we capture a firm's holistic digital transformation rather than isolated initiatives. After identifying occurrences of keywords in annual reports, we do not merely count them (a frequency measure). Instead, using entropy weighting to measure digitalization, we account for the diversity and significance of the terms. Entropy weighting is a method used in textual analysis to assign weights to terms (words or phrases) based on their distribution across documents. The idea is to give greater weight to terms that are more informative (i.e., terms that are unevenly distributed across documents) and less weight to terms that appear uniformly across all documents. Entropy weighting is commonly used in socio-economic research; it enables researchers to determine the degree of uncertainty or variation proportional to the amount of data used in the investigation (Guo et al., 2023; Hao et al., 2023; Ren et al., 2021). When the amount of information is high, the degree of uncertainty is low, the entropy value is low, and vice versa. This method emphasizes terms that are more informative and de-emphasizes terms that appear uniformly. Terms that are merely buzzwords without substantive context would be consistently distributed across our four categories and would thus have a limited impact on our index. This approach mitigates concern that management may use annual reports as a platform for mere talk without substantive action.

The third contribution of this paper is to recognize the complexity and multifaceted nature of green innovation, as highlighted by Barbieri et al. (2020). This complexity requires a departure from

TABLE 10 Regression results with alternative measures—Robustness test.

Panel A: Green patent applications				
Variable	(1) Green_Inno^{app}	(2) Green_Inno^{app}	(3) Green_Inno^{app}	(4) Green_Inno^{app}
Digi$_{t\text{-}1}$	0.1240***			
	(0.0076)			
Digi_MDA$_{t\text{-}1}$		23.4818***		
		(1.7432)		
Intangible$_{t\text{-}1}$			0.0018	
			(0.0014)	
Digi fin index$_{t\text{-}1}$				0.0125
				(0.0829)
Control variables	YES	YES	YES	YES
Constant	−8.1055***	−8.1381***	−8.3185***	−8.3578***
	(0.2950)	(0.2958)	(0.2977)	(0.3376)
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999
Adj-R^2	0.3453	0.3426	0.3355	0.3354
F	580.9963	573.8019	544.0764	543.9110
Panel B: Green patents granted				
Variable	(1) Green_Inno^{grant}	(2) Green_Inno^{grant}	(3) Green_Inno^{grant}	(4) Green_Inno^{grant}
Digi$_{t\text{-}1}$	0.1109***			
	(0.0207)			
Digi_MDA$_{t\text{-}1}$		1.2065*		
		(0.6021)		
Intangible$_{t\text{-}1}$			0.0005	
			(0.0004)	
Digi fin index$_{t\text{-}1}$				0.0092
				(0.0077)
Control variables	YES	YES	YES	YES
Constant	−7.1423***	−10.962***	−7.6083***	−10.5963***
	(0.2965)	(0.3582)	(0.2347)	(0.3842)
Year FE	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999
Adj-R^2	0.3352	0.3243	0.3107	0.3104
F	503.0762	402.2009	372.1084	424.3047

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

single-factor analysis and advocates for a more integrated perspective to fully comprehend the dynamics of green innovation. In our study, we address this complexity by examining green innovation through multiple lenses: national, regional, and internal levels.

At the national level, we explore how the "Internet Plus" strategy shapes the landscape of green innovation, offering insights into the role of government initiatives. Regionally, we examine how differences in local intellectual property rights protection affects green innovation. Internally, we assess the effect of financial

TABLE 11 Results with alternative econometric models—Robustness test.

Panel A

	(1)	(2)	(3)	(4)	(5)	(6)
	Tobit model			Possion model		
Variable	Green_Inno^{app}	Green_Invention^{app}	Green_Utility^{app}	Green_Inno^{app}	Green_Invention^{app}	Green_Utility^{app}
Digitalization$_{t\text{-}1}$	4.4259***	4.9919***	2.5333***	2.2035***	2.7697***	1.5724***
	(0.3360)	(0.3418)	(0.3710)	(0.1674)	(0.1951)	(0.2295)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−11.962***	−13.5966***	−10.4520***	−6.7083***	−9.1194***	−7.5258***
	(0.4584)	(0.4848)	(0.4709)	(0.2412)	(0.2972)	(0.3035)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999
F	259.4781	212.5706	213.5368	——	——	——

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Tobit model			Possion model		
Variable	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}
Digitalization$_{t\text{-}1}$	1.5147***	1.1907	1.0432***	1.0537***	0.9662***	0.5624
	(0.3416)	(1.2173)	(0.2102)	(0.2032)	(0.0745)	(0.4070)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−5.8547***	−4.9285***	−5.8040***	−6.3587***	−7.1332***	−6.7238***
	(0.3392)	(0.2517)	(0.3241)	(0.3475)	(0.2356)	(0.2936)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	16,999	16,999	16,999	16,999	16,999	16,999
F	242.6752	217.3560	212.5014	——	——	——

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

constraints on firms, highlighting the critical role of internal corporate financial health in meeting the capital needs of green innovation.

This holistic approach allows us to understand how each mechanism operates under varying degrees of digitalization, regional differences in intellectual property rights protection, and financial constraints. By integrating these levels, we not only contribute to a broader understanding of green innovation but also provide practical implications for policymakers, business leaders, and academics. The results suggest the need for a coordinated effort across all levels to foster an environment for sustainable, innovative growth. Thus, our study extends the literature by offering a comprehensive view of the factors that drive green innovation. We call for future studies to consider this integrated framework.

6.2 | Implications for practice

The study's findings offer some important implications for corporations and policy makers. First, firms need to recognize the positive effects of digitalization in promoting green innovation. Digitalization should not be viewed merely as a collection of technologies or hardware but rather as a dynamic driver of innovation. We specifically find platform capability has the biggest impact on green innovation and modern information integration capability the least impact. This suggests that firms can prioritize investing in developing robust digital platforms that facilitate information sharing, electronic connectivity, and online partner collaboration (Wang et al., 2022).

However, we also reveal that the positive effect of digitalization on green innovation is less pronounced for firms with financial constraints. This raises a call for support from both local and central

TABLE 12 Regression results with different samples—Robustness test.

Panel A: Green patent applications						
	(1)	(2)	(3)	(4)	(5)	(6)
	Sample that excludes firms without green innovation			Sample that excludes firms located in big cities		
Variable	Green_Inno^{app}	Green_Invention^{app}	Green_Utility^{app}	Green_Inno^{app}	Green_Invention^{app}	Green_Utility^{app}
Digitalization$_{t-1}$	1.6163***	2.2499***	0.1623	4.0155***	3.7234***	1.9671***
	(0.2432)	(0.2444)	(0.2483)	(0.3024)	(0.2704)	(0.2512)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−6.8393***	−8.2961***	−4.8547***	−6.6693***	−5.9906***	−4.7040***
	(0.3375)	(0.3503)	(0.3291)	(0.3796)	(0.3211)	(0.3202)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	9997	9997	9997	11,533	11,533	11,533
Adj-R^2	0.2980	0.2720	0.2712	0.3000	0.2700	0.2621
F	279.3090	261.8832	185.3519	347.2221	258.5779	238.8469
Panel B: Green patents granted						
	(1)	(2)	(3)	(4)	(5)	(6)
	Sample that excludes firms without green innovation			Sample that excludes firms located in big cities		
Variable	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}	Green_Inno^{grant}
Digitalization$_{t-1}$	1.7106***	1.3235***	1.0516***	1.5051***	1.2387***	0.1425
	(0.3021)	(0.1705)	(0.3021)	(0.1723)	(0.2208)	(0.1280)
Control variables	YES	YES	YES	YES	YES	YES
Constant	−7.2963***	−6.5396***	−5.6255***	−6.5097***	−5.9705***	−8.1086***
	(0.3405)	(0.3274)	(0.3012)	(0.4793)	(0.3012)	(0.4213)
Year FE	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
N	9997	9997	9997	11,533	11,533	11,533
Adj-R^2	0.3078	0.3172	0.2792	0.3231	0.2865	0.2827
F	307.3007	306.8381	295.3519	296.2012	295.5097	303.8495

Note: The heteroscedastic robust standard deviations are reported in parentheses.
*Significance at the 10% level.
**Significance at the 5% level.
***Significance at the 1% level.

government to establish financial programs for small and financially constrained firms and startups that are committed to green innovation. These programs could include grants, subsidies, tax incentives, or other forms of financial assistance (Li et al., 2021), tailored to enable these firms to overcome financial barriers and invest effectively in digital platforms.

Our result on local IPR protection shows the effect of digitalization on green innovation is stronger in regions with stronger IPR protection. Policymakers should strengthen the legal and regulatory frameworks for intellectual property. This includes not only the creation of laws but also their effective enforcement. Policy makers should also work toward minimizing regional disparities so that firms operating in different regions enjoy a consistent level of IPR protection. This could involve setting national standards or setting up more IPR courts/centers. In addition, government agencies or schools can educate firms about the importance of IPR and how to effectively protect their intellectual property. This could involve providing resources and support for patent filing, as well as guidance on navigating the IPR system. Such support is particularly crucial for small and medium-sized enterprises (SMEs) and startups that may lack the resources and knowledge to adequately protect their green innovations.

Lastly, a national digital economic policy should continue in full swing to maximize the benefits of digitalization while minimizing the related risks. This includes, but is not limited to, promoting the development of artificial intelligence and big data, building a world-class digital infrastructure, enhancing cybersecurity measuresto protect against cyber threats, and promoting the use of digital currencies and financial technologies to enhance financial inclusion.

6.3 | Limitations and future research

Our methodology relies on keywords in annual reports to measure digitalization capabilities. We recognize that these reports serve various roles, including branding and stakeholder communication. There is the potential for firms to use digitalization buzzwords to align with global trends, which might not reflect their actual capabilities. Despite our entropy weighting aiming to add depth and reduce the impact of buzzwords, our index should ideally be viewed alongside other measures to ensure a holistic understanding of a firm's digitalization maturity.

Another limitation of our study is in the measurement of green innovation; not all innovation is patented (Bellstam et al., 2021). An alternative measure to consider is the number of green products introduced to the market. Huang and Li (2017) have used survey data and Mealy and Teytelboym (2022) relied on WTO data to measure green products. These approaches either focus on country-level data or require access to detailed firm-specific product lists, which we do not have. Consequently, our study leans on patent data to represent green innovation and we acknowledge that this might not capture the full breadth of a firm's green innovation.

Future research could examine how the relationship between corporate digitalization and green innovation varies across different industries such as manufacturing, energy, technology, healthcare, and consumer goods. Since industries have varying degrees of digitalization and environmental impact, understanding this relationship could provide more targeted insights for both policymakers and businesses. Comparing specialized industries with conglomerates could reveal how business structure influences the adoption and effectiveness of digitalization in driving green innovation.

ACKNOWLEDGEMENTS

Open access publishing facilitated by Lincoln University, as part of the Wiley - Lincoln University agreement via the Council of Australian University Librarians.

ORCID

Zhaohua Li https://orcid.org/0000-0002-4732-7997

REFERENCES

Acharya, V., & Xu, Z. (2017). Financial dependence and innovation: The case of public versus private firms. *Journal of Financial Economics*, *124*(2), 223–243. https://doi.org/10.1016/j.jfineco.2016.02.010

Annarelli, A., Battistella, C., Nonino, F., Parida, V., & Pessot, E. (2021). Literature review on digitalization capabilities: Co-citation analysis of antecedents, conceptualization and consequences. *Technological Forecasting and Social Change*, *166*, 120635. https://doi.org/10.1016/j.techfore.2021.120635

Barbieri, N., Marzucchi, A., & Rizzo, U. (2020). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones? *Research Policy*, *49*(2), 103901. https://doi.org/10.1016/j.respol.2019.103901

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. (2008). Financing patterns around the world: Are small firms different? *Journal of Financial Economics*, *89*(3), 467–487. https://doi.org/10.1016/j.jfineco.2007.10.005

Bellstam, G., Bhagat, S., & Cookson, J. A. (2021). A text-based analysis of corporate innovation. *Management Science*, *67*(7), 4004–4031. https://doi.org/10.1287/mnsc.2020.3682

Bu, L., Chung, V., Leung, N., Wang, K. W., Xia, B., & Xie, C. (2021). The future of digital innovation in China: Megatrends shaping one of the world's fastest evolving digital ecosystems. In *McKinsey & Company* (Issue October). https://www.mckinsey.com/featured-insights/china/the-future-of-digital-innovation-in-china-megatrends-shaping-one-of-the-worlds-fastest-evolving-digital-ecosystems

Castellacci, F., & Lie, C. M. (2017). A taxonomy of green innovators: Empirical evidence from South Korea. *Journal of Cleaner Production*, *143*, 1036–1047. https://doi.org/10.1016/j.jclepro.2016.12.016

Cenamor, J., Parida, V., & Wincent, J. (2019). How entrepreneurial SMEs compete through digital platforms: The roles of digital platform capability, network capability and ambidexterity. *Journal of Business Research*, *100*, 196–206. https://doi.org/10.1016/j.jbusres.2019.03.035

Cette, G., Nevoux, S., & Py, L. (2022). The impact of ICTs and digitalization on productivity and labor share: Evidence from French firms. *Economics of Innovation and New Technology*, *31*(8), 669–692. https://doi.org/10.1080/10438599.2020.1849967

Chen, Z., Xiao, Y., & Jiang, K. (2023). Corporate green innovation and stock liquidity in China. *Accounting and Finance*, *63*(S1), 1381–1415. https://doi.org/10.1111/acfi.13027

Chwiłkowska-Kubala, A., Cyfert, S., Malewska, K., Mierzejewska, K., & Szumowski, W. (2023). The impact of resources on digital transformation in energy sector companies. The role of readiness for digital transformation. *Technology in Society*, *74*, 102315. https://doi.org/10.1016/j.techsoc.2023.102315

Ding, N., Gu, L., & Peng, Y. (2022). Fintech, financial constraints and innovation: Evidence from China. *Journal of Corporate Finance*, *73*, 102194. https://doi.org/10.1016/j.jcorpfin.2022.102194

Fang, L. H., Lerner, J., & Wu, C. (2017). Intellectual property rights protection, ownership, and innovation: Evidence from China. *The Review of Financial Studies*, *30*(7), 2446–2477. https://doi.org/10.1093/rfs/hhx023

Ghobakhloo, M., Iranmanesh, M., Grybauskas, A., Vilkas, M., & Petraitė, M. (2021). Industry 4.0, innovation, and sustainable development: A systematic review and a roadmap to sustainable innovation. *Business Strategy and the Environment*, *30*(8), 4237–4257. https://doi.org/10.1002/BSE.2867

Gobble, M. M. (2018). Digitalization, digitization, and innovation. *Research-Technology Management*, *61*(4), 56–59. https://doi.org/10.1080/08956308.2018.1471280

Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature*, *57*(1), 3–43. https://doi.org/10.1257/jel.20171452

Gradillas, M., & Thomas, L. D. W. (2023). Distinguishing digitization and digitalization: A systematic review and conceptual framework. *Journal of Product Innovation Management*, 1–32. https://doi.org/10.1111/jpim.12690

Guo, F., Wang, J., Wang, F., Kong, T., Zhang, X., & Cheng, Z. (2020). Measuring China's digital financial inclusion: Index compilation and spatial characteristics. *China Economic Quarterly*, *19*(4), 1401–1418. (In Chinese).

Guo, Q., Geng, C., & Yao, N. (2023). How does green digitalization affect environmental innovation? The moderating role of institutional forces. *Business Strategy and the Environment*, *32*(6), 3088–3105. https://doi.org/10.1002/BSE.3288

Hadlock, C. J., & Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *The Review of Financial Studies*, *23*(5), 1909–1940. https://doi.org/10.1093/rfs/hhq009

Hall, B. H., & Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the economics of innovation* (Vol. 1, Issue 1 C). Elsevier B.V.. https://doi.org/10.1016/S0169-7218(10)01014-2

Hao, X., Li, Y., Ren, S., Wu, H., & Hao, Y. (2023). The role of digitalization on green economic growth: Does industrial structure optimization and green innovation matter? *Journal of Environmental Management*, *325*, 116504. https://doi.org/10.1016/J.JENVMAN.2022.116504

Hart, S. L. (1995). A natural-resource-based view of the firm. *The Academy of Management Review*, *20*(4), 986. https://doi.org/10.2307/258963

Hart, S. L., & Dowell, G. (2011). A natural-resource-based view of the firm: Fifteen years after. *Journal of Management*, *37*(5), 1464–1479. https://doi.org/10.1177/0149206310390219/FORMAT/EPUB

Hottenrott, H., & Peters, B. (2012). Innovative capability and financing constraints for innovation: More money, more innovation? *The Review of Economics and Statistics*, *94*(4), 1126–1142. https://doi.org/10.1162/REST_a_00227

Huang, C. (2017). Recent development of the intellectual property rights system in China and challenges ahead. *Management and Organization Review*, *13*(1), 39–48. https://doi.org/10.1017/mor.2017.2

Huang, J.-W., & Li, Y.-H. (2017). Green innovation and performance: The view of organizational capability and social reciprocity. *Journal of Business Ethics*, *145*(2), 309–324. https://doi.org/10.1007/s10551-015-2903-y

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT press.

Jamwal, A., Agrawal, R., Sharma, M., & Giallanza, A. (2021). Industry 4.0 technologies for manufacturing sustainability: A systematic review and future research directions. *Applied Sciences*, *11*(12), 5725. https://doi.org/10.3390/app11125725

Kaplan, S. N., & Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *The Quarterly Journal of Economics*, *112*(1), 169–215. https://doi.org/10.1162/003355397555163

Kohtamäki, M., Parida, V., Oghazi, P., Gebauer, H., & Baines, T. (2019). Digital servitization business models in ecosystems: A theory of the firm. *Journal of Business Research*, *104*, 380–392. https://doi.org/10.1016/j.jbusres.2019.06.027

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Chapter 6-Law and finance after a decade of research. In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the economics of finance* (Vol. 2, Issue Part A, pp. 425–491). Elsevier (S&T) https://doi.org/10.1016/B978-0-44-453594-8.00006-9

Lenka, S., Parida, V., & Wincent, J. (2017). Digitalization capabilities as enablers of value co-creation in servitizing firms. *Psychology and Marketing*, *34*(1), 92–100. https://doi.org/10.1002/mar.20975

Li, D., & Shen, W. (2021). Can corporate digitalization promote green innovation? The moderating roles of internal control and institutional ownership. *Sustainability*, *13*(24), 13983. https://doi.org/10.3390/SU132413983

Li, Q., Wang, J., Cao, G., & Zhang, J. (2021). Financial constraints, government subsidies, and corporate innovation. *PLoS ONE*, *16*(11 November), e0259642. https://doi.org/10.1371/journal.pone.0259642

Liu, X., Liu, F., & Ren, X. (2023). Firms' digitalization in manufacturing and the structure and direction of green innovation. *Journal of Environmental Management*, *335*, 117525. https://doi.org/10.1016/j.jenvman.2023.117525

Mealy, P., & Teytelboym, A. (2022). Economic complexity and the green economy. *Research Policy*, *51*(8), 103948. https://doi.org/10.1016/j.respol.2020.103948

Ning, J., Jiang, X., & Luo, J. (2023). Relationship between enterprise digitalization and green innovation: A mediated moderation model. *Journal of Innovation and Knowledge*, *8*(1), 100326. https://doi.org/10.1016/j.jik.2023.100326

Nordhaus, W. D. (2021). The double externality of green innovation. In *The economics of collisions and contagions in a crowded world* (pp. 204–224). Princeton University Press. https://doi.org/10.1515/9780691215396-018

Parida, V., Sjödin, D., & Reim, W. (2019). Reviewing literature on digitalization, business model innovation, and sustainable industry: Past achievements and future promises. *Sustainability*, *11*, 391. https://doi.org/10.3390/su11020391

Peng, C., Jia, X., & Zou, Y. (2022). Does digitalization drive corporate green transformation?—Based on evidence from Chinese listed companies. *Frontiers in Environmental Science*, *10*, 1401. https://doi.org/10.3389/FENVS.2022.963878/BIBTEX

Perrotti, D., Verma, P., Srivastava, K. K., & Singh, P. (2021). Chapter 17 - Challenges and opportunities at the crossroads of environmental sustainability and economy research. In P. Singh, P. Verma, D. Perrotti, & K. K. Srivastava (Eds.), *Environmental sustainability and economy* (pp. 345–360). Elsevier. https://doi.org/10.1016/B978-0-12-822188-4.00013-0

Popp, D., Newell, R. G., & Jaffe, A. B. (2010). Chapter 21 - Energy, the environment, and technological change. In B. H. Hall & N. Rosenberg (Eds.), *Handbook of the economics of innovation* (Vol. 2, pp. 873–937). Elsevier (S&T). https://doi.org/10.1016/S0169-7218(10)02005-8

Qi, H., Cao, X., & Liu, Y. (2020). The impact of digital economy on corporate governance: From the perspective of information asymmetry and managers' irrational behavior. *Reformation*, *4*, 50–64. (In Chinese).

Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, *30*(2), 225–246. https://doi.org/10.2307/25148729

Ramirez Lopez, L. J., Aponte, G. P., & Garcia, A. R. (2019). Internet of things applied in healthcare based on open hardware with low-energy consumption. *Healthcare Informatics Research*, *25*(3), 230–235. https://doi.org/10.4258/HIR.2019.25.3.230

Ren, S., Hao, Y., Xu, L., Wu, H., & Ba, N. (2021). Digitalization and energy: How does internet development affect China's energy consumption? *Energy Economics*, *98*, 105220. https://doi.org/10.1016/j.eneco.2021.105220

Rennings, K. (2000). Redefining innovation — Eco-innovation research and the contribution from ecological economics. *Ecological Economics*, *32*(2), 319–332. https://doi.org/10.1016/S0921-8009(99)00112-3

Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *MIS Quarterly*, *27*(2), 237. https://doi.org/10.2307/30036530

Singh, S. K., Del Giudice, M., Jabbour, C. J. C., Latan, H., & Sohal, A. S. (2022). Stakeholder pressure, green innovation, and performance in small and medium-sized enterprises: The role of green dynamic capabilities. *Business Strategy and the Environment*, *31*(1), 500–514. https://doi.org/10.1002/BSE.2906

Sun, Y., & He, M. (2023). Does digital transformation promote green innovation? A micro-level perspective on the Solow paradox. *Frontiers in Environmental Science*, *11*, 1134447. https://doi.org/10.3389/fenvs.2023.1134447

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Teece, D. J., Pisano, G., & Shuen, A. (1986). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533.

Veldkamp, L. (2023). Valuing data as an asset. *Review of Finance*, *27*(5), 1545–1562. https://doi.org/10.1093/rof/rfac073

Wang, Y., Tian, Q., Li, X., & Xiao, X. (2022). Different roles, different strokes: How to leverage two types of digital platform capabilities to fuel service innovation. *Journal of Business Research*, *144*, 1121–1128. https://doi.org/10.1016/j.jbusres.2022.02.038

Wei, Z., & Sun, L. (2021). How to leverage manufacturing digitalization for green process innovation: An information processing perspective.

Industrial Management and Data Systems, *121*(5), 1026–1044. https://doi.org/10.1108/IMDS-08-2020-0459

Whited, T. M., & Wu, G. (2006). Financial constraints risk. *The Review of Financial Studies*, *19*(2), 531–559. https://doi.org/10.1093/rfs/hhj012

Wu, H., Hu, S., & Hu, S. (2023). How digitalization works in promoting corporate sustainable development performance? The mediating role of green technology innovation. *Environmental Science and Pollution Research*, *30*, 22013–22023. https://doi.org/10.1007/s11356-022-23762-7

Xia, L., Gao, S., Wei, J., & Ding, Q. (2022). Government subsidy and corporate green innovation - does board governance play a role? *Energy Policy*, *161*, 112720. https://doi.org/10.1016/j.enpol.2021.112720

Xia, Q., Jiang, C., Yang, C., Zheng, X., Pan, X., Shuai, Y., & Yuan, S. (2019). A method towards smart manufacturing capabilities and performance measurement. *Procedia Manufacturing*, *39*, 851–858. https://doi.org/10.1016/j.promfg.2020.01.415

Yang, T., & Zhang, X. (2022). FinTech adoption and financial inclusion: Evidence from household consumption in China. *Journal of Banking and Finance*, *145*, 106668. https://doi.org/10.1016/j.jbankfin.2022.106668

Yao, L., & Rao, Y. (2009). IPRs protection measure and regional differences in China. *International Business Research*, *2*(3), 108–114. https://doi.org/10.5539/ibr.v2n3p108

Zhao, C. Y. (2021). Digital development and servitization: Empirical evidence from listed manufacturing companies. *Nankai Business. Rev*, *24*, 149–163.

Zhuo, C., & Chen, J. (2023). Can digital transformation overcome the enterprise innovation dilemma: Effect, mechanism and effective boundary. *Technological Forecasting and Social Change*, *190*, 122378. https://doi.org/10.1016/j.techfore.2023.122378

How to cite this article: Fang, L., & Li, Z. (2024). Corporate digitalization and green innovation: Evidence from textual analysis of firm annual reports and corporate green patent data in China. *Business Strategy and the Environment*, *33*(5), 3936–3964. https://doi.org/10.1002/bse.3677

APPENDIX A

TABLE A1 Word list for compiling the digitalization index (English translations).

The English translations of the original Chinese words used to compile the digitalization index.							
Digital technology		Platform		Smart manufacturing		Modern information integration	
Word	Count	Word	Count	Word	Count	Word	Count
Cloud platform	5619	B2B	1752	Unification	32,626	Information center	1030
Cloud services	4415	B2C	1136	Artificial intelligence	12,068	Information sharing	1044
Cloud ecosystem	105	C2B	117	Industrial cloud	117	Informatization	51,080
Cloud computing	12,025	C2C	26	Industrial intelligence	535	Information management	3428
Blockchain	3382	Internet	50	Numerical control	7050	Information system	9179
Big data	32,503	O2O	2739	Intelligent warehousing	630	Information terminal	162
Digitalization	27,232	Internet+	6930	Smart manufacturing	16,371	Information networks	1055
Digital technology	954	Internet business	1230	Intelligent	35,993	Information software	84
Digital control	104	Internet business model	39	Smart factory	1914	Information integration	262
Digital intelligence	127	Internet platforms	1835	Smart technology	1980	Industrial information	500
Digital terminals	46	Internet applications	871	Intelligent control	3403	Industrial communications	62
Digital networks	70	Internet thinking	632	Intelligent troubleshooting	11	Network	2689
Digital marketing	1896	Internet strategy	124	Smart logistics	1270		
Digital communication	111	Internet technology	1713	Smart production	734		
Data center	10,639	Internet mode	101	Smart mobility	549		
Data platform	2408	Internet ecology	196	Intelligent management	699		
Data mining	823	Internet mobile	38	Intelligent systems	699		
Data science	70	Internet marketing	1720	Smart terminals	4823		
Data management	1644	Internet action	2	Intelligent networking	1588		
Data network	195	Internet solutions	64	Smart devices	1398		
Machine learning	492	Industrial internet	948	Future factory	109		
Internet of things	23,294	Industrial Internet of Things	4630	Manufacturing execution system	50		
		E-commerce	21,263	Lifecycle management	1070		
		Electronic commerce	10,429	Mobile intelligence	867		
		Mobile internet	7122	Automatic control	1204		
		Online to offline	39	Automatically detection	423		
		Online and offline	323	Automated production	598		
		Online and offline	6250	Automatic monitoring	132		
				Automatic surveillance	352		
				Virtual manufacturing	40		
				Virtualization	768		
				Integration	1924		
				Integrated control	233		
				Integrated systems	516		
				Integrated solutions	302		
				High-end intelligence	1261		

TABLE A2 Definitions of the variables.

Variable	Definition
Key variables	
$Green\_Inno^{app}$	Natural logarithm of 1 plus the number of total green patent applications
$Green\_Invention^{app}$	Natural logarithm of 1 plus the number of green invention patent applications
$Green\_Utility^{app}$	Natural logarithm of 1 plus the number of green utility model patent applications
$Green\_Inno^{grant}$	Natural logarithm of 1 plus the number of total green patent grants
$Green\_Invention^{grant}$	Natural logarithm of 1 plus the number of green invention patent grants
$Green\_Utility^{grant}$	Natural logarithm of 1 plus the number of green utility model patent grants
Digitalization	Digitalization index. See Section 3.2.1 for details.
Digital technology	Natural logarithm of word frequency on words from digital technology category.
Platform	Natural logarithm of word frequency on words from platform capability category
Smart manufacturing	Natural logarithm of word frequency on words from smart manufacturing capability category
Modern information integration	Natural logarithm of word frequency on words from modern information integration capability category
Policy	Digital economic policy equals 1 for the year 2015 and beyond, and zero otherwise.
Other variables	
Firm size	Natural logarithm of total revenue
SOE	A dummy variable equal to 1 if it is a state-owned enterprise; otherwise 0
Leverage	Total liability/Total assets
Cash	Cash and cash equivalents/Total assets
ROA	Net profit/Total assets
Tobin's Q	Natural logarithm of 1 plus the ratio of firm market value to book value
Top10	A sum of the ownership percentages held by the top 10 shareholders.
R&D	Natural logarithm of 1 plus the R&D expenditure
GDP	Natural logarithm of provincial per capita GDP