

Henao-Rodríguez, Carolina; Lis-Gutiérrez, Jenny-Paola; Angulo-Bustinza, Harold Delfín

Article

Unveiling greenwashing in Colombian manufacturing: A machine learning approach

Research in Globalization

Provided in Cooperation with:

Elsevier

Suggested Citation: Henao-Rodríguez, Carolina; Lis-Gutiérrez, Jenny-Paola; Angulo-Bustinza, Harold Delfín (2024) : Unveiling greenwashing in Colombian manufacturing: A machine learning approach, Research in Globalization, ISSN 2590-051X, Elsevier, Amsterdam, Vol. 8, pp. 1-14, <https://doi.org/10.1016/j.resglo.2024.100196>

This Version is available at:

<https://hdl.handle.net/10419/331123>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

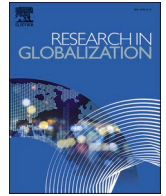
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-nc-nd/4.0/>



Unveiling greenwashing in Colombian manufacturing: A machine learning approach

Carolina Henao-Rodríguez^{a,*}, Jenny Paola Lis-Gutiérrez^a, Harold Delfín Angulo-Bustanza^b

^a Fundación Universitaria Konrad Lorenz, Colombia

^b Universidad Continental del Perú, Perú

ARTICLE INFO

Keywords:

Machine Learning
Greenwashing
Decision Tree
Random Forest
Logistic regression
LASSO
Double machine learning

ABSTRACT

Greenwashing is the misleading use of environmental claims to market non-eco-friendly products, a growing concern as environmentally conscious consumers seek eco-friendly options. This study aims to identify factors that allow the detection of “greenwashing” practices in Colombian companies across various industrial sectors holding certifications such as ISO14001, the Colombian environmental seal, the district environmental excellence program, or international/regional environmental certifications. We employed a quantitative, cross-sectional methodology using three machine learning algorithms (Decision Tree Classification, Random Forest, Logistic Regression) and the double machine learning technique to identify causal effects and significant variables. Additionally, a greenwashing indicator was formulated, considering two variables: (a) Presence of a comprehensive solid waste management plan, and (b) Presence of a water usage and conservation program. Our findings reveal substantial variations in the adoption of environmental management practices across diverse industrial sectors and regions in Colombia. It is noteworthy that our research emphasizes that: (i) companies with an environmental contingency plan and a monitoring program are less likely to engage in greenwashing; (ii) expenses for personnel dedicated to environmental protection activities and the investment and expenses related to air and climate protection also reduce the likelihood of greenwashing; (iii) companies in the manufacturing of non-metallic mineral products sector are more likely to engage in greenwashing. It was also found that a significant proportion of Colombian industrial companies with environmental certifications do not fully embrace sustainable practices, implying limited effectiveness of environmental certifications in the country (56,09%). This research contributes to the taxonomy of greenwashing by delving into various types identified in the literature. It sheds light on selective disclosure, emphasizing that mere certification without tangible actions can constitute greenwashing. The broadened concept of decoupling now includes not only meeting shareholder expectations but also failing to meet those set by certifications. Attention deflection involves diverting focus toward certifications rather than real investments in environmental practices. Deceptive manipulation occurs when certifications create a misleading perception without concrete actions. Additionally, dubious authorizations and labels are revealed, showcasing the vulnerability of eco-labels to fraud. The study enriches existing categories, viewing certifications as potential instruments for greenwashing, reinforcing the need for substantial, measurable actions to validate sustainability claims.

Introduction

Greenwashing, defined as the deceptive utilization of environmental claims to promote products or services that lack genuine eco-friendliness, has garnered increasing attention in recent years. This heightened awareness is due to consumers' growing environmental consciousness and their willingness to pay premiums for sustainable

products. Various studies have delved into greenwashing, offering insights and strategies to combat this phenomenon (Policarpo et al., 2023; Hu et al., 2023; Gregory, 2021).

The term “greenwashing” encompasses multiple forms of misleading communication, spanning direct, indirect, and vicarious practices (Pizzetti et al., 2019). Effectively countering greenwashing requires the development of strategies aimed at raising awareness about the concept

* Corresponding author.

E-mail addresses: lindac.henaor@konradlorenz.edu.co (C. Henao-Rodríguez), jenny.lis@konradlorenz.edu.co (J.P. Lis-Gutiérrez), hanguo@continental.edu.pe (H.D. Angulo-Bustanza).

<https://doi.org/10.1016/j.resglo.2024.100196>

Received 26 October 2023; Received in revised form 26 January 2024; Accepted 27 January 2024

Available online 30 January 2024

2590-051X/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

and educating consumers regarding environmentally responsible behavior (Eng et al., 2021). Nevertheless, the abundance or lack of knowledge about greenwashing can significantly impact consumer attitudes, beliefs, and purchase intentions. The COVID-19 pandemic has further underscored the urgency of environmental concerns, ranking them among the top five critical threats for the next decade (Braga Junior et al., 2019; World Economic Forum, 2022).

This study aims to identify factors that allow the detection of “greenwashing” practices in Colombian companies across various industrial sectors holding certifications such as ISO14001, the Colombian environmental seal, the district environmental excellence program, or international/regional environmental certifications. Greenwashing poses a growing problem, where companies falsely portray their offerings as environmentally sound, eroding consumer trust and fostering deception. By comprehending the factors affecting corporate conduct and gauging the extent of greenwashing across different sectors, policymakers can devise effective measures to tackle this issue.

In this context, the study delves into the prospects for future research as outlined by Santos et al. (2023) and Torelli et al. (2020). These authors underscored the significance of advancing the exploration of greenwashing taxonomy. Additionally, Ghitti et al. (2023) highlighted the challenging empirical nature of quantitatively capturing instances of greenwashing.

Our research seeks to contribute to the solution of the greenwashing problem by shedding light on the factors impacting companies’ post-certification conduct and the prevalence of greenwashing in various domains. Drawing from these insights, policymakers can craft initiatives aimed at enhancing transparency and accountability in environmental certification, bolstering consumer awareness, and motivating companies to adopt sustainable practices.

Our sample comprises companies holding environmental certifications and operating within sectors encompassing water treatment, waste management, environmental programs, and clean development mechanisms. By examining these specific areas, our study aims to offer valuable insights into the factors shaping corporate behavior and the extent of greenwashing within these critical spheres. We employed a quantitative, cross-sectional methodology using the double machine learning technique to identify causal effects and significant variables.

The paper is structured into five sections, beginning with a literature review in Section 1, followed by the methodology in Section 2, presentation of results in Section 3, discussion in Section 4, and concluding remarks in Section 5.

Literature review

The literature review commences by contextualizing the issue of greenwashing, emphasizing the pivotal significance of environmental responsibility and the urgent need to address environmental concerns within the business sphere. It proceeds to delineate various types of greenwashing at the company level, enhancing reader comprehension through concrete examples that illustrate the diverse tactics employed by companies for deceptive practices. Additionally, it explores the empirical and quantitative challenges inherent in quantifying greenwashing. Of note is the focus on the relevance of the greenwashing issue in the Colombian context, emphasizing the prevalence of environmental problems and deceptive practices in the country, thereby providing a localized framework for understanding the phenomenon. Subsequently, the review presents some theoretical approaches to understanding greenwashing.

In the post-COVID era, companies can achieve high returns by implementing environmentally responsible industrial procedures and emphasizing macro marketing strategies that promote green consumption behavior (Helm & Little, 2022). However, the implementation of sustainable development goals (SDG) practices can vary between developed and developing countries (Datta & Goyal, 2022), and companies often engage in greenwashing, misleading consumers by

presenting a false image of sustainability (Pizzetti et al., 2019). Addressing environmental issues necessitates collaborative efforts between firms, governments, and regulatory agencies (Rashidi-Sabet & Madhavaram, 2021).

In this context, numerous companies focus their corporate social responsibility efforts on conspicuous facets while demonstrating subpar performance in less visible domains. Consequently, this selective emphasis results in a manifestation of limited corporate social responsibility, giving rise to deceptive marketing practices. Termed as greenwashing in this scenario, it denotes the disregard for the less observable aspects (Wu et al., 2020). The escalating prevalence of greenwashing over the past two decades underscores its growing importance in the literature (Gatti et al., 2021).

Yang et al. (2020) identified six types of greenwashing present in the literature. The following are descriptions distinguishing each:

- (i) Selective disclosure: Occurs when companies with poor environmental performance excessively disclose information to enhance their sustainability image.
- (ii) Decoupling: Refers to the situation where companies claim to meet the expectations of their shareholders without making real changes in their practices. This greenwashing strategy is often implemented by promoting ambitious projects without sustainable units or sufficient funds.
- (iii) Attention deflection: Involves engaging in symbolic activities designed to divert stakeholders’ attention and conceal corrupt business actions. It includes communication practices such as selective and incorrect disclosure, as well as undefined disclosure to divert attention in corporate social responsibility (CSR) communication.
- (iv) Deceptive manipulation: Also known as deceptive handling, this type involves deceptive conduct where sustainable communication deliberately creates an operation of business practices to promote the perception that a company’s products, goals, and/or policies are environmentally friendly.
- (v) Dubious authorizations and labels: Considers third-party authorizations defining standards as greenwashing, replacing the reliability of an external authorizer with a company’s own rights. Eco-labels, while designed to control greenwashing, can be susceptible to fraud by dishonest manufacturers under certain conditions.
- (vi) Inefficient public voluntary programs: Companies contributing to government-funded voluntary programs may do so for various reasons, but their contribution alone may not lead to environmental improvements. The efficiency of such programs depends on strict regulations and fines imposed on newcomers who fail to comply with standards.

De Freitas Netto et al. (2020) emphasized that greenwashing has a multifaceted nature, leading to different approaches for understanding it. Among these, the following are highlighted:

- (i) Selective disclosure: Where consumers are deceived about a company’s environmental practices by withholding negative information related to environmental behavior.
- (ii) Decoupling behavior: Greenwashing is associated with symbolic actions without concrete pro-environmental actions.
- (iii) Corporate legitimacy theory: It can be distinguished into three types of corporate legitimacy: cognitive, manifested in shared assumptions of the social environment; pragmatic, the result of calculations made by stakeholders of the company; and moral, based on moral judgments about the company. Greenwashing occurs in the realm of pragmatic legitimacy.

Within this perspective, it is worth noting that greenwashing can occur at various levels: at the company level, it involves deceiving

consumers about the environmental practices of the company; while at the product level, it involves misleading consumers about the environmental benefits of a product or service (Delmas & Burbano, 2011).

There are five types of greenwashing at the company level, highlighted by Contreras-Pacheco & Claasen (2017). These are:

- (i) “Dirty Business”. Refers to companies belonging to intrinsically unsustainable economic activities but promote sustainable practices or products that are not.
- (ii) “Ad Bluster”. Companies that, through advertising, seek to divert attention from environmental issues through the exaggeration of achievements.
- (iii) “Political spin”. Firms that influence regulations to obtain benefits that go against sustainability.
- (iv) “It’s the Law, stupid!”. Companies that proclaim environmental achievements or commitments that are required by existing laws.
- (v) “Fuzzy reporting”. Companies that use sustainability reports to distort the truth or project a false image of their environmental responsibility.

In the study conducted by Adamkiewicz et al (2022), four distinct strategies of greenwashing within the fashion industry are delineated. To begin, deceptive claims encompass the exaggeration of environmental benefits, the presence of unclear labels and certifications, and a tendency to spotlight minor advancements while downplaying significant impacts. Secondly, the practice of information omission involves the non-disclosure of harmful substances, neglecting the comprehensive life cycle of products, and shifting responsibility onto consumers. The third strategy, emotional manipulation, encompasses the strategic use of images and narratives, the leveraging of social pressure, and the exploitation of consumer trust in well-established brands. Finally, greenwashing tactics include the strategic utilization of recycled materials, the promotion of return programs without addressing their impact on production, and the selective emphasis on specific aspects of sustainability.

On their part, Dorfleitner & Utz (2023) proposed a concept to determine a greenwashing indicator at the company level, based on the difference between actual and apparent green performance, manifested in five pillars of information: soft Environmental, Social, and Governance (ESG) data, textual self-representation, green marketing expenses, green virtue for apparent green performance, and hard ESG data as a measure of actual green performance.

Yildirim (2023) suggested that, to better understand these deceptive practices, several indicators can be identified, including: (i) promoting products with one or two ecological attributes as entirely sustainable; (ii) lacking evidence to support claims about sustainable products; (iii) the use of unclear environmental expressions; (iv) deceiving consumers about the ecological image of companies and green products; (v) confusing consumers with categories of ecological products; (vi) false ecological labels.

Research has shown that stakeholder pressures, such as those from customers and regulators, can lead firms to engage in greenwashing (Testa et al., 2018). Studies have also suggested that stricter environmental regulations can reduce the prevalence of greenwashing (He et al., 2020).

Ghitti et al. (2023) have demonstrated that quantitatively capturing greenwashing is an empirical challenge. In this regard, the concept of selective disclosure focuses on the incomplete disclosure of environmental metrics, but this approach overlooks the risk of greenwashing related to exaggeration or excessive promise of disclosed metrics.

In Colombia, environmental and greenwashing problems in factories are prevalent, and some companies falsely advertise themselves as “green” institutions (Secretaría Distrital de Ambiente, 2016, 2017; Camelo, 2021). Green taxation, which has been implemented in countries like the Netherlands, Denmark, and Canada, is a potential solution to this issue (Norouzi et al., 2022).

Finally, the attribution theory, in which individuals use three factors to attribute causes to events: Intent, Controllability, and Consistency. In this regard, Szabo & Webster (2021) demonstrated that companies seek to enhance their environmental standings by presenting their environmental efforts to consumers; however, not all eco-friendly marketing claims accurately reflect the environmental behavior of the companies. The findings of this study support this theory and suggest that perceived greenwashing can have a negative impact on consumers’ environmental perceptions.

In conclusion, while the Green Industry promotes sustainable development and environmental protection, industries’ priorities can hinder progress. Collaborative efforts between firms, governments, and regulatory agencies are necessary to address environmental concerns, and the implementation of green taxation can discourage greenwashing and promote environmentally responsible behavior. However, a lack of knowledge about greenwashing can affect consumers’ attitudes and beliefs, and implementing SDG practices can differ between developed and developing countries. Continued research and development are necessary to ensure that environmental claims are accurate and transparent.

Materials and methods

The present study employs a quantitative, cross-sectional methodology to investigate the factors that contributed to greenwashing among manufacturing companies in Colombia during 2020, using three machine learning algorithms. The research specifically focuses on companies that possess ISO14001 certification, the Colombian environmental seal, the district environmental excellence program, or international or regional environmental certification. These certified companies are required to maintain an environmental management system that continuously improves its approach and organizational system procedures in addressing environmental issues (ISO, 2015).

Data

In our investigation, we leveraged data sourced from both the Annual Manufacturing Survey (DANE, 2022a) and the Industrial Environmental Survey (DANE, 2022b). These surveys adhere meticulously to the criteria established by OECD statistical operations, ensuring not only public accessibility but also the meticulous anonymization of the data.

Detailed within the documentation for the “Industry Environmental Survey” (DANE, 2022b) are crucial elements defining the sampling methodology:

- o Unit of Observation: This encompasses industrial establishments nationwide actively involved in manufacturing activities that align with specific parameters.
- o Reporting Unit: It pertains to the entity responsible for furnishing the required information, typically embodied by the professional overseeing process or environmental engineering within the industrial establishment.
 - Unit of Analysis: Encompassing industrial establishments across the nation engaged in manufacturing activities that align with parameters set for the target population.
 - Sampling Unit: This refers to industrial establishments nationwide engaged in manufacturing activities, meeting the specified parameters for the target population.

The sampling approach is characterized by a probabilistic design featuring triple stratification. Its probabilistic nature arises from the fact that each establishment in the sample frame holds a probability greater than zero of being selected. The triple stratification involves the initial partitioning of the universe by regions, followed by industrial divisions, and ultimately, utilizing the Hidiroglou algorithm to create strata of forced and probabilistic inclusion. In certain instances, classification as

forced inclusion is determined solely based on statistical and thematic criteria (Henao, et al., 2023).

This stratification is predicated on the following variables:

- The region of the country to which the establishment belongs (Caribbean, Eastern, Central, Pacific, Bogota, and Amazonia - Orinoquia) (Fig. 1).
- The industrial activity conducted by the company.
- The size of the establishment in terms of employed personnel and industrial production.

The observational unit encompasses industrial establishments in Colombia (7,743), with the sample spanning 681 establishments across 9 productive sectors with at least one of the following environmental certifications: ISO14001, the Colombian environmental seal, the district environmental excellence program, or international/regional environmental certifications. The intricacies of the sample design, as elucidated earlier, permit estimates with an estimated coefficient of variation (sampling error) of less than 5 % at the national level.

We specified the empirical model deduction and its relationship with greenwashing for each target variable. We exclusively worked with companies that possessed environmental certification and evaluated their performance in two basic areas to detect those that did not show environmental responsibility, despite holding certifications that suggested otherwise. Following the proposal by Dorfleitner & Utz (2023), this research considered two specific dimensions to detect greenwashing and formulated a Greenwashing Indicator for Colombian manufacturing companies. Its development is explained below.

The target variables evaluate two facets that an environmentally responsible company should tackle (see Table 1). A dichotomous variable is established, assuming a value of 1 when the establishment holds environmental certification but does not comply with conditions a, or b. This is due to the fact that the criteria for such certifications encompass measures pertaining to water conservation and solid waste management.

- (a) Companies have comprehensive solid waste plans as a planning tool; this variable indicates the existence of a solid waste

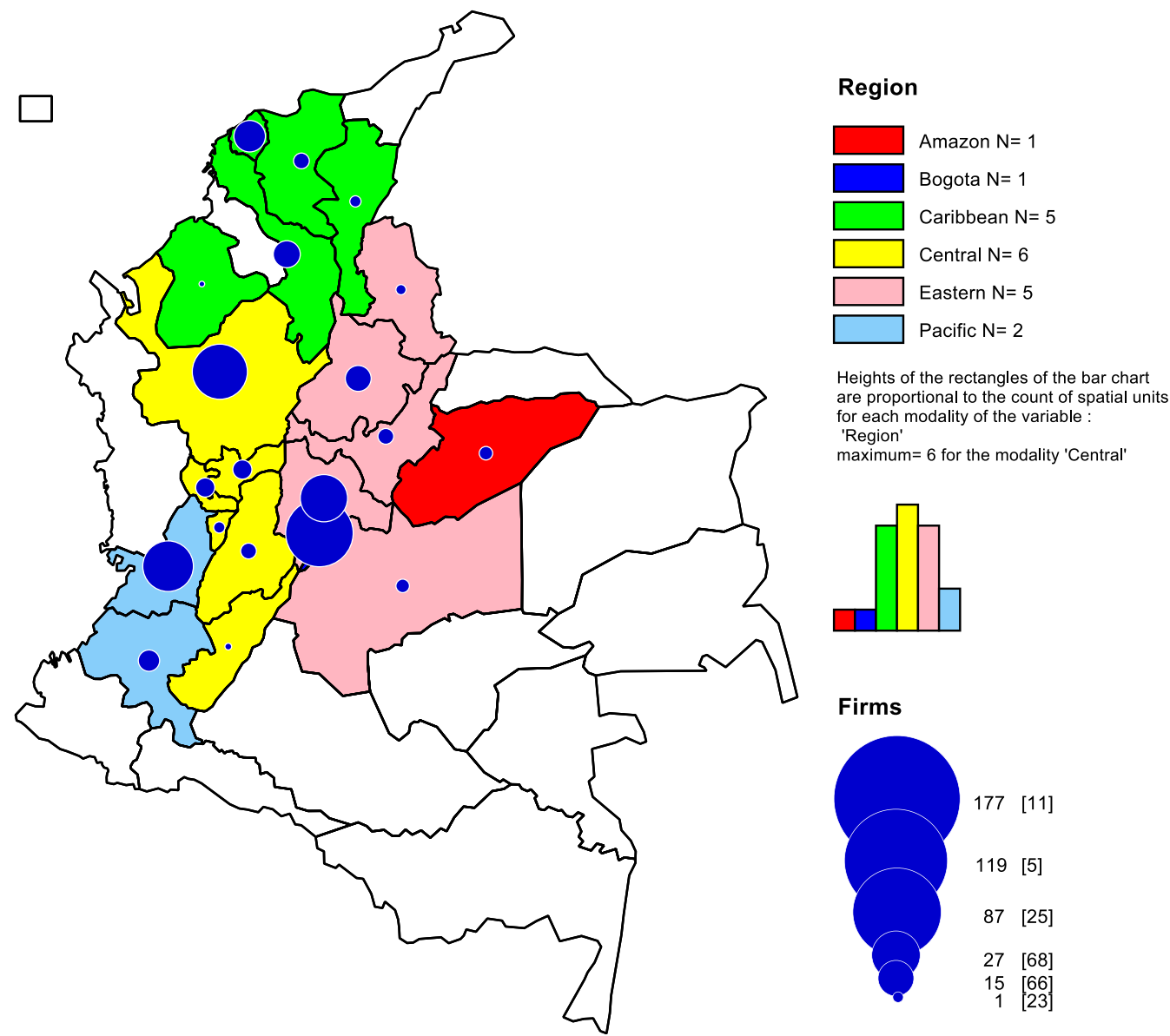


Fig. 1. Geographical Distribution of Sampled Companies. Note: Fig. 1 displays a map of the geographical distribution of the companies that are part of the sample in various regions, such as Amazon, Bogotá, Caribbean, Central, Eastern, and Pacific. Source: Own elaboration using Philcarto (Waniez, 2024).

Table 1

Target variables.

Target Variable	Variable Type	Question
Comprehensive solid waste plans as a planning instrument	Nominal	Comprehensive solid waste plans as a planning instrument 1. Yes, 0. No
Program for efficient use and saving of water	Nominal	Does the establishment have a program for efficient use and saving of water? Yes, 0. No

Note. Adapted from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

management plan. Companies with environmental certification must develop strategies with clear goals and objectives for solid waste management. Firms should responsibly manage waste, demonstrating their environmental commitment and corporate social responsibility.

(b) The company has an efficient water use and savings program. These programs promote the rational and efficient use of water in the company's facilities and production processes. It is crucial to note that some firms with environmental certifications engage in greenwashing practices. These firms publicly announce their environmental responsibility without implementing significant measures of recirculation, water conservation, and wastewater treatment.

The corresponding equations would be as follows:

Variable (a): Presence of a comprehensive solid waste management plan.

a = 1 if the company has a comprehensive solid waste management plan.

a = 0 if the company does not have a comprehensive solid waste management plan.

Variable (b): Presence of a water usage and conservation program.

b = 1 if the company has a water usage and conservation program.

b = 0 if the company does not have a water usage and conservation program.

Target variable:

$$y_i = 1 - \min\{(a_i, b_i)\}$$

In this case, the 'i' corresponds to the observation for each company.

Table 2 displays the predictors that will be used in the machine learning models.

Algorithms

This paper employed a quantitative, cross-sectional methodology using three machine learning algorithms (Decision Tree Classification, Random Forest, Logistic Regression) and the double machine learning technique (Random Forest and Logistic Regression) to identify causal effects and significant variables. The initial model (Decision Tree Classification, Random Forest, Logistic Regression) predicts the impact of an independent variable on a dependent one, while the subsequent model (double machine learning) forecasts the dependent variable based on the independent variable and other factors.

The core premise of double machine learning posits that if the independent variable indeed holds a causal effect on the dependent variable, then the second model should demonstrate superior predictive accuracy compared to the first. This superiority arises from the second model's access to supplementary information provided by the independent variable (Fink et al., 2023).

In the quest to identify significant variables, double machine learning employs a process known as variable selection. This technique sifts through variables to pinpoint the most impactful ones for the

Table 2

Predictor variables.

Category	Question	Notation	Variable Type
Firms level factors	Code of the groups of industrial divisions: Food, beverages, and tobacco	Codigo_actividad	Nominal
	Coking, manufacture of petroleum refining products and nuclear fuel		
	Manufacture of rubber and plastic products		
	Manufacture of chemical products and substances		
	Wood and cork industry, paper manufacturing and publishing and printing activities		
	Non-metallic mineral products industry		
	Metallurgy and manufacture of metal products		
	Textiles, clothing, footwear and leather		
	Other industrial divisions.		
	Fixed assets	Activfi	Scalar
Corporate environmental responsibility	Value Added.	Valagri	Scalar
	Sales value	valorven	Scalar
	Payment of licenses, permits, fees, and environmental fines by the firm	c1acygga	Scalar
	Amount of expenditures made by the company on personnel dedicated to environmental protection activities.	c1acyggg	Scalar
	Total investment and expenses for the Management of mineral and energy resources	c1inygrmye	Scalar
	Amount of investment and expenses for air and climate protection by the firm.	c1iygtort2	Scalar
	Amount investment and expenses for air and climate protection.	c1iygtotr1	Scalar
	Amount investment and expenses for the treatment and disposal of hazardous waste	c1iygtotr3r2	Scalar
	Measurement, control, and analysis to reduce the consumption of non-renewable energy resources	c1mcanrcen	Scalar
	Amount of investment and expenses for the prevention of contaminant infiltration.	ingaprincon	Scalar
	Payments for environmental policies	ppolizamb	Scalar

(continued on next page)

Table 2 (continued)

Category	Question	Notation	Variable Type
Strategic planning	Environmental assessment as a planning tool	c4ga3ip1	Nominal
	Alternative environmental assessment as a planning tool	c4ga3ip2	Nominal
	Environmental management plan as a planning tool	c4ga3ip3	Nominal
	Environmental contingency plan as a planning tool	c4ga3ip4	Nominal
	Environmental monitoring and tracking program as a planning tool	c4ga3ip5	Nominal
	Environmental risk assessment as a planning tool	c4ga3ip6	Nominal
	Cleaner production program as a planning tool	c4ga3ip7	Nominal
	Clean development mechanism (CDM) project	c4ga3ip9	Nominal
	Other planning tool	c4ga3ip10	Nominal
	Amount of investment and expenses research and development (basic, experimental, or applied)	clacygge	Scalar
Geographical location	Geographic region in which it is located.	Region code DANE	Nominal

Note. Adapted from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

machine learning model. In the specific context of double machine learning, variable selection is crucial for singling out the variables exerting a causal effect on the dependent variable.

As mentioned earlier, initially three machine learning algorithms were employed: decision tree, random forests, and logistic regression. Salkind (2007) defines a *Tree Classification* as a machine learning algorithm that establishes a tree- structurer classifier. This tree-structurer classifier uses the Gini Index, which expresses that the smaller the value is, the higher the probability a class has on the samples (Chung-Feng et al., 2021; Ramanujam et al., 2020).

Gini Index (T) = $1 - \sum_{j=1}^n p_j^2$

The random forest algorithm, categorized as ensemble learning, amalgamates the outputs of multiple decision trees to yield a unified result. Each decision tree within the random forest is constructed from a randomly selected subset of the training data. This intentional randomness plays a pivotal role in diminishing the variance associated with individual trees, consequently enhancing the overall precision of the model (Fox et al., 2020).

The sequential process of the random forest can be outlined as follows:

1. Generation of a set of decision trees.
2. Construction of each decision tree using a randomly chosen subset of the training data.
3. Aggregation of the outputs from each decision tree through a voting mechanism to determine the final result.

For the logit model (Zhang, 2023):

$$(d) = \frac{1}{1 + e^{-s}}$$

Assume s as a linear function of several explanatory variables.
 $s = \beta_0 + \beta_1 V_{1k} + \beta_2 V_{2j} + \beta_3 M + \beta_4 V_{3h} + \beta_5 G.$
P(d) is the probability that the dependent variable equals success (take the value of 1). P(d) is a number between zero and one.
 $P(d) = 1/(1 + e^{-(\beta_0 + \beta_1 V_{1k} + \beta_2 V_{2j} + \beta_3 M + \beta_4 V_{3h} + \beta_5 G)})$
$$d_i = \begin{cases} 1 & \text{if } \beta_0 + \beta_1 V_{1k} + \beta_2 V_{2j} + \beta_3 M + \beta_4 V_{3h} + \beta_5 G > 0 \\ 0 & \text{otherwise} \end{cases}$$

V_{1k} represents a vector of factors at the firm level, where k indexes predictors such as economic activity, value-added, fixed asset value, and sales value (Table 2). Industrial firms engaged in polluting economic activities may find it challenging to reduce their environmental impact and may resort to greenwashing. Large companies have incentives to increase short-term profits and attract environmentally conscious consumers without making significant changes because shareholders push for quick profits.

V_{2j} represents the vector of corporate environmental responsibility, and j indexes predictors that show the willingness of the company to invest in the environment. Companies invest in the management of mineral and energy resources, personnel dedicated to environmental protection activities, air and climate protection, treatment and disposal of hazardous waste, reduction in the consumption of non-renewable energy resources, prevention of pollutant infiltration, and environmental policies (Table 2). Another aspect included in this vector is actions focused on water resource conservation, such as the volume of wastewater treated by the facility and savings in the amount of water used.

V₃ is configured as a key component that encompasses a set of fundamental variables within firms' strategic planning. These variables are intrinsically linked to the environmental objectives that form an integral part of the strategy of these organizations. Essentially, the variable V₃ becomes an indicator or specific category that captures and represents the environmental goals and guidelines that the company aims to achieve within the framework of its strategic planning. On its part, h indexes the predictors of environmental goals in planning.

M is the predictor associated with research and development (basic, experimental, or applied) (Table 2). Corporate research and development play a critical role in the fight against greenwashing. Companies that invest more in research manage to develop more sustainable products and mitigate their environmental impact.

G is the predictor associated with geographical location (Table 2). Firms located in regions where environmental oversight is less stringent are more likely to engage in greenwashing. Conversely, in regions where most consumers prioritize environmental consciousness, industrial firms feel pressured to minimize environmental impact and increase transparency.

LASSO (Least Absolute Shrinkage and Selection Operator) minimizes the prediction error subject in the model and measures the complexity through the sum of all absolute values (StataCorp., 2023b). According to Belloni et al. (2016), the plugin implemented for lasso logit i

$$\lambda_{\text{logit}} = \frac{f}{2\sqrt{N}} \Phi^{-1} \left[1 - \frac{1}{2\max\{N, p \ln N\}} \right]$$

Results

Descriptive analysis

The descriptive analysis reveals that out of the 681 environmental companies certified in the sample, 55 companies (8 %) do not invest in personnel dedicated to environmental protection activities. Only 99

companies (14.53 %) invest in preventing water pollution infiltration and have staff devoted to environmental protection activities. Meanwhile, 67 companies (9.8 %) invest in preventing infiltration of water pollution and managing mineral and energy resources. Surprisingly, 391 certified companies (57.4 %), which are more than half of the sample, do not invest in water treatment and energy-saving resources. In addition, 49 companies (7.2 %) do not invest in managing mineral and energy resources or personnel dedicated to environmental protection activities, while 42 (6.2 %) companies do not have staff dedicated to climate protection and environmental management.

Environmental certification assures customers that companies minimize their environmental impact through their actions. However, the descriptive statistics suggest that most industrial companies in Colombia with environmental certification do not adopt sustainable practices. Hence, environmental certifications seem to have limited effectiveness in Colombia.

Machine learning models

Algorithms

Decision tree classification

The data from 681 certified companies were divided into two groups using algorithms, with 80 % of the sample allocated for training purposes. The resulting model, featuring 129 splits, achieved a test accuracy of 0.9, and the decision tree consisted of 5 terminal nodes. Within the certified companies, 23 firms were grouped in a node, characterized by the absence of Comprehensive Solid Waste Plans as a planning instrument or a program for the efficient use and conservation of water. These firms fulfilled three conditions: (i) personnel expenses dedicated to environmental protection activities were less than USD 2378; (ii) they lacked an environmental contingency plan as a planning instrument; (iii) the company's investment and expenses for air and climate protection were less than USD 108 (see Fig. 2).

The decision tree employed the following predictors for classification: (i) personnel expenses dedicated to environmental protection, (ii) the existence of an environmental contingency plan as a planning

instrument, (iii) the presence of an environmental monitoring and tracking program as a planning instrument, (iv) the amount of investment and expenses for air and climate protection by the company, and (v) sales value (refer to Table 3).

According to the algorithm, environmental certifications should vigilantly monitor the presence of a corporate environmental strategy to prevent greenwashing. This study particularly underscores that the investments undertaken by companies and the alignment of environmental goals as an integral part of their strategy are crucial aspects to avoid greenwashing.

Random forest

The dataset was split into two groups, with 80 % of the sample designated for algorithm training. A total of 44 trees were generated by the random forest method, achieving an accuracy level of 0.9. Analysis of the algorithm revealed that the variables contributing to increased node purity included: (i) personnel expenses dedicated to environmental protection activities, (ii) the presence of an environmental contingency plan as a planning instrument, (iii) the investment and expenses related to air and climate protection by the company, (iv) the existence of an

Table 3

Importance of the Variable.

Variable	Relative Importance
c1acyggg	38.7
c4ga3ip4	26.6
c4ga3ip5	10.3
c1iygtotr2	6.8
valorven	6.0
valagri	3.7
c4ga3ip1	2.5
activfi	2.2
c1inygrmye	1.7
c4ga3ip6	1.5

Note. Adapted using JASP©0.16.3 from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

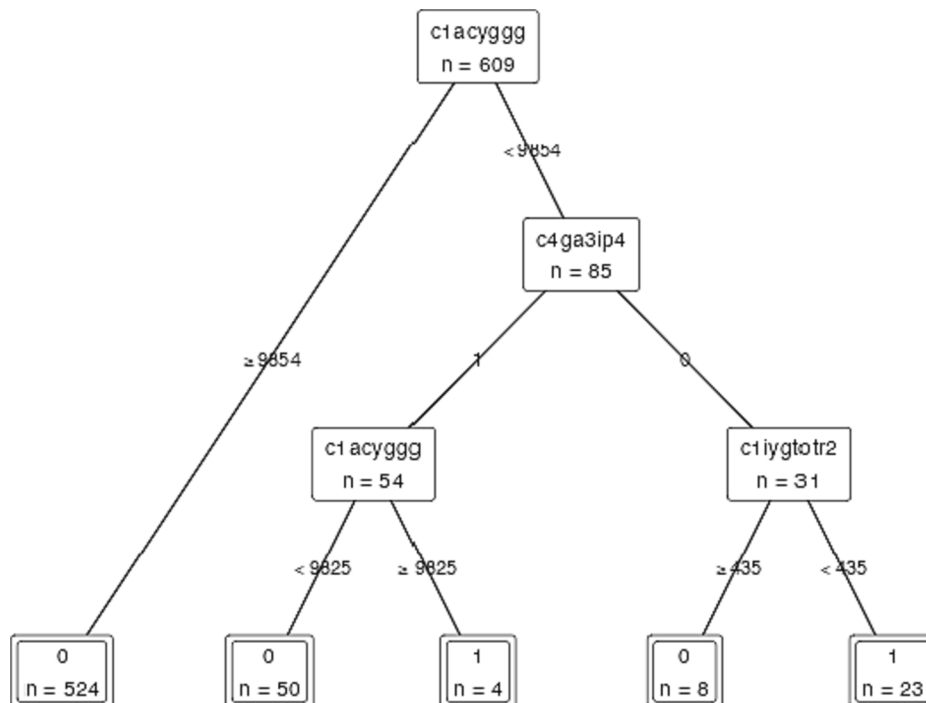


Fig. 2. Decision Tree Plot. Note. Estimation results using JASP©0.16.3. from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

environmental monitoring and surveillance program as a planning instrument, and (v) the payment of licenses, permits, fees, and environmental rights (see Fig. 3).

The results obtained from the random forest analysis suggest that companies investing more in personnel dedicated to environmental protection activities, water and air protection, and the payment of licenses, permits, fees, and environmental rights are less likely to engage in greenwashing, as indicated by the decision tree.

The study emphasizes the crucial role of stringent environmental regulations in Colombia, including the imposition of fines and economic sanctions on polluting companies. By doing so, policymakers can incentivize firms to embrace sustainable practices and reduce their environmental impact, thereby addressing the challenge of greenwashing within the country's manufacturing sector.

The algorithm highlighted that the most significant predictor for Colombian industrial firms engaging in greenwashing is the expenditure on personnel dedicated to environmental protection activities. Having specialized personnel creates opportunities for enhancing environmental management, enabling them to identify and implement measures to reduce resource consumption and minimize waste. Effective management of solid waste and water resources can positively influence the company's image and its relationships with stakeholders.

Moreover, the study revealed that economic obligations and sanctions drive companies to take responsibility for managing their solid waste. Additionally, the algorithm confirmed that research fosters innovation in solid waste and water management.

Logistic regression classification algorithm with LASSO

The data from 681 certified companies were partitioned using algorithms, with 90 % of the sample allocated for training. The logistic regression model with LASSO regularization demonstrated the relationships between greenwashing and the predictors, achieving a classification accuracy of 82 %. Notably, the algorithm unveiled that companies situated in Antioquia, Bolívar, Tolima, Risaralda, or Boyacá, and those engaged in coking, the manufacture of petroleum refining products, nuclear fuel, or the manufacture of rubber and plastic products, or the non-metallic mineral products industry, were more likely to practice greenwashing (Table 4). These results suggest a need for policies that prioritize solid waste and water management in these

Table 4
The logistic regression classification algorithm with LASSO.

Variable	Coefficients
dpto = 63	1.36612
codigo_actividad = 2	0.898777
codigo_actividad = 6	0.745955
c4ga3ip5 = 0	0.357371
dpto = 15	0.323757
dpto = 76	0.233281
codigo_actividad = 3	0.215978
c4ga3ip4 = 0	0.167991
dpto = 5	0.138938
dpto = 13	0.0669605
ingaprincon	1.99E-06
clacygga	1.79E-06
clacyggg	1.30E-06
cliygtotr1	2.36E-07
cliygtotr3r2	8.83E-08
activfi	-1.41E-09
cliygtotr2	-3.56E-07
clinygrmye	-2.65E-06
clacygge	-1.92E-05
ppolizamb	-0.000234
c4ga3ip1 = 0	-0.0590088
dpto = 68	-0.149096
c4ga3ip6 = 0	-0.315262
codigo_actividad = 8	-0.391079
dpto = 11	-0.460875
codigo_actividad = 9	-0.483755
codigo_actividad = 1	-0.526493
c4ga3ip7 = 1	-0.526558
c4ga3ip5 = 1	-0.891547
c4ga3ip9 = 1	-0.936213
c4ga3ip4 = 1	-0.968626

Note. It was adapted using Stata 18 from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

industries. In contrast, companies located in Santander or Bogotá, or involved in the food, beverages, and tobacco industry, or textiles, clothing, footwear, and leather, exhibited a lower likelihood of engaging in greenwashing (Table 4). For a comprehensive environmental management approach, policies should account for these variations and implement targeted interventions addressing the specific challenges faced by

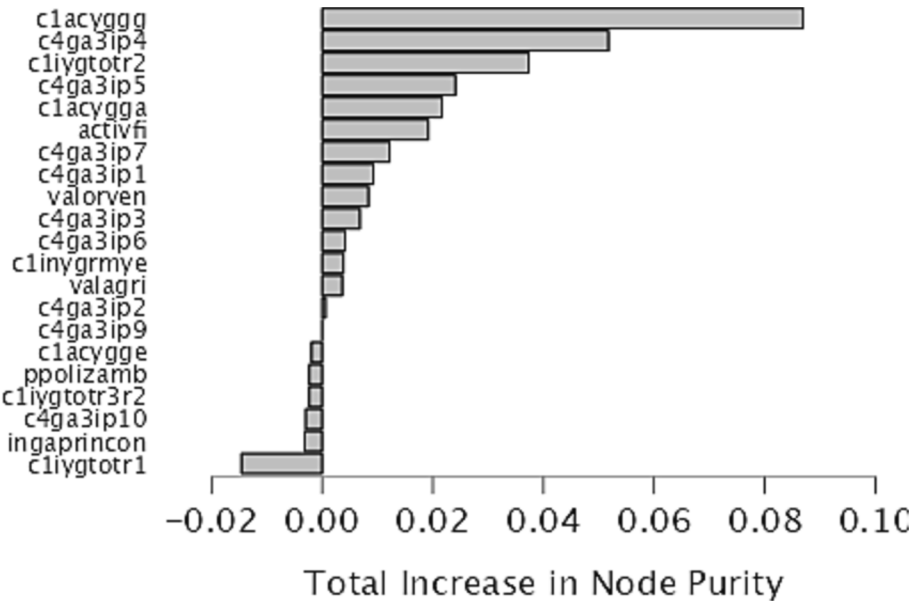


Fig. 3. Total Increase in Node Purity. Note. Ramdon Forest. They were adapted using JASP©0.16.3. from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b).

different industries and regions.

The National Environmental Policy for Integrated Solid Waste and Water Management should adopt a sectoral approach based on these findings. Interestingly, it remains unclear why environmental certifications do not mandate evidence of companies' investments in solid waste management relative to their waste production. For instance, ISO 14001, a certification held by several surveyed companies, commits to pollution prevention. However, some Colombian companies adhering to or implementing this standard do not include solid waste management in their corporate strategy, as revealed by the survey data.

Another variable decreasing the probability of companies practicing greenwashing was the presence of a cleaner production program as a planning instrument and a project for clean development mechanisms (c4ga3ip7 and c4ga3ip9). Cleaner production, a systematic approach to waste reduction, resource conservation, and minimizing industrial environmental impact, offers numerous benefits such as improved environmental performance, reduced operating costs, and enhanced competitiveness. Implementation of a cleaner production program (CPP) can help companies save costs related to waste management and regulatory compliance. The study indicates that CPP adoption in the Colombian manufacturing sector relies on a company's willingness to invest in a program supporting air and climate protection objectives, preventing pollutant infiltration, and decreasing the environmental impact of industrial activities.

These results align with existing literature on environmental management, emphasizing the interconnectedness of various environmental issues and the need for a comprehensive approach to environmental management (Zhu et al., 2018).

The practical implications for industrial companies in Colombia are significant. To enhance water use efficiency, companies should adopt a comprehensive environmental management approach, investing in air and climate protection, preventing pollutant infiltration, and resource management. This approach not only enhances environmental performance but also improves economic performance by reducing water consumption and associated costs.

Double machine learning

The Double/Debiased Machine Learning (DDML) method, employed in this study, combines various machine learning models to predict treatment and outcome variables. It utilizes a cross-fitting procedure to reduce bias and generate more efficient estimates. This technique handles high-dimensional nuisance parameters, is robust to misspecification of machine learning models, and is more efficient than traditional methods, especially with a large number of nuisance parameters (Chernozhukov et al., 2018). Essentially, DDML is a machine learning-based approach to causal inference (Díaz, 2019).

Two types of algorithms, logit and random forest, were used for DDML. Multiple machine learning models were created using different algorithms, generating two global variables—one grouping all predictors other than the treatment variable (X) and the treatment variable itself (D). These models were trained using data defined in the global variables X and D. Finally, the trained models were used to estimate treatment effects.

The Shortstack model, referencing the estimation that combines machine learning models, utilized a logistic regression model to predict the treatment variable (D) and a random forest model to predict the outcome variable (Y). The following steps were taken:

1. Train the logistic regression model on a dataset containing predictor variables X and the treatment variable D.
2. Train the random forest model on a dataset containing predictor variables X and the outcome variable Y.
3. Use the model to predict the treatment variable for each company in the dataset.

4. Use the model to predict the outcome variable for each company in the dataset.
5. Estimate treatment effects using differences between model predictions for companies receiving and not receiving treatment.

In this context, it's important to highlight that the Shortstack model is an effective technique for estimating treatment effects in the presence of high-dimensional data. It is robust to misspecification of machine learning models and capable of producing accurate estimates.

Cross-validation was performed to assess model performance across different data folds, and mean squared error (MSE) was calculated to measure performance in different folds (4). Subsequently, model performance was estimated using three techniques: (i) a robust technique for estimating DDML model performance, which is resilient to misspecification of machine learning models, (ii) a technique that considers all possible combinations of predictor variables to find the specification minimizing MSE, and (iii) the spec(1) replay method that reproduces DDML analysis results using a validation dataset through a cross-fitting procedure. Overall, the analysis results indicated that DDML models were accurate and robust to misspecification.

The three estimation methods mentioned above yielded the following results:

1. The robust estimation method provided coefficients and standard error estimates for the final DDML model.
2. The allcombs method estimated all possible combinations of Y and D learners.
3. The spec(1) replay method reproduced the specification with minimum MSE.

Results from a specification analysis are then presented, where Specification 1, minimizing MSE, is chosen. This is used to select the model that best fits the data.

In the final stacking models, the DDML analysis results suggested that having an environmental contingency plan and an environmental monitoring and tracking program as planning instruments were the two significant variables with a negative effect on firms practicing greenwashing. Meanwhile, having a cleaner production program and a clean development mechanisms (MDL) project showed no significant effect on the probability of firms engaging in greenwashing (MDL) practices (Table 5).

The analysis results underscore the importance of having effective environmental contingency plans and monitoring programs as key strategies to prevent or reduce greenwashing practices in companies. These findings support the relevance of proactive environmental management and the implementation of concrete measures to enhance organizational environmental performance. Such programs can help companies identify areas for improvement and take actions to reduce their environmental impact. Additionally, the DDML analysis results suggest that cleaner production programs and clean development mechanisms (MDL) projects do not have the expected effect on greenwashing.

Double machine learning estimates were then conducted for variables associated with strategic planning that emerged as significant predictors in the previous algorithms. These included: (i) Environmental contingency plan as a planning instrument (c4ga3ip4); (ii) Environmental monitoring and tracking program as a planning instrument (c4ga3ip5); (iii) Cleaner production program as a planning instrument (c4ga3ip7); (iv) Clean development mechanisms (MDL) project (c4ga3ip9).

In order to ascertain the impact of variables associated with corporate environmental responsibility, identified as significant by decision trees and random forests for classification, two dichotomous variables were created. The first variable takes on the value of 1 if expenses for personnel dedicated to environmental protection activities exceeded USD 2378, and 0 otherwise. The second variable assumes the value of 1

Table 5

The DDML model examines variables associated with strategic planning.

Min MSE DDML model, specification 1 y-E[y X] = y-Y1_logit_1 D-E[D X] = D-D1_logit_1						
green1	Coefficient	Robust std. err.	z	P > z	[95 % conf.	interval]
c4ga3ip4	-0.1339252	0.0463691	-2.89	0.004	-0.224807	-0.0430434
_cons	-0.0015544	0.0128804	-0.12	0.904	-0.0267995	0.0236907
c4ga3ip5	-0.1062854	0.0342538	-3.1	0.002	-0.1734217	-0.0391492
_cons	-0.0027582	0.0129198	-0.21	0.831	-0.0280807	0.0225642
c4ga3ip7	-0.0385717	0.0274026	-1.41	0.159	-0.0922799	0.0151364
_cons	-0.0028454	0.0129015	-0.22	0.825	-0.0281319	0.0224411
c4ga3ip9	-0.0524801	0.0275442	-1.91	0.057	-0.1064658	0.0015055
_cons	-0.0069817	0.0117387	-0.59	0.552	-0.0299891	0.0160257

Note. It was adapted using Stata (18) from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b). Source: own elaboration carried out in Stata 18 (StataCorp., 2023a).

if the investment and expenses related to air and climate protection by the company surpassed USD 108, and 0 otherwise. These specific cutoff points were determined based on the outcomes of the classification tree.

The results from the stacking models revealed that, stemming from the DDML analysis, it can be inferred that expenses for personnel dedicated to environmental protection activities and the investment and expenses related to air and climate protection are variables that exhibit a significant correlation and have a negative impact on the likelihood of companies practicing greenwashing (refer to Table 6). These findings underscore the importance of investing in environmental sustainability as a strategic approach to prevent or mitigate greenwashing practices.

A DDML analysis was conducted to examine the economic activities identified by logistic regression using the LASSO regularization algorithm as significant predictors of companies engaging in greenwashing. The findings revealed that only the manufacturing of non-metallic mineral products demonstrated a significant and positive association with the target variable. In contrast, activities related to coking, the manufacture of petroleum refining products and nuclear fuel, as well as the manufacture of rubber and plastic products, did not show a significant relationship (see Table 7).

Companies involved in the manufacturing of non-metallic mineral products also tend to heavily rely on natural resources, such as wood and metals. Additionally, these companies often generate substantial amounts of waste and pollution. Consequently, such companies may be more inclined to partake in greenwashing practices.

The developed indicator for detecting greenwashing enables the identification of the geographical distribution of companies practicing this deceptive strategy, as illustrated in Fig. 4. This tool offers a clear and specific insight into the geographic locations of the companies engaged in greenwashing. This analysis provides an overview of the greenwashing situation in different regions. The Eastern region stands out with the highest average greenwashing percentage, while the Caribbean region has the lowest. The total number of firms engaged in greenwashing across all regions is significant, highlighting the prevalence of this deceptive practice.

Table 6

The DDML model examines variables associated with corporate environmental responsibility.

Min MSE DDML model, specification 1 y-E[y X] = y-Y1_logit_1 D-E[D X] = D-D1_logit_1						
Number of obs personal = 622 Number of obs air1 = 624						
green1	Coefficient	Robust std. err.	z	P > z	[95 % conf.	interval]
personal	-0.2451415	0.0506176	-4.84	0	-0.3443502	-0.1459329
_cons	-0.0051705	0.0124715	-0.41	0.678	-0.0296143	0.0192732
air1	-0.0807041	0.0369273	-2.19	0.029	-0.1530802	-0.008328
_cons	-0.0024589	0.0129181	-0.19	0.849	-0.0277779	0.0228602

Note. It was adapted using Stata (18) from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b). Source: own elaboration carried out in Stata 18 (StataCorp., 2023a).

The results (Table 8) provide valuable insights into the distribution of greenwashing practices across different departments:

- The greenwashing percentage range of 51 % to 75 %, indicating a prevalent occurrence of deceptive practices.
- On average, approximately 56.09 % of firms in the dataset engage in greenwashing, with a total of 382 firms practicing such deceptive tactics.
- Norte de Santander stands out with 54 firms involved in greenwashing, marking it as a department with the highest indicators at 75 %.
- Bogota, the capital, has 100 firms participating in greenwashing, signifying a substantial presence of deceptive practices in this department.
- The distribution of departments based on their greenwashing percentages presents a diverse landscape. Three exhibit minimal greenwashing practices within the 0 %-39 % range, while eleven fall into the 40-59 % range, indicating a moderate level of greenwashing. Furthermore, five are positioned in the 60-75 % range, showcasing a substantial degree of engagement in greenwashing activities.

Discussion

The present study builds upon a growing body of research that has uncovered the prevalence of greenwashing in the Colombian manufacturing industry, as proposed by Camelo (2021), even among firms that hold environmental certifications. This finding is consistent with prior investigations that have exposed the disconnect between a company's outward sustainability image and its actual environmental practices (Pizzetti et al., 2019; Datta & Goyal, 2022). Moreover, the study delves into the factors that influence greenwashing practices, shedding light on five main determinants, including a company's willingness to invest in environmental initiatives and its approach to water and wastewater management. This underscores a recurring theme in

Table 7
The DDML model economic activity.

Min MSE DDML model, specification 1						
y-E[y X] = y-Y1_logit_1						
D-E[D X] = D-D1_logit_1						
Number of obs act2 = 583						
Number of obs act3 = 630						
Number of obs act6 = 624						
green1	Coefficient	Robust std. err.	z	P > z	[95 % conf.	interval]
acti2	0.1288754	0.0934389	1.38	0.168	−0.0542615	0.3120122
_cons	−0.0091187	0.013079	−0.7	0.486	−0.0347531	0.0165158
acti3	0.022981	0.0467359	0.49	0.623	−6.86E-02	0.1145817
_cons	−0.003709	0.0129344	−0.29	0.774	−2.91E-02	0.0216419
acti6	0.1491914	0.0570179	2.62	0.009	3.74E-02	0.2609444
_cons	−0.0026263	0.0129123	−0.2	0.839	−0.027934	0.0226813

Note. It was adapted using Stata (18) from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b). Source: own elaboration carried out in Stata 18 (StataCorp., 2023a).

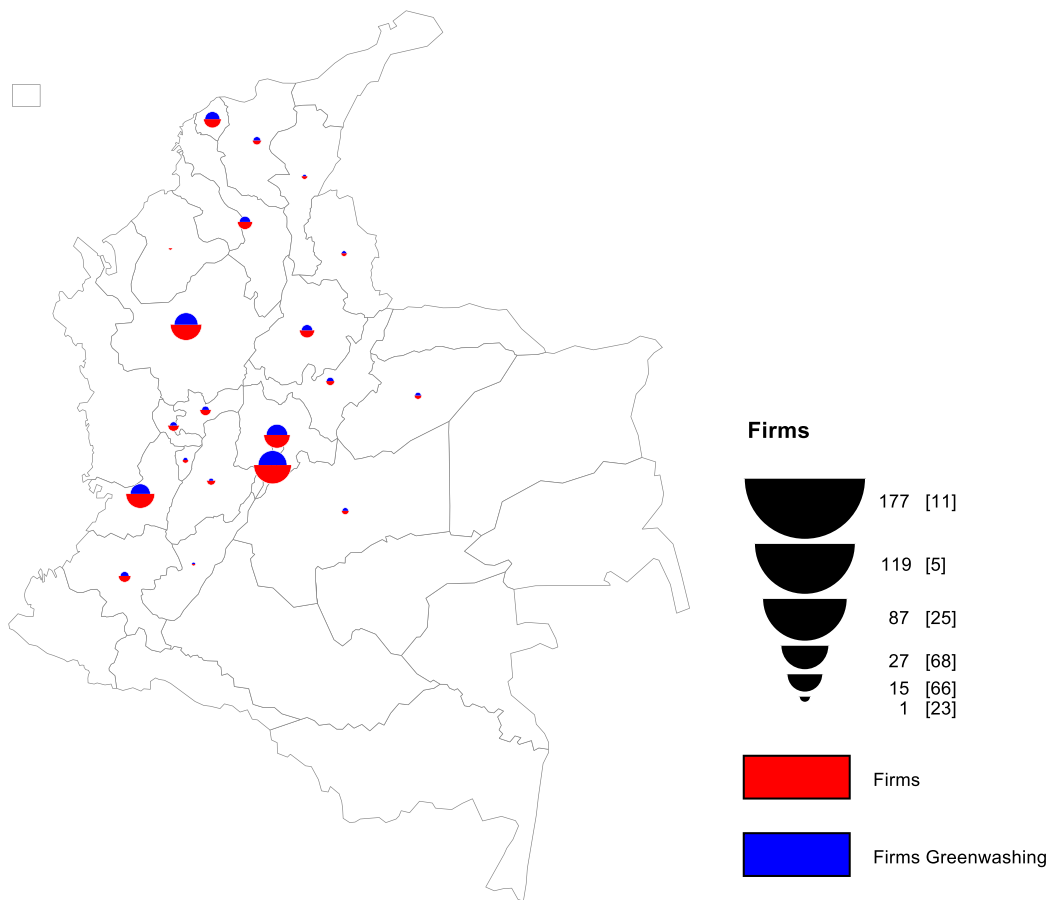


Fig. 4. Total firms in the sample (red) Vs. firms practicing greenwashing (blue) Note: Fig. 1 displays a map of the geographical distribution of the companies that are part of the sample in various regions, such as Amazon, Bogotá, Caribbean, Central, Eastern, and Pacific. Source: Own elaboration using Philcarto (Waniez, 2024). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

research, which suggests that economic interests often take precedence over environmental responsibility (Rashidi-Sabet & Madhavaram, 2021). This research enhances our comprehension of the factors predicting greenwashing in Colombian manufacturing industries, aligning with the approach proposed by De Freitas Netto et al. (2020). Specifically, it follows the theory of corporate legitimacy in the cognitive realm, where environmental certifications obtained by certain companies manifest assumptions about their environmental responsibility, which are shared in the social environment. Consequently, this study is centered on

examining corporate greenwashing, a phenomenon where companies deceive consumers about their environmental practices, as proposed by Delmas and Burbano (2011). Additionally, building on Wu et al.'s (2020) research, this study delves into less observable environmental aspects often neglected by environmentally certified companies. It explores decoupling behavior exhibited by companies, where symbolic actions, like environmental certifications, lack concrete support in terms of strategic planning, investments in human capital for environmental protection, and contributions to air and climate care.

Table 8

Total firms in the sample (red) Vs. firms practicing greenwashing (blue).

Department	Firms in the sample	Region	Percentage of firms in the sample engaging in greenwashing practices	Firms in the sample engaging in greenwashing practices
Antioquia	119	Central	56,90 %	68
Atlántico	40	Caribbean	66,70 %	27
Bogotá	177	Bogotá	56,30 %	100
Bolívar	28	Caribbean	58,60 %	16
Boyacá	9	Eastern	66,70 %	6
Caldas	15	Central	46,70 %	7
Cauca	17	Pacific	47,10 %	8
Cesar	5	Caribbean	40 %	2
Córdoba	1	Caribbean	0 %	0
Cundinamarca	87	Eastern	62,40 %	54
Huila	2	Central	50 %	1
Magdalena	9	Caribbean	66,70 %	6
Meta	7	Eastern	57,10 %	4
Norte de Santander	4	Eastern	75 %	3
Quindío	5	Central	60 %	3
Risaralda	15	Central	46,70 %	7
Santander	27	Eastern	59,30 %	16
Tolima	9	Central	33,30 %	3
Valle del Cauca	98	Pacific	48 %	47
Casanare	7	Amazon	60 %	4
	681		56,09 %	382

Note. It was adapted using Stata (18) from Departamento Administrativo Nacional de Estadística (DANE, 2022a, 2022b). Source: own elaboration carried out in Stata 18 (StataCorp., 2023a).

In alignment with Contreras-Pacheco and Claasen's (2017) typology of corporate greenwashing, this study specifically focuses on companies claiming environmental achievements, such as certifications, yet exhibiting actions inconsistent with these claims. Thus, the study delves into one of the strategies proposed by Adamkiewicz et al. (2022), specifically deceptive claims involving unclear labels and certifications.

Addressing the empirical challenge posed by Ghitti et al. (2023) to quantitatively capture greenwashing, the proposed indicator in this research builds upon the frameworks of Dorfleitner and Utz (2023) and Yildirim (2023). It identifies inconsistencies between real and apparent green performance, utilizing information declared by firms in the biennial industrial environmental survey conducted in Colombia. The objective is to ascertain whether the ecological image projected by certifications truly translates into tangible actions.

While attribution theory has been commonly used to investigate the effects of irresponsible corporate behaviors on consumer trust (Santos et al., 2023), this research applies this theoretical framework to illuminate the behavior of Colombian manufacturing companies with environmental certifications. It considers aspects of intentionality, controllability, and consistency.

The results disclose that, regarding intentionality, companies obtaining these certifications do not always have the intention to mitigate their environmental impact in solid waste management and the use and treatment of water. Furthermore, the study evaluates the control exercised by companies over their environmental practices, emphasizing strategic planning. The findings indicate that possessing an environmental contingency plan and an environmental monitoring and tracking program as planning instruments were the two most significant variables exhibiting a negative effect on greenwashing. Finally, the research explores the consistency between a company's environmental certifications and its actions, revealing that not all companies with environmental certifications align their actions with sustainability.

In light of the growing prevalence of greenwashing and the increasing importance of advancing research on this topic (Gatti et al., 2021), this study addresses future research opportunities highlighted by

Santos et al. (2023) and Torelli et al. (2020), emphasizing the significance of advancing the study of the taxonomy for greenwashing.

The research contributes to the taxonomy identified by Yang et al. (2020) and De Freitas Netto et al. (2020) by delving into the understanding of various types of greenwashing present in the literature. Additionally, it explores additional levels of deceptive environmental communication, such as environmental certifications. Thus, this study contributes to the types of greenwashing as follows:

- (i) **Selective Disclosure:** According to this study, greenwashing can occur when an organization obtains an environmental certification and promotes it as part of its sustainable image. This implies that the issue is not limited solely to the information provided by the company in its sustainability reports, as identified by previous studies, but extends to environmental certifications. In other words, the mere acquisition of certifications without tangible actions can be considered a form of selective disclosure. This suggests that the validity of the certification depends on concrete actions supported by real evidence of sustainable practices, such as the implementation of contingency plans, expenses for personnel dedicated to environmental protection activities, and investment, along with expenses related to air and climate protection.
- (ii) **Decoupling:** This type of greenwashing is broadened, as it not only occurs when meeting the expectations of shareholders but also when failing to meet the expectations set by environmental certifications. This suggests that certification can become an ambitious project without assessing whether sufficient funds are available.
- (iii) **Attention Deflection:** In this type of greenwashing, attention may be diverted toward environmental certification rather than directed at real investments in air, climate, water protection, waste management, and practices that ensure environmental care.
- (iv) **Deceptive Manipulation:** In this case, an environmental certification without concrete actions becomes a tool for deceptive manipulation, promoting a misleading perception that a company's objectives and/or policies are environmentally friendly when, in reality, they are not.
- (v) **Dubious Authorizations and Labels:** The study revealed that in Colombia, companies with eco-labeling do not always make the necessary investments to mitigate their impact on the environment. Despite eco-labels being designed to control greenwashing, this study highlights that they can be vulnerable to fraud.

In this context, the contribution of this research to the taxonomy of greenwashing doesn't stem from the identification of a new type but rather enriches the comprehension of existing categories, as elucidated earlier. This is accomplished by viewing environmental certifications as potential instruments for conducting greenwashing, especially when these certifications and regulatory bodies fall short in ensuring effective monitoring to support genuine environmental actions. The study underscores the importance of not only acquiring certification but also ensuring substantial and measurable actions as backing, ultimately reinforcing its capacity to reflect verifiable sustainable practices.

Another noteworthy contribution of this study is the application of machine learning algorithms to predict environmental planning and management within Colombian industrial firms. This innovative method allows for the identification of specific greenwashing techniques, including unsupported environmental claims. Such insights hold immense value for regulators and investors, aiding them in formulating more efficient strategies to combat greenwashing. Additionally, the study emphasizes the significance of implementing national public policies that involve regular oversight of certified companies, accompanied by penalties and punitive measures for those involved in environmental misconduct. These recommendations align not only with

previous research findings but also with the guidance provided by the Colombian government's environmental agency, as highlighted in the works of Camelo (2021), Norouzi et al. (2022), Testa et al. (2018), He et al. (2020), and Christ & Burritt (2019).

To summarize, the study offers significant contributions, including the identification of specific greenwashing tactics and the novel application of machine learning techniques to detect greenwashing in sustainability reports. It highlights the limitations of environmental certifications in ensuring sustainable practices among Colombian industrial companies and underscores the need for more robust regulatory measures. Implementing periodic monitoring of certified companies, backed by fines and sanctions for environmental transgressions, emerges as a crucial recommendation, in line with both previous research and government reports. Ultimately, the study's findings can inform policies and strategies aimed at curbing greenwashing and promoting genuine environmental responsibility in the Colombian manufacturing sector.

Conclusion

The present study adds to the growing body of research on greenwashing within the Colombian manufacturing industry, revealing its prevalence even among environmentally certified firms. This aligns with earlier investigations highlighting the gap between a company's public sustainability image and its actual environmental practices. The study explores the less observable aspects often neglected by certified companies, focusing on decoupling behavior and deceptive claims. The findings underscore the importance of tangible actions aligned with sustainability rather than mere reliance on certifications.

Specifically, the main finding emphasizes the correlation between companies lacking comprehensive environmental initiatives and those engaging in greenwashing practices. The research, through the proposal of a Greenwashing Indicator for Colombian manufacturing companies and the use of three machine learning algorithms (Decision Tree Classification, Random Forest, Logistic Regression) and the technique for causal inference (double machine learning), showed evidence of a concerning trend in the Colombian industrial landscape, where companies with environmental certifications often do not adhere to sustainable practices. This raises questions about the effectiveness of certifications, urging the implementation of stricter criteria, increased supervision, and continuous evaluation of certification criteria.

This research contributes to the taxonomy of greenwashing by delving into various types identified in the literature. It sheds light on selective disclosure, emphasizing that mere certification without tangible actions can constitute greenwashing. The broadened concept of decoupling now includes not only meeting shareholder expectations but also failing to meet those set by certifications. Attention deflection involves diverting focus toward certifications rather than real investments in environmental practices. Deceptive manipulation occurs when certifications create a misleading perception without concrete actions. Additionally, dubious authorizations and labels are revealed, showcasing the vulnerability of eco-labels to fraud. The study enriches existing categories, viewing certifications as potential instruments for greenwashing, reinforcing the need for substantial, measurable actions to validate sustainability claims.

Also, the developed greenwashing detection indicator has facilitated a nuanced understanding of the geographic distribution of companies employing this deceptive strategy. The results uncover substantial variations, notably with the Eastern region exhibiting the highest average greenwashing percentage and the Caribbean region displaying the lowest indicators. The departmental analysis accentuates the prevalence of greenwashing, as more than half of the participating firms are found to engage in deceptive practices. Distinct findings for Norte de Santander and Bogotá emphasize areas warranting increased attention. The diverse distribution of greenwashing percentages across departments underscores the imperative for tailored approaches to combat this widespread phenomenon, emphasizing the necessity for targeted

measures and heightened awareness to counter deceptive practices in diverse regions and business sectors.

On the other hand, the evaluation of control exercised by companies over their environmental practices revealed that the presence of an environmental contingency plan and a monitoring and tracking program are significant factors that reduce greenwashing practices. This highlights the importance of strategic planning.

Considering that techno-economic viability and sustainability are crucial foundations in the environmental management of companies, the methodological approach employed extended beyond the mere identification of environmental practices. It also encompassed the evaluation of their economic feasibility and their contribution to sustainability. Through the integration of machine learning algorithms, a model was crafted, aiming not only to pinpoint potential instances of greenwashing but also to articulate the economic viability of the adopted environmental practices.

Likewise, based on our study it is possible to suggest some policy recommendations:

- Governments should establish clear and consistent labelling standards that mandate accurate and transparent environmental information disclosure by companies.
- Enhanced enforcement mechanisms, including more severe penalties and sanctions, should deter companies from greenwashing practices.
- Mandated transparency and accountability measures, such as third-party verification and independent audits, can ensure accurate reporting of environmental practices.
- Consumer education initiatives should be initiated to empower individuals in identifying and avoiding greenwashing practices while making informed, eco-conscious purchasing decisions.
- Governments must provide funding and support for sustainability initiatives and research, while simultaneously encouraging sustainable practices across industries.
- Collaborative efforts involving governments, industry stakeholders, and civil society groups are pivotal for the development and implementation of effective policies and initiatives to promote sustainability and combat greenwashing.

Acknowledging the study's limitations is crucial, notably the absence of longitudinal data for evaluating long-term environmental practices. Enhancing the study's robustness could involve integrating comparative analyses or case studies drawn from distinct contexts, thereby bolstering the generalizability of the findings. It's also worth exploring the development of specific tools to assess the authenticity of companies' environmental claims.

Ethical statements

This article does not contain any studies with human participants performed by any of the authors. The dataset used is public and available at <https://microdatos.dane.gov.co/index.php/catalog/724> and <https://microdatos.dane.gov.co/index.php/catalog/746>

Informed consent

Since the authors did not interact with human participants, informed consent was not required. It is clarified that the data used is public and anonymized.

CRediT authorship contribution statement

Carolina Henao-Rodríguez: Writing – review & editing, Writing – original draft, Supervision, Software, Data curation, Conceptualization. **Jenny Paola Lis-Gutiérrez:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Data curation. **Harold Delfin Angulo-Bustanza:** Writing – review & editing,

Writing – original draft, Validation, Investigation, Formal analysis.

Declaration of competing interest

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

References

- Adamkiewicz, J., Kocharńska, E., Adamkiewicz, I., & Łukasik, R. M. (2022). Greenwashing and sustainable fashion industry. *Current Opinion in Green and Sustainable Chemistry*, 38, Article 100710. <https://doi.org/10.1016/j.cogsc.2022.100710>
- Braga Junior, S., Martínez, M. P., Correa, C. M., Moura-Leite, R. C., & Da Silva, D. (2019). Greenwashing effect, attitudes, and beliefs in green consumption. *RAUSP Management Journal*, 54(2), 226–241. <https://doi.org/10.1108/rausp-08-2018-0070>
- Camelo, M. D. F. (2021, February 18). *Secretaría de Ambiente inicia 3 procesos sancionatorios por afectación al río Tunjuelo*. Observatorio Ambiental De Bogotá. <https://oab.ambientebogota.gov.co/secretaria-de-ambiente-inicia-3-procesos-sancionatorios-por-afectacion-al-rio-tunjuelo/>
- Contreras-Pacheco, O. E., & Claassen, C. (2017). Fuzzy reporting as a way for a company to greenwash: Perspectives from the Colombian reality. *Problems and perspectives in management*, 15(2), 525–535. [https://doi.org/10.21511/ppm.15\(si\).2017.06](https://doi.org/10.21511/ppm.15(si).2017.06)
- Chung-Feng, J. K., Chang-Chiun, H., Sung-Han, Y., & Sung-Hua, C. (2021). Integration of multivariate control charts and the decision tree classifier to determine the faults of the quality characteristic(s) of a melt spinning machine used in polypropylene fiber manufacturing. Part II: The application of multivariate control charts and the decision tree classifier to determine the faults of quality characteristic(s). *Textile Research Journal*, 91(21–22), 2567–2580. Doi: 10.1177/00405175211011775.
- Christ, K. L., & Burritt, R. L. (2019). Implementation of sustainable development goals: The role for business academics. *Australian Journal of Management*, 44(4), 571–593. <https://doi.org/10.1177/0312896219870575>
- Departamento Administrativo Nacional de Estadística (2022a). Encuesta Ambiental Industrial - EAI – 2020. <https://microdatos.dane.gov.co/index.php/catalog/746>.
- Departamento Administrativo Nacional de Estadística (2022b). Encuesta Anual Manufacturera 2020 <https://microdatos.dane.gov.co/index.php/catalog/724>.
- Datta, S., & Goyal, S. (2022). Determinants of SDG Reporting by Businesses: A Literature Analysis and Conceptual Model. *Vision: The Journal of Business. Perspective*, 26(4). <https://doi.org/10.1177/09722629221096047>
- De Freitas Netto, S. V., Sobral, M. F., Ribeiro, A. R., & Soares, G. R. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, 32(1), 19. <https://doi.org/10.1186/s12302-020-0300-3>
- Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California management review*, 54(1), 64–87. <https://doi.org/10.1525/cmr.2011.54.1.64>
- Dorfleitner, G., & Utz, S. (2023). Green, green, it's green they say: A conceptual framework for measuring greenwashing on firm level. *Review of Managerial Science*, 1–24. <https://doi.org/10.1007/s11846-023-00718-w>
- Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., & Rodewald, A. D. (2023). A double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, 14(9), 2435–2448. <https://doi.org/10.1111/2041-210X.14186>
- Fox, E. W., Ver Hoef, J. M., & Olsen, A. R. (2020). Comparing spatial regression to random forests for large environmental data sets. *PLoS one*, 15(3), e0229509.
- Gregory, R. P. (2021). When is greenwashing an easy fix? *Journal of Sustainable Finance & Investment*, 1–24. <https://doi.org/10.1080/20430795.2021.1907091>
- Ghitti, M., Gianfrate, G., & Palma, L. (2023). The agency of greenwashing. *Journal of Management and Governance*. <https://doi.org/10.1007/s10997-023-09683-8>
- He, Q., Wang, Z., Wang, G., Zuo, J., Wu, G., & Liu, B. (2020). To be green or not to be green: How environmental regulations shape contractors' greenwashing behaviors in construction projects. *Sustainable Cities and Society*, 63, Article 102462. <https://doi.org/10.1016/j.scs.2020.102462>
- Helm, S., & Little, V. (2022). Macromarketing Our Way to Zero-carbon Future. *Journal of Macromarketing*, 42(2), 262–266. <https://doi.org/10.1177/02761467221088254>
- Henao, C., Lis-Gutiérrez, J. P., Lis-Gutiérrez, M., & Ariza, J. (2023). Determinants of efficient water use and conservation in the Colombian manufacturing industry using machine learning. *Humanities and Social Sciences Communications (pre-print)*.
- Hu, S., Wang, M., Wu, M., & Wang, A. (2023). Voluntary environmental regulations, greenwashing and green innovation: Empirical study of China's ISO14001 certification. *Environmental Impact Assessment Review*, 102, Article 107224.
- ISO. (2015). *Introduction to ISO 14001:2015*. ISO <https://www.iso.org/publication/PUB100371.html>
- JASP Team. (2022). JASP (Version 0.16.3) [Computer software]. <https://jasp-stats.org>.
- Norouzi, N., Fani, M., & Talebi, S. (2022). Green tax as a path to greener economy: A game theory approach on energy and final goods in Iran. *Renewable and Sustainable Energy Reviews*, 156, Article 111968. <https://doi.org/10.1016/j.rser.2021.111968>
- Pizzetti, M., Gatti, L., & Seele, P. (2019). Firms Talk, Suppliers Walk: Analyzing the Locus of Greenwashing in the Blame Game and Introducing 'Vicarious Greenwashing'. *Journal of Business Ethics*, 170(1), 21–38. <https://doi.org/10.1007/s10551-019-04406-2>
- Polcarpo, M. C., Apaolaza, V., Hartmann, P., Paredes, M. R., & D'Souza, C. (2023). Social cynicism, greenwashing, and trust in green clothing brands. *International Journal of Consumer Studies*, 47(5), 1950–1961.
- Ramanujam, R., Zhu, F., Fink, K., Karrenbauer, V. D., Lorscheider, J., Benkert, P., Kingwell, E., Tremlett, H., Hillert, J., & Manouchehrinia, A. (2020). Accurate classification of secondary progression in multiple sclerosis using a decision tree. *Multiple Sclerosis Journal*, 27(8), 1240–1249. <https://doi.org/10.1177/1352458520975323>
- Szabo, S., & Webster, J. (2021). Perceived greenwashing: The effects of green marketing on environmental and product perceptions. *Journal of Business Ethics*, 171(4), 719–739. <https://doi.org/10.1007/s10551-020-04461-0>
- Salkind, N. J. (2007). Classification and regression tree. In *Encyclopedia of measurement and statistics* (Vol. 1, pp. 144–145). SAGE Publications, Inc., Doi: 10.4135/9781412952644.n83.
- Rashidi-Sabet, S., & Madhavaram, S. (2021). A Strategic Marketing Framework for Emerging Out of the Climate Change Social Trap: The Case of Fashion Industry. *Journal of Macromarketing*, 42(2), 267–291. <https://doi.org/10.1177/02761467211058083>
- Santos, C., Coelho, A., & Marques, A. (2023). A systematic literature review on greenwashing and its relationship to stakeholders: State of art and future research agenda. *Management Review Quarterly*, 1–25. <https://doi.org/10.1007/s11301-023-00337-5>
- Secretaría Distrital de Ambiente. (2016, September 29). *Este año 21 empresas han sido sancionadas por contaminar ríos y humedales*. Observatorio Ambiental De Bogotá. <https://oab.ambientebogota.gov.co/este-ano-21-empresas-han-sido-sancionadas-por-contaminar-rios-y-humedales/>
- Secretaría Distrital de Ambiente. (2017, July 19). *20 empresas sancionadas por no cumplir normas Ambientales*. Observatorio Ambiental De Bogotá. <https://oab.ambientebogota.gov.co/20-empresas-sancionadas-por-no-cumplir-normas-ambientales/>
- StataCorp.. (2023). *Stata Statistical Software: Release 18 (Version 18)* [Computer software]. College Station, TX: StataCorp LLC.. <https://www.stata.com>.
- StataCorp.. (2023). *Lasso Intro – Introduction to lasso*. Stata Manuals.
- Testa, F., Boiral, O., & Iraldo, F. (2018). Internalization of environmental practices and institutional complexity: Can stakeholder pressures produce greenwashing? *Journal of Business Ethics*, 147(2), 287–307. <https://doi.org/10.1007/s10551-015-2960-2>
- Testa, F., Miroshnychenko, I., Barontini, R., & Frey, M. (2018). Does it pay to be a greenwasher or a brownwasher? *Business Strategy and the Environment*, 27(7), 1104–1116.
- Torelli, R., Balluchi, F., & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. *Business strategy and the Environment*, 29(2), 407–421. <https://doi.org/10.1002/bse.2373>
- Yang, Z., Nguyen, T. T. H., Nguyen, H. N., Nguyen, T. T. N., & Cao, T. T. (2020). Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review. *Journal of Business Economics and Management*, 21(5), 1486–1507. <https://doi.org/10.3846/jbem.2020.13225>
- Yildirim, S. (2023). Greenwashing: A rapid escape from sustainability or a slow transition? *LBS Journal of Management & Research*, 21(1), 53–63. <https://doi.org/10.1108/LBSJMR-11-2022-0077>
- Waniez, P. (2024). Philcarto [software]. Available at <http://philcarto.free.fr>.
- World Economic Forum. (2022). Executive Summary. *The Global Risk Report 2022 World Economic Forum*, 17(1), 7–10. <https://bit.ly/3A8iPJ9>.
- Wu, Y., Zhang, K., & Xie, J. (2020). Bad greenwashing, good greenwashing: Corporate social responsibility and information transparency. *Management Science*, 66(7), 3095–3112. <https://doi.org/10.1287/mnsc.2019.3340>
- Zhang, D. (2023). Talk the Talk, But Walk the Walk? Female-Managed Firms and Greenwashing Behavior. *Environment and Behavior*, 55(5), 364–394.
- Zhu, Q., Sarkis, J., & Lai, K. (2018). Green supply chain management innovation diffusion and its relationship to organizational improvement: An ecological modernization perspective. *Journal of Cleaner Production*, 170, 1640–1651. <https://doi.org/10.1016/j.jclepro.2017.10.262>