The Impact of National New Generation Artificial Intelligence Innovation and Development Pilot Zones on Corporate Green Innovation: Causal Inference Based on Double/Debiased Machine Learning

Boming Zhao*
School of Economics, Southwest Minzu University, Chengdu, Sichuan, China
**Corresponding Author*

Abstract: In contrast to the growing presence of artificial intelligence, it enables not merely the industrialization of organizations, but also opens up plenty of opportunities within the corporate green innovation sphere. Employing a double/debiased machine learning approach, this study treats the establishment of the National New Generation Artificial Intelligence Innovation and Development Pilot Zones (NAIPZ) as a quasi-natural experiment. By utilizing a comprehensive sample of A-share firms listed in Shanghai and Shenzhen between 2013 and 2023, it rigorously examines how the rise of AI shapes corporate green innovation alongside its specific driving pathways. It conducts both the robustness tests and endogeneity tests. The results of the research are the following: Firstly, the formation of the NAIPZ can be used to improve the green innovation rate of a company and boost the quantity of green patents; Secondly, the mechanism tests indicate that such pilot zones increase the amount of corporate green innovation by relaxing the corporate financing constraints and increasing the corporate R&D investment; Third, heterogeneity tests indicate that these pilot zones are more effective in their impact on state-owned companies and highly-polluting companies. The research conclusions present the micro-level data on the artificial intelligence revolution, and also give the empirical data on the pilot zone optimization and corporate green transformation.

Keywords: Artificial Intelligence; Double/Debiased Machine Learning; Corporate Green Innovation

1. Introduction

Since the Chinese economy has moved into the high-quality development phase following the high-speed development phase, artificial intelligence has become the main catalyst of the industrial reforms and high-quality economic development. The policy of national development in China is extremely clear that we must apply the new model of development. Achieving a sustainable, low-carbon society requires merging the digital and physical economies. This vital integration subsequently paves the way for the synchronized advancement of AI technologies and ecological innovations within China. The topic of how to go all the way with it was addressed in several policy documents published by the national authorities in September 2021, correctly and comprehensively embracing the new development model and performing well in carbon peaking and carbon neutrality with the view to peak carbon dioxide emissions by 2030 and become carbon neutral by 2060 which indicates that China is determined to be green and low-carbon. The enterprises are a micro-subject of green innovation that with the assistance of artificial intelligence technologies in R&D can be assigned greater marginal value in green innovation and are the main ways to reach the aim of the dual carbon.

The national recommendations on the creation of the National Next Generation Artificial Intelligence Innovation and Development Pilot Zones (NAIPZ) were adopted by the appropriate national authorities in 2019. As the first accepted NAIPZ, Beijing had taken an active initiative in piloting the implementation of the policies. As of the end of 2021, there are 18 cities recognized as NAIPZ nationwide considering the development of the east, the middle and the west simultaneously. Its purpose is to encourage a comprehensive infiltration of artificial intelligence and the economy and society by way of technological demonstration and experimenting with policies, and come up with the developmental experience that can be replicated and scaled. In 2023, the size of core artificial intelligence business of China reached nearly 580 billion yuan and more than 4,400 core companies, becoming the second-largest in the world [1]. It shows that the accelerated growth of the artificial intelligence sector and the slow implementation of pilot areas has established an artificial intelligence innovation system through technology, which has become a solid foundation upon which further corporate green innovations will be created.

According to this, the discussion of the impacts and the processes taking place in the development of NAIPZ in terms of corporate green innovation should be conducted in a systematic manner, which will also be very important. The paper will cover and address the following questions: Is there any significant effect of establishing NAIPZ on the green innovation level of the companies under its jurisdiction? In what manner is this policy implemented? What role do financing constraints and investing in R&D play? Is this policy different across companies with different features? They are extremely important to the future of the artificial intelligence phenomenon and the green revolution in society.

2. Literature Review

2.1 Research on Artificial Intelligence

The exploration of artificial intelligence commenced in the mid-twentieth century. The original era focused more on ways that would make a machine think and behave like a human being in thinking and behavior and was dedicated to the field of programming. It took up to the 1990s before the advent of the Internet led to artificial intelligence approaching its practical application. Artificial intelligence in economic field is researched in two dimensions: economic repercussions, and social repercussions. Based on economic consequences, Furman and Seamans (2019) [2] systematically examined existing evidence regarding AI's economic influences, and the authors believed that artificial intelligence has great potential to improve the productivity of businesses and the economy overall. Chen and Qin (2022) [3] adopted the endogenous mediating effect model due to the installation data of industrial robots in 38 countries which proved that artificial intelligence can help promote the inclusive industrial development. Regarding social effects, Michaels et al. (2010) [4] confirmed the polarization impact of ICT technologies using the industry information, i.e., it increases the demand on highly qualified labour and decreases the demand on middle qualified labour. Wang and Dong (2020) [5] adopted the Bartik instrumental variable method to reveal that there is a substitution effect on the labor demand of enterprises when it comes to using robots. In recent studies, environmental governance and green development have become new focuses. Wang and Zhang (2024) [6], it has been affirmed that the use of artificial intelligence and big data technology could be very effective to enhance the environmental performance of corporations depending on the two-way fixed effect model.

2.2 Research on Green Innovation

Green innovation is still in its early stages of forming a unified definition since it is the heart of balancing economic development with environmental benefits, but its connotation and extension have been continuously enriched in academic research. From the perspective of concept, green innovation emphasizes reducing ecological damage and improving resource utilization efficiency, and is mostly measured by data related to green patents. Yang and Chai (2015) [7] emphasized that green innovation is to create commercial value while reducing negative environmental impacts through innovations in dimensions such as technology, processes and products. In terms of measurement methods, Wang and Wang (2021) [8], they measured the number of green patent applications as a key indicator that could be used to depict the strategic innovative behavior of enterprises; Liu et al. (2023) [9], the green innovation efficiency was measured by the rate at which green innovation output is converted into input, it was a more efficient method to present the conversion process of green innovation input and output. It is generally assumed in the academic literature that green innovation can help a business not only in attaining green premium but also enhancing market competitiveness, but it also enables low carbon transformation of industries.

2.3 The Impact of Artificial Intelligence on Green Innovation

The digital economy-green development integration is so strong that the potential of artificial intelligence to promote green innovation is rapidly becoming a frontier topic. Huachuan et al. (2024) [10] highlighted how the development of NAIPZ could be very helpful in enhancing corporate green governance performance through the creation of multi-period difference-in-differences model (DID); the research of Ren and Wang (2025) [11] indicated that the construction of NAIPZ mainly improves corporate green governance performance by reducing corporate operating costs and administrative expenses; using the viewpoint of the innovation value chain, Chao and Shen (2025) [12] built a two-stage model based on the information about Chinese manufacturing listed companies between 2008 and 2022, and demonstrated how the concept of artificial intelligence can be used to make corporate green innovation more effective using such mechanisms as allocation of resources optimization and overall productivity boost. The study of Yu et al. (2025) [13] emphasized that artificial intelligence promotes corporate productivity and corporate green transformation by reducing financing constraints. Such conclusions serve as a sound theoretical basis to implement digital technology and green development, and also present theoretical evidence in favor of conducting a deep analysis of its mechanism.

The recent research has found out that artificial intelligence has a positive influence of green innovation in business in numerous ways. but it still has room to grow. First of all, most of the existing studies have applied the difference-in-difference strategy and its variations, nonetheless, the traditional causal inference models are limited because, on a single hand, there can be a misspecification of the model, but on the other hand, it is simple to encounter the curse of dimensionality and multicollinearity issues with a large number of control variables, which decrease the precision of the outcomes [14]. In addition, most of the current studies on green innovation use the number of green patent applications as an indicator of its growth, and this can be easily influenced by the short-term strategic behaviors of enterprises. Patent applications may also be rejected, resulting in high measurement errors. Finally, most existing studies adopt single-level mediating variables, such as single capital indicators (financing constraints), governance indicators (ESG performance), etc., which are one-sided and fail to fully capture the transmission chain.

According to this, The paper may contribute marginally in several ways: Firstly, it is innovative in the sense that it uses the concept of double/debiased machine learning to evaluate how the NAIPZ policy affects corporate green innovation and provides particular recommendations depending on the evaluation outcomes, which are useful to the government regarding the future planning of artificial intelligence. Second, the number of green patents which are approved under the law is used as an indicator to measure the degree of corporate green innovation, which is more indicative of the substantive innovation of businesses than the number of patent applications. Third, when it comes to mechanism identification, it explores two transmission channels, i.e., finance constraints and R&D expenditure and explains how access to capital improves the amount of innovation investment and strengthens its power.

3. Theoretical Framework and Hypothesis Development

3.1 The Impact of National AI Pilot Zones on Enterprise Green Innovation: A Theoretical Perspective

The NAIPZ can be seen as a single system in terms of its role in promoting and applying the use of artificial intelligence technologies and as a powerful tool in terms of increasing the degree of corporate green innovation in various fields. First of all, the large-scale introduction of highly developed artificial intelligent technology into pilot regions, which is enabled by technology support, will help to lower the cost of companies, improve the efficiency of financial management and eliminate companies of financial constraints on green innovation R&D, ultimately elevating the caliber of enterprise eco-innovation [15]. Secondly, in the context of resource integration, NAIPZ forms an economy of scale through policy guidance, concentration of industries, common green innovation facilities, aids the coordination of the industrial chain, further reduces the barrier and price of corporate green innovation and provides a positive climate to the conversion of enterprises to green ones [16]. Public policy tools, including tax exemptions, research grants, and eco-friendly credit lines, play a pivotal role in accelerating the environmental transition of businesses. Such fiscal and financial stimuli significantly boost a firm's willingness to fund sustainable R&D. Driven by these theoretical premises, the present research advances the following hypothesis:

H1: The implementation of the National New Generation Artificial Intelligence Innovation and Development Pilot Zones exerts a profoundly positive influence on the ecological innovation initiatives of enterprises.

3.2 Underlying Pathways of the NAIPZ's Effect on Sustainable Corporate Breakthroughs

NAIPZ can be very useful in resolving corporate financing constraints. Sufficient funds are required to support green innovation activities in listed companies. It is hard to fund the ongoing R&D expenditure merely with internal sources, and therefore external financing has been a barrier in green innovation. In terms of sources of funding: on the contrary, practical adoption of artificial intelligence technologies may be helpful to enhance the effectiveness of credit decision-making, evaluating the creditworthiness of a company as well as its repayment capacity, reducing information asymmetry between banks and companies, and enhancing access to credit [17]. On the other hand, the supporting policies of the pilot zones, as well as green credit and tax incentives that direct capital towards green and low-carbon industries and provide long-term and low-cost funding to R&D and implement green projects and ease the financing bottlenecks. Should there be less financing limitation, enterprises would be more willing and able to undertake green innovation. Companies are able to assume the risk of large initial investments, lengthy cycles and significant uncertainty associated with green innovation project development. Simultaneously, they could bring new technologies and professionals, advertise the change and implementation of results in green innovation [18] and lastly, improve the level of corporate green innovation.

NAIPZ has the potential to increase the level of R&D investment of enterprises. With the help of policy instruments including government subsidies, NAIPZ is very effective in reducing external financing constraints and financial pressures on enterprises allowing them to use more of their resources in the R&D of green innovation [19]. Innovation ecosystem optimization point of view, the sound industry-university-research cooperation platforms and talent incentive mechanisms in the pilot zones provide a favorable external environment for enterprise R&D, effectively stimulate innovation enthusiasm, and allow businesses to concentrate on green technology R&D and green product innovation, with an ever-increasing R&D investment. R&D investments are financially guaranteed and supported by talents to provide green innovations. helps enterprises carry out more forward-looking technological research, better integrate technology, talent and other resources, improve innovation quality and R&D efficiency [20], and finally lead to a better performance of the corporate green innovation. Building upon the preceding theoretical discussion, this study posits the following hypothesis:

H2: The rollout of the NAIPZ effectively elevates the standard of enterprise eco-innovation, by alleviating the financial constraints on business entities, and also raising the amount of spending on R&D.

Regarding property rights: State-owned enterprises can be engaged in political activities more frequently than non-state-owned companies. They have become one of the key vehicles of implementing national strategies and policies and, therefore, can get more policy assistance and financial aid, have broader external financing channels, respond faster to the NAIPZ policy, and have stronger implementation willingness. Meanwhile, state-owned enterprises have better industry-university-research cooperation platforms and greater talent advantages. Non-state-owned enterprises are mainly private enterprises, which are generally characterized by small scale and fewer mortgaged assets. They face stronger financing constraints, which may increase financing and operating costs and restrict their innovative development [21]. From the perspective of industry pollution characteristics: As environmental pollution has become the focus of social attention, enterprises in heavily polluting industries face stricter market supervision, tighter restrictions on emissions and increased government spending. Construction of NAIPZ is a perfect fit to the green transformation requirements of highly polluting companies as it optimizes production processes via technology [22]. Drawing upon the preceding analytical framework, this study advances the following hypothesis:

H3: The stimulating effects and underlying pathways of the NAIPZ on enterprise eco-innovation exhibit significant heterogeneity, contingent upon the firm's ownership structure and the sector's environmental pollution profile.

4. Research Design

4.1 Model Specification

4.1.1 Double/debiased machine learning model

Based on the theoretical discussion presented above, difference-in-differences and propensity score matching are two common approaches to evaluating policies. However, they have numerous limitations to their assumptions and the features of the data. By leveraging the principle of orthogonalization to partial out the bias introduced by high-dimensional confounders, the double/debiased machine learning (DML) framework demonstrates exceptional proficiency in mapping intricate non-linear correlations across variables, and could provide unbiased estimates of the policy treatment effect with a finite sample. It also addresses issues like overfitting in machine learning by cross-fitting. This paper takes the causal inference perspective of the double/debiased machine learning developed by Chernozhukov et al. (2018) [23] and constructs the following partially linear model:

$$Y_{it}=\theta_0 D_{it}+g(X_{it})+U_{it} \qquad (1)$$

$$E(U_{it}|D_{it},X_{it})=0 \qquad (2)$$

Specifically, Y_{it} acts as the response variable, indicating firm i's level of sustainable innovation in year t. On the other hand, D_{it} is constructed as a policy dummy variable reflecting whether the firm is subject to the NAIPZ intervention; θ_0 is the treatment effect coefficient corresponding to the policy. Additionally, the vector of covariates is represented by X_{it}. The expression $g(X_{it})$ captures the complex, non-linear influence of these controls on ecological innovation, and its empirical approximation, denoted as $\hat{g}(X_{it})$, is derived utilizing advanced machine learning techniques. Finally, U_{it} acts as the idiosyncratic disturbance term, strictly satisfying the zero conditional mean assumption.

Estimating θ_0 directly from this baseline equation risks generating biased estimates in finite samples, given that X_{it} includes confounders correlated with both Y_{it} and D_{it}. Consequently, to partial out these confounding effects, an auxiliary regression is constructed as follows:

$$D_{it}=m(X_{it})+V_{it} \qquad (3)$$

$$E(V_{it}|X_{it})=0 \qquad (4)$$

Where $m(X_{it})$ denotes the nonlinear function of control variables affecting NAIPZ, where advanced machine learning methods are analogously deployed to derive the precise functional form$\hat{m}(X_{it})$. Meanwhile, V_{it} serves as the residual term, strictly adhering to the zero conditional mean assumption. After estimating $\hat{m}(X_{it})$, we construct the residual estimate $\widehat{V_{it}}=D_{it}-\hat{m}(X_{it})$. Then we estimate $\hat{g}(X_{it})$ in the main regression in the same way, and transform the main regression into: $Y_{it}-\hat{g}(X_{it})=\theta_0 D_{it}+U_{it}$. Subsequently we instrument D_{it} using $\widehat{V_{it}}$ to execute the final regression analysis, thereby yielding the coefficient estimator as follows:

$$\widehat{\theta_0}=\Big(\frac{1}{n}\sum_{i\in I,t\in T}\widehat{V_{it}}D_{it}\Big)^{-1}\frac{1}{n}\sum\nolimits_{i\in I,t\in T}\widehat{V_{it}}(Y_{it}-\hat{g}(X_{it})) \qquad (5)$$

Two rounds of machine learning estimation effectively exclude the influence of X_{it} on D_{it} and obtain the net policy effect. In addition, 5-fold cross-fitting is adopted to alleviate problems such as overfitting.

4.1.2 Multi-period difference-in-differences model

Due to the quasi-natural experiment of the NAIPZ policy implementation, where the policy was implemented in a staggered manner, we specify the staggered difference-in-differences empirical equation as:

$$Y_{it}=\alpha_0+\alpha_1 D_{it}+\rho X_{it}+\mu_i+\eta_t+\varepsilon_{it} \qquad (6)$$

In this specification, μ_i and$\eta_t$are incorporated to absorb individual-level and time-specific unobservables, respectively. Additionally, ε_{it} denotes the idiosyncratic disturbance term, while all remaining parameters retain their prior definitions.

4.2 Variable Selection

4.2.1 Dependent variable

Corporate green innovation (Tgreen) refers to the natural logarithm of the total number of accepted green invention and green utility model patents, following the measurement approach of Hu et al. (2023) [24]. In comparison with the green patent application indicator, this indicator may be more indicative of the substantive innovation of enterprises as it does not focus on quantity but rather on the quality of such innovations that can contribute to the effective advancement of technology, as well as support the green development.

4.2.2 Independent variable

In this framework, the primary explanatory variable is the NAIPZ policy indicator (denoted as DID), which is constructed via the multiplicative interaction of the temporal dummy and the municipal treatment dummy ($Time \times Treat$). Specifically, it takes the value of 1 if firm i operates in a designated pilot city during or after the implementation year t, and 0 otherwise.

4.2.3 Mediating variables

(1) Financing constraints (WW) as measured by the WW index. The WW index takes into consideration both the financial attributes of firms, as well as industry-level variables, and therefore has a wider economic relevance. The particular formula used is: WW=-0.091×CF-0.062×DivPos+0.021×Lev-0.044×Size+0.102× ISG-0.035×SG, In this equation, CF denotes the operating cash flow normalized by total assets, while DivPos serves as a binary indicator equaling 1 if the company distributes cash dividends and 0 otherwise. Furthermore, Lev reflects the long-term debt-to-asset ratio, Size is the natural log of total assets, and ISG alongside SG represent the industry-level and firm-level sales growth rates, respectively. Fundamentally, a greater value of the WW index is indicative of deeper financial friction and tighter funding hurdles for the enterprise.

(2) Research and development (RD) investment, calculated as a percentage of research and development spending to operating income.

4.2.4 Control variables

The control variables in this paper as stated in the previous literature include: financial leverage (Lev), defined as the ratio of total liabilities to total assets; operating cash flow (Cfo), measured by scaling net operating cash flows by total assets;

growth of the firm (Growth) which is based on the pace of the growth of operating income; age of the company (Age) which is the time between the date when it was established and the current date; profitability (Roa) which has been defined as the net profit divided by total assets, ownership concentration (Top1) which is the largest shareholder share percentage, natural-log of net fixed asset per employee (PFixA) and natural-log of operating income per employee (Psales).

4.3 Sample Selection and Data Sources

The empirical dataset constructed for this study comprises A-share enterprises publicly traded on both the Shanghai and Shenzhen stock exchanges that will be analyzed within five years. i.e., 2013-2023. The company green patent information and financial data were attained using China Research Data Service Platform (CNDRS) and the financial data based on China Stock Market and Accounting Research (CSMAR). To minimize the influence of abnormal samples, this paper will handle the sample data as follows: (1) remove the finance companies; (2) remove the ST, *ST and PT companies; (3) fill in some of the missing values with interpolation and remove those samples with significant missing values; (4) to mitigate the distortion caused by extreme outliers, all continuous variables are Winsorized at the 1st and 99th percentiles.

Following these screening procedures, a final panel comprising 28,178 valid firm-year observations is retained. Table 1 reports the summary statistics for the principal variables.

Table 1. Descriptive Statistics

Variable	N	Mean	Sd	Min	Max
Tgreen	28,178	0.886	1.055	0	4.043
DID	28,178	0.244	0.430	0	1
Lev	28,178	0.421	0.198	0.062	0.885
Cfo	28,178	0.049	0.065	-0.137	0.238
Growth	28,178	0.160	0.495	-0.834	3.216
Age	28,178	2.924	0.328	1.946	3.584
Roa	28,178	0.036	0.060	-0.230	0.192
Top1	28,178	33.59	14.83	8.350	74.18
PFixA	28,178	12.65	1.140	9.324	15.66
Psales	28,178	13.92	0.803	12.27	16.34

5. Empirical Analysis

5.1 Baseline Regression

The paper employs the partially linear double/debiased machine learning model to examine the effect of the policy on the corporate green innovation due to the NAIPZ policy implementation. In order to have the right split ratio it is set to 1:4, and cross-fitting on both main regression and secondary regression is performed through random forest (rf) algorithm. Table 2 reports the empirical outcomes derived from the baseline regression estimation. Linear control variable functions are used to estimate time fixed effects and individual fixed effects as presented in column (1). The empirical estimation yields a coefficient of 0.138, demonstrating robust statistical significance at the 1% threshold. Quadratic forms of control variables are represented in column (2). The estimated coefficient value of 0.144 is statistically significant at one percentage level. The findings show that the implementation of NAIPZ is positively associated with green innovation of the company and hypothesis H1 can be tested.

Table 2. Baseline Regression Results

Variable	(1)	(2)
	Tgreen	Tgreen
DID	0.138***	0.144***
	(0.016)	(0.016)
Controls	Yes	Yes
Controls_square	NO	Yes
Year FE	Yes	Yes
Firm FE	Yes	Yes
N	28,178	28,178

Notes: Robust standard errors are reported in parentheses. The asterisks ***, **, and * denote statistical significance at the 0.01, 0.05, and 0.10 thresholds, correspondingly. This notation convention is maintained across all following empirical tables.

5.2 Robustness Tests

5.2.1 Replacing the dependent variable

In order to estimate the impacts of the policy on the corporate green innovations, the outcome variable is substituted with the natural logarithm of one plus the volume of green patent filings, and the baseline model is subsequently re-estimated. Columns (1) and (2) of Table 3 report results with and without quadratic terms of control variables respectively. Across all specifications, the estimated parameters maintain robust statistical significance at the 1% threshold, thereby firmly corroborating the validity of the baseline findings.

Table 3. Robustness Test Regression Results (1)

Variable	Replacing the Dependent Variable		Excluding the Influence of Concurrent Policies		
	(1)	(2)	(3)	(4)	(5)
	Tgreen1	Tgreen1	Tgreen	Tgreen	Tgreen
DID	0.150***	0.151***	0.147***	0.153***	0.160***
	(0.017)	(0.017)	(0.016)	(0.017)	(0.018)
gfpolicy	NO	NO	Yes	NO	Yes
depolicy	NO	NO	NO	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Controls_square	NO	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
N	28,178	28,178	28,178	28,178	28,178

5.2.2 Excluding the influence of concurrent policies

There is also a possibility that NAIPZ policy on corporate green innovation may suffer some form of interference in the sampling frame because there are other similar efforts that were initiated at the same time, such as the Green Finance Reform and Innovation Pilot Zone policy (gfpolicy) that was introduced in 2017 and the National Digital Economy Innovation and Development Pilot Zone policy (depolicy) that was launched in 2019. It is made through the generation of dummy variables of each policy and adding them to the regression. The results are shown in columns (3), (4) and (5) of Table 3. After the influence of these two policies and the composite effect were eliminated gradually, it can be clearly seen that DID coefficients are not just statistically significant, but also positive, meaning that the results are extremely robust.

5.2.3 Replacing the machine learning algorithm

To analyze how different methods of machine learning can affect empirical information, the baseline regression will rely on the random forest (rf) algorithm but will be substituted by the gradient boost (gradboost) algorithm and the neural network (nnet) algorithm respectively to estimate the functions and of the double/debiased machine learning again. The DID coefficient is also very significant as shown in Columns (1)-(2) of Table 4. Reassuringly, the deployment of these distinct algorithms does not alter the primary findings; the coefficient continues to clear the 1% significance benchmark.

Table 4. Robustness Test Regression Results (2)

Variable	Replacing the Machine Learning Algorithm		Adjusting the Sample Splitting Ratio		Interactive Model
	(1)	(2)	(3)	(4)	(5)
	gradboost	nnet	Kfolds=3	Kfolds=8	Tgreen
DID	0.189***	0.168***	0.138***	0.139***	0.181***
	(0.017)	(0.021)	(0.016)	(0.016)	(0.019)
Controls	Yes	Yes	Yes	Yes	Yes
Controls_square	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
N	28178	28178	28178	28178	28178

5.2.4 Adjusting the sample splitting ratio

In order to investigate how various

cross-validation folds affect the empirical findings, the original 5-fold cross validation that was used in the regression process is replaced with 3-fold cross-validation and 8-fold cross-validation. More specifically, the split percentage of the samples in regression analysis is modified to 1:2 and 1:7. As shown in columns (3) and (4) of table 4, these regression results are significant and there is a significant positive value of the DID coefficient at a significance level of 1%.

5.2.5 Replacing the machine learning model

Considering that the partially linear model used in the baseline regression artificially sets the independent variable as linear and the control variables as nonlinear, which has certain subjectivity and may affect the estimation results, to facilitate the re-estimation process, a generalized interactive double/debiased machine learning framework is explicitly specified. We formally specify the primary and auxiliary empirical equations as follows:

$$Y_{it}=g(D_{it},X_{it})+U_{it} \tag{7}$$

$$D_{it}=m(X_{it})+V_{it} \tag{8}$$

Table 4 shows regression results in Column (5). The result of the NAIPZ on corporate green innovation after replacing the model does not change, which indicates that findings remain consistent.

5.3 Endogeneity Tests

To rigorously mitigate potential endogeneity concerns stemming from omitted variable bias, sample selection, and reverse causality, we employ a dual-method approach incorporating both instrumental variable (IV) and difference-in-differences (DID) estimators.

5.3.1 Instrumental variable method

In order to relieve the mentioned endogeneity issue, based on the methodology presented by Huang et al. (2019) [25], we instrument the NAIPZ utilizing the historical penetration rate of fixed-line telephones—specifically, the number of landlines per 100 residents—across prefecture-level cities in 1984. As the data is cross sectional; i.e., does not have any observations on more than one year, it will cause problems of hard identification and measurement in the fixed effect model. Thus, the reference to the way of Zhao et al. (2020) [26]. To construct a time-varying instrument for our panel data, we generate an interaction term between a cross-sectional historical anchor—the 1984 municipal landline penetration rate per 100 residents—and a time-series shifter, specifically the lagged national aggregate of internet users. The model is subsequently estimated utilizing a two-stage least squares (2SLS) framework. According to theoretical principles, areas that have been more affected with the historical fixed-line telephones had advantages in the process of digitalization and therefore formed the background of artificial intelligence development, fulfilling the relevance assumption of the instrumental variables. Conversely, the past data in 1984 does not matter with respect to the present rate of corporate green innovation, and due to Internet technology progress, fixed-line phones are no longer the leading means of communication, therefore, the constructed instrument lacks any plausible direct channel to affect firm-level green innovation, which perfectly aligns with the strict exogeneity assumption. The first and second stages regression outputs are reported in columns (1) and (2), respectively, of Table 5. In the first-stage estimation, the instrument yields a coefficient of 0.053($p<0.01$), empirically validating the relevance condition. Diagnostic tests further corroborate the instrument's validity: the Kleibergen-Paap rk LM statistic of 3434.552 robustly rejects the null hypothesis of underidentification at the 1% threshold. Furthermore, the Kleibergen-Paap rk Wald F-statistic stands at 1851.984, overwhelmingly exceeding the Stock-Yogo 10% critical value (16.38), thereby dispelling any concerns regarding weak instruments. The second stage coefficient remains significant at a 1% significance level, so after taking into account the endogeneity issue, the findings are robust.

Table 5. Endogeneity Test Regression Results

Variable	Instrumental Variable Method		Difference-in-Differences Method	
	(1)	(2)	(3)	(4)
	DID	Tgreen	Tgreen	Tgreen
IV	0.053^{***}			
	(0.001)			
DID		0.164^{***}	0.044^{***}	0.043^{***}
		(0.049)	(0.016)	(0.016)
Controls	Yes	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	24167	24167	28178	28178
Kleibergen-Paap rk LM Statistic	3434.552			
	[0.000]			
Kleibergen-Paap rk Wald F Statistic	1851.984			
	{16.38}			

Notes: Values enclosed in square brackets []

represent exact p-values. Furthermore, figures within curly braces {} denote the Stock-Yogo critical thresholds for weak instruments at the 10% significance level.

5.3.2 Multi-period difference-in-differences method

5.3.2.1 Baseline regression of did method

Table 5 reports the baseline empirical estimates derived from the staggered difference-in-differences specification. Controlling exclusively for two-way fixed effects (individual and time) without additional covariates, Column (3) yields a treatment coefficient of 0.044, achieving statistical significance at the 1% threshold. Furthermore, upon introducing the full suite of control variables in Column (4), the estimated effect of the DID indicator remains robustly positive and highly significant ($p < 0.01$).

5.3.2.2 Parallel trend test

To ensure the difference-in-differences findings are credible, this paper also considers the presence of a parallel trend before the policy and its change after the implementation using confidence levels of 95 percent. The effects of the parallel trend test are illustrated in Figure 1. As predicted by the assumption of parallel trend, none of the coefficients is statistically significant before the policy implementation. After policy introduction, the policy effect gradually increases and turns out to be significant positive over time, and therefore the use of the multi-period differences in differences model is justified.

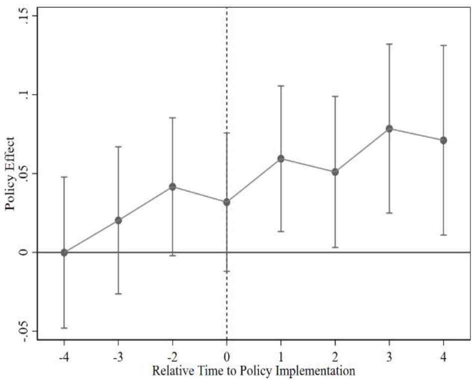

Figure 1. Parallel Trend Test Result

5.3.2.3 Placebo test

To construct the placebo treatment group, we randomly draw firms from the full sample without replacement, assigning the unselected firms to the control group. Following 500 random permutations, the empirical distribution of the estimated placebo coefficients is plotted in Figure 2. The estimated coefficients are tightly clustered at zero and very far away from the real estimated coefficient (0.043) as expected by the placebo test. This confirms that our baseline empirical findings are not spuriously driven by unobservable random shocks, thereby verifying the robust nature of the core difference-in-differences estimates.

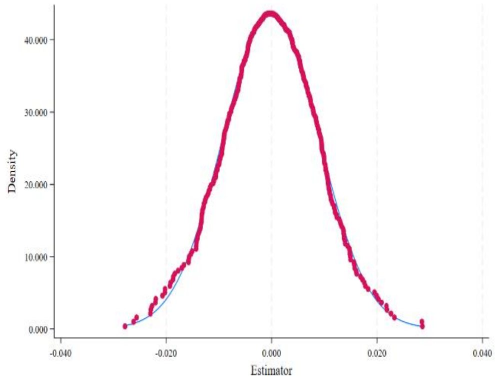

Figure 2. Placebo Test Result

5.4 Mediating Effect

Having established the robust, positive impact of the NAIPZ on corporate green innovation through baseline, robustness, and endogeneity estimations, we proceed to investigate the underlying transmission mechanisms. Guided by the two-step mediation framework articulated by Jiang (2022) [27], we employ a DML approach to explicitly estimate the effect of the policy indicator on the proposed mediating variables.

5.4.1 Financing constraint

Financing constraint is measured by the WW index, where a higher value indicates severer financing constraints. As theorized, the NAIPZ initiative mitigates corporate financial frictions by narrowing bank-firm information asymmetry and directing policy-driven capital allocations. Corroborating this mechanism, Column (1) of Table 6 reports a significantly negative coefficient for the policy indicator, empirically confirming that the establishment of the NAIPZ effectively eases corporate financing constraints. Existing studies point out that internal capital is the core funding source for corporate innovation investment, and tax incentive policies can ease financing constraints and stimulate innovation motivation. On the one hand, the alleviation of financing constraints provides a foundation for the buffer and risk-taking capacity of corporate working capital. On the other hand, it offers long-term stable financial support for green innovation projects, enabling them to bear risks such as large upfront investment and long payback periods [28]. Therefore, the mediating effect of financing constraint exists.

5.4.2 R&D investment

To capture the intensity of corporate innovation efforts, R&D investment is proxied by scaling aggregate R&D expenditures by concurrent operating revenue. From the perspective of transmission mechanism, the NAIPZ policy eases financing constraints and reduces corporate financial pressure through tax incentives, green bonds and other instruments, enabling companies to invest additional money in R&D efforts. As documented in Column (2) of Table 6, the estimated coefficient of 1.258 maintains robust statistical significance at the 1% threshold, thereby supplying compelling evidence that NAIPZ designation substantially boosts corporate R&D investment. Increased investment is favorable to green patents R&D and green projects. As Guo (2018) [29] argues, government innovation subsidies not only directly supplement R&D funds but also send positive signals to attract external resources and build industry-university-research cooperation platforms, providing a favorable external environment for corporate R&D and promoting the transformation of R&D investment into substantive innovation output, thereby improving corporate green innovation. Hypothesis H2 is supported.

Table 6. Mediating Effect Analysis Results

Variable	(1)	(2)
	WW	RD
DID	-0.014^{***}	1.258^{***}
	(0.004)	(0.086)
Controls	Yes	Yes
Controls_square	Yes	Yes
Year FE	Yes	Yes
Firm FE	Yes	Yes
N	28178	28178

5.5 Heterogeneity Analysis

5.5.1 Heterogeneity in industry pollution characteristics

Following the industry classification framework issued by the China Securities Regulatory Commission (CSRC), we partition the sample into heavily polluting industries (HPI) and non-heavy polluting industries (NHPI). Columns (1) and (2) of Table 7 present these split-sample estimation outcomes. The estimated treatment effect for the HPI cohort is robustly positive at the 1% threshold, while the corresponding parameter for the NHPI subgroup demonstrates positive significance at the 5% level. A formal empirical test for the equality of coefficients across the subsamples yields a p-value of 0.004, robustly confirming that the estimated treatment effects diverge significantly between the two cohorts. The implication is that NAIPZ would become much more influential in increasing the level of green innovation of highly polluting organizations. Its cause could be attributed to the increased attention given to highly pollute enterprises and the increased contribution of the government, as well as the application of technologies of artificial intelligence that are able to organize the production process encouraging the increase of enterprises green innovation levels.

Table 7. Heterogeneity Analysis Results

Variable	(1)	(2)	(3)	(4)
	HPI	NHPI	SOE	NSOE
DID	0.180^{***}	0.042^{**}	0.116^{***}	-0.003
	(0.040)	(0.017)	(0.027)	(0.019)
Inter-group Coefficient Difference Test	p-value=0.004^{***}		p-value=0.003^{***}	
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	6754	21424	9385	18793
R^2	0.650	0.738	0.755	0.694

Note: The reported p-values for the cross-group coefficient differences are derived from a bootstrap-based Fisher's test utilizing 1000 replications.

5.5.2 Heterogeneity in enterprise ownership nature

Given the inherent discrepancies in policy endowments and strategic objectives between state-owned enterprises (SOE) and non-state-owned enterprises (NSOE), we partition our A-share sample by ownership structure to conduct a split-sample estimation. As reported in Columns (3) and (4) of Table 7, the estimated parameter for the SOE cohort is 0.116 ($p < 0.01$), whereas the corresponding coefficient for NSOE remains statistically indistinguishable from zero. This pronounced divergence indicates that the NAIPZ initiative is substantially more effective at catalyzing green innovation within SOE. Theoretically, this asymmetry stems from SOE enjoying preferential access to policy support, superior resource allocation, more deeply integrated university-industry collaborative networks, and distinct human capital advantages. Consequently, Hypothesis H3 is firmly corroborated.

6. Conclusions and Implications

By leveraging the establishment of the National New Generation Artificial Intelligence Innovation and Development Pilot Zone (NAIPZ) as an exogenous policy shock, this study constructs a rigorous quasi-natural experimental framework, and it takes data concerning Chinese A-share listed companies of Shanghai and Shenzhen during the time interval of 2013-2023 with the help of double/debiased machine learning model to find out how and why artificial intelligence is created and how it affects business-oriented green innovation. It has been established that the creation of NAIPZ has contributed to the increased volume of corporate green innovation in the pilot zones and the level of substantive innovation in the enterprise significantly. Also, in order to conduct a series of strong tests, endogeneity tests, which are based on the instrumental variable method, difference-in-difference method, have been conducted and the results of them all are the same as those of the analyses above. The mechanisms testing can prove that NAIPZ policy may improve the standard of corporate green innovation performance by minimizing the extent of corporate financing restrictions and increasing the rate of corporate R&D spending. Our cross-sectional heterogeneity analysis reveals that state-owned enterprises (SOE) and heavily polluting sectors exhibit a significantly heightened sensitivity to the policy shock. With such results, this paper would make its policy recommendations, which are:

Then, we should extend the scope of the NAIPZ policy and develop additional pilot zones. According to the results of the study, NAIPZ is very motivating in respect to firm-based green innovations. The increase of the size of the geographical area of pilot zones, the creation of an infrastructure system, and the participation of companies in the role of innovators, and the pilot policy implementation to cities with acute needs of industrialization, and cities with specific categories of unique, specialized, rare and potentially innovative small and medium-size enterprises, as well as encouraging further artificial intelligence and company green governance integration, will be advantageous to a larger number of people.

It should therefore enhance the connection between the government and enterprise, enhance the level of support of NAIPZ, reduce the burden of businesses to finance their own operations and invest more in the research and development sector. Contrastingly, on the other hand, the government has to provide a higher amount of financial aid and encourage the growth of artificial intelligence by applying tax incentives, green bonds, etc., and enhance the clarity of information disclosures and controls, and create the artificial intelligence market systems. Then, in reverse, companies need to facilitate the development of industrial-university-research cooperation within the pilot regions, raise the level of R&D spending, and apply artificial intelligence technologies to speed up the process of transforming and implementing scientific and technological discoveries.

Finally, develop particular strategies that will take into account the heterogeneity of industry and enterprise nature in full. First of all, because of the unique features of the industry, it is necessary to increase the level of support to highly polluting industries and, in particular, to invest in pollution emissions. Government should enhance its control and transparency demands to have the green transformation and improved energy usage efficiency realized. Non-polluting green industries should also be given energy saving, intake reduction and recycling of goods and assets in particular in the non-rigid disclosure system, e.g., ESG reports. Afterwards, with regard to the ownership feature, it is necessary to strengthen government support of the companies in the state-owned enterprises and create a green innovation evaluation system in order to achieve their political objectives; non-state-owned enterprises should be market-led, and green development models and related information disclosure standards must be developed by the industry association to jointly develop green enterprises of green development.

References

[1] Geng Yeqiang, Zhao Huan. Artificial Intelligence Technology and Chinese Enterprises' Export Competition Strategy. Commercial Research, 2025, (05):75-85.

[2] Furman J, Seamans R. AI and the Economy. Innovation Policy and the Economy, 2019, 19(1):161-191.

[3] Chen Dong, Qin Ziyang. Artificial Intelligence and Inclusive Growth: Empirical Findings from the Worldwide

Deployment of Industrial Robots. Economic Research Journal, 2022, 57 (4):85-102.
[4] Michaels G, Natraj A, Van Reenen J. Has Ict Polarized Skill Demand? Evidence from Eleven Countries Over 25 Years. Review of Economics and Statistics, 2010, 96(1):60-77.
[5] Wang Yongqin, Dong Wen. The Impact of Robot Expansion on China's Labor Market: Empirical Analysis Based on Chinese Listed Manufacturing Firms. Economic Research Journal, 2020, 55(10):159-175.
[6] Wang Di, Zhang Yang. How Corporate Digital Transformation Drives Strategic Green Innovation and Environmental Performance of Enterprises. Economic Research Journal, 2024, 59(10):113-131.
[7] Yang Dong, Chai Huimin. An Overview on Motivating Determinants of Enterprises' Green Technological Innovation and Its Influences on Operational Outcomes. China Population, Resources and Environment, 2015, 25(S2):132-136.
[8] Wang Xin, Wang Ying. Mechanism Analysis on Green Credit Regulation and the Promotion of Firms' Green Innovative Capacity. Management World, 2021, 37(06):173-188+11.
[9] Liu Chang, Pan Huifeng, Li Pei, Feng Yaxin. Digital Restructuring's Heterogeneous Effects and Internal Paths on Green Innovation Efficiency within Manufacturing Entities. China Soft Science, 2023, (04):121-129.
[10]Gao Huachuan, Wang Huapu, Dong Zhen. Intelligent Empowerment and Enterprises' Green Innovation: Quasi-experimental Evidence from National AI Pilot Demonstration Zones. Journal of Jiangnan University (Humanities and Social Sciences Edition), 2024, 23(06):55-69.
[11]Ren Xianjing, Wang Feng. Construction Effects of Artificial Intelligence Pilot Zones on Firms' Green Governance Performance. Research on Economics and Management, 2025, 46(06):103-125.
[12]Chao Xiaojing, Shen Lu. Effects of AI Technology on Green Innovation Efficiency in Manufacturing Firms: Evidence from the Innovation Value Chain Framework. Economic Perspectives, 2025, (04):50-67.
[13]Yu Xinxin, Zheng Yi, Chen Chen. Action Mechanism and Realization Path of Artificial Intelligence Empowering Corporate New-quality Productivity. Journal of Technical Economics & Management, 2025, (02):45-51.
[14]Zhang Tao, Li Junchao. Digital Network Infrastructure Construction, Inclusive Green Economic Growth and Regional Imbalance: Empirical Evidence from Double/Debiased Machine Learning Causal Identification. Journal of Quantitative & Technical Economics, 2023, 40(04):113-135.
[15]Li Xueyan, Li Xin. Driving Mechanism and Economic Effect of Artificial Intelligence on Firms' High-quality Development: Evidence from Technology Innovation and Cost Management Dual Dimensions. Journal of Technical Economics & Management, 2025, (11):142-150.
[16]Bian Zuowei. Innovation of Next-Generation Artificial Intelligence and Coupling between Industrial Chain and Innovation Chain: Empirical Test Based on Innovative City Pilot Policy. Journal of Technical Economics & Management, 2026, (01):144-151.
[17]Chen Jin, Peng Gangdong, Han Weidong, Zhang Jisen. Influences of AI on Financing Efficiency in Real Economy: Theoretical Channels and Empirical Analysis. Journal of Central South University (Social Sciences), 2024, 30(06):104-118.
[18]Yang Guozhong, Xi Yuting. Empirical Analysis on Financial Restrictions Faced by Firms in Green Technology Innovation Investment. Industrial Technology Economics, 2019, 38(11):70-76.
[19]Ba Shusong, Wu Lili, Xiong Peihan. Government Subsidy Policy, R&D Spending and Listed Firms' Innovation Output. Statistics & Decision, 2022, 38(05):166-169.
[20]Ni Yuan, Zhang Jian. Cognition of S&T Talent Incentive Policies, Employees' Work Values and Corporate R&D Input. Studies in Science of Science, 2021, 39(04):632-643.
[21]Zhang Xuan, Liu Beibei, Wang Ting, Li Chuntao. Credit Rent-seeking Behavior, Financial Restrictions and Enterprises' R&D Innovation. Economic Research Journal, 2017, 52(05):161-174.
[22]Guo Ye, Su Caizhen, Zhang Yi. Does CSR Information Disclosure Promote Firms' Market Operating Performance? Systems Engineering——Theory & Practice, 2019, 39(04):881-892.
[23]Chernozhukov V, Chetverikov D, Demirer

M, et al. Double/Debiased Machine Learning For Treatment and Structural Parameters. The Econometrics Journal, 2018, 21(1):C1-C68.
[24]Hu Jie, Yu Xianrong, Han Yiming. Does ESG Rating Facilitate Firms' Green Transition? Evidence from Multi-period DID Estimation. Journal of Quantitative & Technical Economics, 2023, 40(07):90-111.
[25]Huang Qunhui, Yu Yongze, Zhang Songlin. Digital Internet Expansion and Productivity Growth in Manufacturing Sectors: Influencing Channels and Evidence from China. China Industrial Economics, 2019, (08):5-23.
[26]Zhao Tao, Zhang Zhi, Liang Shangkun. Digital Economy Growth, Entrepreneurial Behaviors and Urban High-quality Economic Development: Empirical Evidence Based on Prefecture-level Cities in China. Management World, 2020, 36(10):65-76.
[27]Jiang Ting. Mediating and Moderating Mechanisms in Empirical Causal Identification Studies. China Industrial Economics, 2022, (05):100-120.
[28]Ju Xiaosheng, Lu Di, Yu Yihua. Financial Restrictions, Working Capital Governance and Persistence of Firms' R&D Innovation. Economic Research Journal, 2013, 48(01):4-16.
[29]Guo Yue. Signaling Channels of Government R&D Subsidies and Firms' Innovative Behaviors. China Industrial Economics, 2018, (09):98-116.