Asian Journal of Water, Environment and Pollution. 2026;23(2):025430333.
doi: 10.36922/AJWEP025430333

ORIGINAL RESEARCH ARTICLE

How corporate social responsibility improves green transformation technologies inventory output: Evidence from Japanese companies

Jiaming Jiang[1]*, Yu Zhao[2], and Junshi Feng[3]

[1]School of Information and Data Sciences, Nagasaki University, Nagasaki, Japan
[2]Department of Management, School of Management, Tokyo University of Science, Tokyo, Japan
[3]School of International Economics and Trade, Jilin University of Finance and Economics, Changchun, Jilin, China
***Corresponding author:** Jiaming Jiang (jiaming@nagasaki-u.ac.jp)

Received: October 24, 2025; 1st revised: October 31, 2025; 2nd revised: December 19, 2025; Accepted: December 26, 2025; Published online: January 15, 2026

Abstract: Corporate social responsibility (CSR) refers to businesses' contributions to sustainable development. Today, it is expected that economic, social, and industrial structures that have long depended on fossil fuels will ultimately transition to clean energy-driven structures, and that green transformation (GX) will be achieved through an overall transformation of economic and social systems. To the best of our knowledge, this study is the first to examine the relationship between the CSR of Japanese companies and their GX inventory output (GXIO). Using machine learning algorithms, we tested the hypothesis that CSR activities influence GXIO at both the individual company and the overall industry levels. The results of our empirical analysis show that CSR not only affects GXIO but also has significant effects at the individual company and overall industry levels. These results are used to provide decision makers with useful guidelines for integrating social sustainability into procurement procedures.

Keywords: Social responsibility; Green transformation technology; Innovation; Japanese corporations; Statistical analysis; Environment; Patent; Machine learning

1. Introduction

In an economic context, innovation commonly refers to "innovation that produces economic results." Specifically, it refers to creative technological advancement that develops into a product that is accepted in the marketplace, spreads and permeates society, and results in a profitable business. According to Takeishi *et al.*,[1] innovation is the reorganization of the flow of resource mobilization in society, taking the productive forces used in the activity out of their existing cyclical economic path and redirecting them toward new activities. Today, innovation analysis is considered one of the most effective methods for discerning green transformation (GX) trends and objectively demonstrating companies' GX-related efforts and their impact on climate change. The GX Technologies Inventory (GXTI) is a technology inventory that offers a bird's-eye view of GX-related technologies, published in June 2022 by the Japan Patent Office (JPO).

Following the Paris Agreement of 2015, the world has been moving toward carbon neutrality. As noted by Sueyoshi and Goto,[2] Japan has historically faced a variety of environmental problems associated with its industrialization. In recognition of its responsibilities, the Japanese government issued the “2050 Carbon Neutrality Declaration” in 2020. The trend toward a decarbonized Japanese society appears to be rapidly approaching an irreversible state. Recently, based on the “GX Promotion Law” enacted in May of 2023, the GX promotion strategy was approved by the cabinet of ministers, making the movement toward GX clearly evident.

At the same time, a change in the culture of business organizations has been taking place. Various stakeholders are making ever-increasing demands on corporations to increase their contributions to society,[3] putting companies under enormous pressure and holding them accountable for the social and environmental impacts of their policies and practices.[4] Companies have become increasingly aware of the repercussions of their actions on society, fostering the emergence of corporate social responsibility (CSR)[5] as an important factor in the management of their business operations. Japanese companies participating in the GX League, a network of more than 500 Japanese companies, are required to pursue their own emission-reduction efforts and work toward carbon neutrality in their supply chains. It is argued that corporations that are highly concerned about environmental issues, as demonstrated through their GX efforts, will gain the trust of their business partners, consumers, and others, leading to effective corporate branding.

As stated by the Commission of the European Communities (2001), CSR is a concept whereby companies integrate social and environmental concerns in their business operations and in their interaction with their stakeholders on a voluntary basis[6(p6)]. CSR activities can strategically enhance a firm’s value through investment in firm-specific intangible resources, such as organizational innovation capability, human resources, corporate culture and reputation, and trust.[7] In particular, environmental protection has emerged as a significant business issue.[8]

Despite growing interest in corporate sustainability, several important questions remain insufficiently addressed in the academic literature, particularly within the context of Japan. Specifically, it is still unclear whether:

(i) Whether CSR activities actually contribute to firms’ GX inventory output (GXIO);
(ii) how strongly CSR activities influence GXIO, and through which mechanisms;
(iii) Whether the impact of CSR on GXIO differs between the firm level and the broader industry level in Japan.

To address these gaps, our study focuses on Japanese companies and investigates how CSR engagement affects green innovation performance. By examining CSR–GXIO linkages at both the firm and industry levels, we aim to deepen the theoretical understanding of CSR as an intangible resource in Japan and offer practical implications for managers and policymakers seeking to advance GX.

The main contributions of the study include the construction of a database that simultaneously supports the CSR and GXIO of the surveyed companies (The GXIO database is published in Figshare and is publicly available). In addition, our study provides useful guidance for management.

The remainder of the paper is structured as follows: Section 2 reviews the background and methodology-related literature. Section 3 describes the datasets; Section 4 introduces the analytical method; Section 5 reports the study’s results; Section 6 discusses the study’s implications; and Section 7 concludes the paper.

2. Literature review

Montalbán-Domingo *et al.*[9] highlighted the importance of implementing CSR practices to achieve both economic benefits and societal well-being. Santos-Jaén *et al.*[5] showed that the effective development of human resource practices based on strategies oriented toward CSR allows companies to undertake greater and more efficient innovative activities, and implied that CSR-oriented innovation has proven to be a valuable strategy for more efficient enterprise management due to the multiple competitive advantages it generates. Ratajczak and Szutowski[10] empirically demonstrated the positive relationship between CSR and innovation. Extensive research has also established that innovation serves as a positive link between CSR and firm performance.[11,12] Bocquet *et al.*[13] validated the hypothesis that strategic CSR has a positive effect on technological innovation. CSR has evolved from traditional philanthropy into strategic business decision-making and has become a strategic requirement for virtually all businesses, including those in Japan.[14] Dögl and Behnam[15] empirically tested the antecedents and outcomes of corporate environmental responsibility (CER) practices as well as country differences across institutional environments. The results showed that CER practices can positively influence corporate business outcomes. The findings of Li and Lin[16] revealed that slower sustainable total-factor productivity growth in emerging economies is caused by their disadvantages in innovation. Therefore, encouraging innovation might be a breakthrough in stimulating their sustainable development. Chkir *et al.*[17] examined this issue using a newly available comprehensive innovation database across 20 countries and found support for the view that CSR performance fosters innovation, reporting a positive impact of CSR on innovation in civil-law countries. Therefore, CSR-related challenges could become a source of inspiration for novel innovations.[18] Mishra[19] showed that more innovative firms exhibit high CSR performance and concluded that innovative firms benefit from engaging in CSR activities. Furthermore, numerous recent studies have explored the relationship between CSR and green innovation, motivated by the fragmented and inconsistent findings in earlier research. Gürler's[20] results indicate that the strong and positive relationship between CSR and green innovation remains robust across both manufacturing and non-manufacturing industries. Padilla-Lozano and Collazzo[21] examined the impact of CSR and green innovation on firm competitiveness in the manufacturing sector and found that both positively and significantly enhance firm competitiveness. Le[22] investigated the partial relationship between corporate green strategy and sustainable firm performance among small and medium-sized enterprises (SMEs), highlighting the mediating roles of CSR and green innovation in an emerging economy. However, in Japan, existing research tends to address CSR in terms of social contribution, legitimacy, or sustainability practices, while environmental or green innovation is often discussed separately through the lenses of environmental management, sustainability-oriented innovation, or social innovation.[23,24] Consequently, there is a lack of empirical studies that explicitly integrate CSR with green innovation in a unified analytical framework. This fragmentation indicates the need to connect insights from CSR-focused studies with those on environmental and social innovation to fully understand how CSR activities may influence green innovation among Japanese firms.

In recent years, the development of artificial intelligence (AI) and machine learning technologies has made significant progress. Machine learning enables organizations and researchers to build predictive models from assembled datasets.[25] Previous research has tested its contribution in experiments. For example, Cai *et al*.[25] conducted experiments on the Modified National Institute of Standards and Technology dataset (a labeled dataset of handwritten digits) and reported robust contribution scores from both protocols, with larger datasets receiving a greater share of the model. Their experiments also demonstrated the need to combine the two-phase commit (2PC) protocol with a robust model aggregation mechanism to discard low-quality inputs resulting from model poisoning attacks. Schams[26] envisioned a system that models an individual's contributory social capital (CSC), comprising competence, trustworthiness, and social responsibility. The availability of the CSC score can stimulate social behavior and mutual support. Xia *et al.*[27] proposed a sequential ensemble credit scoring model based on a variant of the gradient boosting machine, namely the extreme gradient boosting (XGBoost) algorithm. XGBoost is an optimized version of the gradient boosting machine and can normalize the loss function to reduce model variance. It also reduces modeling complexity and, in turn, the likelihood of model overfitting.[28] Recently, Deng and Lumley[29] proposed a scalable multiple imputation (MI) framework (mixgb) based on XGBoost, subsampling, and predictive mean matching (PMM). To the best of our knowledge, this is the first attempt to implement and evaluate CSR scores using the XGBoost algorithm and to rank CSR activities by their predicted contribution scores. The use of MI to address missing data is becoming increasingly popular, and mixgb is a promising automated approach for large, complex datasets.[29] The results provide an overview of how CSR performance affects GXIO and advance our understanding of which CSR activities are most important for companies. The aim is to gain valuable insight into current corporate CSR and sustainability practices in Japan.

3. Materials

A considerable body of prior research has examined whether CSR performance influences corporate innovation across countries, often comparing developed and emerging economies.[16] However, to provide in-depth insights, we selected Japan as the focus of our analysis. We used the CSR database provided by Toyo Keizai Inc., a leading business publisher in Japan, as one of our principal databases. The Toyo Keizai CSR is the best-known comprehensive CSR database on Japanese companies. Most of the sample companies are listed on the first section of the Tokyo Stock Exchange Market.[8] The study period was from June to October 2022. We focused on the file kankyo (environment) 2023.csv, which contained responses from 1,703 corporations from 34 industry sectors. As the CSR database provided by Toyo Keizai Inc. is based on a questionnaire survey, we selected up to 10 companies from each industry sector that answered at least 50% of the survey questions. For several industries, fewer than 10 companies were selected due to limited responses. The survey questions generally focused on how these companies consider the environment in their production processes, for example, having a medium-term plan for the environmental sector, the reduction of carbon dioxide (or greenhouse gas) emissions, the status of green purchasing initiatives, risk management in the environmental field, and the creation of an environmental management system (EMS).[8] The variables of interest were taken from Goto and Sueyoshi.[8] Finally, we collected 298 responses regarding the sampled companies' CSR activities.

We then searched the GX patent applications of the selected companies on the Japan platform for patent information. The search formulas are publicly available (https://www.jpo.go.jp/e/resources/statistics/gxti.html) and use a combined search strategy of International Patent Classification (IPC) class and keywords. We found 37,476 patent applications for the 298 corporations in the final sample over the period 2000–2022. We assumed that Japanese corporate names are commonly written in both kanji and katakana; therefore, when we searched for a company's patents, we used the company's name in both[*1*].[30] Importantly, we have added two CSR-related variables that have rarely been studied: Supply chain emissions and International Organization for Standardization (ISO) 14001 EMS certifications. The first of these additions was made because a question on supply chain emissions was newly included. With respect to the second addition, the ISO 14001 standard provides guidelines for organizations to continuously improve their environmental performance and demonstrates an organization's commitment to sustainable production processes.[31] It also enables an organization to demonstrate its environmental responsibility to discerning global customers, thereby generating greater interest in its products.[32]

The main limitation of our database lies in the presence of missing values arising from the CSR questionnaire, and the key methodological challenge was to address these missing data using machine learning techniques. Furthermore, although CSR data are available only for 2022, CSR engagement is widely regarded as a relatively stable, long-term organizational attribute rather than a short-term strategic choice.[33] We interpreted the CSR measure as reflecting firms' sustained commitment to social and environmental responsibility associated with cumulative GXIO, rather than as a factor that causes past patenting activity. Table S1 describes our dataset.

Figures 1 and 2 show the relationship between emissions and GXIO, and inputs and GXIO, respectively. In Figure 1, which shows emissions and GXIO by industry, we observed strong contrasts. Several industries showed high emissions but low GXIO. These included petroleum and coal products, mining, electricity and gas, rubber products, iron and steel, and shipping. However, transportation equipment and electrical appliances showed both high emissions and high GXIO, while textiles had low emissions but relatively high GXIO. The other industrial sectors showed both low emissions and low GXIO.

Figure 2 shows the resource input and aggregation of GX patents by industry sectors. Several industries required high inputs with low GXIO, such as petroleum and coal products, mining, and electricity and gas. Some sectors needed low energy inputs but had relatively high GXIO, such as transportation equipment, glass and stone products, non-ferrous metals, electrical appliances, and textiles. Chemistry requires significant water resources input, but has a relatively high GXIO. Most industries had both low energy inputs and low GXIO.

Figure 3 focuses on Japanese companies and shows the CSR of selected companies along with their GXIO. Among the companies shown are those that demonstrated concern for the environment, those that had low emissions and high GXIO, such as Mitsubishi Electric Corporation, Honda Motor Co., Ltd., and Nissan Motor Corporation. On the other hand, some companies had both high emissions or energy inputs and a low GXIO, such as Idemitsu Kosan Co., Ltd., and Mitsubishi Materials Corporation. Idemitsu Kosan Co., Ltd. stands out for remarkably high emissions and energy inputs with a relatively low GXIO. Mitsubishi Materials Corporation and Mitsui Chemicals, Inc. showed high energy input but low GXIO.

4. Analytical method

We proposed the following three hypotheses:

H_1: The GXIO of companies and their industries can be attributed to the CSR activities in which they engage.

H_2: The extent of GXIO in companies and their industries is positively associated with the level of CSR activities in which they engage.

H_3: The impact of CSR activities on GXIO differs between the company level and the industry level.

1 For example, ホンダ (Honda) can also be written as 本田, and トヨタ (Toyota) is sometimes written as 豊田.

 doi: 10.36922/AJWEP025430333

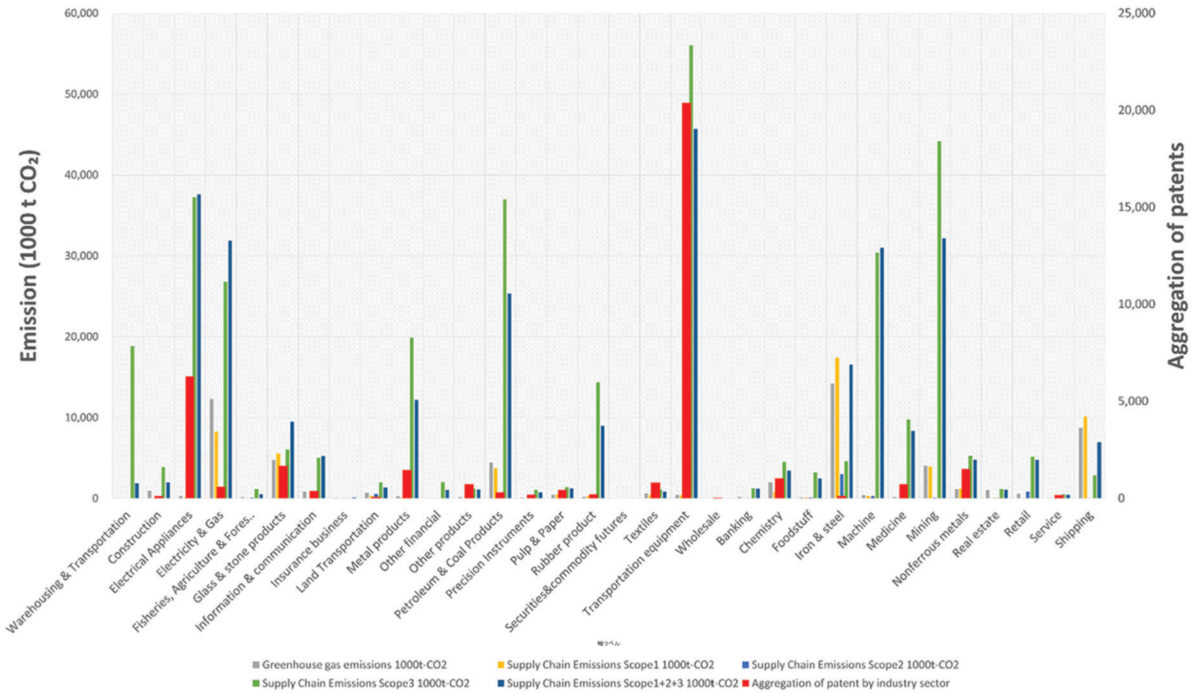

Figure 1. Emissions and aggregation of green transformation patents by industry sector
Abbreviation: CO_2: Carbon dioxide.

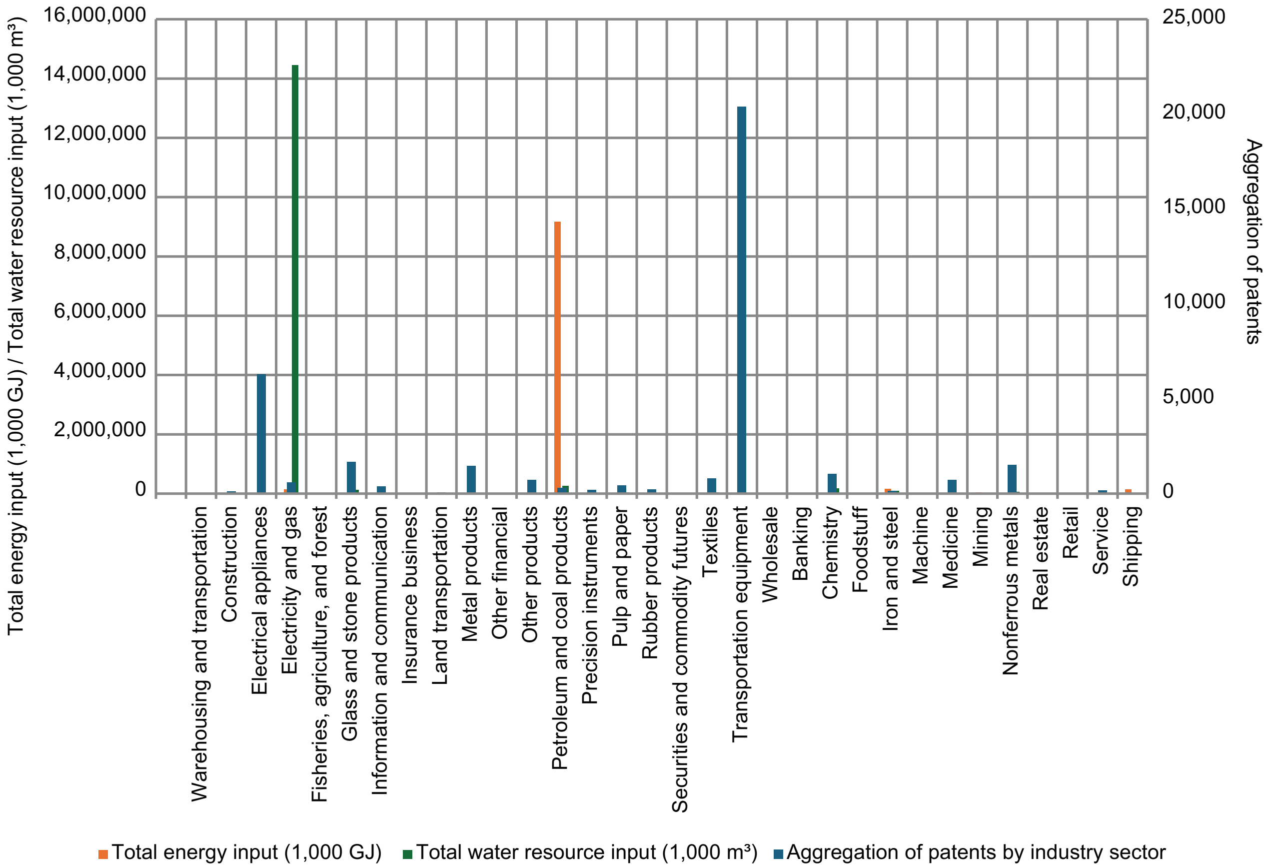

Figure 2. Resource input and aggregation of green transformation patents by industry sector

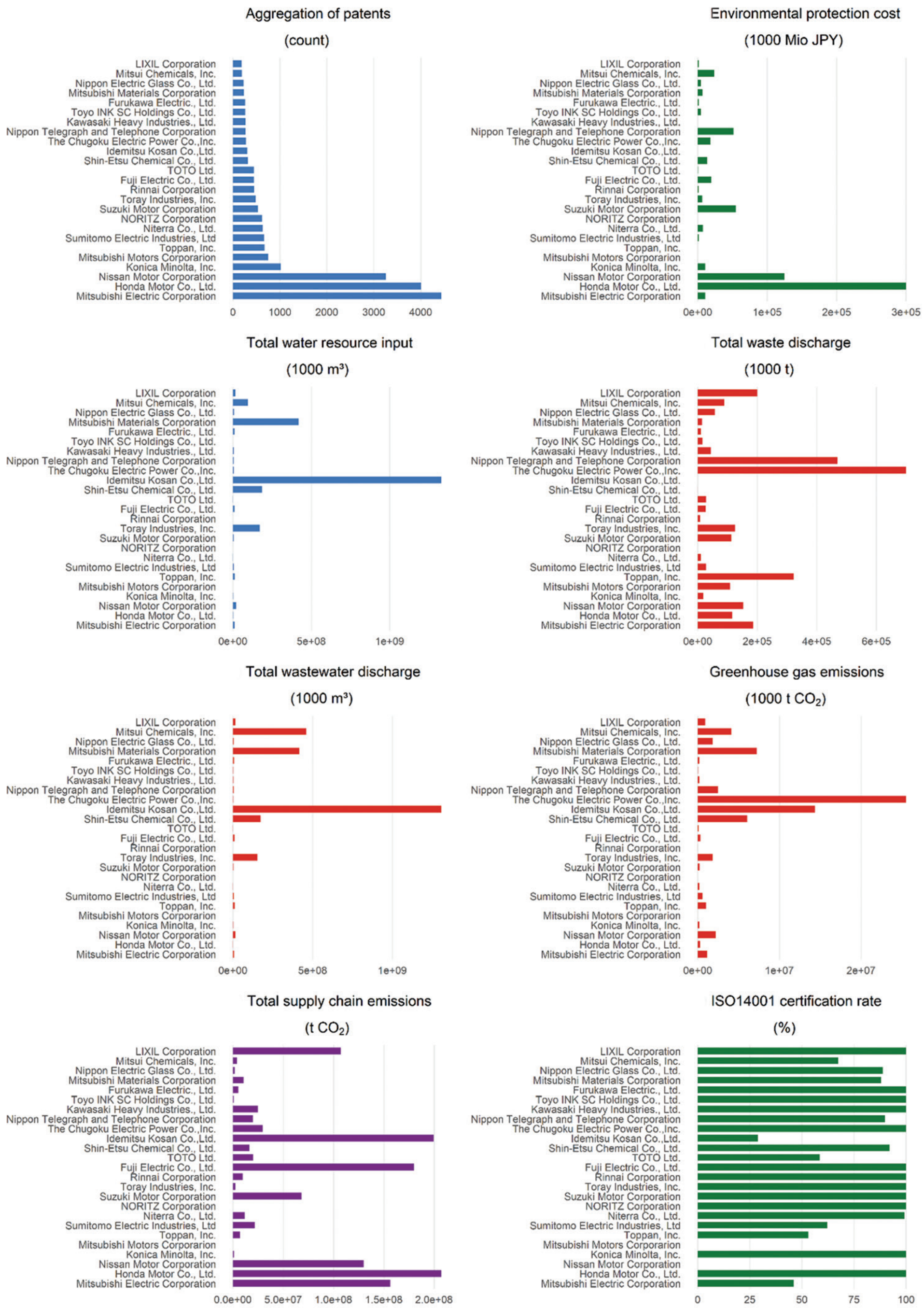

Figure 3. Emissions, energy input, and aggregation of green transformation patents by company
Abbreviation: ISO: International Organization for Standardization.

The primary objective of our analysis is to verify these three hypotheses using machine learning methods. Abbreviations for variable names are summarized in Table A1. We used the number of patents by company (ComPat) and the aggregation of patents by industry sector (IndPat) as target variables to analyze the contribution of the feature variables (the remaining variables excluding the target variable) in individual companies and industry sectors.

Due to missing values in the CSR survey that may affect the accuracy of the prediction, it was necessary to impute them as a pre-processing step for the data. The percentages of missing values for all variables are shown in Figure 4.

We used the mixgb method[29] to impute missing values. The mixgb framework, utilizing XGBoost, subsampling, and PMM, provided a sophisticated approach to handling missing data that outperformed simple imputation and traditional MI techniques. By leveraging the advanced capabilities of XGBoost, mixgb adeptly captures complex non-linear relationships within the data. At the same time, subsampling introduces necessary randomness to model predictions, enhancing the robustness of the imputation process. PMM further refined the approach by ensuring that the variability of the imputed values was not underestimated, thereby reducing bias. This combination made mixgb particularly effective and efficient for large datasets with complex structures, offering a significant improvement in maintaining the integrity and richness of data insights in statistical analysis.

Once the missing values were imputed, we applied an XGBoost regression tree algorithm to estimate the importance of various CSR activity features. We used a grid-search approach to identify the optimal parameter combination using 5-fold cross-validation. The number of boosting iterations used to train the model was set to 100, 200, and 300, with the optimal value being 200. The maximum depth of each decision tree was set to 4, 6, and 8, with the optimal depth being 6. The learning rate was set to 0.01, 0.1, and 0.3, with the optimal rate being 0.01. The minimum loss reduction required to perform a split was set to 0, 0.1, and 0.5, with the optimal value being 0.1. The percentage of features to sample when building each decision tree was set to 0.6, 0.8, and 1.0, with the optimal percentage being 1.0. The minimum instance weight (Hessian) required for a child was set to 1, 3, and 5, with the optimal value being 5. Finally, the percentage of data sampled at each boosting step was set to 0.6, 0.8, and 1.0, with the optimal sampling percentage being 0.6.

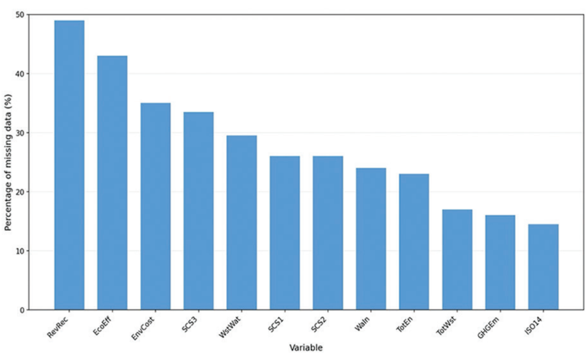

Figure 4. Percentage of missing values for all variables

To validate the selection of XGBoost as our primary analytical method, we conducted a comparative analysis with three baseline models: ordinary least squares linear regression, Lasso regression with L1 regularization, and random forest. All models were evaluated using repeated holdout validation (10 repetitions with 80/20 train–test splits). Table 1 presents the comparison results across three standard performance metrics: coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE).

The results demonstrate that XGBoost consistently outperformed all baseline models in both training and out-of-sample prediction accuracy. Specifically, XGBoost achieved the highest test R^2 of 0.542 (± 0.065), compared to 0.486 (± 0.072) for random forest, 0.261 (± 0.059) for Lasso regression, and 0.247 (± 0.068) for linear regression. XGBoost also achieved the lowest test RMSE (812.6) and MAE (478.5), indicating superior predictive performance.

The substantial improvement of ensemble methods (XGBoost and random forest) over linear models suggests that the relationships between CSR variables and GXIO are inherently non-linear, justifying the use of machine learning approaches. Furthermore, the moderate gap between training and test performance for XGBoost (R^2 difference of 0.282) indicates that the model achieves a reasonable balance between fitting the training data and generalizing to unseen observations.

Table 2 presents detailed evaluation metrics for the final XGBoost model, including 5-fold cross-validation results that demonstrate consistent performance across different data partitions (mean R^2 = 0.542±0.018).

Table 1. Model comparison results

Model	R^2 (train)	R^2 (test)	RMSE (train)	RMSE (test)	MAE (train)	MAE (test)
Linear regression	0.312±0.025	0.247±0.068	986.4±72.3	1042.8±156.2	612.5±48.7	658.3±87.4
Lasso regression	0.298±0.028	0.261±0.059	1002.1±68.9	1018.5±142.8	625.8±45.2	642.1±82.6
Random forest	0.756±0.018	0.486±0.072	587.2±42.1	862.4±124.5	324.6±28.4	512.8±76.3
XGBoost	0.824±0.015	0.542±0.065	498.6±38.7	812.6±118.2	278.4±24.6	478.5±68.9

Note: Values represent mean±standard deviation across 10 repetitions of 80/20 train–test splits.

Abbreviations: MAE: Mean absolute error; RMSE: Root mean squared error; XGBoost: Extreme gradient boosting.

 doi: 10.36922/AJWEP025430333

Table 2. Extreme gradient boosting model evaluation metrics

Evaluation	R^2	RMSE	MAE
Training set	0.824	498.6	278.4
Test set	0.542	812.6	478.5
5-fold cross-validation results			
Fold 1	0.518	835.2	492.1
Fold 2	0.561	798.4	468.3
Fold 3	0.532	821.6	485.7
Fold 4	0.558	802.1	471.2
Fold 5	0.541	806.8	475.2
Mean (±SD)	0.542±0.018	812.8±14.2	478.5±9.8

Abbreviations: MAE: Mean absolute error; RMSE: Root mean squared error; SD: Standard deviation.

5. Analysis results

5.1. Summary of company contributions

Indicators of the contribution scores are summarized in Table 3. The contribution scores are defined as follows:

(i) Gain: Average value of loss reduction brought by a feature. This indicates the contribution to the model's accuracy.

(ii) Cover: Percentage of data points to which a feature is applied. This indicates how important a particular feature is to the overall data.

(iii) Frequency: The number of times a feature was used to split the data. This indicates the frequency of use in the tree's construction.

Figure 5 shows the variable importance visualization.

Based on these variable importance results, we can conclude that IndPat is the most important feature and contributes significantly to the model's prediction accuracy. The number of employees (EmpNo) and supply chain emissions scope 1+2+3 (SCTot) are also highly important features.

Partial dependence plots (PDPs) were created to show how the GXIO of the target variable is related to the top 10 variables; the line graph shows the change in the target variable as each variable's value changes. To improve the readability of the graph, we normalized the feature values to a range of 0–1.

Figure 6 shows that the IndPat and the SCTot of the company contribute most to the value of the PDPs. The PDPs increased significantly as these two variables increased. The third variable, EmpNo, also achieved high PDP values. There were variations in the PDP values of the other variables, such as greenhouse gas emissions

Table 3. Importance of various features at the company level

Feature	Gain	Cover	Frequency
IndPat	0.3357	0.2713	0.2306
EmpNo	0.1908	0.1342	0.1438
SCTot	0.1693	0.1285	0.0732
SCS3	0.0595	0.0989	0.0949
EnvCost	0.0563	0.0680	0.0666
WstWat	0.0304	0.0337	0.0375
GHGEm	0.0291	0.0490	0.0485
TotWst	0.0289	0.0398	0.0614
EcoEff	0.0253	0.0288	0.0452
SCS2	0.0197	0.0497	0.0485
WatIn	0.0163	0.0189	0.0320
RevRec	0.0162	0.0288	0.0379
SCS1	0.0112	0.0206	0.0302
TotEn	0.0067	0.0203	0.0331
ISO14	0.0046	0.0095	0.0166

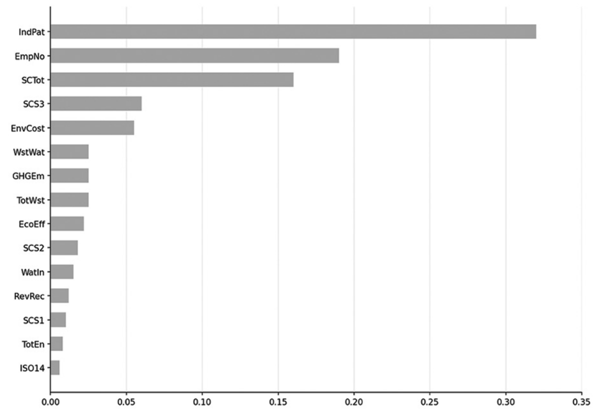

Figure 5. Visualization of the importance of various features at the company level

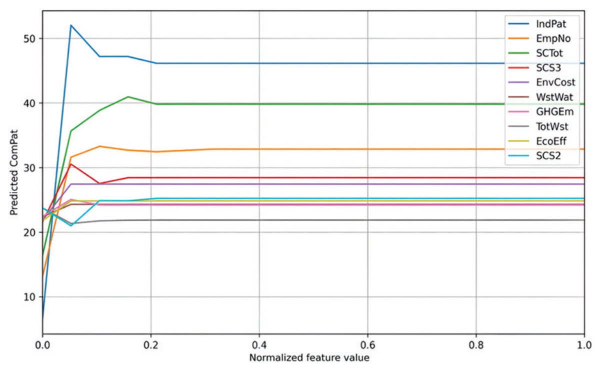

Figure 6. Normalized feature value vs. predicted number of patents by company (ComPat) for different features

(GHGEm), which increased and then decreased as GHGEm increased. Total waste discharge (TotWst) initially had a negative contribution to the PDP value, but turned positive. Environmental protection cost (EnvCost), total wastewater discharge (WstWat), and total economic benefits associated with environmental preservation (EcoEff) consistently contributed to the PDP values, which were positive and significant.

The non-linear patterns revealed by the PDPs merit theoretical interpretation. The interplay between scale effects and efficiency considerations explained the observed variations in GHGEm. At lower emission levels, increases in GHGEm may reflect expanding production activities and organizational growth, which are positively associated with research and development capacity. However, beyond a certain threshold, high emissions may indicate operational inefficiency or reliance on carbon-intensive technologies, both of which are negatively associated with innovative capacity. Firms that successfully extract economic value from existing environmental practices may prioritize short-term returns over long-term research and development investments in breakthrough green technologies, creating path dependencies that inhibit radical innovation. Figure 6 shows that EnvCost has a positive association with ComPat at low levels (approximately 0–0.1), whereas further increases in EnvCost exhibit little or no additional effect on the predicted number of patents. This pattern suggests diminishing marginal returns to environmental cost investments rather than a threshold effect. In other words, initial environmental commitments appear to be sufficient to stimulate innovation-related outcomes, while additional expenditures beyond this range do not translate into further gains. This finding remains consistent with the Porter Hypothesis,[34] which emphasizes the innovation-enhancing role of proactive environmental investments.

Table 4. Importance of various features at the industry level

Feature	Gain	Cover	Frequency
ComPat	0.3115	0.0871	0.0888
EnvCost	0.0844	0.0729	0.1083
SCS2	0.0833	0.0565	0.0455
EcoEff	0.0712	0.0595	0.0555
TotWst	0.0704	0.0818	0.0754
TotEn	0.0679	0.0709	0.0873
SCS1	0.0677	0.0641	0.0448
SCS3	0.0550	0.0658	0.0545
EmpNo	0.0406	0.0975	0.1465
ISO14	0.0333	0.0342	0.0288
GHGEm	0.0275	0.0497	0.0580
WstWat	0.0255	0.0491	0.0414
SCTot	0.0247	0.0703	0.0423
RevRec	0.0206	0.0634	0.0445
WatIn	0.0164	0.0772	0.0785

5.2. Summary of contributions by industries

Indicators of the contribution scores are summarized in Table 4. The definitions are consistent with those in Section 5.1. Figure 7 displays the calculated feature importance. The PDP in Figure 8 shows how the target variable changes as each variable varies.

Figure 8 indicates that ComPat contributes most strongly and positively to IndPat, underscoring the

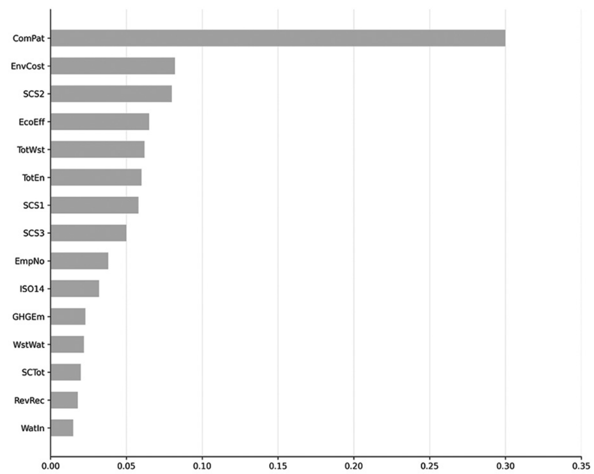

Figure 7. Visualization of the importance of various features at the industry level

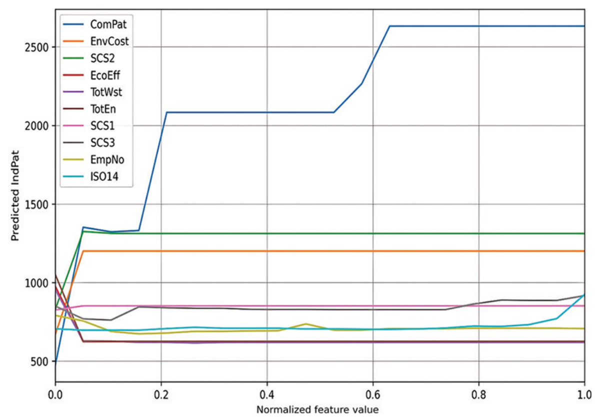

Figure 8. Normalized feature value vs. predicted aggregation of patents by industry sector (IndPat) for different features

cumulative nature of innovation within an industry. Although the PDP exhibits some fluctuations and plateaus at certain ranges, the overall positive relationship suggests that firm-level green innovation outputs aggregate to shape industry-wide innovation performance. Among the CSR-related variables, EnvCost and supply chain emissions scope (SCS) exhibited similar and consistently positive contributions. Their PDP values increased markedly as they rose, before stabilizing at higher levels. This pattern suggests that sustained investments in environmental protection and broader management of supply chain emissions enhance firms' capacity for innovation. In contrast, EcoEff, TotWst, and total energy input displayed relatively high feature importance scores but negatively affected IndPat. The PDP values declined as these variables increased and then leveled off at extreme values. This finding implies that excessive waste generation and energy consumption may reflect inefficiencies that constrain innovation resources. At the same time, short-term economic gains from environmental activities may crowd out longer-term investments in innovation. The effect of firm size, measured by EmpNo, is non-linear: IndPat initially declines and subsequently increases as firm size grows. This pattern suggests that smaller firms may face resource constraints that inhibit innovation, whereas larger firms can eventually exploit scale advantages and organizational capacity to support complex innovation activities. Finally, unlike the firm-level analysis in Section 5.1, ISO14 ranks substantially higher at the industry level in terms of feature importance. As shown in Figure 8, IndPat increased notably when more than 90% of business establishments are ISO 14001 certified. This result indicates that widespread adoption of formal EMSs may create shared standards, knowledge spillovers, and institutional pressure that collectively foster innovation at the industry level.

6. Discussion

Whether CSR affects corporate innovation remains controversial in the literature.[17] Although CSR now appears as an essential dimension of contemporary business activity, the dynamic and practical aspects of developing a CSR orientation within an organization have only recently emerged as a topic in the literature.[35] CSR is a form of international private sector self-regulation aimed at contributing to societal goals of a philanthropic, activist, or charitable nature, with initiatives ranging from voluntary programs to partnerships to reduce the environmental impact of industrial facilities and production methods.[36] CSR has moved from ideology to reality, and it is now seen as necessary for organizations to define their role in society and apply social and ethical standards to their business.[37] Today, it is essential that an organization integrates CSR into its culture and strategy ("corporate DNA"). In Japan, CSR has been implicitly learned and practiced since ancient times in a society that respects the "spirit of harmony." In our study, we sought to explore CSR efforts by Japanese companies to address environmental issues.

As a complementary force, GX is transforming economic, social, and industrial structures that have depended on fossil fuels since the Industrial Revolution into structures powered by clean energy. The GXTI shows how different GX technologies can be categorized and how patent documents related to each category of GX technology can be searched. The GXTI published by the JPO helps Japanese corporate entities and other organizations to objectively identify and explain their GX efforts.

To bridge the gap between CSR efforts of Japanese companies regarding the environment and their GX innovation performance, we searched the GXIO of companies listed in the Toyo Keizai CSR database using a search strategy based on IPC class and keywords. The search strategy was developed after two rounds of discussions (January 6, 2022, and April 6, 2022) in which key technologies were selected based on their potential to reduce GHGEms. In this respect, the study reported here focuses on an entirely new topic.

At present, the models and suggestions available to managers in this area are unclear. Our study aims to fill this research gap using a machine learning methodology. Specifically, we used a decision tree algorithm to analyze the relationship between a company's CSR activities and its GXIO and to test how a company's CSR activities affect its own GXIO and that of the industry sector to which it belongs. We found that CSR activities of companies have different effects at the company and industry levels. In addition, we found that GXIO is driven by environmental awareness. As for ISO 14001 certification—a new CSR variable—it is effective for the industrial sector as a whole.

Our findings contribute to the literature on CSR and green innovation. First, we demonstrate that machine learning approaches, specifically XGBoost, provide superior predictive performance ($R^2 = 0.542$) compared to traditional linear methods ($R^2 = 0.247$) when modeling the CSR–GXIO relationship. This confirms that the mechanisms linking CSR activities to green innovation are inherently complex and non-linear.

Second, our feature importance analysis reveals that existing technological capabilities (proxied by patent stock, gain = 0.312) are the strongest predictor of IndPat, supporting absorptive capacity theory.[38] This finding implies that CSR activities may be most effective when they complement existing research and development capabilities. Third, the significance of environmental conservation costs (EnvCost, gain = 0.084) as the leading CSR-specific predictor provides empirical support for the Porter Hypothesis,[34] suggesting that proactive environmental investments can stimulate green innovation. Fourth, the importance of SCSs (SCS1, SCS2, SCS3) highlights the role of stakeholder pressure in driving GX innovation, consistent with stakeholder theory.[39]

Although it is difficult to conclude definitively that the more companies engage in CSR activities, the more GXIO they produce, our results show that large Japanese companies, especially manufacturers, which consume substantial amounts of environmental resources, require significant energy input, and may even pollute the environment during production, are highly competitive in terms of environmental innovation. Given that the CSR questionnaire focuses mainly on an organization's awareness of environmental protection issues, it seems that although large manufacturers may be a significant cause of environmental damage, they are also aware of the need to reduce this damage. In general, Japanese manufacturers show a high level of social responsibility. In contrast, sectors such as banking and insurance, which require little environmental resource input and produce minimal emissions, exhibit low GXIO. Managers in these sectors are encouraged to make greater efforts to promote GXIO. Such efforts might include setting up an innovation office, seeking guidance from a qualified consulting firm, or formulating a policy or plan to systematically promote various innovation initiatives to help companies within the sector create new value and enhance their competitiveness. It can be argued that a commitment to environmental innovation is likely to result in both increased social respect and increased revenue. Nevertheless, the debate over how to balance economic returns and environmental protection will continue, and our study provides valuable insights.

Today, technical standards and standard-setting organizations are omnipresent and essential to mass production and communications.[40] The standards set forth by the ISO are well known in the high-tech industry. In our search, we found that ISO 14001 certification had a notably positive impact. Although the current standardization system in Japan has contributed greatly to the development of the manufacturing industry and the improvement of people's lives since its establishment after the war, the percentage of ISO 14001-certified companies among Japanese companies is lower than that among major European and American companies.[41] Thus, Japanese companies should make greater efforts to meet certification requirements and intensify their commitment to standard-essential innovation,[42] pushing open innovation beyond national borders as research and development and standardization proceed simultaneously in a global corporate consortium. The newly added variable, SCS, also shows a strong influence. From sourcing raw materials to selling products, supply chain managers need to aggressively pursue new initiatives. Such initiatives might include making the entire supply chain visible, minimizing inventories by introducing just-in-time methods, and implementing drone delivery. Moreover, rapidly advancing AI technology can help reduce supply chain emissions.

Our study is not without limitations. As first pointed out by Lee,[43] the vast majority of CSR research focuses almost exclusively on large, publicly traded corporations. Little is known about what CSR means and how it is implemented in SMEs. The paucity of research on how SMEs can overcome and complement the quantity constraints they face in mobilizing resources for innovation is a serious issue, especially in Japan, where SMEs account for 99.7% of all Japanese firms. Further research is clearly needed to identify the peculiarities of exercising CSR in SMEs and to explore the business impact of CSR among these smaller enterprises. Related research on how the government can put in place policies that promote open innovation in collaboration with large companies and nurture the seeds of innovation is also necessary. In contrast to several expectations, our empirical results indicate that traditional firm characteristics, particularly EmpNo and IndPat, are the most important predictors of GXIO. The direct effects of CSR variables are comparatively weaker. This suggests that CSR engagement alone is insufficient to drive green innovation outcomes but may instead reinforce existing innovation capacities within firms and industries. Finally, as described in Section 5, the EcoEff was shown to negatively affect IndPat in our study, an unexpected result that warrants further investigation.

7. Conclusion

This study examines how CSR influences GXIO among Japanese firms by integrating CSR survey data with green patent information and applying machine learning methods. By analyzing both firm-level and industry-level outcomes, the study provides new evidence on the mechanisms through which CSR supports green innovation in the context of Japan's transition toward carbon neutrality. The results show that CSR does not act as a direct or independent driver of GXIO. Rather, CSR functions as an enabling mechanism that reinforces existing technological capabilities, organizational scale, and environmental awareness. At the firm level, green innovation output is primarily driven by accumulated technological knowledge and firm size, while CSR-related activities—especially environmental protection costs and supply chain emission management—play a complementary role. At the industry level, collective CSR mechanisms, including environmental investment, supply chain coordination, and widespread adoption of ISO 14001 certification, exert a stronger and more systematic influence on GXIO, highlighting the importance of institutionalized standards and coordination effects. Methodologically, the superior performance of XGBoost relative to linear models confirms that CSR–GXIO relationships are complex and non-linear. From a practical perspective, the findings suggest that CSR strategies should be closely aligned with firms' innovation capabilities and supported by industry-wide frameworks. Overall, this study advances understanding of CSR as a catalyst for green transformation by clarifying its role within broader innovation systems.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Jiaming Jiang
Formal analysis: Jiaming Jiang, Yu Zhao
Investigation: Jiaming Jiang, Junshi Feng
Methodology: Jiaming Jiang, Yu Zhao
Writing–original draft: Jiaming Jiang
Writing–review & editing: Jiaming Jiang

Availability of data

The GXIO dataset used in this study is available at https://figshare.com/articles/dataset/Japanese_GX_patents_with_Englishi_title_xlsx/25434724/3

References

1. Takeishi A, Aoshima Y, Karube M. Reasons for innovation: Legitimizing resource mobilization for innovation in the cases of the Okochi Memorial Prize winners. In: *Dynamics of Knowledge, Corporate Systems and Innovation*. Germany: Springer Science and Business Media; 2009. p. 165-189. doi: 10.1007/978-3-642-04480-9_7
2. Sueyoshi T, Goto M. DEA non-radial approach for resource allocation and energy usage to enhance corporate sustainability in Japanese manufacturing industries. *Energies*. 2019;12(9):1785. doi: 10.3390/en12091785
3. Dobele AR, Westberg K, Steel M, Flowers K. An examination of corporate social responsibility implementation and stakeholder engagement: A case study in the Australian mining industry. *Bus Strat Environ*. 2014;23(3):145-159. doi: 10.1002/bse.1775
4. Epstein MJ, Roy MJ. Implementing a corporate environmental strategy: Establishing coordination and control within multinational companies. *Bus Strat Environ*. 2007;16(6):389-403. doi: 10.1002/bse.545
5. Santos-Jaén JM, Madrid-Guijarro A, García-Pérez-de-Lema D. The impact of corporate social responsibility on innovation in SMEs: The mediating role of debt terms and human capital. *Corp Soc Responsib Environ Manag*. 2021;28(4):1200-1215. doi: 10.1002/csr.2125
6. European Commission. *Promoting a European Framework for Corporate Social Responsibility*. COM Final. p. 366; 2001. p. 6. Available from: https://eur-lex.europa.eu/legal-content/en/txt/?uri=celex:52001dc0366 Last accessed on 2026 Jan 13].
7. Suto M, Takehara H. Employee-oriented corporate social responsibility, innovation, and firm value. *Corp Soc Responsib Environ Manag*. 2022;29(4):765-778. doi: 10.1002/csr.2232
8. Goto M, Sueyoshi T. Sustainable development and corporate social responsibility in Japanese manufacturing companies. *Sustain Dev*. 2019;28(4):844-856.

doi: 10.1002/sd.2035

9. Montalbán-Domingo L, García-Segura T, Sanz-Benlloch A, Pellicer E, Torres-Machi C, Molenaar K. Assessing social performance of construction companies in public-works procurement: Data envelopment analysis based on the benefit of the doubt approach. *Environ Impact Assess Rev*. 2022;96:106844. doi: 10.1016/j.eiar.2022.106844
10. Ratajczak P, Szutowski D. Exploring the relationship between CSR and innovation. *Sustain Account Manag Policy J*. 2016;7(2):295-318. doi: 10.1108/sampj-07-2015-0058
11. Margolis JD, Walsh JP. Misery loves companies: Rethinking social initiatives by business. *Adm Sci Q*. 2003;48(2):268-305. doi: 10.2307/3556659
12. Smith, NC. 'Consumers as drivers of corporate social responsibility'. In: Crane A, editors. *The Oxford Handbook of Corporate Social Responsibility*. Oxford: OUP Oxford; 2008. p. 181-302. doi: 10.1093/oxfordhb/9780199211593.003.0012
13. Bocquet R, Le Bas C, Mothe C, Poussing N. CSR, innovation, and firm performance in sluggish growth contexts: A firm-level empirical analysis. *J Bus Ethics*. 2017;146:241-254. doi: 10.1007/s10551-015-2959-8
14. Eweje G, Sakaki M. CSR in Japanese companies: Perspectives from managers. *Bus Strat Environ*. 2015;24(7):678-687. doi: 10.1002/bse.1894
15. Dögl C, Behnam M. Environmentally sustainable development through stakeholder engagement in developed and emerging countries. *Bus Strat Environ*. 2015;24(6):583-600. doi: 10.1002/bse.1839
16. Li J, Lin B. The sustainability of remarkable growth in emerging economies. *Resour Conserv Recycl*. 2019;145:349-358. doi: 10.1016/j.resconrec.2019.01.036
17. Chkir I, El Haj Hassan B, Rjiba H, Saadi S. Does corporate social responsibility influence corporate innovation? International evidence. *Emerg Mark Rev*. 2021;46:100746. doi: 10.1016/j.ememar.2020.100746
18. Wagner M. Innovation and competitive advantages from the integration of strategic aspects with social and environmental management in European firms. *Bus Strat Environ.* 2009;18(5):291-306. doi: 10.1002/bse.585
19. Mishra DR. Post-innovation CSR performance and firm value. *J Bus Ethics*. 2017;140(2):285-306. doi: 10.1007/s10551-015-2676-3
20. Gürler HE. The impact of corporate social responsibility on green innovation: Do industry, data type and region matter? A meta-analysis. In: *Environment, Development and Sustainability*. Germany: Springer; 2024. doi: 10.1007/s10668-024-05277-y
21. Padilla-Lozano CP, Collazzo P. Corporate social responsibility, green innovation and competitiveness: Causality in manufacturing. *Compet Rev*. 2022;32(7):21-39. doi: 10.1108/CR-12-2020-0160
22. Le TT. How do corporate social responsibility and green innovation transform corporate green strategy into sustainable firm performance? *J Clean Prod*. 2022;362:132228. doi: 10.1016/j.jclepro.2022.132228
23. Doi M. Legitimacy of corporate social innovation: Legitimizing resource mobilization for corporate social innovation in the case studies in Japan. *J Innov Manag*. 2020;17:101-117. doi: 10.24677/riim.17.0_101
24. Genba K. Environmental innovation for sustainable corporate growth. *Res Technol Plann*. (玄場公規. 持続可能な企業成長を実現する環境イノベーション (< 特集> イノベーションとサステイナビリティ). 研究 技術 計画). 2009;23(3):220-226. doi: 10.20801/jsrpim.23.3_220
25. Cai H, Rueckert D, Passerat-Palmbach J. *2CP: Decentralized Protocols to Transparently Evaluate Contributivity in Federated Learning Environments*. [arXiv. Preprint]; 2020. doi: 10.48550/arXiv.2011.07516
26. Schams S, Hauffa J, Groh G. Analyzing a user's contributive social capital based on activities in online social networks and media. In: *Companion Proceedings of the the Web Conference 2018 (WWW'18 Companion)*; 2018. p. 1457-1464. doi: 10.1145/3184558.3191593
27. Xia Y, Liu C, Li Y, Liu N. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Syst Appl*. 2017;78:225-241. doi: 10.1016/j.eswa.2017.02.017
28. Chang YC, Chang KH, Wu GJ. Application of extreme gradient boosting trees in credit risk assessment models for financial institutions. *Appl Soft Comput*. 2018;73:914-920. doi: 10.1016/j.asoc.2018.09.029
29. Deng Y, Lumley T. Multiple imputation through XGBoost. *J Comput Graph Stat*. 2023;33(2):352-363. doi: 10.1080/10618600.2023.2252501
30. Jiang J, Zhao Y, Feng J. A dataset of Japanese corporates' green transformation technology patents. *F1000Res*. 2024;13:288. doi: 10.12688/f1000research.147798.1
31. Salim HK, Padfield R, Hansen SB, *et al*. Global trends in environmental management system and ISO 14001 research. *J Clean Prod*. 2018;170:645-653. doi: 10.1016/j.jclepro.2017.09.017
32. Darnall N, Carmin J. Greener and cleaner? The signaling

accuracy of U.S. voluntary environmental programs. *Policy Sci*. 2005;38(2-3):71-90.
doi: 10.1007/s11077-005-6591-9

33. Rosecká N, Machek O, Stasa M, Kubíček A. Long-term orientation and corporate social responsibility in small and medium-sized enterprises: The role of strategy formation mode. *Soc Responsib J*. 2024;20(4):825-842.
doi: 10.1108/SRJ-07-2023-036
34. Porter ME, Van der Linde C. Toward a new conception of the environment-competitiveness relationship. *J Econ Perspect*. 1995;9(4):97-118.
doi: 10.1257/jep.9.4.97
35. Lindgreen A, Swaen V, Johnston WJ. Corporate social responsibility: An empirical investigation of U.S. organizations. *J Bus Ethics*. 2009;85(S2):303-323.
doi: 10.1007/s10551-008-9738-8
36. Rondinelli DA, Berry MA. Environmental citizenship in multinational corporations: Social responsibility and sustainable development. *Eur Manage J*. 2000;18(1):70-84.
doi: 10.1016/S0263-2373(99)00070-5
37. Lichtenstein DR, Drumwright ME, Braig BM. The effect of corporate social responsibility on customer donations to corporate-supported nonprofits. *J Mark*. 2004;68(4):16-32.
doi: 10.1509/jmkg.68.4.16.42726
38. Cohen WM, Levinthal DA. Absorptive capacity: A new perspective on learning and innovation. *Adm Sci Q*. 1990;35(1):128-152.
doi: 10.2307/2393553
39. Freeman RE. The stakeholder approach. In: *Strategic Management: A Stakeholder Approach*. Cambridge: Cambridge University Press; 2010.
40. Jiang J, Goel RK, Zhang X. IPR policies and determinants of membership in standard setting organizations: A social network analysis. *Netnomics*. 2020;21:129-154.
doi: 10.1007/s11066-020-09144-6
41. Jiang J. Review of recent development in empirical literature on technological standards. *J Humanit Soc Sci Okayama Univ*. 2018;46:185-206.
42. Jiang J, Zhang X. Essential patents and knowledge position: A network analysis on the basis of patent citations. *Standards*. 2021;1(2):90-104.
doi: 10.3390/standards1020009
43. Lee MP. A review of the theories of corporate social responsibility: Its evolutionary path and the road ahead. *Int J Manag Rev*. 2008;10(1):53-73.
doi: 10.1111/j.1468-2370.2007.00226.x

Appendices

A1. Alternative model specifications

To assess the sensitivity of our findings to model specification, we estimated extreme gradient boosting (XGBoost) models with alternative hyperparameter configurations (Table A2). The results demonstrate remarkable consistency: The test R^2 ranged from 0.498 to 0.542 across specifications, and the top three predictors (number of patents by company [ComPat], environmental protection cost [EnvCost], and supply chain emission scopes [SCSs]) remained stable regardless of the hyperparameter settings. This consistency suggests that our main findings are robust to alternative model specifications.

We examined the robustness of our findings by estimating models with different feature sets (Table A3). When patent-related variables were excluded, EnvCost emerged as the leading predictor, confirming its importance among CSR-specific variables. The substantial drop in predictive performance when excluding patent variables (R^2 from 0.542 to 0.387) underscores the critical role of existing technological capabilities in predicting GXIO, consistent with absorptive capacity theory.

A2. Sensitivity analyses

We tested the sensitivity of our results to the outlier removal threshold (Table A4). The variable importance rankings remained consistent across all specifications, with ComPat and EnvCost consistently emerging as the top two predictors. While predictive performance varied somewhat (R^2 ranging from 0.478 to 0.558), the qualitative conclusions regarding the key drivers of GXIO are robust to the choice of outlier treatment.

To ensure the choice of missing-value imputation method does not drive our findings, we re-estimated our models using alternative approaches (Table A5). The multi-imputation framework (mixgb) method yielded the highest predictive performance ($R^2 = 0.542$), validating our methodological choice. Notably, the variable importance rankings remained identical across all imputation methods, confirming that our substantive conclusions are robust to the handling of missing data.

Table A1. Abbreviations for variable names

Abbreviations	Definitions
IndNm	Industry sector
IndPat	Aggregation of patents by industry sector
CompNm	Company name
ComPat	Number of patents by company
EmpNo	Number of employees
EnvCost	Environment protection cost
TotEn	Total energy input
WatIn	Total water resource input
GHGEm	Greenhouse gas emissions
TotWst	Total waste discharge
WstWat	Total wastewater discharge
EcoEff	Total economic benefits associated with environmental preservation
RevRec	Revenue amount of recycling, etc.
SCS1	Supply chain emissions scope 1
SCS2	Supply chain emissions scope 2
SCS3	Supply chain emissions scope 3
SCTot	Supply chain emissions scope 1+2 + 3
ISO14	ISO 14001 certification rate

Note: SCS is divided into Scopes 1–3 to reflect structurally different emission sources with distinct managerial relevance, while total SCS is additionally included to represent overall supply chain emissions and preserve result interpretability.

Table A2. Sensitivity to hyperparameter settings

Configuration	Max_depth	Eta	Test R^2	Top three variables (gain)
Baseline (optimal)	6	0.01	0.542	ComPat, EnvCost, SCS2
Shallow trees	4	0.01	0.518	ComPat, EnvCost, SCS2
Deep trees	8	0.01	0.536	ComPat, EnvCost, SCS2
Higher learning rate	6	0.1	0.524	ComPat, EnvCost, SCS1
Conservative	4	0.01	0.498	ComPat, EnvCost, SCS2

Table A3. Sensitivity to feature selection

Feature set	N variables	Test R^2	Top predictor	Gain
Full model	15	0.542	ComPat	0.312
Excluding patent variables	13	0.387	EnvCost	0.156
CSR variables only	10	0.324	EnvCost	0.189
Environmental variables only	8	0.298	EnvCost	0.215

Abbreviation: CSR: Corporate social responsibility.

Table A4. Sensitivity to outlier treatment

Outlier threshold	N (after removal)	Test R^2	Variable importance rank
90^{th} percentile	256	0.558	ComPat > EnvCost > SCS2
95^{th} percentile (baseline)	270	0.542	ComPat > EnvCost > SCS2
99^{th} percentile	281	0.512	ComPat > EnvCost > SCS2
No removal	284	0.478	ComPat > EnvCost > SCS1

Table A5. Sensitivity to the missing-value imputation method

Imputation method	Test R^2	Top three variables
Mixgb (baseline)	0.542	ComPat, EnvCost, SCS2
Mean imputation	0.486	ComPat, EnvCost, SCS2
Median imputation	0.479	ComPat, EnvCost, SCS2
KNN imputation (k=5)	0.521	ComPat, EnvCost, SCS2
Multiple imputation (MICE)	0.534	ComPat, EnvCost, SCS2

Abbreviation: KNN: K-nearest neighbors.