

The Impact of Carbon Emissions Trading on Corporate Environmental Performance--A Hybrid Study Based on PSM-DID and Dual Machine Learning

Mengqi Chen *, Shixing Chen, Jingyi You

School of Economics and Management, East China Jiaotong University, Nanchang, China, 330013

* Corresponding Author Email: 15210307091@163.com

Abstract. In the context of "dual carbon", the green economy has gradually become the mainstream trend of economic development. As the main force of social production, the environmental governance behaviour of enterprises is crucial to China's economic transformation. Based on the panel data of Chinese A-share listed companies from 2011 to 2021, this study empirically examines, analyses the mechanism of action, and analyses the heterogeneity of enterprises using the PSM-DID model and the dual machine learning DML model. It is found that the carbon trading pilot policy significantly and positively affects firms' environmental performance, especially for non-state-owned enterprises, non-high-tech and manufacturing industries. In terms of spatial distribution, enterprises in the eastern and central regions are more significantly affected by the carbon trading pilot policy. In addition, the study also finds that executives' green awareness plays a partial mediating role in carbon trading's influence on firms' environmental performance through mediation effect analysis. Finally, the article puts forward recommendations for the government, enterprises and executives to promote the further application and effect of the carbon trading pilot policy in the enhancement of corporate environmental performance.

Keywords: Carbon Emissions Trading, Environmental Performance, PSM-DID, Dual Machine Learning.

1. Introduction

Global attention to climate change has grown in recent years, with the Paris Agreement paving the way for global climate cooperation. China will increase its nationally owned contribution, strive to reach peak carbon dioxide emissions by 2030, and aim to achieve carbon neutrality by 2060. Carbon emissions trading, a key policy tool, has been operating smoothly in China's national carbon emissions trading market since its piloting in 2011, covering emissions of more than 5.1 billion tonnes per year on average, or more than 40 per cent of the country's total emissions.

Since its implementation, the carbon emissions trading pilot policy has received extensive academic attention, and the research mainly focuses on the following aspects: firstly, using quantitative analysis methods such as multi-period double-difference models to explore the impact of the carbon trading policy on enterprise value and its intrinsic mechanism. Wang et al. (2019), taking 30 provinces and cities in China as samples, showed that the carbon trading policy reduces energy consumption while moreover promoting China's green low-carbon economy in China [1]; Secondly, the operational mechanism of carbon trading is studied from the aspects of carbon market construction and carbon trading system design and carbon quota formulation. Wang Jiaming innovatively proposes the H-CSW-DEA model to measure the initial allocation of urban carbon emission quotas [2].

Executive green cognition is an important driver of corporate green performance through green behaviour. Some scholars have introduced strategic cognition theory into the study of corporate green behaviour and explored the relationship between corporate executives' environmental awareness and energy-saving and emission reduction behaviours. For example, Malone and Ward argued that the stronger the eco-cognition of corporate executives, which consists of two factors: environmental knowledge and perception of environmental problems, the more positive their behaviour [3].

In addition, more and more scholars have begun to apply dual machine learning to the study of carbon emissions trading, which is mainly applied to carbon trading price prediction. Hao Yan (2020) used machine learning methods to construct multiple carbon trading price prediction models from the perspectives of analytical prediction, interval prediction and multi-step prediction [4]. In addition to the popular carbon trading price prediction, Li Yujie (2022) uses the random forest method in machine learning to explore the influencing factors of the executive green cognition, reveals the relationship between each influencing factor and the executive green cognition, and evolves the executive green cognitive and predict the trend of green perception of executives The study will be carried out in the future [5].

In summary, the academic community has not yet reached a unified conclusion on whether there is an interactive relationship between carbon emissions trading pilot policies and corporate environmental performance, and the mediating role played by executives' green perceptions. This paper analyses the environmental management activities of enterprises from the micro level under the perspective of carbon trading, and discusses the role of carbon trading in executives' green perceptions, and then the further impact on the environmental performance of enterprises. Secondly, machine learning techniques have been used in the literature mainly for carbon price prediction, but machine learning is used in causal inference. This paper utilises a dual machine learning model to empirically test, mechanism and analyse heterogeneity with higher accuracy.

2. Content of the study

2.1. Research hypotheses

To address the research question, we formulate the hypothesis:

H1: Carbon emissions trading pilot policies have a positive effect on corporate environmental performance.

H2: Carbon emissions trading pilot policy can improve corporate environmental performance through executives' green perceptions.

2.2. Modelling

2.2.1 DID model construction

$$CEG_{i,t} = \alpha + \beta_1 did_{i,t} + \gamma Controls + \sum id + \sum year + \sum Prov \times year + \varepsilon_{i,t} \quad (1)$$

In equation (1), the $CEG_{i,t}$ represents corporate environmental performance; $did_{i,t}$ represents the implementation of carbon emissions trading policy; Controls represents the control variables; the model also incorporates individual fixed effects (id), time fixed effects (Year), and province-time interaction fixed effects (Prov*year); $\varepsilon_{i,t}$ represents the random perturbation term.

Intermediation modelling:

$$EA_{i,t} = \beta_0 + \beta_1 did_{i,t} + \gamma Controls + \sum id + \sum year + \sum Prov \times year + \varepsilon_{i,t} \quad (2)$$

$$CEG_{i,t} = \delta_0 + \delta_1 did_{i,t} + \delta_2 EA_{i,t} + \gamma Controls + \sum id + \sum year + \sum Prov \times year + \varepsilon_{i,t} \quad (3)$$

Where $EA_{i,t}$ is the mediating variable, representing the green cognition of the executives of listed companies; β_1 indicates the degree of influence of the carbon emission right trading pilot policy on the green cognition of the executives of listed companies; δ_1 denotes the degree of influence of carbon emissions trading pilot policy on corporate environmental performance after the introduction of mediating variables.

2.2.2 PSM-DID model construction

However, the impact of carbon emission trading on the environmental performance of enterprises cannot be directly realized through the comparison of DID double scores. There are two main reasons for this:

(1) As a quasi-natural experiment, it is difficult to avoid the problems caused by self-selection bias and time and individual differences, and the selection of listed companies for carbon trading pilots is not random, so there may be sample selection bias, resulting in not meeting the prerequisites of difference-in-difference natural experiments.

(2) Due to the small number of listed companies participating in carbon emission trading, there may be bias in regression with the full sample, so the estimation results strictly depend on the selection of the control group [6, 7].

Therefore, in this paper, the PSM-DID model is constructed:

$$ATT = E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 1) = E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 0, P(X)) = [E(y_{1i}|D_i = 1, T = 1) - E(y_{0i}|D_i = 0, T = 1, P(X))] - [E(y_{1i}|D_i = 1, T = 0) - E(y_{0i}|D_i = 0, T = 0, P(X))] \quad (4)$$

Of these, the y_{1i} and y_{0i} denote potential outcome variables when individual i receives and accepts the treatment, respectively; D_i denotes whether individual i receives the treatment or not, taking the value of 0 or 1; T denotes time, taking the value of 0 or 1; $P(X)$ denotes the propensity score function, a function on the observable covariate X ; $E(\cdot)$ denotes conditional expectation.

2.2.3 Dual machine learning original model construction

This paper refers to Chernozhukov et al. (2018), Zhang Tao et al and Wu Han approach, a semiparametric regression model is built first [8, 9]:

$$Y_{i,t} = \theta_0 D_{i,t} + g_0(X) + U_v \quad (5)$$

$$E(U|X, D_{i,t}) = 0 \quad (6)$$

Where Y is the observed value of the dependent variable CEG, DID is the disposal variable taking the value of (0, 1), i, t denotes the individual firm, year respectively; X is the control variable, and θ_0 is the coefficient of disposition we wish to obtain, which needs to be estimated by machine learning methods $\hat{g}_0(X)$ specific form; U is the error term, with conditional mean $E(U|X, D_{i,t})$ is 0. The model can be specifically represented as:

$$Y_{i,t} = \theta_0 D_{i,t} + g_0(\text{Size, Lev, Cashflow, ATO, Indep, Growth, FirmAge, Board}) + U \quad (7)$$

2.3. Selection of variables

2.3.1 Choice of explanatory variables: corporate environmental performance CEG

This paper borrows from Li Xiaomei, Li Manman established corporate environmental governance index system, set a three-level index system of corporate environmental performance performance, scores according to: 0 points for no description, 1 point for qualitative description, 2 points for quantitative description. The total CEG score is the sum of each item, and the higher the total score, the better the enterprise's environmental performance is considered to be [10]. The details are shown in Table 1.

Table 1. Breakdown of environmental performance indicators for enterprises

Level 1 indicators	Secondary indicators	Tertiary indicators	Scoring criteria
Corporate environmental performance	Government regulatory disclosure	Is a key pollution monitoring unit	(0, 1, 2)
	Environmental pollution control efforts	Emission reduction and management of waste gases	(0, 1, 2)
		Wastewater abatement and management	(0, 1, 2)
		Dust and fume management	(0, 1, 2)
		Solid waste utilisation and disposal	(0, 1, 2)
		Treatment of noise, light pollution, radiation, etc.	(0, 1, 2)
		Cleaner production implementation	(0, 1, 2)
		special environmental operation	(0, 1)
		Environmental Incident Response Mechanism	(0, 1)
		Pollutant emission compliance	(0, 1)
	Pollutant Release Information	Wastewater discharge	(0, 1, 2)
		COD emissions	(0, 1, 2)
		SO2 emissions	(0, 1, 2)
		CO2 emissions	(0, 1, 2)
		Smoke and dust emissions	(0, 1, 2)
		Industrial solid waste generation	(0, 1, 2)
	disclosure vehicle	Annual reports of listed companies	(0, 1)
		Social Responsibility Report	(0, 1)
		Environmental reporting	(0, 1)
	Environmental management evaluation	Environmental Philosophy	(0, 1)
		environmental objective	(0, 1)
		Environmental management system	(0, 1)
		Environmental Education and Training	(0, 1)
		"Three simultaneous" system	(0, 1)
		ISO14001 certified or not	(0, 1)
		ISO9001 certified or not	(0, 1)

2.3.2 Choice of explanatory variables: policy pilots DID

According to the provincial and municipal development and reform commissions, ecological bureaus and carbon emission rights exchange related information to match the listed companies involved in the pilot and extract the time of their participation in the carbon trading pilot policy, if the listed company is involved in the carbon trading pilot, the time (time) from the year it starts the pilot to 2021 is taken as 1, otherwise it is taken as 0; at the same time, if the listed company is a pilot

(treat) company, then all the years of the company corresponding to treat take the value of 1, otherwise take 0; make $DID = time \times treat$, we get the explanatory variable DID, which indicates that the enterprise has implemented the carbon trading policy in that year, and takes the value of 1, otherwise it takes 0.

2.3.3 Control variable selection

Considering the policy's performance on the environmental performance of listed companies, we selected the following indicators as control variables, as shown in Table 2.

Table 2. Definitions of selected variables

Variable type	variable name	variable label
explanatory variable	Corporate environmental performance performance	CEG
explanatory variable	Policy pilots	DID
control variable	Company size	Size
	gearing	Lev
	Cash flow ratio	Cashflow
	Total asset turnover	ATO
	Percentage of independent directors	Indep
	Growth rate of operating income	Growth
	Years of incorporation	FirmAge
	Number of directors	Board
intermediary variable	Executive Environmental Awareness	EA

2.4. Data sources

This paper selects A-share listed companies from 2011 to 2021 as the research sample, and the data of corporate financial data and related indicators of corporate environmental performance are from Wind database (Wind) and CSMAR database (CSMAR). Referring to Duriau et al. (2007) on the measurement of executive perception, the text analysis of listed companies' annual reports is conducted, and relevant keywords are selected for word frequency statistics, to construct listed companies' executives' green perception, and to measure the green concern of the corporate management's decision-making.

3. Impact of carbon emissions trading on corporate environmental performance

3.1. Analysis of PSM-DID modelling results

Whether to be selected as a pilot city for carbon pilot trading policy is non-random, to ensure the robustness of the estimation results, DID double differencing and PSM-DID propensity score matching-double differencing are performed respectively, and the results are as follows Table 3 The results are shown in Table 3 below.

Columns (1)-(3) use the DID double difference model: column (1) does not include control variables; column (2) includes control variables; and column (3) includes individual and year fixed effects and province-time interaction fixed effects in the model. Columns (4)-(5) show the regression results using the PSM-DID model: column (4) without control variables; column (5) with control variables and fixed effects. The coefficients of the explanatory variables of the five models, DID, on the environmental performance of enterprises (CEG) are all significantly positive at the 1% level, indicating that the carbon emissions trading policy can significantly affect the environmental performance of enterprises. Among them, the coefficient of enterprise size is significantly positive, indicating that larger enterprises are more affected by carbon trading policy and have better feedback for the policy, which is in line with reality.

Table 3. Regression results

variant	(1) CEG	(2) CEG	(3) CEG	(4) CEG	(5) CEG
DID	3.682*** (13.85)	2.09*** (8.52)	1.52*** (3.35)	1.476*** (0.001)	1.518*** (0.001)
Size		2.18*** (46.89)	0.894*** (6.07)		0.894*** (0.000)
Lev		-3.92*** (-12.76)	-0.878 (-1.61)		-0.878 (0.108)
Cashflow		10.221*** (0.758)	0.814 (1.34)		0.814 (0.182)
ATO		0.862*** (13.49)	0.162 (0.63)		0.162 (0.528)
Indep		0.001 (0.10)	0.0189 (1.20)		0.019 (0.229)
Growth		-0.505*** (-3.67)	-0.220** (-2.20)		-0.220** (0.028)
FirmAge		1.540*** (10.63)	2.343** (2.48)		2.343** (0.013)
Board		0.665** (2.19)	-1.061* (-1.93)		-1.061* (0.054)
constant term (math.)	8.316*** (157.17)	-45.881*** (-38.03)	-16.699*** (-3.74)	8.404*** (0.000)	-16.699*** (0.000)
individual fixed effect	clogged	clogged	be	be	be
Year fixed effects	clogged	clogged	be	be	be
Province/time fixed effects	clogged	clogged	be	be	be
sample size	18193	18193	18193	18195	18193
R ² value	0.0104	0.1683	0.7605	0.720	0.723

3.2. Analysis of dual machine learning results

The machine learning method of random forest is applied for regression and the results are shown in Table 4 shown in Table 4, are significant at the 5 per cent level, indicating that the carbon emissions trading policy has a significant positive impact on corporate environmental performance.

Table 4. Dual machine learning random forest regression

	DID	Z-score	SE	P-value	N
Random Forest	1.220**	2.40	0.297	0.016	18193

3.3. Dual machine learning robustness test

3.3.1 Changing machine learning methods

Replacement of machine learning algorithms, replacing the Random Forest algorithm previously used as a predictor with LASSO regression, gradient boosting, and neural networks, to explore the possible impact of the prediction algorithms on the conclusions of this paper.

The regression results are as follows Table 5 The regression results are shown in Table 5. All three methods are significant at 1% level; the regression coefficient of the gradient boosting tree is the largest, but the standard error is also the smallest, while the standard error of the neural network regression reaches 43.7%, and the corresponding regression coefficient is the smallest, with the value of 1.135. The regression results of the three methods are significant despite their differences, which indicates that the regression has robustness.

Table 5. Dual machine learning regression results

	LASSO	Gradient Boosting	Neural Network
DID	1.661***	2.348***	1.135***
Z-score	5.48	8.15	2.60
SE	0.303	0.288	0.437
P-value	0.000	0.000	0.000
N	18193	18193	18193

3.3.2 Changing sample proportions

Changing the sample split ratios of the dual machine learning model from the previous 1:4 to 1:2 and 1:7 to explore the possible impact of the sample split ratios on the conclusions of this paper. See Table 6 The regression results are still significant at the 5% level.

Table 6. Changing Scale DML Regression

	(1) 1:2	(2) 1:7
DID	1.311**	1.493**
Z-score	2.76	2.77
SE	0.474	0.538
P-value	0.006	0.006
N	18193	18193

3.3.3 Removal of parallel policy effects

Changes in firms' environmental performance are also confounded by other contemporaneous policies. In this paper, a dummy variable for did2 is constructed, which takes the value of 1 after being located in the pilot city and conducting the pilot, and 0 for the rest, and is added to the regression analysis. The results of the DML regression analysis are presented in Table 7, all of which are significantly positive at the 1 per cent level.

Table 7. DML regression results

	DID	SE	Z-score	P-value	N
CEG	1.595***	0.520	3.07	0.002	18183

4. Mediating Effects of Green Perceptions of Executives

In order to explore the impact mechanism of carbon emissions trading policy on corporate environment and performance, and to consider that the pilot carbon emissions trading policy may have an impact on corporate environmental performance by influencing executives' green perceptions, this paper explores in-depth the role played by executives' green perceptions of listed companies in the enhancement of corporate environmental performance by the carbon rights trading policy.

Table 8. Intermediation effect regression results

variant	ln_EA	CEG
DID	0.190*** (2.83)	
ln_EA		2.157*** (36.55)
Controls	be	be
fixed effect	be	be
SE	0.067	0.059
sample size	16213	16213

The regression results are shown in Table 8 shows, the carbon emissions trading pilot policy has a significant positive effect on executive sustainability awareness. At the same time, executive sustainability awareness also has a significant positive effect on corporate environmental performance ECG. This indicates that executive sustainability awareness mediates the relationship between carbon emissions trading policy and corporate environmental performance.

5. Heterogeneity test

5.1. Tests for heterogeneity in the nature of property rights

In order to investigate the relationship between the carbon emissions trading pilot policy and corporate environmental performance (CEG) under different property rights, the enterprises are divided into state-owned enterprises and non-state-owned enterprises, and regressions are conducted separately, and the results are shown in Table 9. The environmental performance of state-owned enterprises is more significantly positive at the 10 per cent level, while the environmental performance of non-state-owned enterprises is significantly positive at the 5 per cent level, which is a more significant response to the policy. The reason may lie in the fact that non-state-owned enterprises hope to get the government's support by responding positively to the policy to reduce the cost of the enterprise; while state-owned enterprises because of its nature has a higher degree of importance to energy conservation and emission reduction, the policy implementation of the promotion of the space is small, the impact of the policy is relatively small.

Table 9. DML Heterogeneity Tests (Nature of Firm, Region)

	state-run	non-municipal	eastern part	western part	central section
DID	1.207*	1.298**	1.075**	3.062	2.719***
SE	0.706	0.614	0.534	2.163	1.112
Z-score	1.71	2.11	2.01	1.42	2.79
P-value	0.087	0.034	0.044	0.157	0.005
N	7746	10437	12972	2293	2918

Significance: *** p<0.01, ** p<0.05, *p<0.1

5.2. Tests for spatial heterogeneity

After grouping, regression analyses were carried out for the eastern, western and central groups, and the regression results are shown in Table 9 The regression results are shown in Table 9. Except for the western region, which is not significant, the eastern and central regions are significantly positive at the 5% and 1% levels, respectively. Carbon trading pilot provinces and cities are also mostly concentrated in the eastern region, and their enterprises have high capital maturity and better systems, so the impact is significant; while in the central region, there are a large number of heavy industrial enterprises, and their energy demand and carbon emissions are larger than other enterprises, so the improvement of enterprise environmental performance is more significant.

5.3. Tests for industry heterogeneity

According to the classification standard for the manufacturing industry in the National Economic Industry Classification (GB/T 4754-2017) published by the National Bureau of Statistics, the manufacturing industry is 1, and the non-manufacturing industry is 0. For the classification of high-tech industries, Guo Lei, Xiao Shufang, Li Xuejing et al. method, the high-tech industry is 1 and the non-high-tech industry is 0 [6].

The environmental performance of high-tech industries is significantly positive at the 5 per cent level, while that of non-high-tech industries is significantly positive at the 1 per cent level, as shown in Table 10. Shown. The reason for this is that in order to improve the environmental performance of non-high-tech industries, they will gradually begin to carry out technological innovation and research and development, and establish a more complete environmental pollution control and environmental

management system; while high-tech industries generally have a higher awareness of energy saving and emission reduction, and most of the industry has adopted high-tech technology to reduce environmental pollution, environmental performance is limited to improve the space, so the significance of the policy implementation is slightly less significant than that of the non-high-tech industries. Therefore, the significance of the policy is slightly lower than that of non-high-tech industries.

The results of the regressions for whether the industry is a manufacturing industry are shown in Table 10. Manufacturing is significant at the 1 per cent level, while non-manufacturing is not. The reason for this may be that manufacturing firms with higher emissions are subject to stricter restrictions, have stricter regulation and enforcement, and are required to be more transparent about their pollution emission information.

Table 10. DML Heterogeneity Tests (Industry)

	High-tech industries	Non-high-tech industries	service industry	non-manufacturing industry
DID	1.512**	2.216***	1.636***	0.898
SE	0.623	0.744	0.528	0.740
Z-score	2.43	2.98	3.09	1.21
P-value	0.015	0.003	0.002	0.225
N	8601	9582	11908	6275

Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

6. Conclusion

In the context of "dual carbon", this study conducted empirical tests, mechanism analysis and heterogeneity analysis using PSM-DID model and dual machine learning DML model. It is found that the carbon trading pilot policy has a significant positive impact on the environmental performance of enterprises. The mediating effect analysis found that executives' green awareness played a partial mediating role in carbon trading affecting corporate environmental performance. The paper also conducts a heterogeneity analysis and finds that carbon trading has a more significant impact on non-state-owned enterprises, non-high-tech and manufacturing industries. In terms of spatial distribution, enterprises in the eastern and central regions are more significantly affected by the carbon trading pilot policy.

In this regard, we put forward the following suggestions: the government should actively promote carbon trading, achieve differentiated management, dynamic implementation, and at the same time increase incentives and set up a complete risk prevention and control system to motivate enterprises to devote themselves to green development; corporate executives should profoundly cultivate green awareness and practice green behaviours, so as to enable enterprises to take a new step forward in green development; state-owned enterprises should play a moderating role in the carbon market, implement the policy content, and promote green technological innovation; non-state-owned enterprises should embark on the road to green development by enhancing the green awareness of executives, technological innovation, green transformation, etc. policy content and promote green technological innovation, and non-state-owned enterprises should embark on the road of green development by enhancing the green awareness of executives, technological innovation and green transformation.

References

- [1] Han Wang, Zhoupeng Chen, Xingyi Wu, Xin Nie. Can a carbon trading system promote the transformation of a low-carbon economy under the framework of the porter hypothesis?-Empirical analysis based on the PSM-DID method [J]. Energy Policy, 2019, 129.

- [2] Wang, J.M.. Research on Carbon Emission Quota Allocation in China's Resource-Based Cities-Based on H-CSW-DEA Model [J]. Statistical Research, 2023, 40 (09): 31-44.
- [3] MALONE M P, WARD M P. Let's hear from the people: an objective scale for the measurement of ecological attitudes and knowledge [J]. American Psychologist, 1973, 28 (7): 583-586.
- [4] Hao Yan. Research on carbon trading price prediction method and application based on machine learning [D]. Northeast University of Finance and Economics, 2022.
- [5] Li Yujie. Analysis and Prediction of Influential Factors of Green Technology Innovation Based on Random Forest Model [D]. Shanxi Normal University, 2022.
- [6] Guo Lei, Xiao Shufang, Li Xuejing et al. Non-executive Employee Equity Incentives and Innovation Output - Empirical Evidence Based on Listed High-Tech Firms in China [J]. Accounting Research, 2019 (07): 59-67.
- [7] Zhou Chang, Cai Haijing, Liu Meijuan. The financial effect of carbon emission trading on micro enterprises: PSM-DID test based on "Porter's hypothesis" [J]. Journal of Finance and Economics, 2020 (03): 68-77.
- [8] Zhang Tao, Li Junchao. Network Infrastructure, Inclusive Green Growth and Regional Disparities - Causal Inference Based on Dual Machine Learning [J]. Research on Quantitative and Technical Economics, 2023, 40 (04): 113-135.
- [9] Wu Han. Evaluation of carbon trading pilot policy based on dual machine learning [D]. Southwestern University of Finance and Economics, 2024.
- [10] Li Xiaomei, Li Manman. Green Investment and Corporate Environmental Governance under Carbon Neutral Goal - A Mediation Effect Test Based on Technological Innovation [J]. China Science and Technology Forum, 2022, (09): 118-127+138.