

Article

Logistics Performance and the Three Pillars of ESG: A Detailed Causal and Predictive Investigation

Nicola Magaletti ¹, Valeria Notarnicola ¹, Mauro Di Molfetta ¹, Stefano Mariani ¹ and Angelo Leogrande ^{1,2,*}

¹ LUM Enterprise S.r.l., 70010 Casamassima, Italy; magaletti@lumenterprise.it (N.M.); notarnicola@lumenterprise.it (V.N.); dimolfetta@lumenterprise.it (M.D.M.); s.mariani@lumenterprise.it (S.M.)

² Dipartimento di Management, Finanza e Tecnologia, LUM University Giuseppe Degennaro, 70010 Casamassima, Italy

* Correspondence: leogrande.culture@lum.it

Abstract

This study investigates the complex relationship between the performance of logistics and Environmental, Social, and Governance (ESG) performance, drawing upon the multi-methodological framework of combining econometrics with state-of-the-art machine learning approaches. Employing Instrumental Variable (IV) Panel data regressions, viz., 2SLS and G2SLS, with data from a balanced panel of 163 countries covering the period from 2007 to 2023, the research thoroughly investigates how the performance of the Logistics Performance Index (LPI) is correlated with a variety of ESG indicators. To enrich the analysis, machine learning models—models based upon regression, viz., Random Forest, k-Nearest Neighbors, Support Vector Machines, Boosting Regression, Decision Tree Regression, and Linear Regressions, and clustering, viz., Density-Based, Neighborhood-Based, and Hierarchical clustering, Fuzzy c-Means, Model-Based, and Random Forest—were applied to uncover unknown structures and predict the behavior of LPI. Empirical evidence suggests that higher improvements in the performance of logistics are systematically correlated with nascent developments in all three dimensions of the environment (E), social (S), and governance (G). The evidence from econometrics suggests that higher LPI goes with environmental trade-offs such as higher emissions of greenhouse gases but cleaner air and usage of resources. On the S dimension, better performance in terms of logistics is correlated with better education performance and reducing child labor, but also demonstrates potential problems such as social imbalances. For G, better governance of logistics goes with better governance, voice and public participation, science productivity, and rule of law. Through both regression and cluster methods, each of the respective parts of ESG were analyzed in isolation, allowing us to study in-depth how the infrastructure of logistics is interacting with sustainability research goals. Overall, the study emphasizes that while modernization is facilitated by the performance of the infrastructure of logistics, this must go hand in hand with policy intervention to make it socially inclusive, environmentally friendly, and institutionally robust.

Academic Editor: Giada La Scalia

Received: 13 November 2025

Revised: 4 December 2025

Accepted: 15 December 2025

Published: 18 December 2025

Citation: Magaletti, N.; Notarnicola, V.; Di Molfetta, M.; Mariani, S.; Leogrande, A. Logistics Performance and the Three Pillars of ESG: A Detailed Causal and Predictive Investigation. *Sustainability* **2025**, *17*, 11370. <https://doi.org/10.3390/su172411370>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: logistics performance index (LPI); environmental social and governance (ESG) indicators; panel data analysis; instrumental variables (IV) approach; sustainable economic development

JEL Classification: C33; F14; O18; Q56; M14

1. Introduction

In the globalized world of today, logistics systems' productivity and resilience are essential drivers of competitiveness at the national level as well as of economic development and sustainability. The empirical organization of supply chains, developments in technology and global trade intensification have brought the performance of logistics to the forefront of both economic policy and corporate decision-making. In parallel to these developments has been the rise of the Environmental, Social, and Governance (ESG) paradigm as the leading framework used to evaluate sustainable economic performance, transcending conventional financial measurements to consider broader societal and environmental consequences [1,2].

Justification for the study. Despite the increasing importance of both the logistics performance perspective and the environmental, social, and governance (ESG) framework in the design of sustainable economic systems, the intersection between the two areas remains largely [3,4]. Currently, existing literature focuses predominantly on logistics performance as economic infrastructure, with most economic studies being carried out on a firm level, whereas, conversely, most existing studies on ESG frameworks focus mostly on a firm level paradigm, thereby largely ignoring their systemic dynamics within the country's economic infrastructure. Currently, this systemic divide also leaves a huge knowledge gap, particularly with increasing recognition being accorded to the roles of logistics systems, as either facilitators or barriers, within environmental sustainability, as well as social well-being, and within governance dynamics [3]. Knowledge within this intersection is also highly required, particularly since logistics infrastructure designs significantly impact issues such as energy consumption, carbon emissions, resource use efficiency, working conditions, inclusive supply chains, and transparency in governance [4]. In addition, most global commitments, such as the United Nations Sustainable Development Goals, significantly depend on the assurance of sustainable logistics systems, so empirical studies within this field, particularly within the intersection of systemic dynamics within logistics infrastructure design, within ESG frameworks within different countries, remain largely uncharted.

In the midst of these twin evolutions, a recurring and relatively unexamined question sits at its core:

- How do the interactions between the quality of logistics performance and each of the ESG pillars vary by country?

In contrast to the expanding real-world applicability of both ESG and logistics globally, academic work connecting the two is relatively rare. Most research on the Logistics Performance Index (LPI) targets economic metrics like trade levels, industrial competitiveness, and infrastructure quality [4], whereas ESG scholarship is typically centered around firm-level sustainability, ethical investment practices, and policy at a high level [5]. Consequently, our knowledge base is missing a systematic exploration of how logistics capabilities impact environmental sustainability, social fairness, and governance quality at the country level. That is a stark deficiency, given how essential sustainable logistics has become to attainment of the United Nations Sustainable Development Goals (SDGs) [1]. This study has as its objective bridging that gap through a data-driven examination of how disaggregated ESG indicator variables correlate with logistics performance. In contrast to research using composite ESG indices, however, the research takes a disaggregated framework and looks at how infrastructure and efficiency in operations independently impact environmental (E), social (S), and governance (G) dimensions [2]. The research question is simple but fundamental:

- Does better logistics performance systematically have a positive impact on ESG results—and if so by which mechanisms?

This research advances the frontier of merging sustainable development and logistics research by offering practical lessons for both governments and MNCs. The key strength of this research is that it employs multiple methods, combining econometric analysis with ML. Endogeneity in this research has been controlled using instrumental-variable panel-data regression techniques, namely 2SLS and G2SLS, on a balanced panel of 163 countries from 2007 to 2023. This tackles endogeneity by improving the accuracy of results by mitigating the challenges posed by unobserved variables and reverse causality. To complement the econometric model, this research applies both supervised machine learning algorithms (Random Forests, k-Nearest Neighbors, Support Vector Machines) and unsupervised clustering algorithms (Density-Based, Fuzzy C-Means, Hierarchical, Model-Based, Neighborhood Clusters). This dual modeling approach not only provides robustness testing rigor but also identifies nonlinear behaviors and hidden patterns that are not accounted for or apparent in traditional statistical modeling. The ever-deepening integration of ML applications in the sustainability literature has led to greater predictive precision and the detection of patterns in high-dimensional data [6,7]. The dual-methodology approach thus strengthens internal validity and enhances generalizability, in line with prevailing research perspectives that intertwine ML and econometric analysis in the realm of ESG studies [8]. An important aspect of the analysis involves using ESG factors decomposed across the environmental, social, and governance pillars, rather than considering overall ESG scores, to identify the individual relationships of each factor with the Logistics Performance Index (LPI). Environmental factors are measured based on pressures exerted by emissions, air pollution, and land use; social factors are measured based on factors such as education access, service delivery, income levels, and child labor; and governance factors are measured based on the rule of law, regulatory quality, and innovation. This type of analysis has not been carried out in the literature with the depth and rigor reported here [6,7]. The conclusive evidence indicates that the ESG and logistics nexus involves various aspects of LPI improvements, including significant positive outcomes and challenges in environmental sustainability, tangible social impacts, applicable risks, and significant improvements in governance. However, effective logistics may exacerbate environmental or social inequalities in the absence of strengthened regulatory protection. Descriptive observations from all the collective data confirm the imperative of congruent policies that align with the evolution of logistics and the universal principles of ESG goals.

Study purpose. This research aims to develop a paradigm that establishes empirical links between logistics performance metrics and the Environmental, Social, and Governance (ESG) aspects of Sustainable Development. Recent studies indicate that national logistics performance closely correlates with both social and environmental aspects [3]. Although the importance of the world's logistics capabilities as basic determinants of economic competitiveness, trade efficiency, and development is well-appreciated, their relationship with Sustainable Development remains uncharted. There still remains a chasm in knowledge regarding the independent and cumulative roles of different logistics capabilities, particularly with regard to their effects on ESG metrics, given recent debates stratified by their effects on Sustainable Development. Studies on the effects of logistics performance on sustainability suggest that the G20 nations' sustainability levels remain largely driven by their logistics performance, underscoring the importance of joint policy formulation [9]. This assumption's importance lies in its effort to bridge this chasm by exploring determinants of ESG, along with the interaction between logistical efficiency and ESG dimensions, as defined in the existing literature. Empirical studies suggest that ESG practices in logistics, environmental compliance, social responsibility, and corporate governance exert independent as well as cumulative effects at the firm and macroeconomic levels [4]. Similarly, improvements in ESG capabilities,

particularly as applicable to Small, Medium, and Large Enterprises, remain identified as an effective approach for sustainable and human-oriented practices [2]. However, the use of digital technologies, along with Industry 4.0 technologies, in logistics remains effective in integrating urban and corporate logistics strategies with ESG dimensions [10]. Therefore, within this background, this proposal's subsequent analysis shall conduct a rigorous, sequential, bi-variate analysis of the Logistics Performance Index's variables, prepared through thorough verification against detailed ESG variables, within the realm of its governance units spanning 2007–2023, totaling 163 governance units. Specifically, the analysis will focus on: (1) the correspondence between improved logistics capabilities and environmental stress, as opposed to environmental efficiency (e.g., customs clearance and lead-time reliability); (2) the social determinants as antecedents of logistics capabilities, shedding light on education, working conditions, demographics, and accessibility issues (e.g., ease of arranging shipments and customer satisfaction); (3) the governance quality, with a focus on its constitutive aspects: enabling institutions, regulatory systems, scientific productivity, as well as enabling governance structures (e.g., tracking and tracing capabilities and supply chain transparency). Using instrumental variables panel regression analyses and sophisticated machine learning models, this analysis will examine the two-way relationships between logistics capabilities and sustainable development.

Study Purpose and Research Themes. This proposed research aims to develop a theoretical framework that explains the relationship between Logistics Capabilities Metrics and Environmental, Social, and Governance (ESG) factors within Sustainable Development. The emerging literature shows a strong correlation between national logistics capabilities metrics and social or environmental issues [3]. Although the correlation between global logistics capabilities metrics and Sustainable Development remains uncharted, the general implication is that logistics capabilities significantly affect economic competitiveness, trade, and Sustainable Development. However, there seems to be limited insight into the independent or cumulative aspects of such logistics capabilities, particularly as a Sustainable Development factor. With regard to the literature established in the precedent of existing literature on logistics capabilities metrics/sustainability, more recent literature asserts that the Sustainability path within G20 nations is largely dependent on their capabilities, thus establishing that policy as collectively decisive [9]. By implication, that this literature fills a much-needed aspect, within existing Sustainable literature, that might explore this correlation between determinants of ESG factors, through measures of Logistics Efficiency established within pertinent literature, as emergent empirical notions demonstrate that ESG Logistics, defined through necessary environmental, social, or governance protocols, remain collectively independent, with negative, positive, notions within the macroeconomic paradigm [4]. Improving ESG Capabilities, particularly within the realm of small-to-large-scale businesses, demonstrates a push towards more Sustainable, People-centric paradigms [2]. Simultaneously, technology such as Industry 4.0 remains effective within the strategic paradigm of transforming urban, corporate, or logistics infrastructure, as defined through the resultant ESG paradigm [10]. With this background, the proposed analysis will conduct a meticulous, sequential, multivariate analysis of Logistics Performance Index variables, cross-checked against more refined ESG variables, focusing on a dataset comprising 163 governance units for the years 2007–2023. To increase the validity of the results, posterior predictive checks and robustness analyses will be used. This will allow for issues of model specification parsimony and sensitivity issues that might qualify the conclusions. More specifically, this analysis will focus on the interaction between improved logistics capabilities under environmental stress and environmental efficiency, as captured by variables such as customs simplicity, customs clearance, and lead time

reliability. Social-influence variables, such as education, working conditions, demographics, accessibility, ease of arranging shipments, and customer satisfaction, will be considered antecedents of logistics capabilities. More specifically, the knowledge generated by this analysis of education and working conditions might serve as a basis for formulating personnel management policies through train-and-develop programs or by establishing benchmark standards for laboratory practices, thus further reinforcing the social component of ESG. Governance quality will be analyzed through its constituent parts, including governance structures, governance frameworks, scientific productivity, and facilitative governance frameworks, which comprise tracking and tracing capabilities and transparency. By combining instrumental-variable panel regression with more sophisticated machine-learning analytics, this analysis will explore the two-way interaction between logistics capabilities and sustainable development.

Study Hypotheses. With the aforementioned research questions as the background, this study formulates three inclusive hypotheses that provide direction for the analysis, thereby aligning the conceptual framework with the methodology. These hypotheses assume that the correlation between logistics performance metrics and sustainability outcomes is complex, interacting with environmental, social, and governance factors within the ESG framework [11].

H1. *Logistics performance shows a systematic relationship with mixed environmental effects, reflecting trade-offs between development and the environment. This hypothesis argues that improvements in logistics infrastructure can minimize resource use and some types of pollutants, but simultaneously increase other pollutants, such as GHG emissions. The existing literature suggests that ESG innovations focused on logistics and transportation can improve environmental efficiency while addressing new environmental pressures, such as increased energy use and GHG emissions [1,12]. Using disaggregated measures of environmental effects, such as air, GHG emissions, heat stress, and land use, this research will examine the impact of environmental stresses and efficiencies as forces behind changes in the Logistics Performance Index [11].*

H2. *Socio-economic variables significantly and diversely affect logistics performance. This hypothesis assesses the influence of education, basic service accessibility, demographics, working conditions, and income distribution on logistics performance. Evidence confirms that socio-economic variables, such as employee education, fair working conditions, and service accessibility, affect logistics efficiency [1]. Social determinants, such as education, access to basic services, demographics, working conditions, and income distribution, create inequality in human capital, working conditions, or both, affecting the efficiency of global logistics.*

H3. *Improving governance quality promotes a positive outcome on logistics performance. This hypothesis assumes that a high-quality institution, with attributes of proper regulation, the rule of law, efficient administration, and scientific strength, fosters a supportive environment that facilitates the establishment of a sound, modern, and trustworthy logistics infrastructure. Empirical evidence shows that sound governance principles or regulations can enhance ESG practices and sustainable development across nations [12,13]. These hypotheses collectively form the focal point of this analysis, through which the rest of this report will explore the relationships that exist between logistical performance and the environmental, social, and governance aspects of sustainable development.*

The research is organized as follows. Section 2 reviews the existing literature, identifying the main conceptual frameworks and empirical findings to date. Section 3 presents the data sources, sample characteristics, and the econometric and machine learning methodologies employed. Sections 4–6 are dedicated, respectively, to the analysis of the relationships between LPI and the Environmental, Social, and Governance

components, detailing both the regression-based and clustering-based results. Section 7 concludes with a discussion of policy implications, limitations, and directions for future research. Furthermore, Appendix A presents the hyperparameter settings of the regression algorithms, Appendix B presents the hyperparameter settings of the clustering algorithms, Appendix C presents the summary statistics of the environmental (E) indicators, Appendix D presents the summary statistics of the social (S) indicators, and Appendix E presents the summary statistics of the governance (G) indicators.

2. Literature Review

The existing literature presents informed but incomplete insights into the interrelation between ESG outcomes and logistic performance tending to lack the level of systemic integration and granularity desired by this study. The research by [4,5] has as its main objective assessing the financial impact of adopting ESG in the case of logistic firms but does not reveal its investigation to wider systemic interactions unfolding from country-wide metrics such as the Logistics Performance Index (LPI). While suggesting that the impact of ESG schemes is mediated by logistic performance and economic results, ref. [14] does fail to differentiate the ESG pillars and does not treat direct causality, a concern treated by this research. The issue of ESG challenges and opportunities in the post-COVID-19 context is broached by [2,15], albeit in a way failing to integrate results systematically to transportation efficiency metrics such as the LPI. In the same spirit, research by [1,16] analyzes ESG's impact on competitiveness and on stock performance but falls short of considering logistic infrastructure as country-wide driver of sustainability. Refs. [10,17] deal with smart and digitalized logistic as ESG enablers and participate in thematic add-ons short of adopting serious quantitative research practices like in the research presented here. The effect on firm performance of green logistic action is demonstrated by [18–20], the latter focused on the dimension of ESG transparency but both are subject to micro perspectives. The use of technology is analyzed by [21,22], and [23] but short of structural embedding of country-wide logistic performance in ESG effect. The research by [24] generalizes ESG discourse to maritime and seaport logistic industries but fails to systematically analyze environmental, social, and governance dimensions separately vis-à-vis the LPI as it does in this study.

Refs. [25,26] acknowledge transport and logistic firms to be influenced by ESG but reduce ESG to aggregate scores and fail to identify pillar-specific effects as identified here. Refs. [27,28] discuss communication and perception dimensions of ESG in the logistic sector but fail to attain econometric robustness. Refs. [29,30] discuss impact of ESG on supply chains but by a generalized application by qualitative methods and non-dynamic panel data methods or by using machine learning algorithms. Refs. [31,32] include governance variables like board diversity but fail to capture how the impact of logistic infrastructure performance on ESG is systematically captured. Refs. [33–35], and associate ESG and operation efficiency and productivity in the supply chain but to firm-specific or industry-specific studies and to system levels in countries by using LPI. Ref. [36] associate climate policy uncertainty and logistic stock returns and ESG scores but fail to include pillar disaggregation. Refs. [37,38] document sustainable optimization of the logistic industry but fail to document how optimization practices are associated with larger ESG systems in countries. Ref. [39] calculate competitiveness on efficiency of the logistic sector but their work does not systematically rule out environmental and social spillovers identified here. Ref. [40] discuss digitization and benefits to ESG and [41] discuss sustainable infrastructure but both fail to utilize instrumental variable panel data methods or machine learning regressions.

Research by [42,43] focuses on sustainability and governance in logistics companies but lacks generalizability at a country level. Refs. [44,45] design ESG assessment models

but work primarily at conceptual or firm levels and lack the cross-country and long-dimensioned data included in this study. Refs. [46,47] connect ESG to credit risk at the firm level but do not conceptualize the firm as a fundamental unit of analysis as they do so. Refs. [48,49] acknowledge the role supply chain digitalization plays in improving ESG but do not systematically tie it to LPI measurements. Refs. [50,51] emphasize the predictive ability of sustainability initiatives and ESG outcomes but fail to discuss drivers exclusive to the logistics sector at the country level. Refs. [52,53] equate ESG with efficiency at the terminals and ports and get close to LPI issues but keep to a sectorial scope. Refs. [54,55] discuss procurement benefits and circular economy models but fail to consider logistics performance as a systemic driver. Together, this study is the first to combine both econometric and machine learning approaches to reveal LPI to be a first-order determinant of ESG outcomes and not a secondary measure and to do so across countries, filling gaps in existing research.

3. Data and Methodology

One of the main methodological difficulties faced in the current research stems from the non-existence of a continuous historical time series of the Logistics Performance Index (LPI). The available LPI data intermittently between the period of 2007–2023 pose a number of missing values by country and year and thereby complicate the creation of a full and balanced panel dataset adequate to perform rigorous econometric and machine learning analysis. In a bid to overcome this problem and maintain the consistency and integrity of the data's longitudinal form, a polynomial-regression-based interpolation scheme was utilized. Polynomial fitting was used to fill in missing values on a country-wise basis to rebuild realistic historical traces of the LPI values and avoid risks of injecting spurious biases using simpler linear interpolation methods. The methodology is informed by existing research suggesting the benefits of using imputation as well as advanced interpolation methods in LPI research ranging from genetic algorithm-based weights to imputation methods using regression [56]. The second core analytic decision concerns ESG disaggregation. In contrast to keeping ESG as a combined or aggregate indicator, the research systematically breaks up the model into its three pillars—Environmental (E), Social (S), and Governance (G)—and studies the interrelation of LPI across each of these dimensions in turn. The pillar-wise design allows a finer and more detailed understanding of how the interactions between logistics performance and sustainability outcomes unfold than has been the case with prior research which tended to work with ESG as a uniform block. The research design is aligned with contemporary research underlining the different and diverging influence of a particular ESG dimension on firm and sector performance [4,57]. In keeping with the research question's adverseness to simplicity, the analytic design follows both conventional econometric and sophisticated ML approaches. The econometric analysis was conducted by using Instrumental Variables (IV) panel regressions comprising both Two-Stage Least Squares (2SLS) and Generalized Two-Stage Least Squares (G2SLS) models to rigorously contend with endogeneity issues and ascertain causal interpretation of the estimated coefficients. Complementarily to the above, machine learning methodologies were implemented in both the regression and clustering tasks—utilizing Random Forest, k-Nearest Neighbors, Support Vector Machines, Decision Tree Regression, Boosting Regression, and Lasso in the case of the former and Density-Based Clustering, Fuzzy c-Means, Model-Based Clustering, Neighborhood Clustering, Random Forest Clustering, and Hierarchical Clustering in the case of the latter. The interplay between the econometric and machine learning models facilitates both the verification of outcomes by means of different methodological perspectives and the determination of nonlinear and latent patterns likely to pass under the radar of conventional regression analysis. These combined methodological options

respond to the requirements of data constraints but also intensify the robustness, exhaustiveness, and novelty of the research's empirical contribution to the extant literature on the topic of logistic performance and sustainable development (Figure 1).

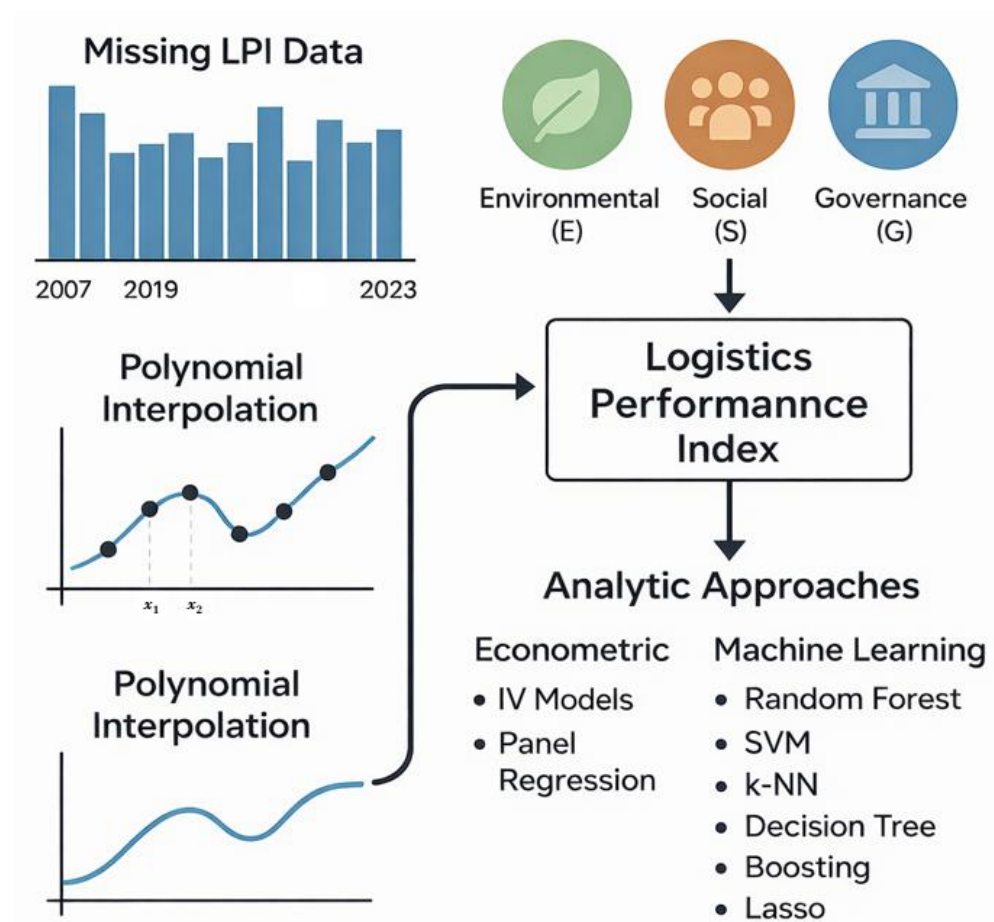


Figure 1. Overview of the Research Design and Analytical Framework. Note: the diagram illustrates the full research workflow, including the reconstruction of missing LPI values through polynomial interpolation, the integration of Environmental, Social, and Governance indicators, and the application of both econometric and machine learning approaches for the analysis.

Study model. In this case, the proposed research will use a multi-method design to investigate the relationship between logistics performance and the different dimensions of Environmental, Social, and Governance (ESG). In this design, the Logistics Performance Index (LPI) will be the dependent variable, with the environmental, social, and governance dimensions as determinants. Such designs have been used previously in other studies that investigated the effects of different dimensions of ESG issues on the quality of institutions and the economic aspects of different countries [58]. To specify the nature of this link, the analysis resorts to Instrumental Variable (IV) panel fixed-effect regression models, namely Two-Stage Least Squares (2SLS) and generalized (2SLS). These econometric models address endogeneity, missing variables, and reverse causality. Finally, the models assess the specific drivers of environmental, social, and governance variables on logistics performance across 163 countries spanning 2007–2023. Previous studies have shown that IV models, as well as panel regression models, may effectively examine the links between logistics performance, innovation, and environmental issues [59,60]. Apart from this established framework of causality, the study uses more complex machine learning models to examine nonlinear correlations, thereby improving predictability and the ability to identify hidden dynamics across nations. Regression

models (Random Forest, Support Vector Machines, k-Nearest Neighbors, Decision Trees, Boosting, Lasso, or Elastic Net) will be used for the analysis of accuracy, while the application of clustering models (DBSCAN, Fuzzy C-Means, Hierarchical, Model-based, Neighborhood-based, or Random Forest) will identify structural variations, more specifically within models linked with different nations. Applying such models aligns with recent improvements in predictive analytics, where machine learning algorithms were rigorously tested for feature selection and model accuracy assessment in logistics models. By embracing this convergence, analyses that treat ESG variables as discrete will seek to identify their net impact on the logistics industry as a whole, delivering valuable insights into any correlations within the realm of sustainability studies. These analyses will also position the logistics industry as a key environmental agent, demonstrating the positive impact that increased adoption of best practices can have on minimizing resource use. Case analyses will also demonstrate the use of freight consolidation, routing, and other solutions that address industry improvement as a tool for environmental remediation, as defined by [59] and subsequent studies such as [60].

Study analysis. This empirical analysis combines econometric identification with predictions derived from machine learning models, focusing on the impact of environmental, social, and governance (ESG) factors on the logistics performance of 163 different nations from 2007 through 2023. Similar studies combining multiple models into a single methodology were recently used in tandem with other studies involving artificial intelligence models to achieve more realistic results through the intersection of economic models with AI models [61]. Using instrumental variables (IV) panel regression analysis, the empirical results show a twofold, double-edged, but mostly negative implication of environmental variables for every Logistics Performance Index (LPI). More specifically, greenhouse gas (GHG) emissions, agricultural value added, air pollutants, and extensive agricultural use are positively or negatively associated with logistics efficiency. This illustrates that environmental and governance variables often have both positive and negative implications for economic performance, depending on the relevant environmental conditions and economic factors [62]. Specifically, variables such as water accessibility, sanitation facilities, aging, education, and increasing elementary education enrollment rates reflect modest negative adjustments, whereas child labor reflects higher LPI levels. However, income inequality has strong negative effects on logistics activity, suggesting that stronger social development, with reduced income inequality, facilitates efficient value chain management. By contrast, predictions from machine learning models indicate that IV estimates of environmental stress, education, and demographics remain applicable, valid, and accurate. Integration between the two models results in more robust models with enhanced predictability, a methodological improvement supported by previous studies on machine learning-based predictions of logistics performance. Clustering analyses identify distinct elements within each country, defined by attributes such as air pollutants, extreme temperatures, and agricultural intensity, and determine distinct loci for each country. Thus, the empirical results for this topic show that multifaceted logistics sustainability prevails, implying that the well-balanced evolution of logistics must address, alongside environmental enhancement, improved social conditions and proper governance [62].

Limitations. Across all analyses, the dataset includes 163 countries. Although such a large cross-country dataset helps derive corresponding cross-country correlations with relative ease, this inevitably affects the level of granularity available to inspect individual national settings. To address this tension, the necessity of accounting for national cross-country variability is carefully explained in this manuscript through complementary cross-country analysis strategies, including cluster and machine-learning algorithms.

4. Environmental Sustainability and Logistics Efficiency: A Multi-Method Analysis Using IV Regressions, Predictive Algorithms, and Clustering

This section examines the interplay between the Environmental (E) component of the ESG framework and the Logistics Performance Index (LPI) using a two-methodological framework involving Instrumental Variable (IV) panel models and machine learning (ML) models. IV models eliminate issues of endogeneity and enable causal inference of how environmental indicators such as PM2.5, nitrous oxide emissions, heat exposure levels, and agricultural land cover are determinative of logistics performance. This framework is a following of [63], in which they emphasize controlling for environmental-economic interactions when measuring LPI, and particular emphasis on the dimensions of green innovation, renewable energy, and global integration. ML models—such as used by [64] in environmental hazard predictions—are applied to best achieve predictive power and to compare the relative effect of environmental variables. The clustering methods following [65], who used functional regression-based clustering of air pollution data, identify latent country profiles through shared environmental-logistics patterns and add richness to the ensuing analysis.

4.1. Causal Estimation of Environmental Determinants of Logistics Performance Within the ESG Framework

This section investigates the impact of environmental and land use variables on the Logistics Performance Index (LPI) across 163 countries from 2007 to 2023. Using fixed-effects two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) models, the analysis addresses endogeneity by employing a broad set of instrumental variables. Key factors examined include nitrous oxide emissions, PM2.5 pollution, extreme heat exposure, agricultural land share, and agricultural value added. The results reveal that environmental degradation and land use dynamics significantly influence logistics performance, underscoring the need to integrate environmental considerations into logistics development strategies aligned with ESG objectives.

Specifically we have estimated the following model:

- $X_{it} = Z_{it}\Pi + v_{it}$ (First Stage)
- $Y_{it} = X_{it}\beta + \mu_{it}$ (Second Stage)
- $Y_{it} = LPI_{it}$
- $X_{it} = \{\text{NOE, PM25AE, HI35, ALPA, AFFVA}\}$
- $Z_{it} = \{\text{ACFTC, PSMWS, PSMS, LEBT, FRT, PA65A, LRAT, SEP, GEET, CET, LFPRT, CODCDMPN, MRU5, HB, POA, ISL20, GI, PHRNPL, AAGRPCI, IUI, GDPG, PSHWNP, RFMLFPR, SLRI, STJA, RLE, NM}\}$
- $i = 163$
- $t = [2007; 2023]$

The results are indicated in the following Table 1.

Table 1. Environmental Stressors and Logistics Performance: An IV Panel Data Analysis.

Dependent Variable	LPI
Endogenous	NOE PM25AE HI35 ALPA AFFVA
Instruments	ACFTC PSMWS PSMS LEBT FRT PA65A LRAT SEP GEET CET LFPRT CODCDMPN MRU5 HB POA ISL20 GI PHRNPL AAGRPCI IUI GDPG PSHWNP RFMLFPR SLRI STJA RLE NM
Observation	using 2771 observations

Times	17					
Countries	163					
	Fixed-effects TSLS			G2SLS random effects		
Variable	Coefficient	Std. Error	z-Statistic	Coefficient	Std. Error	z-Statistic
const	4.360 **	1.806	2.414	4.303 **	1.814	2.372
NOE	0.003 ***	0.001	3.045	0.003 ***	0.001	3.049
PM25AE	−0.109 **	0.043	−2.528	−0.109 **	0.043	−2.510
HI35	0.008 ***	0.002	3.035	0.008 ***	0.002	3.022
ALPA	−0.005 **	0.002	−2.397	−0.005 **	0.002	−2.347
AFFVA	0.083 ***	0.021	3.825	0.083 ***	0.021	3.834
Statistics	SSR = 1880.24			SSR = 2790.55		
	sigma-hat = 0.849(df = 2603)			sigma-hat = 1.004 (df = 2765)		
	R-squared = corr(y, yhat) ² = 0.000			R-squared = corr(y, yhat) ² = 0.000		
	Included units = 163			Included units = 163		
	Time-series length: min = 17, max = 17			Time-series length: min = 17, max = 17		
	Wald chi-square(5) = 33.861 [0.0000]			Wald chi-square(5) = 34.012 [0.0000]		
	Null hypothesis: The groups have a common intercept			sigma-hat(within) = 0.849		
	Test statistic: F(162, 2603) = 14,978.8 [0.0000]			sigma-hat(between) = 25.712		

Note. *** indicates significance at the 1% level; ** indicates significance at the 5% level.

This research focuses on the factors underlying the Logistics Performance Index (LPI) across 163 countries over 17 years, using panel data comprising 2771 observations. The research focuses on the significance of environmental factors, including nitrous oxide emissions, exposure to air pollution (PM2.5), exposure to extreme heat, and land use factors such as the share of agricultural land and the value added by agriculture. The conceptual approach uses the framework recently developed in the context of the relationship between environmental factors and logistics networks [59]. The issue of endogeneity is addressed using fixed-effect two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) with a large set of instruments based on living standards, health status, demographics, governance factors, education factors, and overall economic factors. This research follows the literature by emphasizing the importance of accounting for variations in LPI determinants across different locations [66] and for endogenous variables in cross-sectional analyses in the context of logistics research [67]. For both specifications, the results are significant and robust. All five endogenous variables are significant determinants of LPI; however, the relatively small values suggest that logistics performance depends on various factors not covered in this research, such as environmental and land use factors. Nitrous oxide release emerges as a significant positive determinant of LPI, indicating that improvements in logistics infrastructure often accompany the increase in industry and pollution levels in those regions; this agrees with other evidence in support of the idea that the increase in logistics infrastructure often conflicts with the increase in pollution levels in regions [68,69]. The exposure level of the population to air pollution shows a negative association with LPI levels, in line with other research in this context, suggesting that increased levels of air pollution exert negative pressure on workers' productivity in industries and on transport infrastructure in regions [70]. The population exposure level in regions to extreme heat data (HI35) shows significant positive effects on LPI values. This indicates that technological adaptation in regions may accompany effective advanced infrastructure in the logistics sector in regions prone to high levels of temperature data. Agricultural land use shares exhibit negative significance in LPI values. This indicates that agrarian regions often possess ineffective infrastructure in the logistics sector, whereas data from overall value added in agriculture indicates significant positive significance in LPI values; this

approaches research done in this context through support of research in support of commercialized regions through the increase in colder chains infrastructure in regions [71–73]. The collective data indicate complex conflicts between the expansion of infrastructure in the logistics sector and regions under environmental pressure. Although enhanced logistics improvements are linked to diversification and the commercialization of agriculture, they may also lead to environmental deterioration. The results highlight the need for a link between logistics enhancement and environmental preservation.

Causality. The fixed-effects two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) applications allow a causally robust interpretation of the correlation between environmental variables and logistics performance. Leveraging a dense set of instrumental variables that influence environmental and land use patterns but plausibly exogenous to the domain of logistics performance, the analysis manages to evade common issues of endogeneity like omitted variable bias and reverse causality. The methodology is aligned with recent empirical work which has utilized the TSLS and G2SLS framework to separate causal effects in the presence of complicated interdependencies and confounders, particularly in studies of environmental and economic performance [74]. Similarly, in environmental quality and green logistics as well, ref. [75] demonstrated how two-stage estimation methods are influential in capturing delicate interactions between logistics performance and sustainability outcomes across a variety of economies. Consequently, the positive effects of nitrous oxide emissions and agricultural value added and the negative effects of PM2.5 air pollution and agricultural share of land are causal effects and pure associations. The research is thus more policy-relevant because it means environmental quality and land management directly impact a country's ability to perform logistics. The low R-squared values do however reveal that even though remarkable influence is exerted by these variables on logistics performance, they capture only a fraction of the complicated determinants driving it.

Impact of the results within the E-Environmental Component within the ESG model. Empirical evidence elucidates a two-side and multifaceted relationship between environmental consequences and the performance of logistics. While on the first side, improved LPI scores are typically associated with greenhouse emissions such as nitrous oxide evidencing the environmental impact of widespread transport, warehouse operation, and industrial production. This presents a time-tested trade-off in development-environment terms: more developed infrastructure of logistics produces a superior level of economic development but also accelerates environmental degradation if it is uncontrolled. More contemporary research has identified systems of logistics such as third-party and heavy goods-associated systems as prominent producers of emissions unless practices of sustainability are implemented [76]. Environmental degradation per se as well as air pollution (exposure to PM2.5) on the other hand negatively impinges on the efficiency and dependability of logistics. Pollution reduces productivity by labor, makes transport flows difficult and damages public health all of which impair the efficiency and dependability of logistics. Apart from environmental degradation per se, exposure to climate extremes such as hot days also underscores building climate-resilient systems of logistics. Adaptive practices such as green chains of supply, energy-efficient services and products as well as eco-friendly infrastructure are necessary to render logistics operations climate-resilient to climate risks. All of the above solutions are now increasingly implemented by models of logistics worldwide ranging from electric fleets and renewable sources to tracking emissions by blockchain in the supply chain [77]. The relationship between land use and logistics also confirms the role of the environment. Land economies with a high share of agricultural land have weaker performance of logistics while economies commercialized with sustainable land management are capable of developing stronger infrastructure of logistics. This is a part of a general transition towards a

sustainable phase change in the development of logistics whereby firms are increasingly viewing green logistics as a source of competitive power to avoid the costs of emissions and to enhance resilience as opposed to a constraint [78]. Overall, incorporating strong environmental concerns into planning logic of logistics is now a requirement and not a choice but necessary to become competitive in the long term. Aligning LPI developments to Environmental pillar of ESG requires proactive investment in green logistics, regulatory transformation and sustainable innovation to ensure development of logistics complements and does not compromise global environmental goals.

4.2. Environmental Determinants of Logistics Efficiency: Evidence from Machine Learning Analysis Under ESG Standards

This section explores the application of various machine learning regression algorithms to predict the Logistics Performance Index (LPI) based on environmental and land use variables. Models such as Boosting Regression, Decision Tree Regression, k-Nearest Neighbors, Linear Regression, Random Forest, Lasso, and Support Vector Machine (SVM) are compared using standard performance metrics including MSE, RMSE, MAE, MAPE, and R^2 . The analysis identifies Random Forest Regression as the most robust model, offering the best trade-off between accuracy and generalizability. Further, variable importance measures from Random Forest highlight the critical role of environmental factors in shaping logistics performance across countries and over time (Table 2).

Table 2. Comparative Performance of Machine Learning Models in Predicting Logistics Performance.

Statistics	Boosting Regression	Decision Tree Regression	k-Nearest Neighbors Regression	Linear Regression	Random Forest Regression	Lasso	Support Vector Machine
MSE	668.052	435.315	596.462	603.118	464.679	606.449	842.876
MSE (scaled)	1.333	1.03	0.955	1.472	0.922	1.452	1.556
RMSE	25.847	20.864	24.423	24.558	21.556	24.626	29.032
MAE/MAD	13.713	8.824	8.57	14.267	10.264	14.032	9.458
MAPE	229.26%	182.74%	150.97%	284.71%	181.05%	287.14%	24.52%
R^2	0.111	0.234	0.272	0.069	0.29	0.074	0.049

Comparing the results of different algorithms provides concrete evidence of the suitability of various predictive models for forecasting Logistics Performance Index (LPI) values. Among the various algorithms, such as Boosting Regression, Decision Tree Regression, k-Nearest Neighbors (k-NN), Linear Regression, Random Forest Regression, Lasso Regression, and Support Vector Machine (SVM), Random Forest Regression proves to be the most balanced or stable algorithm [79,80]. This algorithm not only provides the highest ' R^2 ' of 0.29 but also proves to be the most effective model at explaining the difference in values (0.29), although the overall explained variability accounts for only 29 per cent. This high ' R^2 ' value indicates greater adaptability to the nonlinearities prevalent in larger data samples that constitute the global research landscape in logistics [81]. Moreover, across various measures of prediction error, Random Forest Regression again proves more effective than other models. This algorithm yields values of 464.679 for mean squared error (MSE) and 21.556 for root mean squared error (RMSE), whereas Decision Tree Regression yields slightly lower values of 254.149 for MSE. However, this algorithm provides greater stability or resistance to overfitting and hence proves more effective in terms of generalizability or universal applicability [79]. Decision Tree Regression and k-NN Regression again provide slightly lower values in terms of mean absolute errors (MAE) but are otherwise hampered by inherent drawbacks of high sensitivity towards

noise and complexity arising out of depth in Decision Tree Regression [80], while k-NN Regression gets hampered by the hassles in scaling up and sparsity in data samples [81]. The data of SVM Regression again gets hampered by inconsistency in various parameters, such as high mean absolute percentage errors (MAPE) and lower values in terms of MSE and 'R²' measures due to appropriateness of choice of kernels in this context [81], although this again gets commonly exhibited in various other studies in comparison among various algorithms irrespective of parameters [81,82]. Linear Regression, Lasso Regression, and Boosting Regression again get completely hampered in terms of overall underperformances in various parameters due to domination by nonlinearities prevailing in the data samples in terms of interrelations of various environmental factors in constituting global governance parameters and various other economic factors in overall domains of the global arena in the realm of logistics [83].

Applying the Random Forest Regression we have the following results as showed in Table 3:

Table 3. Variable Importance Metrics for Predicting Logistics Performance.

Variables	Mean Decrease in Accuracy	Total Increase in Node Purity	Mean Dropout Loss
NOE	277.497	114.677	23.130
PM2.5AE	224.074	107.476	21.434
ALPA	294.265	98.796	23.223
HI35	237.642	77.966	21.182
AFFVA	16.990	30.634	17.586

Applying the Random Forest process to the specified dataset unveiled pertinent information on the relative importance of explanatory variables to predict the Logistics Performance Index (LPI). The three importance metrics of Mean Decrease in Accuracy, Total Increase in Node Purity, and Mean Dropout Loss all recognize a core group of predictors key to determining the performance of both countries and times. The evidence shows agricultural land (ALPA) with a maximum Mean Decrease in Accuracy of 294.265 and is hence the most predictive variable on prediction accuracy. Removing or permuting ALPA causes the most harm to the performance of the Random Forest model. ALPA also tops Total Increase in Node Purity at a value of 98,796,892. This indicates how ALPA makes decision nodes purer with each split in the forest and contributes to its key determination of distinguishing better and worse performing logistics (Figure 2).

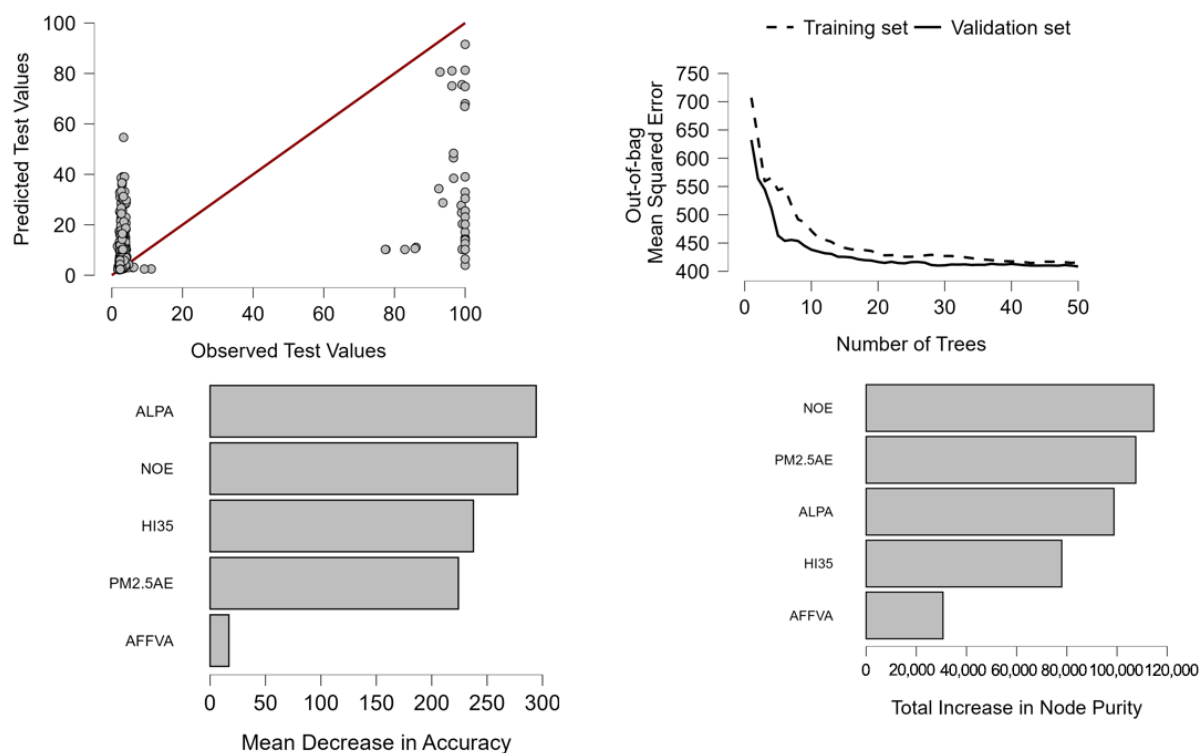


Figure 2. Random Forest Analysis of Environmental Drivers of Logistics Performance. The red diagonal line represents the 1:1 reference line (perfect agreement), where predicted values equal observed test values. Points lying on this line indicate ideal model performance, while deviations from the line reflect prediction errors and model bias.

The Mean Increase in Node Purity and Mean Decrease in Accuracy values reveal that Nitrous oxide emissions (NOE) and PM2.5 air pollution exposure (PM2.5AE) are major predictors of logistics performance. NOE's Mean Decrease in Accuracy of 277.497 and exceptionally high Mean Increase in Node Purity of 114,677,766 rank the variable the second most important feature after ALPA, while PM2.5AE's Mean Decrease in Accuracy of 224.074 and substantial Mean Increase in Node Purity also rank high. The results demonstrate that environmental degradation, as reflected in the release of greenhouse gases and air pollution, significantly influences logistics. This aligns with current trends in Random Forest analysis, in which nitrous oxide emerges as a significant predictor in environmental modeling [84,85]. The Heat Index above 35 °C (HI35) demonstrates significant predictive capacity in this regard, registering a Mean Increase in Node Purity of 77,966,120 and a Mean Decrease in Accuracy of 237.642. The rising significance of the Heat Index indicates that climate-related factors are exerting a growing—and profound—influence on the efficiency of logistics systems operating in extreme climates. Conversely, the “value added by agriculture, forestry, and fisheries” (AFFVA) variable shows little predictive power in this context. This variable's Mean Increase in Node Purity of 30,634,277 and Mean Increase in Node Purity compared unfavorably with other factors but favorably with Mean Dropout Loss values, which placed ALPA and NOE first and highest in losses, followed by high rankings from PM2.5AE and HI35. This indicates that sectoral contributions are less significant than environmental factors. Dropout losses support this interpretation, wherein both ALPA and NOE exhibit the highest losses, followed by PM2.5AE and HI35 in descending order [86]. The results of the Random Forest analysis indicate that environmental factors are the primary determinants of interpretations of logistics improvements, while sectoral contributions are insignificant in

this regard. This indicates that there must be synchronization in logistics improvements and environmental adaptation.

4.3. Identifying Country Profiles: A Cluster Analysis of LPI and Environmental Indicators

This section explores the clustering of countries based on environmental factors influencing the Logistics Performance Index (LPI) within the ESG framework. Using six different clustering algorithms—including Density-Based, Fuzzy C-Means, Hierarchical, Model-Based, Neighborhood, and Random Forest clustering—we assess model quality through key metrics such as Dunn Index, Silhouette score, Pearson’s gamma, and entropy. The goal is to identify homogeneous groups that reveal distinct patterns between environmental variables and logistics performance. Among the evaluated methods, Density-Based Clustering emerges as the most robust, offering well-separated, compact, and interpretable clusters that deepen understanding of the environmental dimension’s impact on LPI outcomes (Table 4).

Table 4. Comparative Evaluation of Clustering Algorithms for Environmental Impacts on Logistics Performance.

Metric	Density-Based	Fuzzy C-Means	Hierarchical	Model-Based	Neighborhood	Random Forest
Maximum diameter	0.508	0.778	0.000	0.763	0.243	0.763
Minimum separation	1.000	0.008	0.184	0.000	0.026	3.16×10^{-5}
Pearson’s γ	0.482	0.261	1.000	0.000	0.682	0.029
Dunn index	1.000	0.009	0.247	0.001	0.035	0.000
Entropy	0.000	0.709	0.266	0.940	0.941	0.695
Calinski-Harabasz index	0.099	0.060	0.161	0.000	1.000	0.000
R ²	0.000	0.280	0.547	0.241	1.000	0.207
AIC	1.000	0.615	0.000	0.455	0.000	0.509
BIC	1.000	0.594	0.000	0.432	0.000	0.493
Silhouette	0.476	0.128	0.537	0.063	0.414	0.000

The analysis of the results reveals that the Density-Based approach achieves the best structural quality, and that the best-balanced clustering sizes are obtained either with model-based or k-means clustering. The rank skill indicates that the Density-Based approach has the best structural quality, followed by hierarchical clustering. The Pearson correlation and Silhouette indices further support this result regarding structure quality. Although the maximum Diameter index values are not optimal, this does not strongly affect the overall quality of the approach. The Density-Based approach produces three clusters of sizes 2517, 238, and 8 with eight additional noise samples. This result indicates that although the clustering scheme works well overall, the resulting structure is unbalanced. For fuzzy c-means clustering methods, this structure balances relatively well compared with hierarchical clustering methods. The model-based and k-means clustering methods result in the most balanced structure. Based on the overall analysis above, the result again confirms that the Density-Based approach has the best structural quality compared to other methods. The k-means-based models are the most balanced and have the highest stability in terms of balanced clustering sizes. The overall analysis indicates that Density-Based clustering possesses the best structure quality compared to other approaches. The best-balanced clustering size structure can be obtained from either model-based or k-means clustering methods (Table 5).

Table 5. Comparison of Clustering Algorithms by Cluster Size Distribution and Structural Stability.

Clustering Algorithms	Size										
	Noisepoints	1	2	3	4	5	6	7	8	9	10
Density-Based	8	2517	8	238							
Fuzzy C-Means		206	131	215	159	29	587	519	218	636	71
Hierarchical		2319	90	61	18	11	4	8	216	22	22
Model-Based		694	550	238	287	255	266	271	210		
K-Means		111	290	88	217	468	60	978	222	158	179
Random Forest		1432	225	114	70	95	173	130	196	239	97

Based on the analysis, the Density-Based clustering approach shall be discounted due to the formation of highly polarized clusters, in which 90% of the data points are confined to a single cluster. This defeats the purpose of understanding the relationship between environmental variables and their impact on the Logistics Performance Indicator. The Model-Based clustering analysis gives outcomes that are helpful in understanding the inter-relationships between environmental (E) factors studied in this research work: nitrous oxide emissions, exposure to PM2.5 air pollutants, heat stress exposure, proportion of land under agricultural use, value addition in agriculture, and Logistics Performance Indicator in terms of Environmental, Social, and Governance factors. Clustering based on mean values would allow sectorial profiles to be generated based on intensity levels. With eight clusters in the analysis, LPI values are mostly aggregated around the normalized mean, and the largest difference would come from factors in environmentally challenging situations. This means that no single environmental factor, but multiple factors, are responsible for LPI values. For instance, in Cluster 1, there are moderate levels of emissions and PM2.5 concentrations combined with below-average LPI values, signifying circumstances in which environmental pressures may hamper logistics systems. Conversely, in Cluster 2, there are clean environmental circumstances combined with slightly negative LPI values, signifying that effective logistics are not necessarily possible in excellent environmental conditions. However, in Cluster 3, there are the most stressful environmental circumstances caused by high levels of PM2.5 exposure and hot conditions combined with the highest LPI values signifying developed regions in which effective logistics are possible in areas with dirty environmental circumstances. Meanwhile, in Clusters 4 through 8, there are minimal levels of pollution and negligible use of arable land, yet relatively equivalent LPI values, suggesting that slight variations in environmental conditions do not affect logistical capability. The Model-Based clusters summarize that, in fact, the article's conclusion regarding environmentally nonlinear LPI magnitudes holds true (Table 6).

Table 6. Standardized Mean Values of Environmental and Logistical Indicators Across the Eight Model-Based Clusters.

Cluster	LPI	NOE	PM2.5AE	HI35	ALPA	AFFVA
1	−0.045	0.621	0.304	−0.298	−0.141	−0.412
2	−0.011	−0.559	−0.646	−0.312	0.193	0.576
3	0.169	0.684	0.606	3.250	−0.033	−0.125
4	−0.002	−1.709	5.289×10^{-5}	−0.323	-1.735×10^{-7}	0.442
5	−0.002	-5.890×10^{-4}	5.289×10^{-5}	−0.303	-1.735×10^{-7}	6.433×10^{-4}
6	−0.002	−0.760	5.289×10^{-5}	−0.305	-1.735×10^{-7}	−0.119
7	−0.002	0.952	5.289×10^{-5}	−0.291	-1.735×10^{-7}	−0.483
8	−0.002	0.706	5.289×10^{-5}	−0.311	-1.735×10^{-7}	0.166

The shape of the mixing probabilities distribution sheds further light on the significance of each component in capturing the correlation between Logistics Performance Index (LPI) and the environmental (E) component of the ESG structure. The probabilities indicate the number of observations associated with each component and allow us to identify which forms of environmental-logistics correlation the majority of observations in the data are distributed across. Component 1 has a probability of 0.247 and is the most common correlation type. This corresponds to about one out of every five observations in the dataset and asserts that the environmentally related factors and LPIs pertaining to this component capture the most common correlation pattern in the overall relationship between environmental pressures and LPI. Component 2, with probability 0.202, also shows a relatively abundant data representation and indicates the presence of another common correlation pattern in the data, through which environmental factors affect LPI. The remaining components, with probabilities of 0.075–0.104, capture less abundant data patterns and highlight more specific correlations between environmental factors and LPIs in smaller country groups. However, the existence of all components underlines that the correlation pattern between LPI's and environmental factors is not homogenous in various contexts but depends in each particular case upon levels of release of polluting gases into the atmosphere from industry and transport, levels of air pollution caused after those releases by both gases and solid matter wastes from those releases through various meteorological factors like temperature and wetness levels, and structure parameters of agriculture in each country. The entire mixture probability distribution shows, in essence, that there are, in the overall correlation structure of the LPI-ESG-Environment model, not only the dominant but also other relatively rare patterns that are essential for covering the variability of this correlation phenomenon worldwide. This pattern further underscores the multidimensional nature of LPI-ESG-Environment correlation patterns, highlighting the overall interrelationships between environmental sustainability and LPIs (Table 7).

Table 7. Mixing Probabilities of the Eight Components in the LPI-ESG-Environment Mixture Model.

Components	Mixing Probability
Component 1	0.247
Component 2	0.202
Component 3	0.086
Component 4	0.104
Component 5	0.092
Component 6	0.096
Component 7	0.098
Component 8	0.075

The analysis of the standardized means for each component sheds further light on how different components of the environmental variables are linked to the values of the Logistics Performance Index (LPI). The most interesting result is Component 3, with the LPI well above the average. This particular component has moderate GHG emissions and PM_{2.5} exposure levels, high heat stress, and high weights for agricultural factors. This particular combination shows that quite high levels of logistical strength can coexist with strong levels of environmental pressure, further establishing the nonlinear relationship between environmental sustainability and logistical effectiveness. The remaining components are negative in terms of the LPI values. This shows that each component has underperformed in terms of average logistical capability. Component 1 shows relatively lower levels of GHG emissions and air pollution, along with reduced LPI levels and strong levels of agricultural factors. This further suggests that there may be restrictions on overall

logistical efficiency in the particular economy, due more to structural factors than to environmental pressures. Component 2 further provides evidence that the values resulting from environmental factors are not linear. This component shows relatively higher GHG emissions alongside relatively higher levels of both temperature and humidity. The next components, 4 through 8, appear negligible in terms of both PM2.5 levels and Heat Index 35 levels, in combination with considerable levels of agricultural land use. This further indicates overall particular levels of logistical capacity in each of these economies, with relatively more homogeneous particular levels of environmental factors. This further suggests that there may not be substantial nonlinear variations in particular levels of LPI, in combination with relatively marginal variations in particular levels of individual factors (Table 8).

Table 8. Component-Wise Standardized Means of Logistical and Environmental Factors in the Mixture Model.

Means	LPI	NOE	PM2.5AE	HI35	ALPA	AFFVA
Component 1	−0.298	−0.150	−0.410	0.296	0.609	−0.063
Component 2	−0.311	0.197	0.556	−0.620	−0.524	0.010
Component 3	3.250	−0.033	−0.125	0.606	0.684	0.169
Component 4	−0.323	-1.735×10^{-7}	0.449	5.289×10^{-5}	−1.709	−0.002
Component 5	−0.303	-1.735×10^{-7}	6.433×10^{-4}	5.289×10^{-5}	-5.890×10^{-4}	−0.002
Component 6	−0.305	-1.735×10^{-7}	−0.126	5.289×10^{-5}	−0.748	−0.002
Component 7	−0.291	-1.735×10^{-7}	−0.483	5.289×10^{-5}	0.952	−0.002
Component 8	−0.311	-1.735×10^{-7}	0.163	5.289×10^{-5}	0.710	−0.002

Figure 3 provides a summary of the model-based clustering outcomes with respect to both identifying the number of clusters and visualizing the data structure in the data space. Figure 3A illustrates the change in Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores and Within-Cluster Sum of Squares (WSS) values with respect to the increase in the number of clusters. The red spot in the figure marks the point with the lowest BIC score; this indicates that the best compromise between data fit and model complexity is achieved with the model including eight clusters. This indicates that this number of factors provides sufficient information about the data structure pattern without sacrificing accuracy through overfitting or underfitting. The plot of the WSS indicates a gradual reduction in values until reaching the point corresponding to the medium number of clusters, beyond which the values vary, though in a smooth pattern. This pattern corresponds to data with a complex structure that cannot be adequately represented by a few clusters. Figure 3B provides information regarding the pattern of the eight clusters in the projected feature space. The clusters are well differentiated in this representation and express complex configurations in most cases. This color representation indicates that each component differs in predetermined regions of the data point space. This indicator also shows regions of differentiation in most components, with some appearing sparser than others. This provides information regarding the complexity of the patterns of various environmental and logistical factors represented through the model. The structure that emerges from this representation provides information on the effectiveness of the clustering analysis through the model's grouping of profiles in conformity with the analysis's objectives (Figure 3).

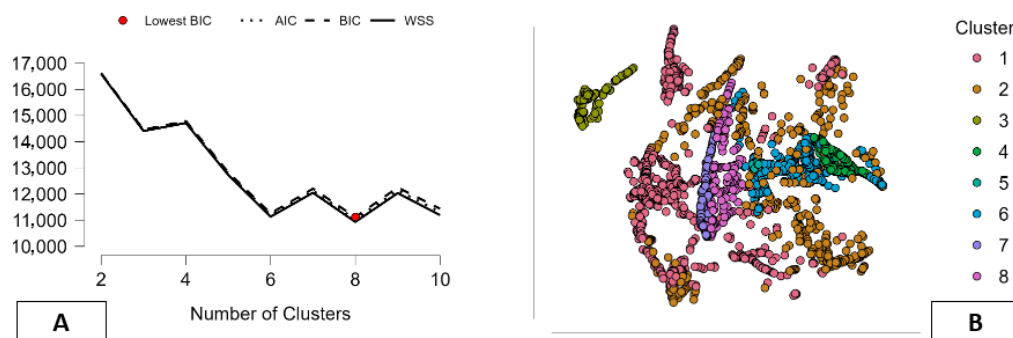


Figure 3. Model-Based Clustering Evaluation and Visualization of Cluster Structure. Note: Panel (A) displays changes in AIC, BIC, and WSS values across models with varying numbers of clusters, with the lowest BIC indicating an optimal solution at eight clusters. Panel (B) illustrates the spatial arrangement of the eight identified clusters in the projected feature space, highlighting distinct and complex structural patterns within the data.

Figure 4 depicts the standardized means for each variable across the eight clusters formed using the model-based approach, thereby clarifying the distinct environmental-logistical configurations within each grouping. The LPI values show small variation, indicating minimal differences among the clusters in LPI. However, environmental factors exhibit significant variation. The nitrous oxide emissions (NOE) and PM2.5 exposure values show both positive and negative standardisations, indicating clusters with high levels of pollution and others that are relatively clean, based on environmental conditions. The Heat Index (HI35) showed the highest discrimination value, with one cluster recording a substantially high value, indicating higher heat stress in this grouping than in others. The other clusters recorded values close to the overall average. The Agricultural land share (ALPA) and Agricultural value added (AFFVA) recorded moderate discrimination values. The figure illustrates that overall group formation is driven by environmental factors rather than LPI values, indicating that the majority of the model's variation is attributable to these factors (Figure 4).

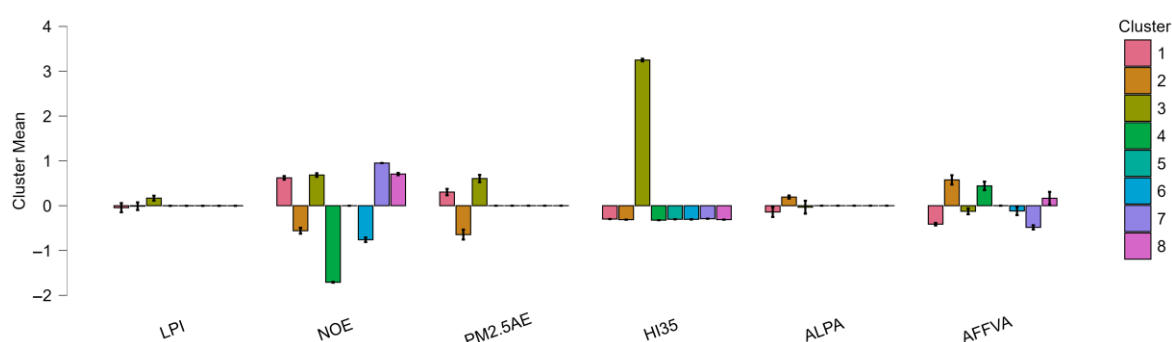


Figure 4. Standardized Variable Means Across the Eight Model-Based Clusters. Note: This figure displays the standardized means of key environmental and logistical variables across the eight clusters derived from the model-based clustering approach. LPI values show minimal variation across clusters, whereas environmental variables—particularly NOE, PM2.5AE, and HI35—exhibit strong differentiation. HI35 demonstrates the greatest discrimination, with one cluster showing notably elevated heat stress relative to others. ALPA and AFFVA present moderate variation among clusters. Overall, the figure highlights that cluster differentiation is primarily driven by environmental factors rather than LPI values.

Figure 5 summarizes a pairwise scatter plot matrix of six standardized variables across eight clusters derived from a model-based clustering analysis. Each subplot presents the relationship between two variables using colored ellipses, noting the probability distribution of each component. The combined results indicate that environmental factors such as NOE, PM2.5 exposure levels, and HI35 are the key determinants in distinguishing the various clusters. Clusters are arranged in well-defined regions based on these factors, especially in the NOE and PM2.5 exposure level plots, where clear separations occur between high and low emission factors. However, the most significant determinant in this analysis is the HI35 variable, which shows one component with abnormally high levels of heat stress compared to other factors. Conversely, LPI shows minimal variation between factors and groups, yet remains clustered around the standard mean in terms of standardization. The result indicates that variations in logistical performance do not significantly account for the observed pattern in the data and supports the primary idea of relying on the environmental factors outlined in this analysis. The agricultural factors of ALPA and AFFVA provide more information, though they exhibit less-defined edges in the figure, suggesting minute variations in land use or in the overall economic role of agriculture. This analysis indicates that the figure presents a complex clustering pattern across multiple dimensions, in which environmental factors account for overall variation, except for LPI, which is significant yet supplementary in the overall context of Environmental Social Governance (Figure 5).

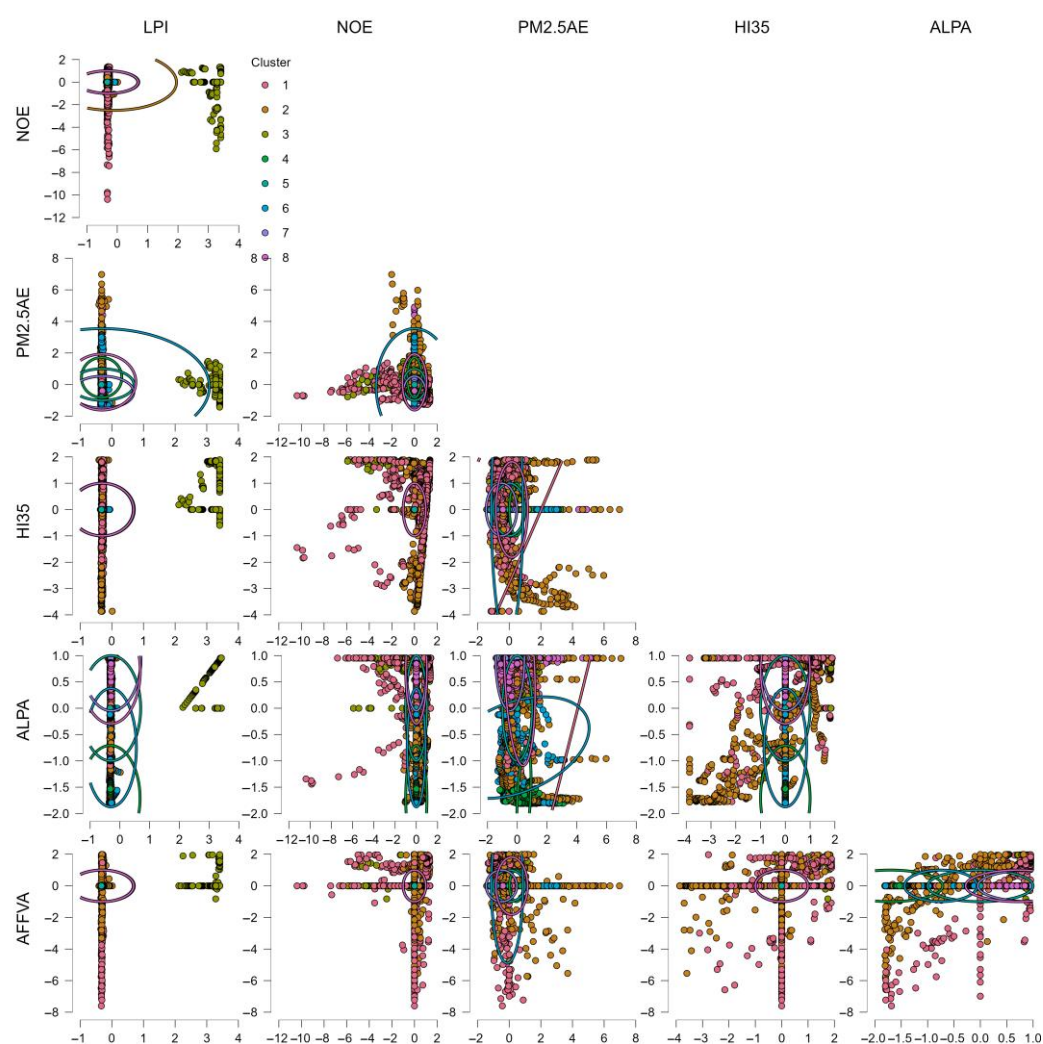


Figure 5. Pairwise Scatter Plot Matrix of Standardized Variables Across Eight Model-Based Clusters. Note: This figure presents a pairwise scatter plot matrix of six standardized variables across eight

clusters derived from the model-based clustering analysis. Each subplot illustrates the bivariate relationship between variables, with colored ellipses representing the probabilistic distribution of each cluster component. The environmental variables—NOE, PM2.5AE, and HI35—demonstrate the strongest discriminatory power, producing distinct separations between high- and low-emission clusters. HI35 is the most influential factor, with one cluster exhibiting markedly elevated heat stress. In contrast, LPI shows minimal variation across clusters, clustering closely around the standardized mean and contributing little to group differentiation. Agricultural variables (ALPA and AFFVA) reveal moderate but less sharply defined patterns, suggesting subtle variations in land use and agricultural economic contribution. Overall, the matrix highlights that environmental factors predominantly drive cluster formation, whereas logistical performance indicators play a supplementary role within the broader ESG context.

5. Exploring the Interaction Between Social Factors and LPI in an ESG Context

This part examines the causality between the Logistics Performance Index (LPI) and the Social (S) pillar of the ESG framework in 163 nations from the period 2007 to 2023. Employing two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) techniques, the research looks at how important social variables like water and sanitation accessibility, education, population structure, income distribution and labor conditions influence the efficiency of logistics. Accounting for endogeneity by using a comprehensive set of instrumental variables, the outcomes show social development drivers to be important influencers of logistic performance and prove why socially inclusive approaches are required to boost supply chain systems everywhere.

5.1. Analyzing the S-Social Component's Impact on Logistics Performance

This section explores the relationship between the Logistics Performance Index (LPI) and the Social (S) pillar of the ESG model. Using fixed-effects two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) methods, the study investigates how social factors—such as access to basic services, education, income distribution, labor market conditions, and demographic structures—impact logistics performance. The results reveal that improvements in social indicators can have both positive and negative effects on LPI, highlighting the intricate connections between human development, equity, and logistics efficiency within a sustainable growth framework.

We have estimated the following model:

- $X_{it} = Z_{it}\Pi + v_{it}$ (First Stage)
- $Y_{it} = X_{it}\beta + \mu_{it}$ (Second Stage)
- $Y_{it} = LPI_{it}$
- $X_{it} = \{\text{PSMWS PSMS PA65A SEP CET POA ISL20}\}$
- $Z_{it} = \{\text{IUI GDPG PSHWNP RFMLFPR SLRI STJA RLE NM CO2E NOE PM25AE GHGLUCF EILPE REC FFEC EU CDD}\}$
- HDD HI35 SPEI LST PD LWS ALPA FPI AFFVA MST AFWT TMPA ASFD ASNRD}
- $i = 163$
- $t = [2007; 2023]$.

Results are indicated in Table 9.

Table 9. Impact of Social Factors on Logistics Performance: Fixed-Effects TSLS and G2SLS Estimates.

Y	LPI					
Endogenous	PSMWS PSMS PA65A SEP CET POA ISL20					
Instruments	IUI GDPG PSHWNP RFMLFPR SLRI STJA RLE NM CO2E NOE PM25AE GHGLUCF EILPE REC FFEC EU CDD HDD HI35 SPEI LST PD LWS ALPA FPI AFFVA MST AFWT TMPA ASFD ASNRD					
T	17					
N	163					
Observations	2771					
	Fixed-effects TSLS			G2SLS random effects		
	coefficient	std. error	z	coefficient	std. error	z
Constant	14.203 ***	0.931	15.25	14.213 ***	0.929	15.28
PSMWS	−0.012 *	0.006	−1.832	−0.012 *	0.006	−1.865
PSMS	−0.048 ***	0.013	−3.502	−0.048 ***	0.013	−3.514
PA65A	−0.046 **	0.022	−2.095	−0.046 **	0.022	−2.096
SEP	−0.364 **	0.181	−2.012	−0.363 **	0.180	−2.013
CET	1.695 ***	0.400	4.232	1.694 ***	0.399	4.236
POA	0.029 ***	0.008	3.305	0.029 ***	0.008	3.299
ISL20	−1.596 ***	0.370	−4.307	−1.595 **	0.370	−4.311
Statistics and Tests	SSR = 1043.01			SSR = 2755.43		
	sigma-hat = 0.633 (df = 2601)			sigma-hat = 0.998 (df = 2763)		
	R-squared = corr(y, yhat) ² = 0.002			R-squared = corr(y, yhat) ² = 0.002		
	Included units = 163			Included units = 163		
	Time-series length: min = 17, max = 17			Time-series length: min = 17, max = 17		
	Wald chi-square(7) = 71.986 [0.0000]			Wald chi-square(7) = 72.366 [0.0000]		
	Null hypothesis: The groups have a common intercept			sigma-hat(within) = 0.633		
	Test statistic: F(162, 2601) = 27,100.4 [0.0000]			sigma-hat(between) = 26.654		

Note. *** indicates significance at the 1% level; ** indicates significance at the 5% level; * indicates significance at the 10% level. Coefficients marked with asterisks are statistically significant at the corresponding confidence levels.

This research examines the determinants of the Logistics Performance Index (LPI) of 163 countries over a period of 17 years using fixed-effects two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) models with random effects. The framework includes a broad range of instruments capturing economic, demographic, governance, and environmental data. One key finding from the research has a direct bearing on the Social (S) component of the ESG framework. The endogenous variables—i.e., access to safely managed drinking water (PSMWS) and sanitation services (PSMS), elderly population percentage (PA65A), primary school enrollment (SEP), employment of children (CET), prevalence of overweight adults (POA), and income share held by the poorest 20% (ISL20)—all are dimensions of social development considered essential. The correlation unfolds as follows: More widespread provision of simple services like water and sanitation is somewhat counterintuitively negatively related to the variable of logistics performance. While statistically robust, however, the effect is small and implies high-performance social service provision may be related to more rigorous regulatory systems or greater operational costs marginally impacting the effectiveness of logistics. Demographic issues are also seen: a larger percentage of aging population and more enrollment in schools is negatively related to LPI. This may represent the effects of changing labor market fundamentals, whereby aging societies and higher education enrollment fewer youth in the workforce temporarily limit the labor available to the heavily labor-intensive industries like logistics. The opposite effect is identified in the case

of the prevalence of child labor (CET), which has a strong positive effect on LPI—a worrying indicator. This indicates improving the performance of logistics in less developed economies may depend partly on exploitative employment arrangements. This has a fundamental social sustainability issue at its core: efficiency gains at the expense of youth welfare and human rights are unacceptable if it goes against the core tenet under the Social pillar of ESG. Equally, the positive effect of overweight prevalence (POA) on the variable of logistics performance is likely a reflection of deeper patterns of economic prosperity and consumerism requiring more sophisticated systems of logistics. This also has social concerns related to modern lifestyles and unjust food systems. The negative correlation of income inequality (ISL20) and the variable of logistics performance is a fundamental finding. In economies in which the bottom 20% of the population possess less income, logistics systems look less efficient. More economic inequality contributes to fragmented markets, stagnant mobility, and lower human capital, all of which contribute to less smooth logistics operations. From an ESG-Social stance, this result confirms that more inclusive economic development bolsters better-performing logistics and supply chain systems. The extensive range of tools utilized—and range of indicators including internet penetration and rule of law, female labor force participation and governance—also highlight social and institutional environments as the determinants of the performance of logistics. More robust social structures, improved legal protections and more inclusive labor markets are not social goods alone but also efficiency enablers of global supply chain operations. In general, this examination makes it evident that social development underpins the performance of logistics. Education, services provision, equality of condition, labor quality and provision of health services all play important parts. Logistics infrastructures policies to enhance them must be strongly integrated with social investment plans to guarantee progress in the area of logistic infrastructures does not happen at the expense of the development of humanity but hand in hand with it and in full coherence with ESG-S objectives.

Causality. The causal identification strategy employed—fixed-effects TSLS and G2SLS with a rich instrument set—permits a strong identification of the causal impact of social variables on the Logistics Performance Index (LPI). The coefficients imply the causal influence of variations in social development indicators on logistics performance and do not simply correlate with it. In particular, better access to safely managed water (PSMWS) and sanitation (PSMS), a larger elderly population percentage (PA65A), and increased school enrollment (SEP) are causally associated with a marginally declining LPI, possibly through augmented regulatory costs or labor force shortages. More troublingly, the causal positive effect of child labor (CET) on LPI illustrates how, in certain settings, improving the efficiency of logistics depends on unsustainable and ethically challenged forms of labor. The causal negative effect of income inequality (ISL20) on LPI also shows how more equal income distribution facilitates the efficiency of the logistics system. Significantly, the instrumental variables technique enhances the causal assertions by reducing endogeneity generated by reverse causality or missing variable bias. Nevertheless, low R^2 values signify how social variables have statistically significant causal impacts but account for a minimal share of overall variance in the performance of the logistics system and argue in favor of combining social interventions with more general economic and infrastructural reforms.

Overall impact of the S-Social component within the ESG model. The evidence presents unequivocal empirical proof that the Social (S) pillar of the ESG framework has a causal and sizable yet multifaceted effect on the performance of logistics. Social improvements in indicators have a positive or negative impact on the Logistics Performance Index (LPI), highlighting the subtle tradeoff between operational efficiency and human development. The provision of fundamental services such as safely managed drinking water (PSMWS) and sanitation (PSMS), demographic transitions like population aging (PA65A), and

increased enrollment in schools (SEP) are causally linked to declines of minor magnitude in the performance of logistics, probably indicative of increased regulatory costs or labor shortage. The worrying causal positive effect of child labor (CET) on LPI also indicates the persistence of socially unsustainable patterns supporting the efficiency of logistics in some economies. The positive causal effect of overweight prevalence (POA) on LPI also shows stronger consumer-led logistic requirements, while income inequality (ISL20) has a negative effect on logistic efficiency and highlights the importance of equalized growth. Although the causal evidence is statistically strong because a rich list of instrumental variables was used, the low values of R^2 reveal a minimal share of variance explained by social variables. Summing up, the development of logistic performance has to be coordinated with socially sustainable development policies completely aligned with ESG-S principles.

5.2. Machine Learning Estimation of Socio-Economic Impacts on Logistics Performance

This section applies machine learning methods to estimate the relationship between socio-economic variables and the Logistics Performance Index (LPI). Several algorithms—including Boosting, Decision Trees, Random Forests, and Support Vector Machines—are evaluated based on normalized performance metrics. The K-Nearest Neighbors (KNN) algorithm emerges as the most accurate and robust model, achieving the lowest prediction errors and the highest explanatory power. Further analysis identifies key social predictors, such as school enrollment, overweight prevalence, and child labor incidence, highlighting the critical influence of human development factors on logistics performance. These results underline the complex interplay between social structures and logistic efficiency (Table 10).

Table 10. Comparison of Machine Learning Algorithms for Predicting Logistics Performance Based on Socio-Economic Factors.

Metric	Boosting	Decision Tree	K-Nearest Neighbors	Linear	Random Forest	Regularized Linear	SVM
MSE	0.617	0.110	0.000	0.451	0.007	0.642	0.708
MSE (scaled)	0.568	0.091	0.000	0.822	0.056	0.777	1.000
RMSE	0.643	0.099	0.000	0.470	0.005	0.664	0.724
MAE/MAD	0.776	0.140	0.000	0.727	0.277	0.857	0.316
MAPE	0.763	0.172	0.000	1.000	0.290	0.955	0.000
R^2	0.211	0.793	1.000	0.092	0.950	0.103	0.000

This cluster is seen to represent the overall or “baseline” population. Cluster 2, though having very small number of observations (8), also has a very high silhouette value of 0.791 as a testament to good clustering and separation between groups. The average values confirm positive NOE (+0.423), very low PM2.5 exposure levels (−2.623), very low agricultural land usage (−2.766), and high value added from agriculture, forestry, and fishing (+0.843). This proves that Cluster 2 consists of countries or regions with high productivity in terms of agriculture and good air quality despite relatively high nitrous oxide emissions [87]. Cluster 3 with 238 has a high Heat Index 35 (+3.250), indicating extreme exposure to hot air and heat stress, with associated positive departures of NOE (+0.684) and PM2.5 exposure (+0.606). The silhouette value of 0.523 indicates good but imperfect separation of the groups. This group appears to represent countries or regions with both high exposure to heat and air pollution levels as per conclusions drawn in recent semi-supervised PM2.5 clustering and air pollution patterns by region by [88,89]. Within the quality of clustering, the silhouette values range from 0.382 to 0.791 across groups and are representative of an acceptable but imperfect data partitioning. The within-cluster sum of squares is very high on Cluster 1 (12,160.403), as a marker of data variability internally in the group and is very low on Cluster 2 (3.617), as an indication of the

closeness of the small group. Generally, the model is capable of separating groups at the extremes of data distribution but a majority of the data fall into a very large heterogeneous core group [87].

Using the K-Nearest Neighbors (KNN) algorithm to forecast the Logistic Performance Index (LPI) on the basis of socio-economic and demographic variables produces results both statistically robust and informative in terms of substance. Primary school enrollment (SEP) is the most significant predictor identified by feature importance assessment expressed as mean dropout loss (28.085), followed by adult overweight prevalence (POA, 26.403) and child labor (CET, 26.196). Other variables, such as access to safely managed sanitation services (PSMS), population percentage aged 65 and above (PA65A), income share of the lowest 20% (ISL20), and percentage of population with access to safely managed drinking water services (PSMWS), are also contributory but to a lesser magnitude. These results imply educational level, labor and public health indicators are fundamental determinants of logistic capacities at the national level (Figure 6).

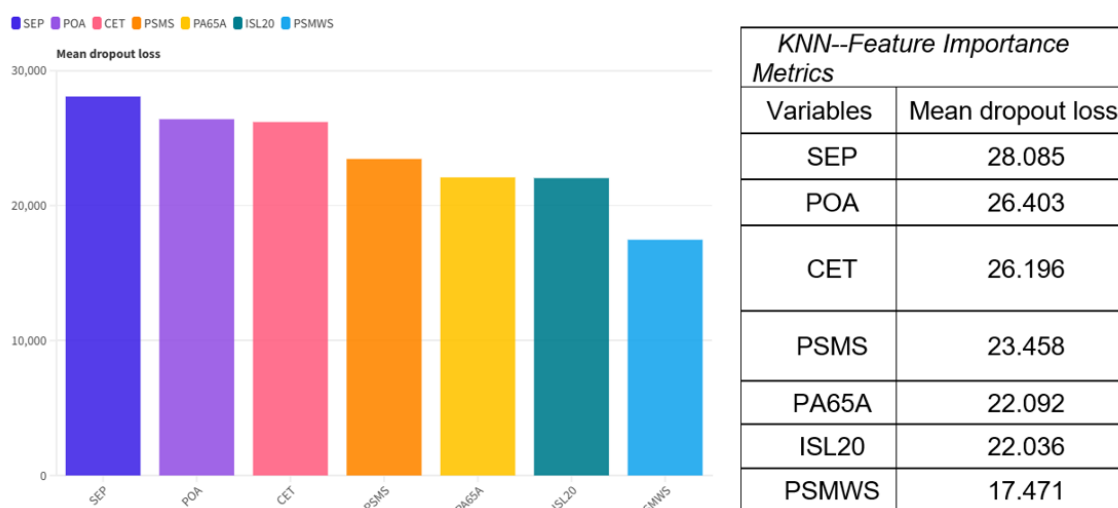


Figure 6. KNN Feature Importance Analysis for Socio-Economic Predictors of Logistics Performance.

The additive feature attribute analysis of the test dataset better represents the effects of single predictors on the model's predictions. In all scenarios, the base prediction, the model's prediction when particularized feature effects are removed, is a fixed value of 10.241. Deviations from the baseline represent the subtle interactions among variables: School enrollment (SEP) has a consistent strong positive effect on LPI predictions everywhere, especially in cases 1 to 4. Contrariwise, access to drinking water services (PSMWS) consistently has a negative effect, especially in cases 2, 3, and 4, and represents a mediated association with logistic performance by other infrastructural or governance variables. The negative effects of overweight prevalence (POA) and child labor (CET) also demonstrate the adverse effect of labor market distortions and healthcare on logistic efficiency. These inferences are consistent with recent studies using SHAP (Shapley Add ExPlanations), which demonstrate the capacity of the technique to identify the marginal effect of predictors on models with a high degree of complexity [90,91]. Overall, the KNN model not only makes good LPI predictions but also allows better interpretation by quantifying the marginal effects of key socio-economic variables, in a manner analogous to the SHAP-based explanations used in the prediction of the attrition of employees and diagnostics in healthcare [91,92]. These inferences demonstrate interdependencies

between logistic output and human development indicators and represent the significance of social policy considerations in logistic performance maximization plans (Figure 7).



Figure 7. Additive Feature Contributions to LPI Predictions Using K-Nearest Neighbors (KNN).

5.3. Clustering to Verify the Relationship Between LPI and the S-Social Component of the ESG Model

This study examines the predictive correlation between the Logistics Performance Index (LPI) and a range of socio-economic and demographic variables using machine learning regression methods. Comparing different algorithms using normalized performance measurements highlights K-Nearest Neighbors (KNN) as the optimal technique to capture the underlying variance in logistics performance. Not only does KNN perform better in terms of predictive precision, but it also provides innovative insights into relative importance values of important social variables like education, health, and labor conditions. The investigation underscores how socio-economic development indicators play a pivotal role in determining logistics outcomes, thus supporting socially inclusive logistics approaches in the ESG framework (Table 11).

Table 11. Normalized Performance Metrics for Clustering Algorithms: Predicting LPI with Socio-Economic Variables.

Metric	Density-Based	Fuzzy C-Means	Hierarchical	Model-Based	Neighborhood-Based	Random Forest
Maximum diameter	1.000	0.072	0.063	0.967	0.061	0.081
Minimum separation	1.000	0.029	0.216	0.000	0.056	0.033
Pearson's γ	0.527	0.000	0.870	0.179	0.538	0.056
Dunn index	1.000	0.043	0.314	0.000	0.081	0.043
Entropy	0.000	0.752	0.340	1.000	0.899	0.693
Calinski-Harabasz index	1.000	0.001	0.002	0.002	0.004	0.001
R ²	0.000	0.351	0.642	0.627	1.000	0.494
AIC	1.000	0.593	0.000	0.008	0.000	0.569
BIC	1.000	0.593	0.000	0.008	0.000	0.569
Silhouette	1.000	0.115	0.926	0.370	0.963	0.000

Based on normalized performance measurements, Neighborhood-Based Clustering is the most suitable out of the methods considered. This is evident in better performance on a set of core clustering validity measurements. Notably, it has the best R^2 value with a higher percentage variance explained compared to other methods. Moreover, it has a high Silhouette score, reflecting good internal cohesion and good separation between groups—properties of paramount importance to measuring the quality of a clustering structure [93]. In addition to that, its strategically low maximum diameter and acceptable minimum separation values further attest to Neighborhood-Based Clustering to effectively minimize within-cluster dispersion and maintain different groups separated. Though it fails to achieve the best AIC and BIC values to evaluate model simplicity and goodness of fit, its performance remains competitive considering the merit of structural clarity and interpretability to clustering analysis [93]. Density-Based Clustering approaches, for example, despite having best scores on maximum diameter and Dunn index scores, register poor Silhouette values and weaker R^2 values and demonstrate weaker model robustness in the respective setting of this type of application [94]. Likewise, while targeted metrics have good performance by Random Forest Clustering, it does not outperform consistently on all dimensions. Although it has good performance on certain dimensions of the clustering problem, its stability and interpretability are unstable on different datasets [95]. Neighborhood-Based Clustering therefore has the best trade-off among the considered methods between separation and compactness and model explanatory power and stability. Overall performance also means it is best suited to applications requiring consistent group distinction as well as internal consistency to exist and best used in the setting of the current investigation (Table 12).

Table 12. Socio-Economic Characterization of Clusters Affecting Logistic Performance.

Cluster	1	2	3	4	5	6	7	8	9	10
Size	564	244	218	434	68	409	76	352	237	169
Explained proportion within-cluster heterogeneity	0.168	0.105	0.108	0.153	0.041	0.108	0.064	0.063	0.117	0.072
Within sum of squares	1.153	718.762	745.245	1.053	284.004	744.270	442.164	433.747	801.223	493.387
Silhouette score	0.227	0.161	0.219	0.194	0.378	0.239	0.235	0.450	0.204	0.430
Center LPI	−0.302	−0.315	−0.324	−0.312	3.309	−0.299	−0.327	−0.277	−0.311	3.233
Center PSMWS	−0.077	0.164	2.259	−0.088	−0.042	−0.515	1.799	−0.673	−0.127	−0.632
Center PSMS	0.668	0.183	0.741	−1.210	−2.134	−0.063	−2.193	0.633	0.090	0.207
Center PA65A	−0.624	−0.466	−0.165	−0.078	1.284	−0.121	1.423	−0.220	2.149	−0.252
Center SEP	−0.253	−0.209	−0.691	−0.662	−1.185	0.883	−2.489	1.200	0.242	0.358
Center CET	−0.495	−0.520	−1.018	−0.478	−0.783	0.723	−1.620	1.748	0.028	0.556
Center POA	−0.360	1.542	−0.600	−0.400	−0.580	−0.505	−0.423	1.168	0.147	−0.217
Center ISL20	−0.545	−0.474	−1.035	−0.416	−0.978	0.751	−1.560	1.603	0.114	0.685

Applying Neighborhood-Based Clustering to the chosen socio-economic and demographic variables confirms a significant splitting of the dataset into ten groups with different profiles by logistic performance and corresponding indicators of human development. The silhouette values are mostly average but confirm acceptable cohesion among the groups, with cluster groups 8 and 10 sharing the highest internal consistency (0.450 and 0.430, respectively), suggesting consistency in relatively homogeneous patterns in the data [96]. The explained percentage of heterogeneity among the groups further confirms adequacy in the model, as in Cluster 5, the low percentage of heterogeneity (0.041) and a high cluster center LPI value (3.309) pick out a distinctive group with high logistic performance. Clusters 5 and 10 are indeed the most differentiated structural groups and show much higher Logistic Performance Index values compared to other

groups with central values around negative LPis [97]. A look at the cluster centers picks out significant socio-economic contrasts. The groups found to have a higher LPI values are predominantly marked by improved coverage in terms of sanitation (high scores on PSMS), relatively higher proportions of elderly population (PA65A), improved coverage of safely managed drinking water (PSMWS), and more balanced income distribution (ISL20). The groups found to have low LPI centers (now classified as groups 3 and 7) are marked by negative performance in all of the above dimensions combined with increased prevalence of child labor (CET) and decreased enrolment in schools (SEP), suggesting structural weaknesses [96]. Surprisingly, Cluster 8 has a positive logistic profile even though it has low scores on water service indicators, implicating the hypothesis that education and income distribution may in this group make up deficits in infrastructure. These patterns amplify the importance of the inclusion of socio-economic dimensions in clustering methods in the case of logistics and infrastructure evaluation, as shown in previous examples of clustering in supply chain and logistic environments [95]. Overall, the results demonstrate that logistic performance is closely intertwined with broader social determinants, including education access, labor market conditions, health outcomes, and basic service provision, confirming the multi-dimensional nature of logistics capacity within national and regional contexts [97].

Results are showed in Table 13.

Table 13. Cluster means.

	LPI	PSMWS	PSMS	PA65A	SEP	CET	POA	ISL20
Cluster 1	−0.495	−0.545	−0.302	−0.624	−0.360	0.668	−0.077	−0.253
Cluster 2	−0.520	−0.474	−0.315	−0.466	1.542	0.183	0.164	−0.209
Cluster 3	−1.018	−1.035	−0.324	−0.165	−0.600	0.741	2.259	−0.691
Cluster 4	−0.478	−0.416	−0.312	−0.078	−0.400	−1.210	−0.088	−0.662
Cluster 5	−0.783	−0.978	3.309	1.284	−0.580	−2.134	−0.042	−1.185
Cluster 6	0.723	0.751	−0.299	−0.121	−0.505	−0.063	−0.515	0.883
Cluster 7	−1.620	−1.560	−0.327	1.423	−0.423	−2.193	1.799	−2.489
Cluster 8	1.748	1.603	−0.277	−0.220	1.168	0.633	−0.673	1.200
Cluster 9	0.028	0.114	−0.311	2.149	0.147	0.090	−0.127	0.242
Cluster 10	0.556	0.685	3.233	−0.252	−0.217	0.207	−0.632	0.358

6. Governance and Logistics Performance: An Empirical Assessment Within the ESG Framework

The chapter examines the interconnection between governance quality and logistics performance in the ESG framework. By using fixed-effects two-stage least squares (TSLS), generalized two-stage least squares (G2SLS), machine learning models, and clustering methods on data from 163 countries between the period 2007–2023, the study documents how five key indicators of governance—government effectiveness, regulatory quality, political stability, rule of law, and scientific innovation—affect the Logistics Performance Index (LPI). The findings highlight the importance of robust, transparent, and accountable institutions to underpin efficient logistics systems but also the multifaceted and dynamic character of governance impacts on global supply performance.

6.1. The Role of Institutional Governance in Shaping Logistics Efficiency: An ESG Perspective

This section analyzes the causal impact of governance quality on logistics performance within the ESG framework, using an instrumental variables (IV) panel data approach. Drawing on a balanced dataset of 163 countries from 2007 to 2023, and applying fixed-effects TSLS and G2SLS estimators, the study isolates the effects of key governance

dimensions—such as government effectiveness, regulatory quality, voice and accountability, and rule of law—on the Logistics Performance Index (LPI). By addressing potential endogeneity and omitted variable bias, the analysis provides robust evidence that governance factors are not merely correlated with, but causally linked to, sustainable logistics performance under the ESG model.

- $X_{it} = Z_{it}\Pi + v_{it}$ (First Stage)
- $Y_{it} = X_{it}\beta + \mu_{it}$ (Second Stage)
- $Y_{it} = LPI_{it}$
- $X_{it} = \{\text{GEE RQE ESRPS VAE STJA PSAOV RLE}\}$
- $Z_{it} = \{\text{IUI CO2E NOE PM25AE GHGLUCF EILPE REC FFEC EU CDD HDD HI35 SPEI LST PD LWS ALPA FPI AFFVA MST AFWT TMPA ASFD ASNRD}\}$
- $i = 163$
- $t = [2007; 2023]$.

Results are synthesized in Table 14.

Table 14. Causal Effects of Institutional Governance on the Logistics Performance Index (LPI).

Y	LPI					
Endogenous	GEE RQE ESRPS VAE STJA PSAOV RLE					
Instruments	IUI CO2E NOE PM25AE GHGLUCF EILPE REC FFEC EU CDD HDD HI35 SPEI LST PD LWS ALPA FPI AFFVA MST AFWT TMPA ASFD ASNRD					
T	17					
N	163					
Observations	2771					
	G2SLS random effects			Fixed-effects TSLS		
	coefficient	std. error	z	coefficient	std. error	z
const	11.911 ***	0.588	20.24	11.929 ***	0.593	20.09
GEE	0.015 ***	0.0023	6.585	0.015 ***	0.002	6.544
RQE	$-5.51554 \times 10^{-6} **$	2.38212×10^{-6}	−2.315	$-5.52359 \times 10^{-6} **$	2.40369×10^{-6}	−2.298
ESRPS	−0.035 ***	0.009	−3.634	−0.035 ***	0.009	−3.632
VAE	0.543 ***	0.137	3.956	0.546 ***	0.138	3.946
STJA	0.025 ***	0.006	4.153	0.025 ***	0.006	4.121
PSAOV	$9.78199 \times 10^{-7} **$	4.06429×10^{-7}	2.407	$9.77849 \times 10^{-7} **$	4.10115×10^{-7}	2.384
RLE	0.282 **	0.110	2.560	0.283 **	0.111	2.543
Statistics And Tests	SSR = 2713.7			SSR = 1072.64		
	sigma-hat = 0.991 (df = 2763)			sigma-hat = 0.642 (df = 2601)		
	R-squared = $\text{corr}(y, \hat{y})^2 = 0.009$			R-squared = $\text{corr}(y, \hat{y})^2 = 0.009$		
	Included units = 163			Included units = 163		
	Time-series length: min = 17, max = 17			Time-series length: min = 17, max = 17		
	Wald chi-square(7) = 72.355 [0.0000]			Wald chi-square(7) = 71.052 [0.0000]		
	sigma-hat(within) = 0.642			Null hypothesis: The groups have a common intercept		
	sigma-hat(between) = 30.771			Test statistic: $F(162, 2601) = 26,449.2$ [0.0000]		

Note. *** indicates significance at the 1% level; ** indicates significance at the 5% level. Coefficients marked with asterisks are statistically significant at the corresponding confidence levels.

The results prove insightful and straightforwardly strong on the premier role of governance in supporting logistics performance. Government Effectiveness (GEE) has a direct and very powerful impact on Logistics Performance Index (LPI), both with a coefficient of around 0.0152 and a significance of the 1% level. This shows competent, transparent, and efficient governments support the development of logistic systems in

better managing infrastructures, providing services, and putting in force policies—a correlation also found in research on governance and economic development across regions [98,99]. Regulatory Quality (RQE), to our surprise, has a very low but negative and statistically significant impact. Although the impact size is low, the result may signify a case of excessive restraint by overly burdensome or ill-designed regulatory systems to exact unintended costs or frictions on logistic operations in environments with excessive bureaucracy or controls stifling innovation and flexibility [100]. The Economic and Social Rights Performance Score (ESRPS) is inversely associated with the performance of logistics. The result means a structural problem: in nations heavily concentrated on generous social protections, regulatory heaviness or resource redistributive mechanisms may unwittingly limit investments or operational efficiencies core to logistic networks. It highlights a thin line between social progress and logistic efficiency [101]. Voice and Accountability (VAE) has a strong and significant correlation with LPI and indicates nations with freer media, better civic engagement, and accountable government have better logistic performance because they have better visibility, are more responsive to the marketplace, and less corrupt [102]. Science and Technical Journal Articles (STJA) grow LPI positively and show a role of innovation, research power, and tech creation to advance efficient and modern logistic fields. Political Stability and Absence of Violence (PSAOV), though modeled using very small coefficients, has a positive and a statistically significant correlation with LPI and means stability brings safety and predictability to assure local and foreign supply chains and thus avert risks and operational interruption—an inference supported by research on the impact of political governance on the economy in different nations [102]. Finally, Rule of Law (RLE) shows a positive and tangible impact in supporting the argument that solid legal institutions, protection of property rights, and adherence to contracts are essential pillars to support good logistic networks [100]. Statistically, the importance of the models is reflected in the Wald chi-square statistics as very high and confirm the joint appropriateness of the variables included. Although the R-squared values of approximately 0.0093 are low and express that variables of governance alone describe a relatively small percentage of the overall variance in logistic performance, their effect is statistically relevant and economically considerable. The use of a range of environmental indicators as a basis includes a range of indicators such as CO₂ emissions, exposure to PM_{2.5}, consumption of energy, and climate variables, which adds richness to the model. While they are secondary to the objectives here but are relevant to any broader consideration of environmental issues and are remnants from our investigation of the environmental systemic shocks to the governance and the standards of logistics, their inclusion makes identification stronger by capturing risks on a higher level indirectly affecting governance and logistic environments. These results confirm good governance as a building block to the efficient logistic standards. Stable and efficient government, transparent and innovative as it is, will empower nations to build and maintain efficient logistic chains integral to competitiveness in a globalized economy. The results also alert, however, social and regulatory ambitions to be developed thoughtfully to avoid unwanted trade-offs with efficiency of operation. In the ESG framework, research supports that the Governance (G) pillar is not an ancillary variable but a direct determinant of infrastructure quality and efficiency of economy in logistics and sustainable development [99,101].

Causality. The results strongly confirm the causal association with the performance of logistics and the quality of governance in the ESG framework. The empirical approaches using an instrumental variables (IV) panel data methodology—fixed-effects two-stage least squares (TSLS) and generalized two-stage least squares (G2SLS) estimators—are used to address concerns on endogeneity problems such as reverse causation and specification of a relevant variable. The used econometric approaches are

consistent with recent research on causal inference using IVs in data setups with a high level of data complexity [6,103]. The robust coefficients on governance metrics such as Government Effectiveness (GEE), Voice and Accountability (VAE), and the Rule of Law (RLE) provide strong evidence to confirm the premise that improvements to governance institutions are linked to but do not merely correlate with improving logistics performance. The careful selection of the instrumental set of variables such as environmental and macro-structural drivers (e.g., CO₂ emissions and PM_{2.5} exposure and energy consumption), removes confounding exogenous variation in governance quality and thereby bolsters identification. The subject modeling selections are consistent with stronger and distributionally robust IV estimation methods now suggested in available literature [104]. This provides support to the overall finding that institutional efficiency, accountability, transparency, and stability are key drivers to efficiency in logistics in any confounding macroeconomic environments. Although the relatively low R-squared values indicate that governance contributes partially to variance in logistics outcomes, the strongly significance Wald chi-square statistics confirm the combined significance of the governance predictors. The study thus presents robust causal evidence supporting the inclusion of governance reforms as a key prescription to leverage logistics systems under the ESG framework [6,103].

Overall effects of G-Governance elements in the ESG framework on Logistic Performance Indicators. The evidence confirms the Governance (G) pillar of the ESG model as having a causal and pivotal impact on country differences in logistics performance. Applying an instrumental variables (IV) panel data framework with fixed-effects TSLS and G2SLS estimators to avoid endogeneity issues, the research highlights the effect on governance quality independently. Results indicate higher government effectiveness, rule of law, and voice and accountability are significantly and positively related to improved logistics outcomes, whereas excessively complicated regulatory environments and redistributive social policy options on occasion may bring in unnecessary inefficiencies. While governance on its own accounts for a relatively small percentage of the variance in logistics performance, its effect exists and is both statistically and economically significant. These results are consistent with existing research on the importance of good governance practices improving both logistics capacity and financial performance in logistics firms and markets—particularly ESG-aware markets [4,5]. In addition, they complement research evidence that ESG integration, and, in particular, good governance mechanisms to facilitate it, may represent a performance catalyst even in financial markets—highlighting the strategic importance of governance to investor confidence and sectoral returns [1]. Overall, the research demonstrates transparent, stable, and efficient institutions as essential drivers to sustainable and competitive logistics systems and confirms the centrality of the Governance pillar of the ESG model to public policy and infrastructure and also to private sector logistics strategy and investor behavior [4,5].

6.2. Machine Learning Regressions LPI and G-Governance

In a range of multiple regression algorithms including Boosting Regression, Decision Tree Regression, k-Nearest Neighbors Regression, Linear Regression, Random Forest Regression, and Support Vector Machine Regression, models were systematically compared against a range of statistical performance metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). Of the models tested, the k-Nearest Neighbors (k-NN) Regression algorithm consistently outperformed its peers with the best values for MSE (215.583), RMSE (14.683), and MAE (5.779), and the relatively high R^2 as a result of 0.619. These combined to confirm k-NN Regression as having the best ability to minimize prediction error while also having the ability to

maintain a high percentage of variance of the response variable. The same has been seen in uses of k-NN to predict solar radiation and cryptocurrency prices where it maintained competitive accuracy and stability [105,106]. While the Support Vector Machine (SVM) model had an anomalous low MAPE (18.16%), its severely low R^2 value (0.024) highlights a serious lack of explanatory power and consequently makes it inadvisable to maintain robust predictive model suitability in the case. The result concurs with issues identified in other areas of research using SVM as its sensitivity to data distribution has resulted in unstable estimates despite low error values [107]. While Random Forest Regression had the best value of R^2 (0.628), it had marginally higher error values than k-NN and was subsequently unable to outrank it in terms of predictive ability. These trade-offs demonstrate typical practice using ensemble learning whereby reductions in variance might result in increases in small levels of bias [106]. Combined, these results confirm k-Nearest Neighbors Regression to have the most optimal mixture of both minimalism on both sides and optimally ensuring both correctness as well as generalizability. As such, it is hereby proposed as the best methodology to be adopted in predictive work in datasets of similar behavior (Table 15).

Table 15. Comparative Performance of Regression Algorithms for Predicting Logistics Performance.

	Boosting Regression	Decision Tree Regression	k-Nearest Neighbors	Linear Regressions	Random Forest Regression	Support Vector Machine
MSE	710.124	395.86	215.583	646.107	327.09	681.308
MSE (scaled)	1.198	0.759	0.425	1.488	0.415	1.689
RMSE	26.648	19.896	14.683	25.419	18.086	26.102
MAE/MAD	13.847	6.537	5.779	14.92	8.665	7.702
MAPE	212.44%	128.57%	133.26%	294.11%	145.91%	18.16%
R^2	0.16	0.384	0.619	0.065	0.628	0.024

Within the proposed study, the k-Nearest Neighbors (k-NN) regression model was used to probe the impact of the “Governance” (G) component of the ESG framework on the Logistic Performance Indicator (LPI). The research used a range of governance-focused predictors to include Government Effectiveness Estimate (GEE), Regulatory Quality Estimate (RQE), Economic and Social Rights Performance Score (ESRPS), Voice and Accountability Estimate (VAE), Scientific and Technical Journal Articles (STJA), Political Stability and Absence of Violence Estimate (PSAOV), and Rule of Law Estimate (RLE). Feature importance was constructed using mean dropout loss metrics to reveal STJA (29.515) and VAE (28.538) as the most vital variables on the predictive capability of the model. This result implies the importance of elements associated with scientific output as well as with participatory governance as key drivers in the governance dimension on logistic system efficiency—a finding concordant with recent research on the contributions to investment climates and institutional performance from innovation and democratic accountability [100,108]. ESRPS (23.916) and RLE (20.574) also proved to have substantial importance to identify the instrumental role of leveled-up rights protection and legal pillars. This concurs with existing work pointing to the prediction power of legal-institutional variables in performance modeling in industries such as infrastructure and building [109]. GEE (20.056), RQE (17.422), and PSAOV (16.924), on the other hand, had comparatively low but non-zero impacts. These outcomes highlight the non-uniformity of the governance dimension whereby all governance indicators do not have equal impact on logistic performance. Specifically, the empirical data highlight the disparate influence of knowledge production and accountability mechanisms compared with more conventional governance metrics (Table 16). These add to a finer-grained comprehension of the “G” component’s operationalization of ESG-led logistic performance models and

impart strategic insights into policy design and institution building to augment logistic system capability through governance reforms [100,108].

Table 16. Governance Predictors and Their Influence on LPI: Mean Dropout Loss Analysis.

Variables	Mean Dropout Loss
STJA	29.515
VAE	28.538
ESRPS	23.916
RLE	20.574
GEE	20.056
RQE	17.422
PSAOV	16.924

Note. Mean dropout loss (defined as root mean squared error (RMSE)) is based on 50 permutations.

Following the global feature importance analysis, additive explanation outputs were utilized to dissect the individual contributions of each governance-related predictor toward the Logistic Performance Indicator (LPI) across specific test cases within the k-Nearest Neighbors (k-NN) regression model framework. The base score (i.e., the predicted outcome without the influence of any predictors) remained constant at 10.678 across all instances, allowing for a direct comparison of feature impacts. The use of additive interpretability methods is increasingly recognized as essential for understanding the nuanced behavior of ML models, particularly in k-NN and SVM contexts [110]. Across all five cases analyzed, the Economic and Social Rights Performance Score (ESRPS) consistently exhibited the most substantial negative contributions, with reductions ranging from -16.067 to -14.116 . This indicates a strong inverse relationship between perceived human rights performance and logistic efficiency under the conditions observed—possibly reflecting a trade-off between social equity measures and operational productivity in constrained institutional environments. Simultaneously, the Voice and Accountability Estimate (VAE) demonstrated large positive contributions (ranging from approximately $+7.964$ to $+9.862$), reaffirming its pivotal role as a driver of logistic performance within the governance dimension. The findings support existing literature that links participatory governance with improved infrastructure and service delivery outcomes [111]. Scientific and Technical Journal Articles (STJA) presented a more nuanced pattern, occasionally contributing positively (e.g., $+0.773$ in Case 3) or negatively (e.g., -5.932 in Case 5), suggesting a context-dependent influence, potentially moderated by other institutional or sector-specific factors not captured in the model. Such complex and dynamic relationships are often uncovered through interpretable ML frameworks in health and policy analytics, where variable interactions depend heavily on contextual moderators [112]. Conversely, Government Effectiveness Estimate (GEE) and Regulatory Quality Estimate (RQE) exhibited minor, mostly near-zero impacts on the predicted LPI values, with a notable exception in Case 5, where GEE contributed positively ($+2.120$) and RQE negatively (-1.209). This implies that governance efficacy and regulatory oversight may exert influence only under specific institutional or structural conditions [110]. Political Stability and Absence of Violence (PSAOV) and Rule of Law Estimate (RLE) consistently produced modest effects, albeit with variability in direction and magnitude, highlighting their secondary but non-trivial role. Overall, the additive explanations reinforce the existence of a differentiated structure within the Governance component of the ESG model, where participatory governance (captured through VAE) emerges as the primary positive driver, while human rights considerations (ESRPS) represent a critical constraint. This nuanced insight emphasizes the necessity of selective governance interventions, tailored not merely to

improve aggregate institutional scores but to strategically enhance the most impactful subdimensions for logistic system optimization (Table 17).

Table 17. Additive Feature Contributions to LPI Predictions Using k-NN Model (Governance Dimension).

Case	Predicted	Base	GEE	RQE	ESRPS	VAE	STJA	PSAOV	RLE
1	2.180	10.678	−0.686	−0.046	−16.907	9.827	−0.116	−0.529	−0.042
2	2.370	10.678	0.019	−0.002	−16.008	9.862	−1.659	−0.550	0.029
3	6.203	10.678	0.493	−0.006	−14.116	7.964	0.773	0.172	0.245
4	2.370	10.678	−0.387	−0.015	−16.067	9.473	−0.832	−0.417	−0.062
5	2.503	10.678	2.120	−1.209	−0.858	−1.044	−5.932	−0.239	−1.014

6.3. Clustering Governance Profiles and Their Impact on Logistics Performance

This section examines the relationship between governance quality and logistics performance through an advanced clustering analysis. Using a comprehensive dataset spanning 163 countries from 2007 to 2023, multiple clustering algorithms—including Density-Based, Fuzzy C-Means, Hierarchical, Model-Based, Neighborhood, and Random Forest clustering—were compared across several internal and external validation metrics. Among these, Neighborhood Clustering demonstrated superior performance, achieving the highest R^2 and Calinski-Harabasz scores, alongside strong compactness and separation properties, reflecting the method’s effectiveness in identifying stable and interpretable clusters [112]. The application of Neighborhood Clustering revealed ten distinct clusters characterized by varying governance and logistics performance profiles. Some clusters, particularly those with high government effectiveness, regulatory quality, and voice and accountability, were associated with better logistics outcomes—a pattern consistent with prior spatial and regional analyses emphasizing the link between governance infrastructure and logistics development in economically integrated zones [113]. However, other clusters showed that strong governance indicators alone do not always guarantee superior logistics performance, suggesting the presence of additional mediating factors such as technological capacity, regional integration, or socio-economic disparities. This observation reinforces the need for interpretability and contextual sensitivity in unsupervised learning applications to ensure that model outputs reflect real-world complexities and policy-relevant dynamics [112]. Overall, the clustering analysis underscores the complex, multifaceted relationship between the Governance (G) pillar of ESG and the Logistics Performance Index (LPI), highlighting that institutional quality interacts with a broader set of structural and operational variables to shape outcomes (Table 18).

Table 18. Comparative Evaluation of Clustering Algorithms for Governance and Logistics Performance Analysis.

Metric	Density-Based	Fuzzy c-Means	Hierarchical	Model-Based	Neighborhood	Random Forest
Maximum diameter	0.447	0.740	0.000	0.791	0.057	1.000
Minimum separation	0.997	0.126	0.981	0.061	0.149	0.000
Pearson’s γ	0.805	0.368	1.000	0.283	0.588	0.000
Dunn index	0.492	0.046	1.000	0.000	0.110	0.001
Entropy	0.000	0.674	0.095	0.764	0.668	0.490
Calinski-Harabasz index	0.179	0.248	0.221	0.159	1.000	0.000
R^2	0.130	0.425	0.392	0.297	1.000	0.000
AIC	0.699	0.273	0.327	0.578	0.000	1.000
BIC	0.672	0.258	0.333	0.578	0.000	1.000
Silhouette	0.787	0.328	0.704	0.463	0.598	0.000

The comparative assessment of clustering models was established using multiple internal and external validation indices, such as Maximum Diameter, Minimum Separation, Pearson's γ , Dunn Index, Entropy, Calinski-Harabasz Index, R^2 , AIC, BIC, and Silhouette Score. Each index highlights different clustering performance aspects and therefore provides a multi-aspect basis for model selection [114]. From the assessment, Neighborhood Clustering had better overall performance on most of the key indices. It had the best R^2 value (0.702), which shows the best explanatory strength compared to other models, and had an extraordinary Calinski-Harabasz Index (721.077), indicating exceptional cluster closeness and distinctiveness. In addition to that, Neighborhood Clustering had a good Silhouette Score (0.250), which indicates relatively cohesive clustering structure. The low values of AIC and BIC also reveal high model simplicity and fit and are a strong aspect conducive to its applicability in real-life scenarios requiring simplicity of models [115]. Hierarchical Clustering also had competitive performance, but especially outshines on Pearson's γ (0.618) and the Dunn Index (0.064), revealing good intra-cluster coherence and inter-cluster separation. Nonetheless, its relatively low R^2 and high values of information criteria in comparison to Neighborhood Clustering might confine it to a second option in scenarios requiring maximal predictive stability. The above finding aligns with other comparative research evincing the trade-offs associated with hierarchical approaches [116]. In contrast to the above findings, algorithms like Random Forest Clustering and Model-Based Clustering had multiple shortcomings. Random Forest Clustering had the poorest Silhouette Score (−0.170) and relatively weak R^2 (0.267), which indicates less cohesive cluster formation and weaker explanatory power. Although Model-Based Clustering had relatively good performance on some indices, its negative Silhouette Score (−0.030) is a concern as it questions the clarity of cluster interpretation, a concern commonly espoused when model assumptions fail to match data structure [114]. Density-Based Clustering and Fuzzy c-Means Clustering had varied performance results. While a good Silhouette and Pearson's γ were returned by the Density-Based Clustering model, it had a weaker R^2 and Calinski-Harabasz Index and thus less optimal cluster structures. Fuzzy c-Means Clustering, although having a relatively high Calinski-Harabasz Index, reflected poor cohesion (Silhouette Score = 0.120) and separation (Dunn Index = 0.004), revealing weaker clustering behavior [115]. Overall, on a balanced comparison of cohesion, separation, model fit, and predictive power, Neighborhood Clustering proves to be the best fit algorithm best suited to start with the dataset. The fact that it has been found superior on more than one dimension validates its suitability, especially in applications requiring structural simplicity, model stability, and explanatory power [115,116].

So applying Neighborhood Clustering we have the following results as showed in Table 19.

Table 19. Governance and Logistics Performance: Cluster Characterization via Neighborhood Clustering.

Cluster	1	2	3	4	5	6	7	8	9	10
Size	27	347	236	375	20	9	85	491	385	796
Explained proportion within-cluster heterogeneity	0.036	0.096	0.202	0.128	0.019	0.016	0.074	0.140	0.102	0.189
Within sum of squares	235.437	631.869	1.332.939	848.618	128.765	102.540	488.693	922.706	675.294	1.247.106
Silhouette score	0.391	0.342	0.263	0.149	0.458	0.684	0.315	0.155	0.317	0.243
Center LPI	−0.052	−0.282	3.251	−0.311	−0.276	−0.277	−0.307	−0.302	−0.314	−0.312
Center GEE	−0.851	1.256	0.415	−0.128	0.923	0.398	0.011	0.711	−1.064	−0.534
Center RQE	−0.144	−0.127	0.237	−0.107	2.339	15.237	−0.125	−0.164	−0.051	−0.051
Center ESRPS	−0.480	1.169	0.657	0.159	0.830	−0.000	−0.178	0.543	−1.892	−0.185

Center VAE	−0.744	1.378	0.075	−1.335	0.794	−1.503	0.705	0.520	−0.639	−0.059
Center STJA	0.732	−1.684	−0.117	0.986	−0.415	−0.697	−0.154	−0.205	0.482	0.207
Center PSAOV	−6.510	0.371	0.273	0.012	5.663	−0.853	0.190	0.045	−0.170	−0.126
Center RLE	−0.512	0.190	−0.117	−0.271	0.198	−0.236	4.396	−0.044	−0.246	−0.229

Note. The Between Sum of Squares of the 10 cluster model is 15,546.03. The Total Sum of Squares of the 10 cluster model is 22,160.

The clustering analysis, which was used with the aim to investigate the association of governance indicators and the Logistic Performance Indicator (LPI), demonstrates a sophisticated and subtle form across ten different groups of observations. Each of the groups is identified not merely by size but also by distinctive governance and institutional profiles as indicated by the cluster centers [96]. The most distinctive group is Cluster 3 with its very high LPI center (3.251) and positive centers of GEE (0.415), RQE (0.237), ESRPS (0.657), VAE (0.075), and PSAOV (0.273). The configuration indicates better improvements in different dimensions of governance—such as government effectiveness, regulatory quality, performance in terms of human rights, and stability in politics—are aligned with much superior logistic performance. The relatively high silhouette score of Cluster 3 (0.263) also confirms its internal consistency. These developments are in line with larger research evidence connecting governance quality with better logistic outcomes when governance is combined with technological and administrative advancement [117]. In contrast, most other groups (Clusters 1, 2, 4, 5, 6, 7, 8, 9, and 10) have negative LPI centers and thus depict inferior logistic performance. In all such groups, governance indicators are often both negative and extremely polarized. In Cluster 2, despite positive centers of GEE (1.256), ESRPS (1.169), and VAE (1.378), the LPI center is negative (−0.282). The inconsistency implies that even though indicators of governance are good-looking, other underlying variables such as quality of infrastructure or geographical location unexplained in the model may damp down logistic efficiency [118]. Cluster 5 and Cluster 6 are of particular concern. Cluster 5 has the maximum RQE center (2.339) and a maximum PSAOV center (5.663), which demonstrates high regulatory quality and stability in politics. The corresponding LPI center is still negative (−0.276), which indicates a mismatch between governance improvement and logistic outcomes possibly caused by lag effects or sectoral inefficiencies. Cluster 6 has a small group ($n = 9$) with a maximum silhouette score (0.684), which demonstrates very good internal constancy. Although the group has high stability in politics and relatively neutral profiles of governance indicators, its LPI center is negative (−0.277), showing that even under highly homogeneous circumstances logistic performance is poor. This result confirms that rule of law by itself is insufficient to guarantee logistic success but requires complementarity by economic or infrastructural variables [96]. Cluster 7 presents a remarkable pattern with a strongly positive RLE (4.396) but yet a negative LPI (−0.307), indicating the rule of law as vital but insufficient in itself to guarantee logistic success (Table 20). Generally, the clustering solution documents that rule of law variables are vital but exert complicated and mediated effects on logistic performance. The finding emphasizes the necessity to pursue a multidimensional approach towards logistic success models by complementing rule of law reforms with focused investments in infrastructure, education, and a diversified economy [117,118].

Table 20. Governance Profiles and Their Logistic Outcomes: Cluster Mean Comparisons.

	LPI	GEE	RQE	ESRPS	VAE	STJA	PSAOV	RLE
Cluster 1	−0.480	−0.851	−0.052	−6.510	−0.512	−0.144	0.732	−0.744
Cluster 2	1.169	1.256	−0.282	0.371	0.190	−0.127	−1.684	1.378
Cluster 3	0.657	0.415	3.251	0.273	−0.117	0.237	−0.117	0.075

Cluster 4	0.159	−0.128	−0.311	0.012	−0.271	−0.107	0.986	−1.335
Cluster 5	0.830	0.923	−0.276	5.663	0.198	2.339	−0.415	0.794
Cluster 6	-1.098×10^{-7}	0.398	−0.277	−0.853	−0.236	15.237	−0.697	−1.503
Cluster 7	−0.178	0.011	−0.307	0.190	4.396	−0.125	−0.154	0.705
Cluster 8	0.543	0.711	−0.302	0.045	−0.044	−0.164	−0.205	0.520
Cluster 9	−1.892	−1.064	−0.314	−0.170	−0.246	−0.051	0.482	−0.639
Cluster 10	−0.185	−0.534	−0.312	−0.126	−0.229	−0.051	0.207	−0.059

The cluster means analysis presents differentiated governance and logistic performance profiles among the ten identified groups. The best mean LPI of Cluster 2 (1.169) is accompanied by strong positive governance indicators such as Government Effectiveness (GEE = 1.256) and Voice and Accountability (VAE = 1.378), even with as yet marginally negative Regulatory Quality (RQE = −0.282). These are in line with previous work on the strong performance of groups with high civic engagement and high social capital as regards filling up logistic indices [118]. Similarly, Cluster 5 also presents a high LPI mean (0.830), supported by outstanding scores in Economic and Social Rights Performance (ESRPS = 5.663) and Regulatory Quality (RQE = 2.339) and shows the significance of improved protection of rights and better regulation as drivers to advanced logistic performance—akin to patterns seen in general policy and general supply chain clustering research [96], (Figure 8).

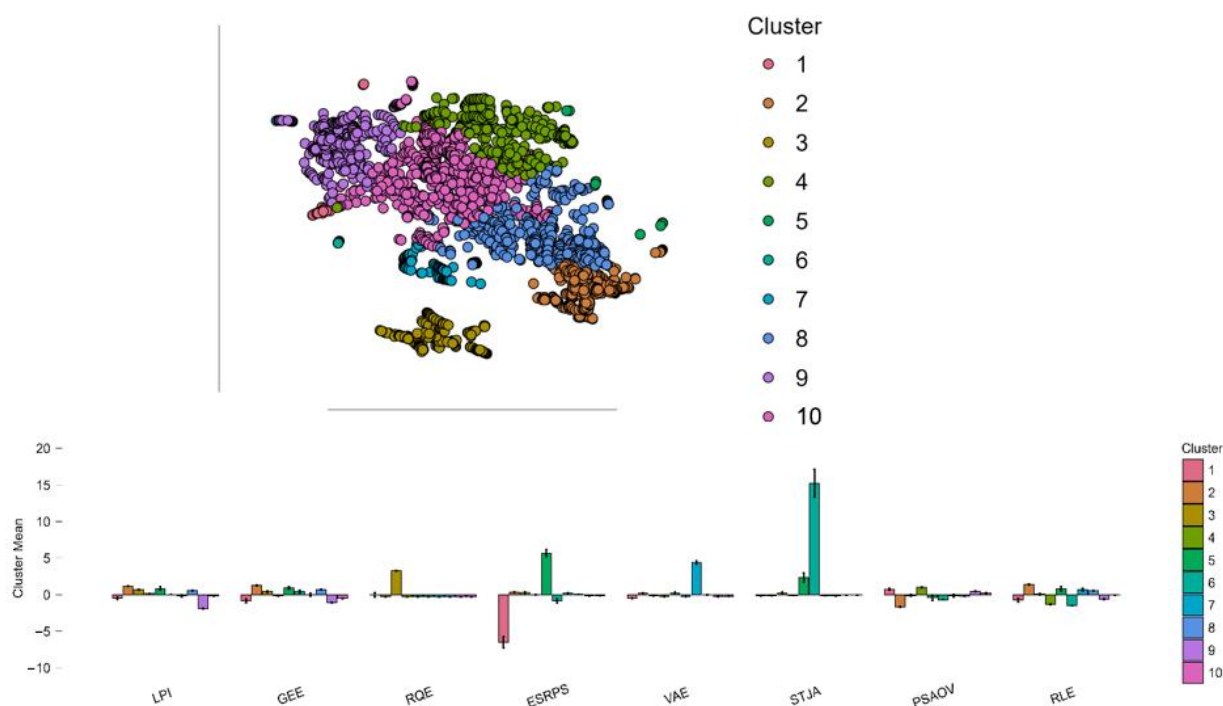


Figure 8. Distribution and Cluster-Wise Means of Governance and Logistics Performance Indicators (LPI) Across Ten Groups.

While Cluster 3 has the best RQE (3.251), it has a marginally high increase in LPI (only 0.657), which validates the fact that the quality of regulation by itself, without concomitant increases in other dimensions of governance, does not holistically optimize logistic performance [119]. Cluster 6 has the maximum divergence in Scientific and Technical Journal Articles (STJA = 15.237), but its LPI mean is near zero (almost), and it shows that scientific production as desirable may have to be accompanied by better governance to contribute meaningfully to logistic system improvements. Clusters 1, 9, and

10 are characterized by low means in LPI (-0.480 , -1.892 , and -0.185) and by overall poor governance indicators. In particular, Cluster 9 has the pessimistic profile with strongly negative values on all the variables GEE, RQE, ESRPS, VAE, and RLE and demonstrates the synergies of low governance to logistic low efficiency [119]. Interestingly, Cluster 7 positions itself strongly positive on RLE ($RLE = 4.396$) but has a negative LPI mean (-0.178), and it shows that legal structures by themselves are insufficient to propel logistic performance because other dimensions of governance are missing. Finally, Cluster 8 is a comparatively balanced configuration with a relatively high LPI (0.543) and mean scores on all dimensions of governance and has a more integrated model of governance [96]. All of these combined outcomes verify the multifaceted character of logistic performance as a phenomenon in which discrete governance elements make a non-uniform contribution and require synergistic enhancement to make noticeable improvements [118], (Figure 9).

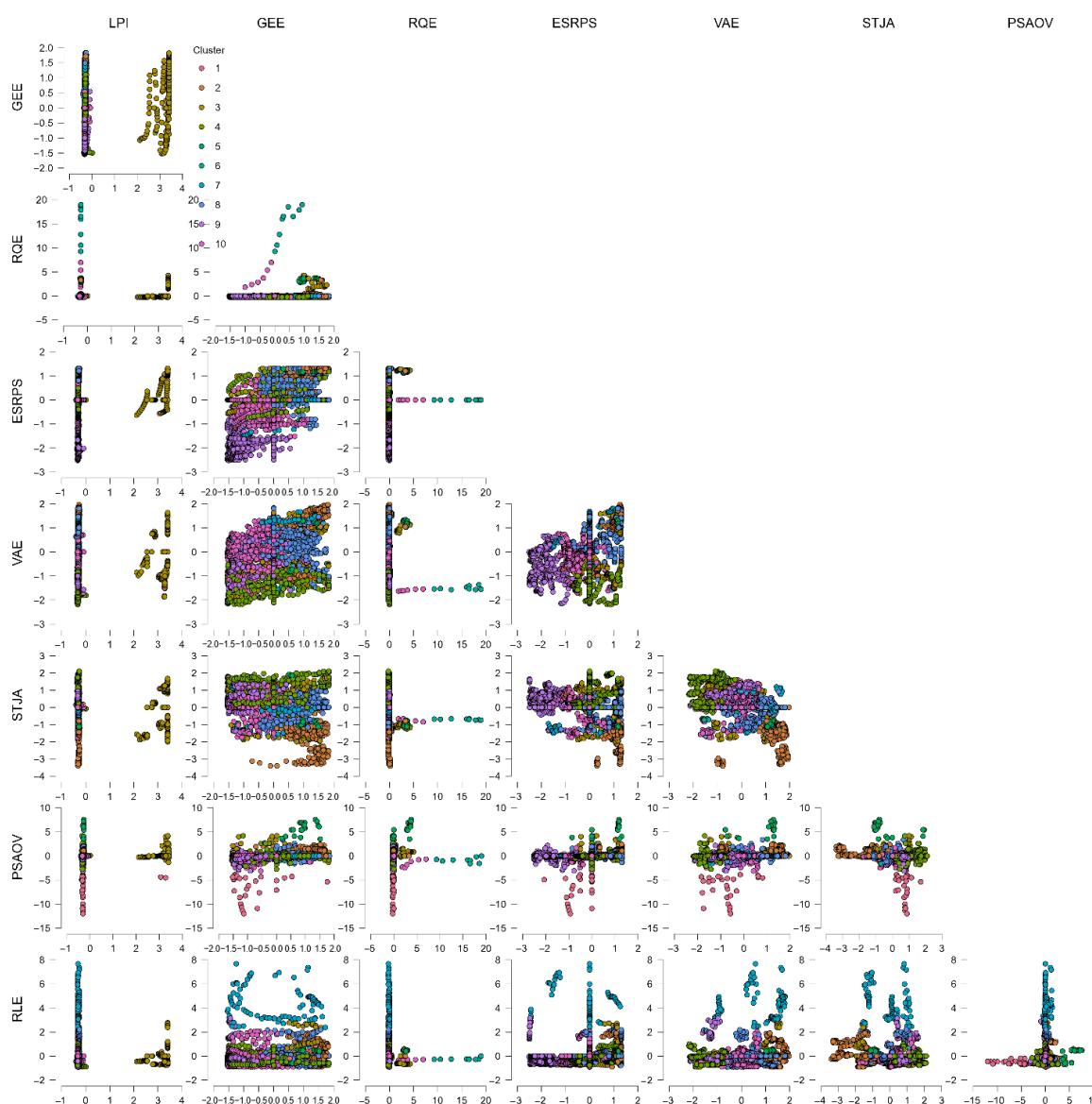


Figure 9. Pairwise Relationships Among Governance and Logistics Indicators by Cluster.

7. Policy Implications

The results of this research have vast implications for policymakers seeking to coordinate improvements in Logistics Performance with overall Environmental, Social, and Governance factors. The outcomes of this research reveal strong correlations between

improvements in the Logistics Performance Index (LPI) and key factors in the Environmental, Social, and Governance domain, thereby validating current idioms regarding the essential role of logistics infrastructure in Sustainable Development Policies [120]. From a governance perspective, the strong correlations between LPI and other factors such as Government Effectiveness (GEE) and Regression Estimate of Regulation Quality (RQE) highlight the need for more open and effective governance in logistics [121]. From a social perspective, the study shows that effective logistics networks support overall achievement of social rights, as evidenced by Economic and Social Rights Performance Scores (ESRPS). This indicates that effective logistics networks contribute positively to the access of necessary products and services in society, underscoring the strategic need to align logistics network investment with clear social aims. An environmental perspective analysis indicates that effective improvements in the sector should be aligned with robust environmental regulatory frameworks and incentives that neutralize the negative environmental impacts of supply chains [120]. The clustering analysis also illustrates the diversified patterns of country-level ESG and logistics achievements, thus emphasizing the need for differentiated approaches [96]. Countries with weaker LPI and poor ESG outcomes should focus on institutional and infrastructural improvements, whereas more successful countries should focus on further improving environmental and social sustainability dimensions [5]. The significance of scientific achievement (STJA) in improving logistics achievement also draws attention in this study, suggesting that innovation and research strategies should align with or follow overall logistics improvements [4]. Finally, political stability and overall values (PSAOV) become essential factors in ensuring effective logistics systems, underscoring the need for intersectoral governance and policymaking.

Theoretical implications. Several theoretical implications of this analysis emerge as a consequence of its results. Among them, the most central implication lies within the assumption of ‘two-sided logistics’ that not only captures its enabling dimensions but is also sensitive to any negative externality that might emerge as a consequence of deregulation [122]. From a theoretical perspective, this report clearly shows that, through the use of appropriate causal analysis, logistics, as well as sustainability, move beyond traditional linear or unidirectional notions. Rather, they assume multiple-layered, structurally defined trade-off forms that transform and evolve in specifically dynamic ways [123]. The application of instrumental variables with machine learning tools in logistics and sustainability pushes beyond conventional descriptive or correlation-based literature, opening theoretically innovative dimensions in defining logistics and sustainability as theoretically robust. Logistics, on the one hand, can be regarded as enabling infrastructure for achieving logistics sustainability, whereas, on the other hand, negative externalities may emerge with unregulated logistics [10]. Empirical confirmations of the Logistics Performance Index (LPI) as well as its improvements that contribute, along with their effects, towards achieving logistics sustainability exist. Improvements in LPI will enhance social sustainability, such as better education or the absence of child labor in some industries and countries. However, LPI, along with logistics, may cause negative social imbalances or inequalities that may call for structural and meaningful corrective measures [43]. Moreover, environmental sustainability may emerge due to the promotion of clean technologies along with more effective resource management. However, negative environmental imbalances, such as pollution, may emerge from such LPI, along with their effects that may call for structural and meaningful remedies [123]. Thus, theories recognize the application of “two-sided logistics sustainability” that relies heavily on, as well as being dependent on, some extremely specific assumptions: that logistics may enhance social sustainability in multiple aspects (improving educations or eliminating child labor in some industries) or may generate

social inequities requiring attention, along with contributing towards environmental sustainability in some ways, however, with the possibility of environmental pollutions [10,122]. Further, this analysis refutes the existing theoretical consideration that takes a marginal stance concerning the role of environmental and demographic variables in the field of logistics. However, the causal dynamics of air pollutants, GHG, heat stress, and logistics efficiency identify the importance of theoretically formulating logistics systems as units that respond to environmental dynamics [124]. This goes beyond the existing theoretical focus that considers infrastructure, cost, and trading volume. An extra theoretical implication relates to the formulation of ESG variables. By employing a disaggregated framework, there would be the possibility of proving that the interaction of the three variables of ESG with logistics system dynamics differs in ways that would not be recognized via the aggregate metric [43,124]. Finally, in light of the above, the interaction of machine learning with econometric insights within the methodology is a material theoretical component, clearly showing that large-scale sustainability dynamics are, in fact, nonlinear, such that country-level profiles must be factored in for their effective probing. In other words, this theoretical exercise again vindicates the need for hybrid studies in the theory of supply chain dynamics and sustainability [122,123]. Moving ahead, a practical policy implication would be that carbon pricing policies must be adopted with a focus on keeping corporations on their toes in terms of carbon emissions, along with providing incentives for more investments in green technology. At the same time, green freight routes may significantly enhance the efficiency of logistics along with its positive effects on the environment [10].

Practical implications. A number of the results have implications that are important for policymakers, decision-making bodies in international forums, investment circles, and managers responsible for logistics. Recent studies suggest that improvements in logistics do not always yield positive outcomes for environmental, social, and governance (ESG) factors, requiring policy measures to reconcile efficiency with sustainability [125]. However, it seems entirely valid that improvements in logistics will increasingly necessitate more focused strategies, as greater efficiency may increase costs in terms of sustainability. With regard to environmental aspects, stronger links in the logistics system exhibit a strong positive correlation with NO_x emissions, whereas air pollutants, specifically particulate matter (PM_{2.5}), negatively affect efficiency. These findings align with existing empirical insights that link logistics practices with material environmental spillovers, such as higher NO_x levels, when effective sustainability-mitigating strategies are not considered [19,126]. This further explains that any project linked to logistics practices must effectively address comprehensive environmental protection, including decarbonization, environmentally friendly transportation, ecologically safe warehouses, and climate-proof infrastructure, as mentioned by [19]. Socially, the topic highlights the importance of incorporating human development considerations into logistics planning. Specifically, the positive causal link between Child Labor (CET) and the Logistics Performance Index (LPI) across different societies highlights that some logistics improvements may have stemmed from problematic social behaviors. These results confirm that logistics improvement must operate within the parameters of responsibility, so that the development of the value chain occurs in a fair, humane manner [126]. Also, the negative impacts of demographics, such as aging societies or increasing school enrollment, on logistics efficiency indicate that the adoption of effective active labor policies, as well as the promotion of automated technology, would be necessary as a buffer against changes in the unavailability of the workforce [127]. Thus, social logistics and sustainable logistics practices must be seen as structural aspects that ensure the supply chain's resilience, rather than mere afterthoughts. With regard to governance, nations with better quality governance, defined as better regulation, rule of law, accountability, as

well as scientific capabilities, show better logistics capabilities [125]. This explicitly indicates that betterment in the Logistics Performance Index requires improvements in governance quality [19]. Machine learning analyses provide additional helpful tools for decision-makers. Prediction algorithms such as Random Forests or K-Nearest Neighbors (KNN) may help governments and companies predict logistics efficiency. In addition, cluster analysis helps identify different national profiles. This will allow global organizations and development finance institutions to develop strategies grounded in concrete environmental or structural issues [126]. In conclusion, the implications of the mentioned dimensions are that Sustainable Logistics requires that the processes of infrastructure development, environmental issues, social aspects, and the quality of institutions must move forward together. Thus, logistics transformation must be incorporated into broader sustainable development paradigms so that efficiency for both the economy and the environment can be achieved [125,127].

8. Discussion

In this regard, the analysis illustrates the complex but intricate link between logistics performance and the trilogy of E(S)G issues with sufficient evidence that shows that improvements in the Logistics Performance Index (LPI) positively correlate with environmental, social, as well as governance issues, although with different degrees of intensity [4,124]. Regardless of the models, be it economics, AI, or cluster models, the fact remains that logistics efficiency is both driven as well as a driver of sustainability, with the LPI being the mediator that facilitates the link between environmental sustainability, social issues, and good governance [43]. Environmentally, the findings indicate a strong trade-off among economic development, resource use, and environmental conditions. There is a positive correlation between higher LPI rankings and greater GHG emissions, such as nitrogen oxides, indicating that as economic development occurs, associated carbon-intensive practices often increase. This observation aligns with the traditional diseconomies observed in early-stage industrialization, where increased transport infrastructure, warehousing, and freight activity lead to higher associated GHG emissions [15]. At the same time, environmental degradation, particularly air pollutants such as PM2.5, significantly costs the logistics system. Air pollutants contribute to poor logistics performance, leading to reduced productivity and logistical infrastructure disruptions [43].

It is further found that there is a positive correlation between extreme heat exposure and LPI, which is explained by the efforts of hot-region countries that invested significantly in resilient logistics structures to overcome inefficiencies caused by weather conditions [124]. Land use factors further interact with this aspect, such that a higher percentage of agricultural land is associated with poorly developed logistics infrastructure, but higher value addition in agriculture is associated with better logistics efficiency. This difference indicates that subsistence agriculture does not enhance logistics efficiency, but a more commercialized agricultural sector promotes investment in the Cold Chain, exports, and transport [4]. References confirm the importance of environmental variables, pinpointing agricultural land, nitrous oxide, PM2.5 air quality, and heat stress as the top determinants of LPI [43]. Taken together, this literature collectively suggests that, contrary to the topic's periphery, environmental issues play a fundamental, defining role within the efficiency of logistics, finding that any improvement within this realm will necessarily require addressing issues of environmental degradation as well as those of climate change [15]. Socially, the implication of this finding is that there is a complex interaction of positive and negative effects, pointing towards a convergence of logistics performance with human development. Higher LPI scores indicate improved education, as evident through higher education enrollment and fewer children in the workforce, as the workforce, processes, and technology required in logistics systems depend on human

knowledge, competencies, and technology-enabled human resource capabilities, thus their development in societies with positive education dynamics [4]. However, income inequality appears to be a robust negative factor for logistics performance, suggesting that societies with income inequality will suffer from inefficient human resource allocation, unconsolidated service value, and inadequate infrastructure accessibility. Simultaneously, other dimensions, such as an aging population and imbalanced access to basic services, appear as weak negative determinants of logistics efficiency. Such results indicate that, despite human development, logistics development may increase inequalities in societies, which may further strain them if proper social policies do not channel appropriate attention [15]. The results confirm that logistics performance extends beyond techno-infrastructure aspects, encompassing social cohesion, human capital development, and equal opportunities for economic participation [124]. In terms of governance, the results are more definitive, being strongly positive. There is a positive correlation between higher governance, as indicated by more effective governments, the rule of law, the quality of regulation, scientific productivity, and logistics performance. This occurs because the enabling effects of governance shape proper regulations, customs policies, the enforceability of contracts, and stable environments that attract investment in logistics infrastructure [43]. Voice and accountability, combined with active scientific production, further enhance the enabling effects of innovation-modified logistics systems [124]. Similarly, machine learning models identify governance variables as robust predictors of LPI, with cluster analyses affirming that nations with strong governance, as identified via systematic clustering, position closely with strong logistics practices [4]. These indicate that governance as a consideration is essential, as it promotes efforts toward logistics modernization, allowing better alignment between the scale of logistics development and ESG principles [15]. In any case, the conclusions of this research indicate that logistics performance is firmly positioned at the center of the dynamics of ESG. Environmental considerations call for proper management via greener innovation as well as more climate-resilient infrastructure, with the social factor pinpointing that logistics systems always remain strong within more equitable, more educated societies, contrary to effects that generate reduced efficiency within more inequitable, poorly developed societies, with governance proving as a cornerstone much more closely associated with improved logistics performance [43,124]. See Table 21.

Table 21. Summary of ESG Components and Their Principal Relationships with Logistics Performance.

ESG Component	Main Relationships Identified	Direction of Effects	Interpretation
Environmental (E)	Nitrous oxide emissions rise alongside improvements in logistics performance; PM2.5 air pollution reduces logistics efficiency; exposure to extreme heat correlates positively with LPI; larger agricultural land share is associated with weaker logistics systems; greater agricultural value added improves logistics performance.	Mixed effects: both positive and negative depending on the indicator.	Logistics modernization is linked to higher emissions, showing development–pollution trade-offs; environmental degradation harms logistics, while commercialized agriculture and climate-adapted systems support efficiency.
Social (S)	Higher education levels and increased school enrollment correspond with stronger logistics performance; reductions in child labor correlate with higher LPI; income inequality negatively	Generally positive for education-related variables and negative for inequality or demographic strain.	Social conditions shape logistics capabilities through human capital quality, workforce stability, and access to services, but unequal or

	influences logistics performance; aging populations and gaps in basic services exert moderate negative effects.		aging societies face structural barriers to logistics efficiency.
Governance (G)	Government effectiveness, rule of law, regulatory quality, and scientific research productivity all show strong positive correlations with logistics performance; better governance systems support modern customs procedures and transparent institutional environments.	Strongly positive.	Governance quality is a foundational enabler of logistics performance, reinforcing institutional stability, innovation, and efficient regulatory environments that support logistics modernization.

9. Conclusions

This research undertakes a systemic examination of the multifaceted relationship between ESG outcomes and logistics performance and makes a valuable addition to the available research by combining econometric panel data methods with machine learning algorithms. Contrary to prior research and its tendency to typically discuss logistics and sustainability as two distinct areas or to limit itself to aggregate indices, this research breakingly examines the ESG dimensions separately and analyzes how infrastructures of logistics are interwoven with and influence each pillar in a large sample of countries during a long period. Empirical estimates derived from instrumental variable (IV) regressions demonstrate systematically how a higher Logistics Performance Index (LPI) is related to multiple aspects of sustainable development. In the environmental pillar area, better logistics performance exhibits a twofold character: in addition to promoting resource efficiency and mitigating certain types of pollution, it also correlates with higher levels of greenhouse gas emissions and thus with environmental dimensions as a consequence of infrastructure expansion and industry development. In the social pillar area, better logistics performance correlates with better education, less child labor, and wider accessibility to basic services but risks causing negative effects related to inequalities as well. In terms of governance, more robust logistics systems are found to support better institutional quality and more scientific productivity, more robust rule of law and more participative governance arrangements. The use of machine learning models, i.e., of Random Forest and k-Nearest Neighbors algorithms by applying them to regression and Neighborhood-Based and Density-Based clustering to unsupervised modeling, supports and confirms the results of the econometric models. These methodologies confirm both the predictive power of key ESG indicators but also reveal latent data structures and enhance the multifaceted interconnection between logistic capabilities and targets of sustainable development. The clustering analysis in particular identifies the presence of diverging country profiles where certain groups of countries achieve both better logistics performance and better ESG outcomes at the same time and others are caught in a vicious circle of low efficiency in logistics and weak sustainability indicators. Most importantly, the research indicates that while the development of logistics is a necessary condition to modernize the economy and integrate into the global economy, it does not necessarily translate into good ESG outcomes. Unless complemented by policies on environmental protection, social inclusion, and good governance, gains in the performance of logistics would risk making existing sustainability issues worse. The findings therefore highlight the imperatives of coordinated policy schemes to align investments in logistics with ESG priorities to guarantee that investments in improving infrastructures are used to increase economic efficiency as much as to bring about fair, resilient, and sustainable development. In conclusion, it underlines the importance of logistics systems as more than technical or economic enablers but as key drivers of larger sustainability pathways. Future studies will

need to delve deeper into causal processes by which the interplay between logistics and ESG results occurs with possibly more detailed data by region or industry and further expanding the methodology to dynamic machine learning methods and causal inference models. Policymakers and international organizations need to acknowledge that investment in sustainable logistic infrastructures is a strategic means towards fulfilling the United Nations Sustainable Development Goals and enabling a shift towards a more sustainable and environmentally responsible global economy.

Author Contributions: Conceptualization, N.M., V.N., M.D.M., S.M. and A.L.; Methodology, N.M., V.N., M.D.M., S.M. and A.L.; Software, N.M., V.N., M.D.M., S.M. and A.L.; Validation, N.M., V.N., M.D.M., S.M. and A.L.; Formal analysis, N.M., V.N., M.D.M., S.M. and A.L.; Investigation, N.M., V.N., M.D.M., S.M. and A.L.; Resources, N.M., V.N., M.D.M., S.M. and A.L.; Data curation, N.M., V.N., M.D.M., S.M. and A.L.; Writing – original draft, N.M., V.N., M.D.M., S.M. and A.L.; Writing – review & editing, N.M., V.N., M.D.M., S.M. and A.L. All authors have read and agreed to the published version of the manuscript.

Funding: The proposed work has been developed within the framework of the project “Logistics 4.0” (Regional call of Apulia for aid under exemption No. 17 of 30/09/2014—BURP No. 139 supplement of 06/10/2014 and subsequent amendments—TITLE II, CHAPTER 2 OF THE GENERAL REGULATION, “Notice for the submission of projects promoted by Large Enterprises pursuant to Article 17 of the Regulation”).

Data Availability Statement: We used data from the World Bank database at the links below: <https://esgdata.worldbank.org/?lang=en> (accessed 4 January 2025) and <https://lpi.worldbank.org/> (accessed 2 January 2025).

Conflicts of Interest: Nicola Magaletti, Valeria Notarnicola, Mauro Di Molfetta, Stefano Mariani, and Angelo Leogrande were employed by the company LUM Enterprise S.r.l. The authors declare no conflict of interest.

Abbreviations

LPI	Logistic Performance Index
AAGRPCI	Annualized average growth rate in per capita real survey mean consumption or income, total population (%)
ACFTC	Access to clean fuels and technologies for cooking (% of population)
AFFVA	Agriculture, forestry, and fishing, value added (% of GDP)
AFWT	Annual freshwater withdrawals, total (% of internal resources)
ALPA	Agricultural land (% of land area)
ASFD	Adjusted savings: net forest depletion (% of GNI)
ASNRD	Adjusted savings: natural resources depletion (% of GNI)
CDD	Cooling Degree Days
CET	Children in employment, total (% of children ages 7–14)
CO2E	CO ₂ emissions (metric tons per capita)
CODCDMPN	Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total)
EILPE	Energy intensity level of primary energy (MJ/\$2017 PPP GDP)
ESRPS	Economic and Social Rights Performance Score
EU	Energy use (kg of oil equivalent per capita)
FFEC	Fossil fuel energy consumption (% of total)
FPI	Food production index (2014–2016 = 100)
FRT	Fertility rate, total (births per woman)
GDPG	GDP growth (annual %)
GEE	Government Effectiveness: Estimate
GEET	Government expenditure on education, total (% of government expenditure)
GHGLUCF	GHG net emissions/removals by LUCF (Mt of CO ₂ equivalent)
GI	Gini index

HB	Hospital beds (per 1000 people)
HDD	Heating Degree Days
HI35	Heat Index 35
ISL20	Income share held by lowest 20%
IUI	Individuals using the Internet (% of population)
LEBT	Life expectancy at birth, total (years)
LFPRT	Labor force participation rate, total (% of population ages 15–64) (modeled ILO estimate)
LRAT	Literacy rate, adult total (% of people ages 15 and above)
LST	Land Surface Temperature
LWS	Level of water stress: freshwater withdrawal as a proportion of available freshwater resources
MRU5	Mortality rate, under-5 (per 1000 live births)
MST	Mammal species, threatened
NM	Net migration
NOE	Nitrous oxide emissions (metric tons of CO ₂ equivalent per capita)
PD	Population density (people per sq. km of land area)
PHRNPL	Poverty headcount ratio at national poverty lines (% of population)
PM2.5AE	PM2.5 air pollution, mean annual exposure (µg/m ³)
POA	Prevalence of overweight (% of adults)
PSAOV	Political Stability and Absence of Violence/Terrorism: Estimate
PSHWNP	Proportion of seats held by women in national parliaments (%)
PSMS	People using safely managed sanitation services (% of population)
PSMWS	People using safely managed drinking water services (% of population)
REC	Renewable energy consumption (% of total final energy consumption)
RFMLFPR	Ratio of female to male labor force participation rate (%) (modeled ILO estimate)
RLE	Rule of Law: Estimate
RQE	Regulatory Quality: Estimate
SEP	School enrollment, primary (% gross)
SLRI	Strength of legal rights index (0 = weak to 12 = strong)
SPEI	Standardized Precipitation–Evapotranspiration Index
STJA	Scientific and technical journal articles
TMPA	Terrestrial and marine protected areas (% of total territorial area)
VAE	Voice and Accountability: Estimate

Appendix A. Hyper Parameters of Regression Algorithms

Table A1. Support Vector Machine Hyperparameters.

Category	Option	Setting
Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Training and Validation Data-Sample	20% for validation data
Training Parameters	Weights	Linear
	Degree (for polynomial kernel)	3
	Gamma parameter	1
	r parameter	0
	Tolerance of termination criterion	0.001
	Epsilon	0.01
	Scale features	Enabled
	Set seed	1
Costs of Constraints Violation	Costs settings	Optimized
	Max. violation cost	5

Table A2. Regularized Linear Regression Hyperparameters.

Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Training and Validation Data-Sample	20% for validation data
Training Parameters	Penalty	Lasso
	Include intercept	Enabled
	Scale features	Enabled
	Set seed	1
Lambda (λ) Settings	Selection	Optimized
	Fixed value (if selected)	1 (not selected)
	Largest λ within 1 SE of min	Disabled

Table A3. Random Forest Regression Hyper parameters.

Split Preferences	Holdout Test Data-Sample	20% of all data
	Training and Validation Data-Sample	20% for validation data
Training Parameters	Training data used per tree	50%
	Features per split	Auto
	Scale features	Enabled
	Set seed	1
Number of Trees	Tree selection	Optimized
	Maximum number of trees	100

Table A4. Linear Regression Hyperparameters.

Category	Option	Setting
Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Add generated indicator to data	Disabled
	Test set indicator	None (not selected)
Training Parameters	Include intercept	Enabled
	Scale features	Enabled
	Set seed	1

Table A5. K-Nearest Neighbors Regression Hyperparameters.

Category	Option	Setting
Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Add generated indicator to data	Disabled
	Test set indicator	None (not selected)
Training and Validation Data	Validation Sample	20% for validation data
	K-fold	Disabled
	Leave-one-out	Disabled
Training Parameters	Weights	Rectangular
	Distance	Euclidean
	Scale features	Enabled
	Set seed	1
Number of Nearest Neighbors	Selection Method	Optimized
	Max. nearest neighbors	10
	Fixed nearest neighbors	Disabled

Table A6. Decision Tree Regression-Hyperparameters.

Category	Option	Setting
Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Add generated indicator to data	Disabled
	Test set indicator	None (not selected)

Training and Validation Data	Validation Sample	20% for validation data
	K-fold	Disabled
	Leave-one-out	Disabled
Training Parameters	Min. observations for split	20
	Min. observations in terminal node	7
	Max. interaction depth	30
	Scale features	Enabled
	Set seed	1
Tree Complexity	Penalty Type	Optimized
	Max. complexity penalty	1
	Fixed complexity penalty	Disabled (value: 0.01 grayed out)

Table A7. Boosting Regression Hyperparameters.

Category	Option	Setting
Data Split Preferences	Holdout Test Data-Sample	20% of all data
	Add generated indicator to data	Disabled
	Test set indicator	None (not selected)
Training and Validation Data	Validation Sample	20% for validation data
	K-fold cross-validation	Disabled
Training Parameters	Shrinkage	0.1
	Interaction depth	1
	Minimum observations in node	10
	Training data used per tree	50%
	Loss function	Gaussian
	Scale features	Enabled
	Set seed	1
Number of Trees	Tree selection	Optimized
	Maximum number of trees	100
	Fixed number of trees	Disabled (value: 100 grayed out)

Appendix B. Hyper Parameters of Clustering Algorithms

Table A8. Density-Based Clustering hyper parameters.

Parameter	Value	Description
Epsilon neighborhood size	2	Maximum distance to include points in a point's neighborhood (ϵ)
Min. core points	5	Minimum number of points required to form a core point
Distance	Normal	Type of distance used (likely Euclidean)
Scale features	Enabled	Features are scaled (normalized or standardized)
Set seed	Disabled	No seed set for result reproducibility

Table A9. Fuzzy C-Means Clustering hyper parameters.

Category	Parameter	Value	Description
Algorithmic Settings	Max. iterations	25	Maximum number of iterations allowed during optimization
	Fuzziness parameter	2	Degree of fuzziness in fuzzy clustering (e.g., Fuzzy C-Means)
	Scale features	Enabled	Features are scaled (standardized or normalized)
	Set seed	Disabled	No random seed set for reproducibility
Cluster Determination	Determination method	Optimized according to BIC	Number of clusters determined by Bayesian Information Criterion (BIC)
	Max. clusters	10	Maximum number of clusters to consider in optimization
	Clusters (Fixed)	3 (disabled)	Fixed cluster number is not used

Table A10. Hierarchical Clustering hyper parameters.

Parameter	Value	Description
Epsilon neighborhood size	2	Maximum distance to include points in a point's neighborhood (ϵ)
Min. core points	5	Minimum number of points required to form a core point
Distance	Normal	Type of distance used (likely Euclidean)
Scale features	Enabled	Features are scaled (normalized or standardized)
Set seed	Disabled	No seed set for result reproducibility

Table A11. Model-based Clustering hyper parameters.

Parameter	Value	Description
Center type	Means	Type of cluster center used (centroids)
Algorithm	Hartigan-Wong	Algorithm variant used for clustering (K-Means method)
Distance	Euclidean	Distance metric used for clustering
Max. iterations	25	Maximum number of iterations allowed
Random sets	25	Number of random initializations for better clustering
Scale features	Enabled	Features are scaled (standardized or normalized)
Set seed	Disabled	No random seed set for reproducibility
Cluster determination	Optimized (BIC)	Number of clusters determined using Bayesian Information Criterion (BIC)
Max. clusters	10	Maximum number of clusters to evaluate
Fixed clusters	Disabled (3 shown)	Fixed number of clusters not selected

Table A12. Neighborhood-Based.

Parameter	Value	Description
Model	Auto	Automatically selects the best clustering model
Max. iterations	25	Maximum number of iterations for model fitting
Scale features	Enabled	Features are scaled (standardized or normalized)
Set seed	Disabled	No seed set for reproducibility
Cluster determination	Optimized (BIC)	Number of clusters selected based on Bayesian Information Criterion (BIC)
Max. clusters	10	Maximum number of clusters to evaluate
Fixed clusters	Disabled (3 shown)	Fixed number of clusters not used

Table A13. Random Forest Clustering hyper parameters.

Parameter	Value	Description
Model	Auto	Automatically selects the best clustering model
Max. iterations	25	Maximum number of iterations for model fitting
Scale features	Enabled	Features are scaled (standardized or normalized)
Set seed	Disabled	No seed set for reproducibility
Cluster determination	Optimized (BIC)	Number of clusters selected based on Bayesian Information Criterion (BIC)
Max. clusters	10	Maximum number of clusters to evaluate
Fixed clusters	Disabled (3 shown)	Fixed number of clusters not used

Appendix C. E-Environmental Summary Statistics

Table A14. Descriptive statistics of environmental (E) variables used in the analysis.

	LPI	NOE	PM2.5AE	HI35	ALPA	AFFVA
Valid	2771	2771	2771	2771	2771	2771
Missing	0	0	0	0	0	0
Mode	100.000	−21.265	5.179	67.170	100.000	83.890
Median	2.760	−21.265	4.830	67.170	72.900	83.890

Mean	10.854	−21.265	5.177	67.169	65.608	83.909
Std. Error of Mean	0.497	1.688	0.054	0.330	0.686	0.154
95% CI Mean Upper	11.828	−17.955	5.283	67.817	66.954	84.211
95% CI Mean Lower	9.880	−24.576	5.071	66.521	64.262	83.607
Std. Deviation	26.155	88.873	2.852	17.387	36.134	8.104
95% CI Std. Dev. Upper	26.862	91.277	2.929	17.857	37.111	8.324
95% CI Std. Dev. Lower	25.484	86.593	2.779	16.941	35.207	7.897
Coefficient of variation	2.410	−4.179	0.551	0.259	0.551	0.097
MAD	0.380	0.000	1.250	0.000	27.100	0.000
MAD robust	0.563	0.000	1.853	0.000	40.178	0.000
IQR	0.940	26.786	2.210	0.000	64.900	0.000
Variance	684.069	7.898	8.133	302.294	1.305	65.683
95% CI Variance Upper	721.575	8.331	8.579	318.868	1.377	69.284
95% CI Variance Lower	649.425	7.498	7.721	286.985	1.239	62.356
Skewness	2.986	−4.217	2.322	−1.261	−0.650	−3.304
Std. Error of Skewness	0.047	0.047	0.047	0.047	0.047	0.047
Kurtosis	6.984	26.820	7.759	3.475	−1.067	18.588
Std. Error of Kurtosis	0.093	0.093	0.093	0.093	0.093	0.093
Shapiro–Wilk	0.336	0.551	0.796	0.754	0.826	0.483
<i>p</i> -value of Shapiro–Wilk	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Range	99.810	1.044.803	23.950	100.000	99.800	77.688
Minimum	0.190	−944.893	1.110	0.000	0.200	22.312
Maximum	100.000	99.910	25.060	100.000	100.000	100.000
25th percentile	2.460	−21.265	3.380	67.170	35.100	83.890
50th percentile	2.760	−21.265	4.830	67.170	72.900	83.890
75th percentile	3.400	5.521	5.590	67.170	100.000	83.890
25th percentile	2.460	−21.265	3.380	67.170	35.100	83.890
50th percentile	2.760	−21.265	4.830	67.170	72.900	83.890
75th percentile	3.400	5.521	5.590	67.170	100.000	83.890
Sum	30.076	−58.926	14.345	186.125	181.800	232.511

Table A15. Covariances of environmental variables.

	NOE	PM2.5AE	HI35	ALPA	AFFVA	LPI
NOE	7.898	−26.637	−119.086	5.641	−36.715	−23.566
PM2.5AE	−26.637	8.133	−7.922	−22.657	−2.142	−3.410
HI35	−119.086	−7.922	302.294	245.815	17.297	86.509
ALPA	5.641	−22.657	245.815	1.305	88.102	209.274
AFFVA	−36.715	−2.142	17.297	88.102	65.683	11.491
LPI	−23.566	−3.410	86.509	209.274	11.491	684.069

Table A16. Correlations of environmental variables.

	NOE	PM2.5AE	HI35	ALPA	AFFVA	LPI
NOE	1.000	−0.105	−0.077	0.002	−0.051	−0.010
PM2.5AE	−0.105	1.000	−0.160	−0.220	−0.093	−0.046
HI35	−0.077	−0.160	1.000	0.391	0.123	0.190
ALPA	0.002	−0.220	0.391	1.000	0.301	0.221
AFFVA	−0.051	−0.093	0.123	0.301	1.000	0.054
LPI	−0.010	−0.046	0.190	0.221	0.054	1.000

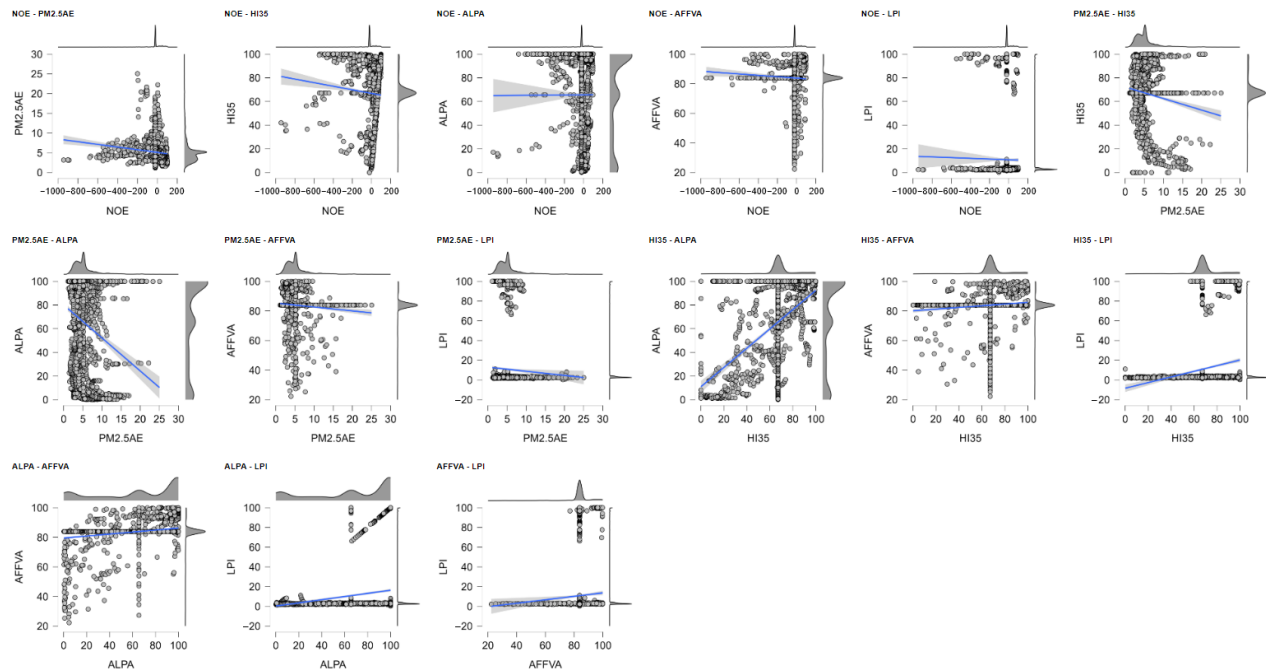


Figure A1. Pairwise relationships and distributional patterns among environmental (E) variables and the Logistics Performance Index (LPI).

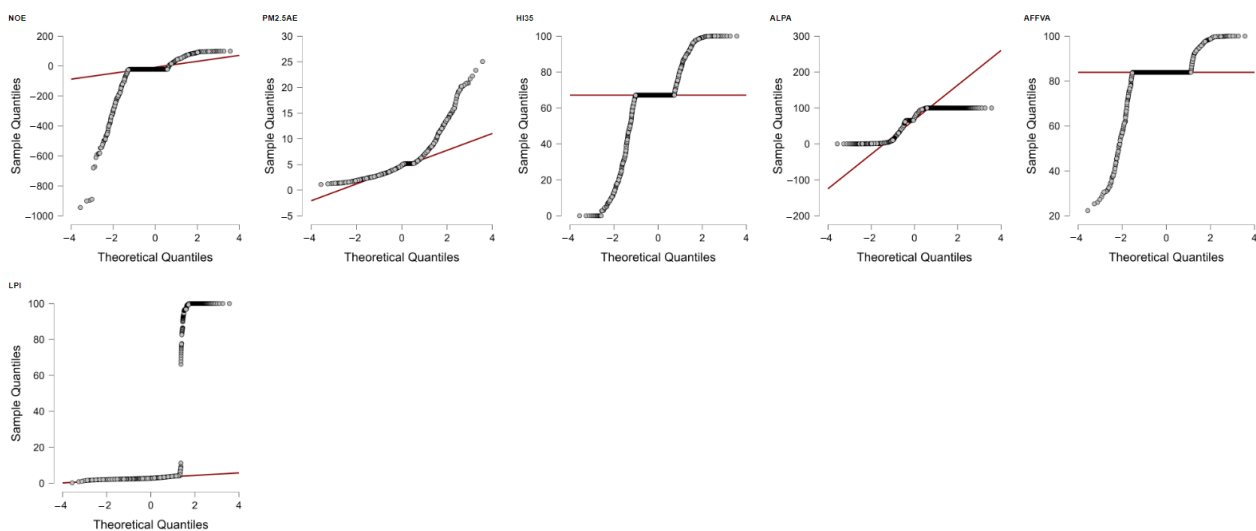


Figure A2. Quantile–quantile (Q–Q) plots assessing the distributional properties of environmental (E) variables and the Logistics Performance Index (LPI). The red line represents the theoretical normal distribution reference line. Deviations of the empirical quantiles from this line indicate departures from normality in the environmental (E) variables and the Logistics Performance Index (LPI).

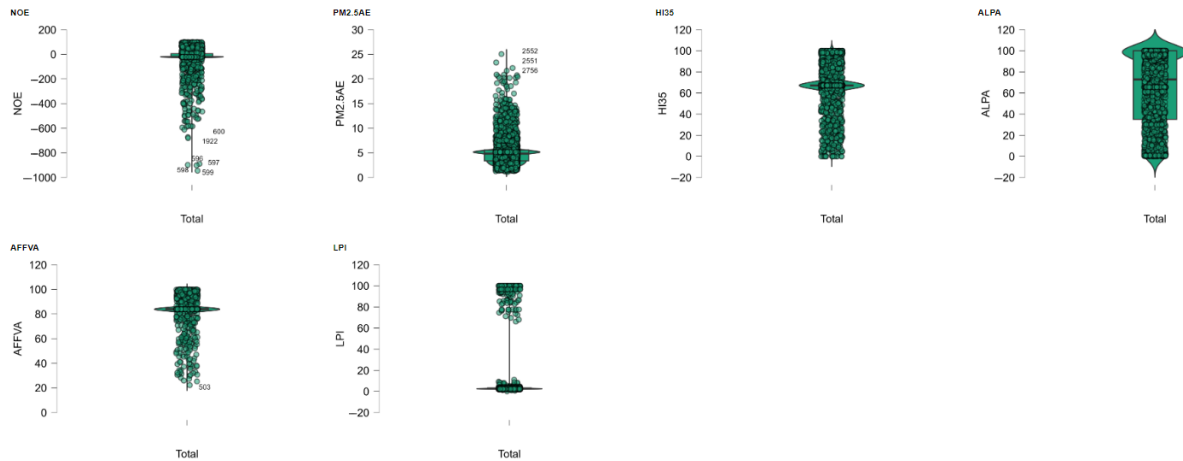


Figure A3. Boxplots illustrating the distribution of environmental (E) variables and the Logistics Performance Index (LPI).

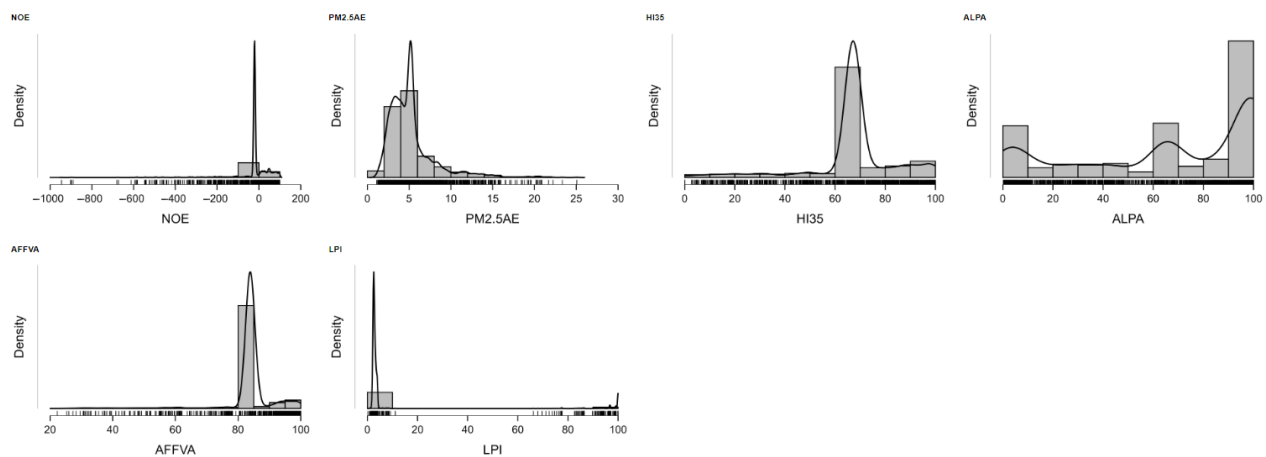


Figure A4. Distribution plots with kernel density estimates illustrating the distributions of environmental (E) variables and the Logistics Performance Index (LPI).

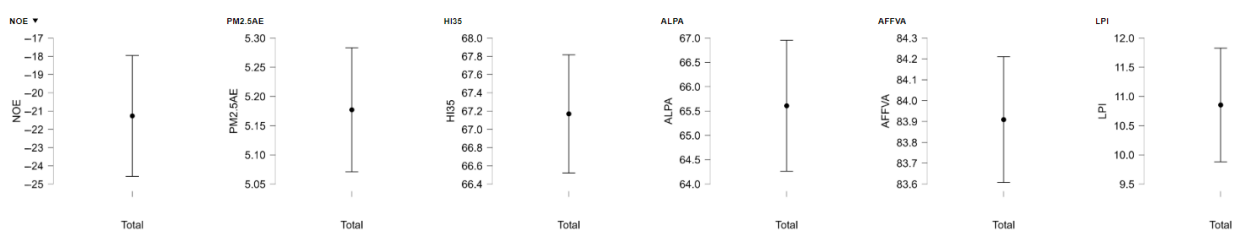


Figure A5. Interval plots showing the mean values and confidence intervals of environmental (E) variables and the Logistics Performance Index (LPI).

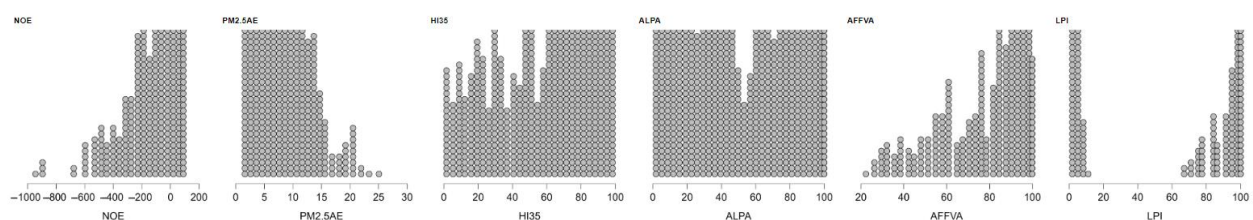


Figure A6. Dot plots illustrating the distribution of environmental (E) variables and the Logistics Performance Index (LPI).

Appendix D. S-Social Summary Statistics.

Table A17. Descriptive Statistics.

	LPI	PSMWS	PSMS	PA65A	SEP	CET	POA	ISL20
Valid	2771	2771	2771	2771	2771	2771	2771	2771
Missing	0	0	0	0	0	0	0	0
Mode	100.000	2.500	70.660	7.346	−0.176	−0.115	21.092	−0.046
Median	2.760	8.700	75.890	5.759	−0.175	−0.256	20.800	−0.089
Mean	10.854	10.523	70.649	7.346	−0.176	−0.115	21.081	−0.046
Std. Error of Mean	0.497	0.188	0.368	0.102	0.018	0.019	0.218	0.019
95% CI Mean Upper	11.828	10.892	71.370	7.547	−0.141	−0.079	21.508	−0.010
95% CI Mean Lower	9.880	10.154	69.928	7.146	−0.212	−0.152	20.654	−0.083
Std. Deviation	26.155	9.899	19.355	5.380	0.950	0.983	11.467	0.978
95% CI Std. Dev. Upper	26.862	10.167	19.879	5.526	0.976	1.009	11.777	1.005
95% CI Std. Dev. Lower	25.484	9.645	18.859	5.242	0.925	0.958	11.173	0.953
Coefficient of variation	2.410	0.941	0.274	0.732	−5.389	−8.514	0.544	−21.161
MAD	0.380	5.700	10.464	2.491	0.635	0.644	7.526	0.686
MAD robust	0.563	8.451	15.514	3.693	0.941	0.955	11.158	1.017
IQR	0.940	9.400	23.747	5.737	1.279	1.352	14.860	1.381
Variance	684.069	97.998	374.634	28.946	0.902	0.966	131.488	0.957
95% CI Variance Upper	721.575	103.371	395.175	30.533	0.952	1.019	138.698	1.010
95% CI Variance Lower	649.425	93.035	355.661	27.480	0.857	0.917	124.829	0.909
Skewness	2.986	2.056	−1.021	1.417	−0.582	0.424	0.612	0.153
Std. Error of Skewness	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.047
Kurtosis	6.984	5.374	0.578	1.875	0.053	−0.490	0.238	−0.539
Std. Error of Kurtosis	0.093	0.093	0.093	0.093	0.093	0.093	0.093	0.093
Shapiro–Wilk	0.336	0.768	0.919	0.873	0.970	0.966	0.971	0.985
<i>p</i> -value of Shapiro–Wilk	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Range	99.810	68.400	99.177	28.880	4.933	4.716	63.750	4.800
Minimum	0.190	2.500	7.345	0.100	−3.313	−2.591	0.000	−2.548
Maximum	100.000	70.900	106.522	28.980	1.620	2.125	63.750	2.252
25th percentile	2.460	2.500	60.459	3.688	−0.714	−0.823	12.531	−0.741
50th percentile	2.760	8.700	75.890	5.759	−0.175	−0.256	20.800	−0.089
75th percentile	3.400	11.900	84.206	9.425	0.566	0.529	27.391	0.640
25th percentile	2.460	2.500	60.459	3.688	−0.714	−0.823	12.531	−0.741
50th percentile	2.760	8.700	75.890	5.759	−0.175	−0.256	20.800	−0.089
75th percentile	3.400	11.900	84.206	9.425	0.566	0.529	27.391	0.640
Sum	30.076	29.159	195.768	20.356	−488.409	−319.845	58.416	−128.115

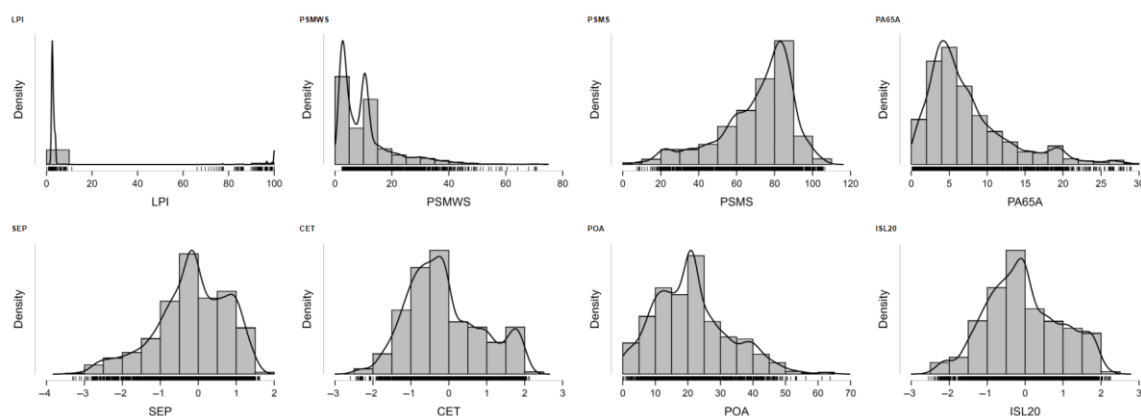


Figure A7. Histograms with kernel density estimates illustrating the distributions of social (S) indicators and the Logistics Performance Index (LPI).

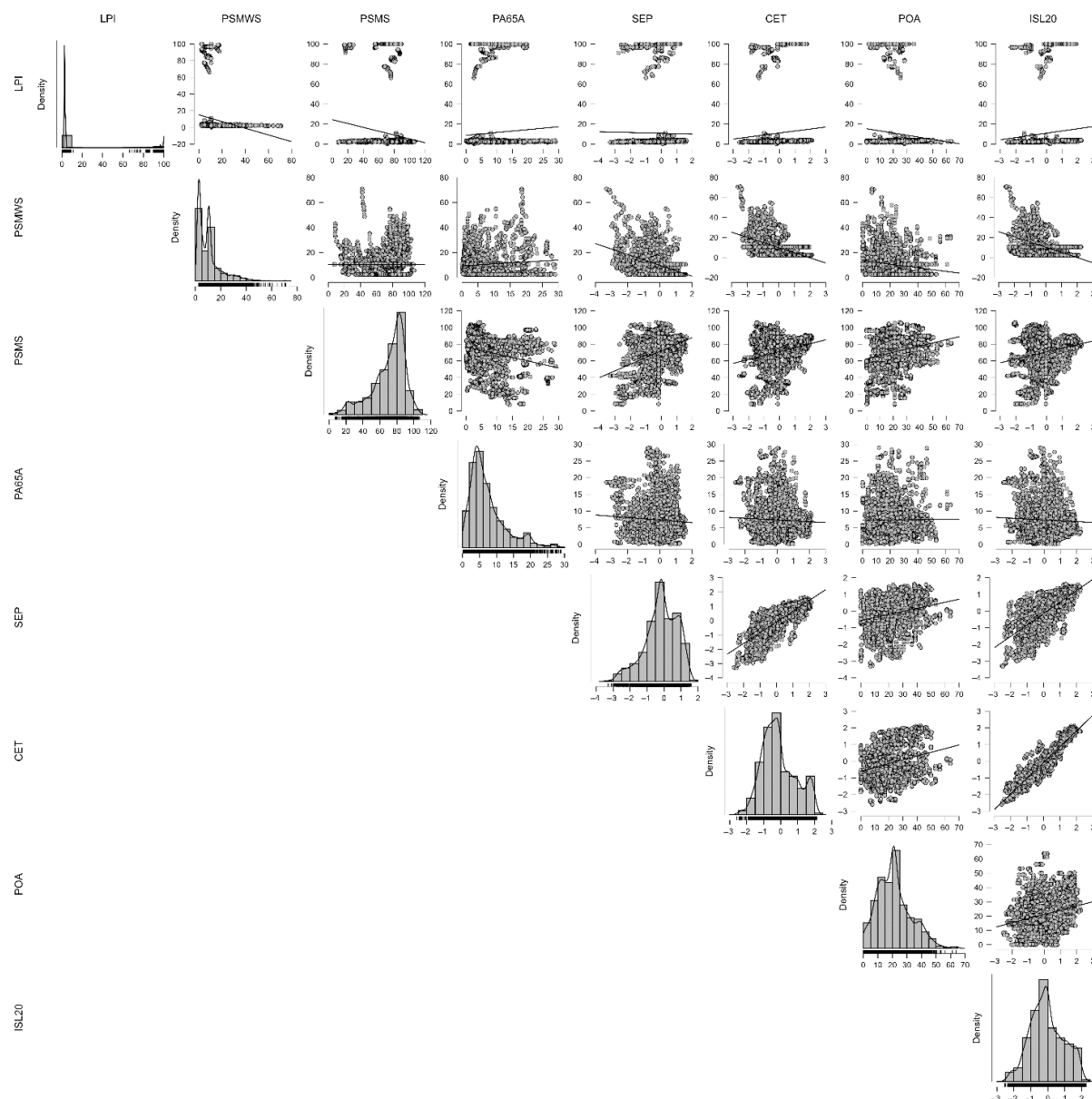


Figure A8. Correlation plot illustrating pairwise relationships among social (S) indicators and the Logistics Performance Index (LPI).

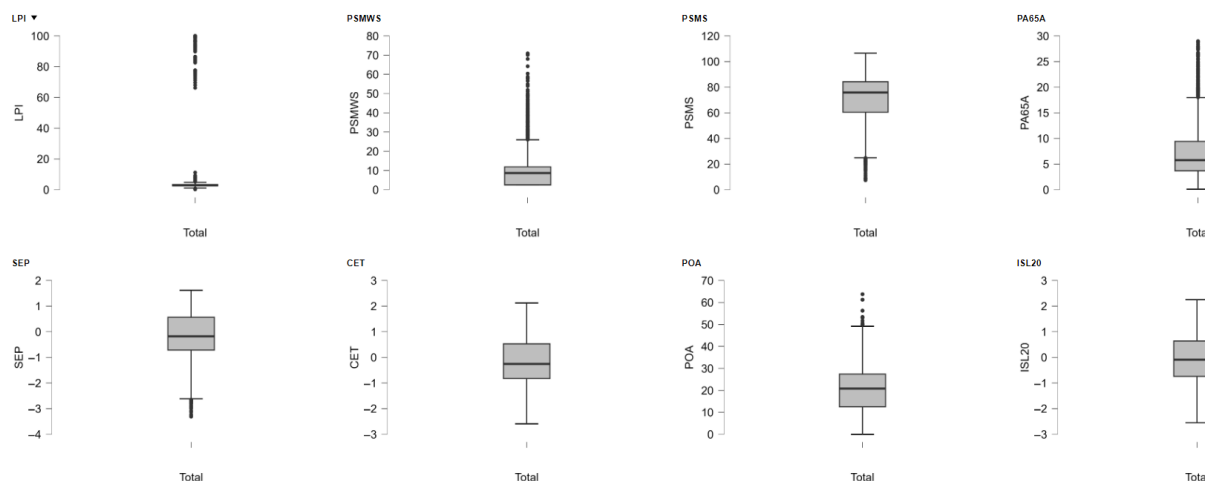


Figure A9. Box plots illustrating the distribution of social (S) indicators and the Logistics Performance Index (LPI).

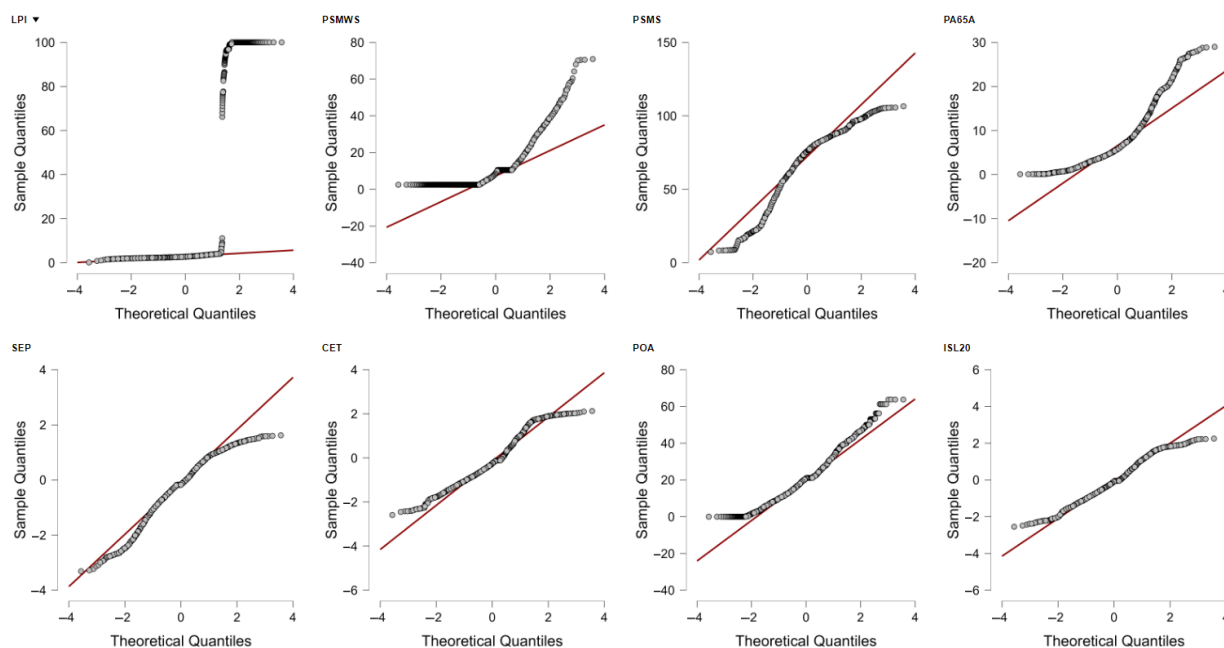


Figure A10. Quantile–quantile (Q–Q) plots assessing the distributional properties of social (S) indicators and the Logistics Performance Index (LPI).

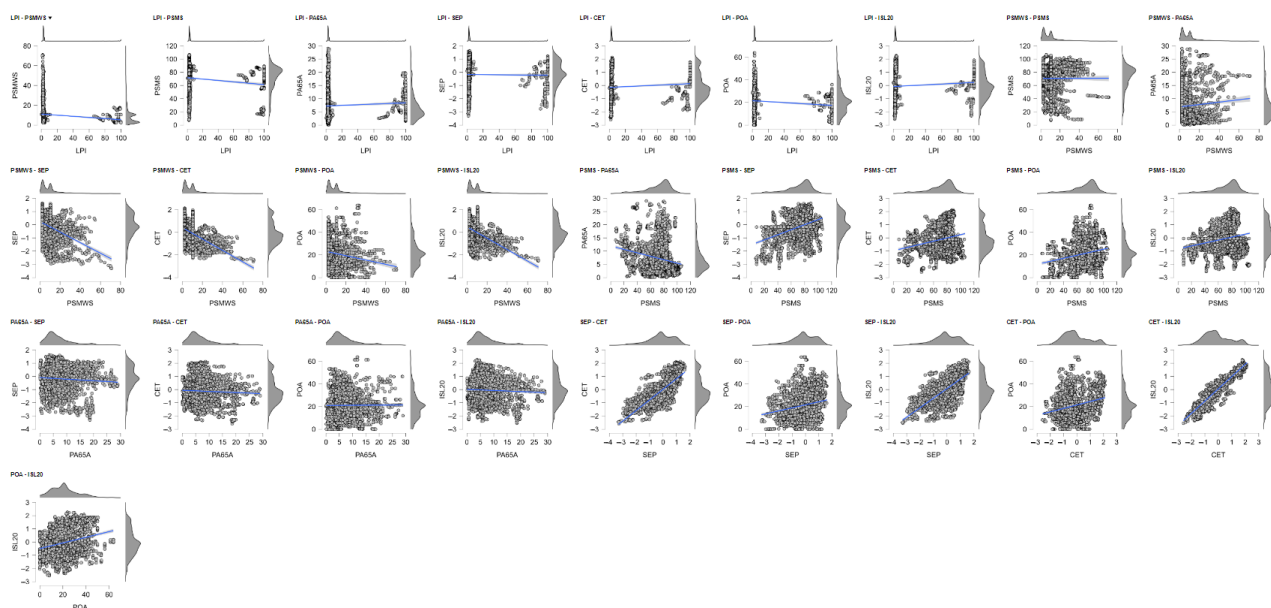


Figure A11. Scatter plots illustrating pairwise relationships among social (S) indicators and the Logistics Performance Index (LPI).

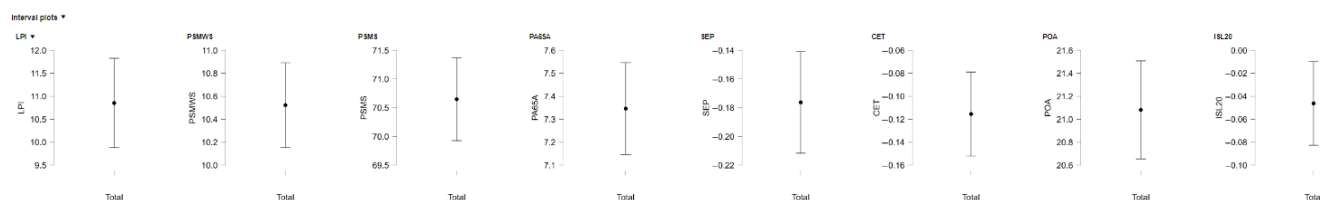


Figure A12. Interval plots showing mean estimates and confidence intervals of social (S) indicators and the Logistics Performance Index (LPI). The points represent sample means, while the vertical lines denote confidence intervals, providing insight into the central tendency and variability of each social component across the dataset.

Table A18. Covariance matrix of social (S) indicators and the Logistics Performance Index (LPI).

	LPI	PSMWS	PSMS	PA65A	SEP	CET	POA	ISL20
LPI	684.069	−39.318	−71.754	8.334	−0.339	1.931	−28.347	2.105
PSMWS	−39.318	97.998	0.399	4.457	−3.922	−4.956	−18.778	−4.947
PSMS	−71.754	0.399	374.634	−24.405	7.209	4.557	49.932	4.168
PA65A	8.334	4.457	−24.405	28.946	−0.350	−0.248	0.497	−0.241
SEP	−0.339	−3.922	7.209	−0.350	0.902	0.725	2.395	0.652
CET	1.931	−4.956	4.557	−0.248	0.725	0.966	2.941	0.899
POA	−28.347	−18.778	49.932	0.497	2.395	2.941	131.488	2.807
ISL20	2.105	−4.947	4.168	−0.241	0.652	0.899	2.807	0.957

Table A19. Correlation matrix of social (S) indicators and the Logistics Performance Index (LPI).

	LPI	PSMWS	PSMS	PA65A	SEP	CET	POA	ISL20
LPI	1.000	−0.152	−0.142	0.059	−0.014	0.075	−0.095	0.082
PSMWS	−0.152	1.000	0.002	0.084	−0.417	−0.509	−0.165	−0.511
PSMS	−0.142	0.002	1.000	−0.234	0.392	0.240	0.225	0.220
PA65A	0.059	0.084	−0.234	1.000	−0.069	−0.047	0.008	−0.046
SEP	−0.014	−0.417	0.392	−0.069	1.000	0.776	0.220	0.702
CET	0.075	−0.509	0.240	−0.047	0.776	1.000	0.261	0.935
POA	−0.095	−0.165	0.225	0.008	0.220	0.261	1.000	0.250
ISL20	0.082	−0.511	0.220	−0.046	0.702	0.935	0.250	1.000

Appendix E. G-Governance Summary Statistics

Table A20. Descriptive statistics of governance (G) indicators and the Logistics Performance Index (LPI).

	LPI	GEE	RQE	ESRPS	VAE	STJA	PSAOV	RLE
Valid	2771	2771	2771	2771	2771	2771	2771	2771
Missing	0	0	0	0	0	0	0	0
Mode	100.000	45.761	16.506.	67.641	−0.137	25.367	912.876	0.543
Median	2.760	45.761	8.307	67.641	−0.137	26.437	912.876	0.422
Mean	10.854	45.766	16.506	67.641	−0.137	25.367	918.963	0.543
Std. Error of Mean	0.497	0.562	1.410	0.469	0.018	0.171	3.626	0.011
95% CI Mean Upper	11.828	46.868	19.273	68.562	−0.101	25.702	8.030	0.564
95% CI Mean Lower	9.880	44.665	13.739	66.721	−0.174	25.032	−6.192.	0.521
Std. Deviation	26.155	29.572	74.273	24.706	0.970	8.989	190.909	0.578
95% CI Std. Dev. Upper	26.862	30.372	76.282	25.374	0.997	9.232	196.073	0.594
95% CI Std. Dev. Lower	25.484	28.814	72.368	24.072	0.945	8.758	186.012	0.563
Coefficient of variation	2.410	0.646	4.500	0.365	−7.064	0.354	207.745	1.065
MAD	0.380	26.861	8.199	15.044	0.756	5.062	13.807	0.139
MAD robust	0.563	39.825	12.156	22.305	1.121	7.504	20.471	0.207
IQR	0.940	53.576	16.340	27.922	1.505	11.438	26.629.000	0.289
Variance	684.069	874.520	$5.517 \times 10^{+9}$	610.394	0.942	80.803	$3.645 \times 10^{+10}$	0.334
95% CI Variance Upper	721.575	922.468	$5.819 \times 10^{+9}$	643.861	0.993	85.233	$3.844 \times 10^{+10}$	0.353
95% CI Variance Lower	649.425	830.231	$5.237 \times 10^{+9}$	579.481	0.894	76.711	$3.460 \times 10^{+10}$	0.318
Skewness	2.986	0.077	13.755	−0.715	−0.015	−0.688	−2.333	3.636
Std. Error of Skewness	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.047
Kurtosis	6.984	−1.196	221.461	−0.002	−0.881	0.577	45.528	16.473

Std. Error of Kurtosis	0.093	0.093	0.093	0.093	0.093	0.093	0.093	0.093
Shapiro–Wilk	0.336	0.944	0.141	0.894	0.977	0.957	0.478	0.609
<i>p</i> -value of Shapiro–Wilk	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Range	99.810	99.783	$1.427 \times 10^{+6}$	94.525	4.034	49.571	$3.740 \times 10^{+6}$	4.964
Minimum	0.190	0.217	1.000	5.475	−2.259	−5.258	$−2.290 \times 10^{+6}$	0.018
Maximum	100.000	100.000	$1.427 \times 10^{+6}$	100.000	1.775	44.313	$1.449 \times 10^{+6}$	4.982
25th percentile	2.460	18.005	166.000	61.593	−0.909	19.500	−17.033	0.254
50th percentile	2.760	45.761	8.307	67.641	−0.137	26.437	912.876	0.422
75th percentile	3.400	71.581	16.506	89.516	0.596	30.938	9.596	0.543
25th percentile	2.460	18.005	166.000	61.593	−0.909	19.500	−17.033	0.254
50th percentile	2.760	45.761	8.307	67.641	−0.137	26.437	912.876	0.422
75th percentile	3.400	71.581	16.506	89.516	0.596	30.938	9.596	0.543
Sum	30.076	126.817	$4.574 \times 10^{+7}$	187.434	−380.677	70.291	$2.546 \times 10^{+6}$	1.504

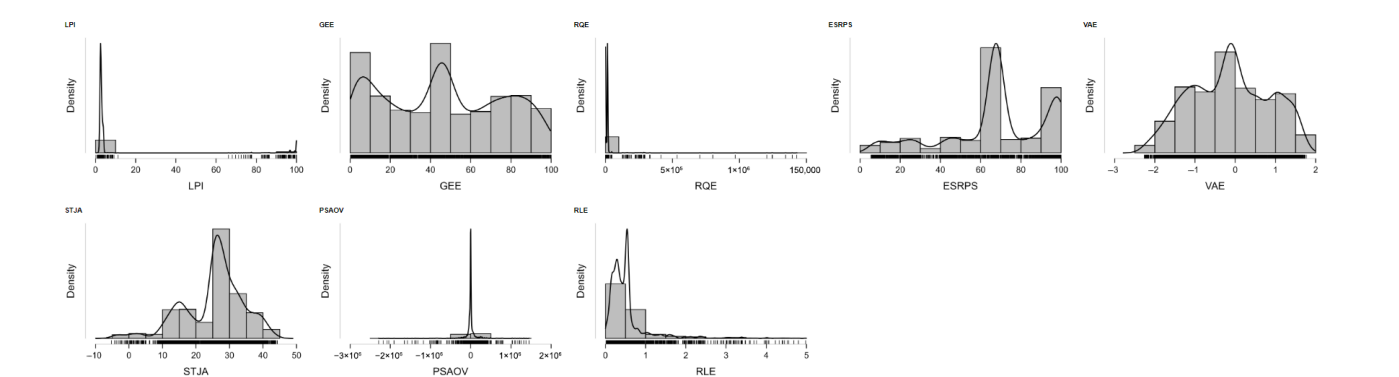


Figure A13. Histograms with kernel density estimates illustrating the distributions of governance (G) indicators and the Logistics Performance Index (LPI).

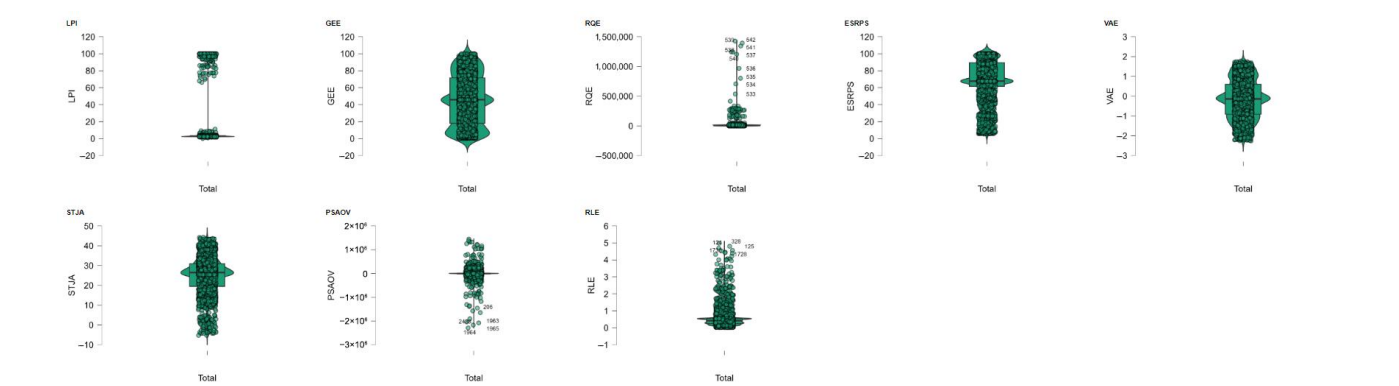


Figure A14. Violin plots illustrating the distribution of governance (G) indicators and the Logistics Performance Index (LPI).

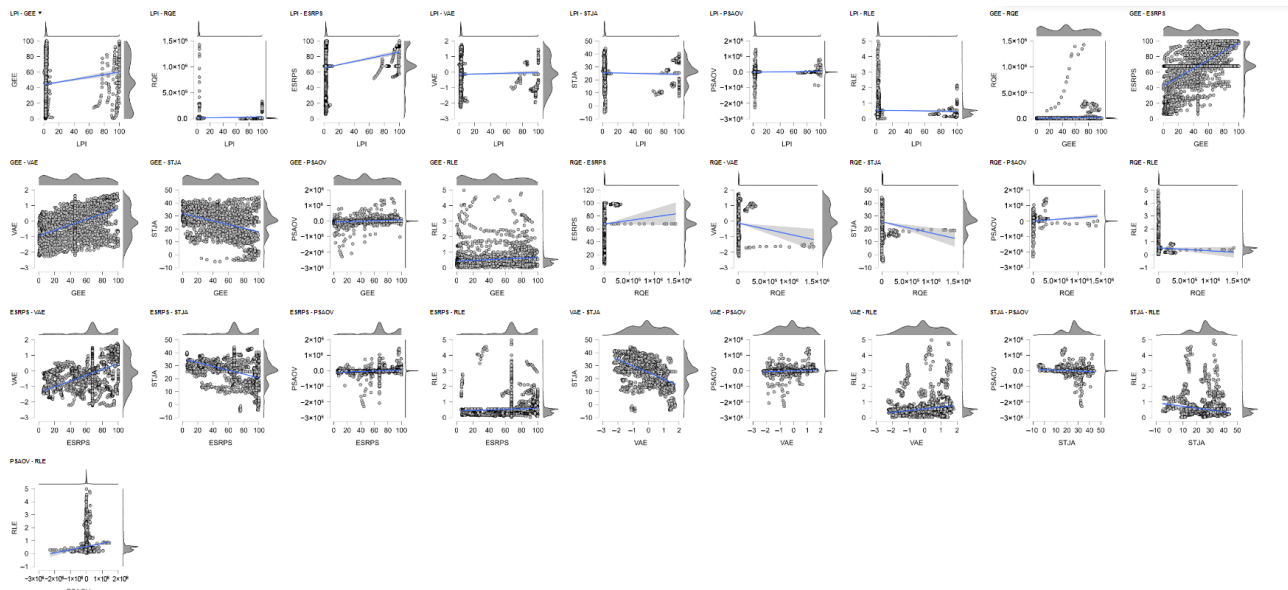


Figure A15. Scatter plots illustrating pairwise relationships among governance (G) indicators and the Logistics Performance Index (LPI).

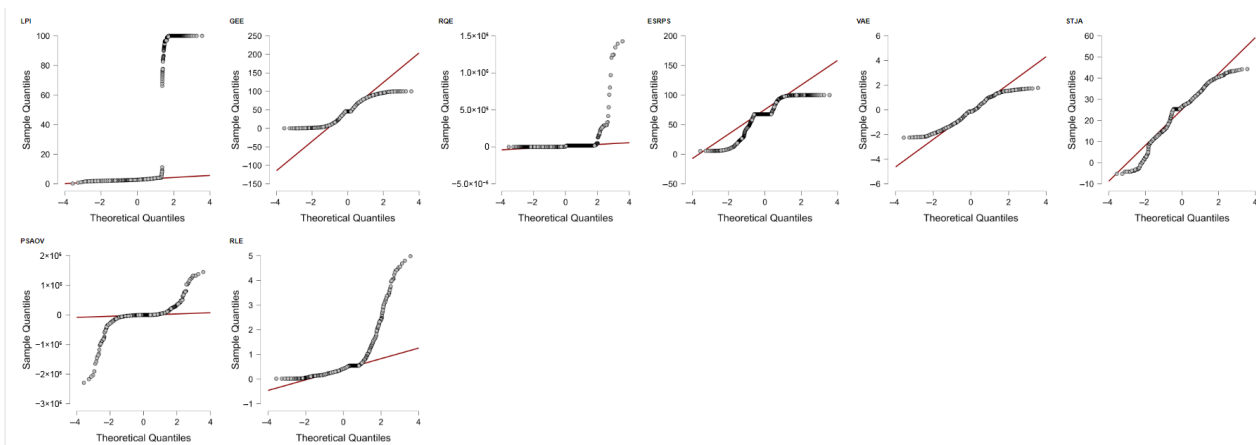


Figure A16. Quantile–quantile (Q–Q) plots assessing the distributional properties of governance (G) indicators and the Logistics Performance Index (LPI).

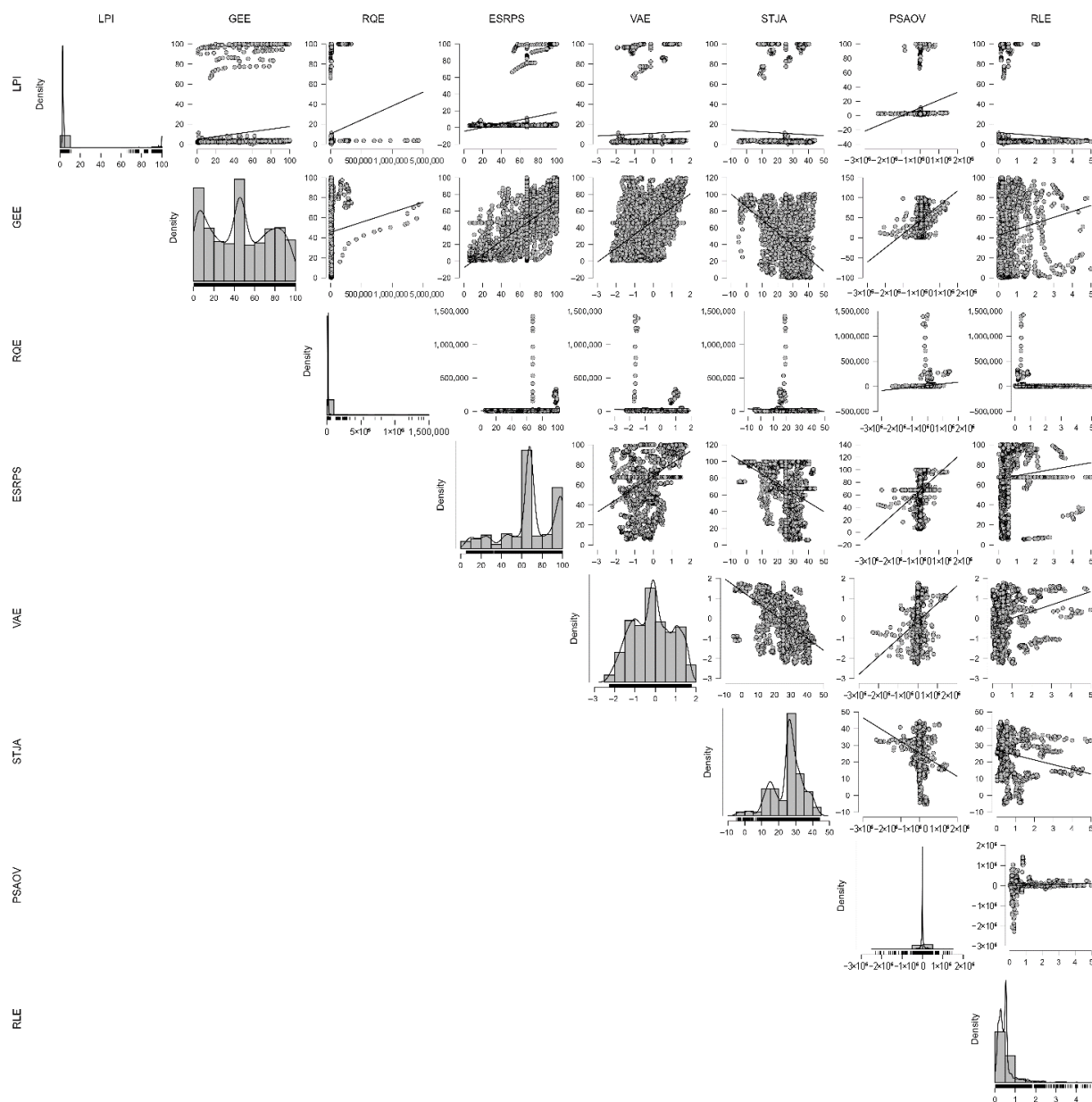


Figure A17. Correlation plot illustrating pairwise relationships among governance (G) indicators and the Logistics Performance Index (LPI). Grey dots represent individual observations, grey shaded areas indicate the marginal distributions of each variable, and the solid black line denotes the fitted linear trend highlighting the direction and strength of the pairwise relationships among governance (G) indicators and the Logistics Performance Index (LPI).

References

1. Rodionova, M.; Skhvediani, A.; Kudryavtseva, T. ESG as a booster for logistics stock returns—Evidence from the us stock market. *Sustainability* **2022**, *14*, 12356.
2. Tsang, Y.P.; Fan, Y.; Feng, Z.P. Bridging the gap: Building environmental, social and governance capabilities in small and medium logistics companies. *J. Environ. Manag.* **2023**, *338*, 117758.
3. Larson, P.D. Relationships between logistics performance and aspects of sustainability: A cross-country analysis. *Sustainability* **2021**, *13*, 623.
4. Nenavani, J.; Prasuna, A.; Siva Kumar, S.N.V.; Kasturi, A. ESG measures and financial performance of logistics companies. *Let. Spat. Resour. Sci.* **2024**, *17*, 5.
5. Lee, E.S. Evaluation of the Impact of ESG Practices on Financial Performance in Korean Small and Medium Logistics Companies. *Asia-Pac. J. Conver. Res. Interchange (APJCRI)* **2024**, *10*, 237–248.

6. Binzaiman, F.; Edhrabooh, K.M.; Alromaihi, M.; AlShammari, M. Predicting Environmental, Social, and Governance Scores with Machine Learning: A Systematic Literature Review. In Proceedings of the 2024 5th International Conference on Data Analytics for Business and Industry (ICDABI), Zallaq, Bahrain, 23–24 October 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 117–122.
7. Gupta, A.; Sharma, U.; Gupta, S.K. The role of ESG in sustainable development: An analysis through the lens of machine learning. In *Proceedings of the 2021 IEEE International Humanitarian Technology Conference (IHTC), Virtual, 2–4 December 2021*; IEEE: Piscataway, NJ, USA; pp. 1–5.
8. Ali, H.; Zafar, M.B. The ESG Code: A Multi-Method Review of Ai in Sustainable Finance. SSRN 5205753. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5205753 (accessed on 15 January 2025).
9. Harsono, M.I. Logistics Performance and Other Factors as Antecedents of the Sustainability Performance of G-20 Countries. *Int. J. Econ. Res. Financ. Account.* **2023**, *2*, 252–270.
10. Barykin, S.E.; Strimovskaya, A.V.; Sergeev, S.M.; Borisoglebskaya, L.N.; Dedyukhina, N.; Sklyarov, I.; Sklyarova, J.; Saychenko, L. Smart city logistics on the basis of digital tools for ESG goals achievement. *Sustainability* **2023**, *15*, 5507.
11. Zhang, W.; Wei, Z.; Ge, L.; Zhang, Y.; Xu, G. How Does ESG Performance Matter for Corporate Sustainability Performance? Evidence from China. *Sustainability* **2025**, *17*, 1684.
12. Wang, Q.; Zhang, Y.; Li, Y.; Wang, P. ESG performance and green innovation in commercial banks: Evidence from China. *PLoS ONE* **2024**, *19*, e0308513.
13. Jílková, P.; Kotěšovcová, J. ESG performance and disclosure: National composite indicators for monitoring sustainable growth conditions in the EU-27. *TEM J.* **2023**, *12*, 1845.
14. Park, B. The Impact of ESG Frameworks on Economic Performance: The Mediating Role of Logistics Performance and Liner Shipping Connectivity. *J. Korea Port Econ. Assoc.* **2023**, *39*, 163–190.
15. Juvvala, R.; Sangle, S.; Tiwari, M.K. Post-Covid challenges and opportunities: Rethinking ESG performance in the logistics sector. *Int. J. Prod. Res.* **2025**, *63*, 1256–1274.
16. Fan, M.; Tang, Y.; Qalati, S.A.; Ibrahim, B. Can logistics enterprises improve their competitiveness through ESG in the context of digitalization? Evidence from China. *Int. J. Logist. Manag.* **2025**, *36*, 196–224.
17. Leogrande, A. Integrating ESG Principles into Smart Logistics: Toward Sustainable Supply Chains. Available online: <https://ssrn.com/abstract=5022211> (accessed on 10 January 2025).
18. Kim, J.; Kim, M.; Im, S.; Choi, D. Competitiveness of E Commerce firms through ESG logistics. *Sustainability* **2021**, *13*, 11548.
19. Kim, D.; Na, J.; Ha, H.K. Exploring the impact of green logistics practices and relevant government policy on the financial efficiency of logistics companies. *Heliyon* **2024**, *10*, e30916.
20. Engelhardt, N.; Ekkenga, J.; Posch, P. ESG ratings and stock performance during the COVID-19 crisis. *Sustainability* **2021**, *13*, 7133.
21. Zhang, M.; Yang, W.; Zhao, Z.; Pratap, S.; Wu, W.; Huang, G.Q. Is digital twin a better solution to improve ESG evaluation for vaccine logistics supply chain: An evolutionary game analysis. *Oper. Manag. Res.* **2023**, *16*, 1791–1813.
22. Bo, P. The Impact of Digital Technology Application on Logistics Enterprise ESG Performance in VUCA Environment: Base on the Moderated Mediation Model. *J. Roi Kaensarn Acad.* **2024**, *9*, 1530–1548.
23. Moreira, O.J.; Rodrigues, M.C.M. Sourcing third party logistics service providers based on environmental, social and corporate governance: A case study. *Discov. Sustain.* **2023**, *4*, 36.
24. Dos Santos, M.C.; Pereira, F.H. ESG performance scoring method to support responsible investments in port operations. *Case Stud. Transp. Policy* **2022**, *10*, 664–673.
25. Pham, T.N.; Tran, P.P.; Le, M.H.; Vo, H.N.; Pham, C.D.; Nguyen, H.D. The effects of ESG combined score on business performance of enterprises in the transportation industry. *Sustainability* **2022**, *14*, 8354.
26. Šulentić, T.; Rakić, E.; Kavran, K.M.Z. ESG management-the main factors of sustainable business in the postal logistics sector. In Proceedings of the First International Conference on Advances in Traffic and Communication Technologies, Sarajevo, Bosnia-Herzegovina, 26–27 May 2022; p. 9.
27. Błaszczuk, A.; Le Viet-Błaszczuk, M. The role of social media marketing of ESG in warehouse logistics. *Zesz. Naukowe. Organ. I Zarządzanie/Politech. Śląska* **2024**, 49–66.
28. Lee, J.W.; Lee, H.S. An Analysis of ESG keywords in the logistics industry using SNA methodology: Using news article and sustainable management report. *Korea Trade Rev.* **2022**, *47*, 121–132.
29. Stan, S.E.; Țițu, A.M.; Mănescu, G.; Ilie, F.V.; Rusu, M.L. Measuring Supply Chain Performance from ESG Perspective. SSRN 5093491. Available online: <https://reference-global.com/2/v2/download/pdf/10.2478/kbo-2023-0026> (accessed on 15 January 2025).

30. Gündoğdu, H.G.; Aytekin, A.; Toptancı, Ş.; Korucuk, S.; Karamaşa, Ç. Environmental, social, and governance risks and environmentally sensitive competitive strategies: A case study of a multinational logistics company. *Bus. Strategy Environ.* **2023**, *32*, 4874–4906.
31. Shakil, M.H.; Munim, Z.H.; Zamore, S.; Tasnia, M. Sustainability and financial performance of transport and logistics firms: Does board gender diversity matter? *J. Sustain. Financ. Invest.* **2024**, *14*, 100–115.
32. Chien, F. The role of corporate governance and environmental and social responsibilities on the achievement of sustainable development goals in Malaysian logistic companies. *Econ. Res.-Ekon. Istraživanja* **2023**, *36*, 1610–1630.
33. Kudryavtseva, T.; Rodionova, M.; Skhvediani, A. Event Study on the Stock Performance: The Case of US Logistics Companies. In *International Scientific Conference “Digital Transformation on Manufacturing, Infrastructure & Service”*; Springer Nature: Cham, Switzerland, 2022; pp. 218–229.
34. Lee, J.; Lee, J.; Lee, C.; Kim, Y. Identifying ESG trends of international container shipping companies using semantic network analysis and multiple case theory. *Sustainability* **2023**, *15*, 9441.
35. Yang, F.; Chen, T.; Zhang, Z.; Yao, K. Firm ESG Performance and Supply-Chain Total-Factor Productivity. *Sustainability* **2024**, *16*, 9016.
36. Altın, F.G.; Gürsoy, S.; Doğan, M.; Ergüney, E.B. The Analysis of the Relationship Among Climate Policy Uncertainty, Logistic Firm Stock Returns and ESG Scores: Evidence from the TVP-VAR Model. *İstatistik Araştırma Derg.* **2023**, *13*, 42–59.
37. Chiang, K.L. Delivering Goods Sustainably: A Fuzzy Nonlinear Multi-Objective Programming Approach for E-Commerce Logistics in Taiwan. *Sustainability* **2024**, *16*, 5720.
38. Zeng, H.; Li, R.Y.M.; Zeng, L. Evaluating green supply chain performance based on ESG and financial indicators. *Front. Environ. Sci.* **2022**, *10*, 982828.
39. Zheng, D.; Wang, T. Supply chain resilience, logistics efficiency, and enterprise competitiveness. *Financ. Res. Lett.* **2025**, *79*, 107335.
40. Shen, Y.; Ma, J.; Wang, W. Supply chain digitization and enterprise ESG performance: A quasi-natural experiment in China. *Int. J. Logist. Res. Appl.* **2024**, *28*, 1956–1978.
41. Borisova, V.; Pechenko, N. Sustainable Development of Logistic Infrastructure of the Region. In *E3S Web of Conferences, Proceedings of the International Scientific Forum on Sustainable Development and Innovation (WFSDI 2021)*, Yekaterinburg, Russia, 10–11 July 2021; Volume 295, p. 01042.
42. Govindan, K.; Karaman, A.S.; Uyar, A.; Kilic, M. Board structure and financial performance in the logistics sector: Do contingencies matter? *Transp. Res. Part E: Logist. Transp. Rev.* **2023**, *176*, 103187.
43. Mutambik, I. Digital Transformation as a Driver of Sustainability Performance—A Study from Freight and Logistics Industry. *Sustainability* **2024**, *16*, 4310.
44. Yu, K.; Wu, Q.; Chen, X.; Wang, W.; Mardani, A. An integrated MCDM framework for evaluating the environmental, social, and governance (ESG) sustainable business performance. *Ann. Oper. Res.* **2024**, *342*, 987–1018.
45. Sun, X.; Kuo, Y.H.; Xue, W.; Li, Y. Technology-driven logistics and supply chain management for societal impacts. *Transp. Res. Part E Logist. Transp. Rev.* **2024**, *185*, 103523.
46. Kanno, M. Does ESG performance improve firm creditworthiness? *Financ. Res. Lett.* **2023**, *55*, 103894.
47. Wu, M.; Xie, D. The impact of ESG performance on the credit risk of listed companies in Shanghai and Shenzhen stock exchanges. *Green Financ.* **2024**, *6*, 199.
48. Skhvediani, A.E.; Gutman, S.S.; Rodionova, M.A.; Perfilova, J.A. Being green as an instrument for increasing firm value: Case of US transport and logistics companies. *Int. J. Logist. Syst. Manag.* **2024**, *47*, 105–124.
49. Tian, L.; Tian, W.; Guo, M. Can supply chain digitalization open the way to sustainable development? Evidence from corporate ESG performance. *Corp. Soc. Responsib. Environ. Manag.* **2025**, *32*, 2332–2346.
50. Das, A. Predictive value of supply chain sustainability initiatives for ESG performance: A study of large multinationals. *Multinatl. Bus. Rev.* **2024**, *32*, 20–40.
51. Burcă, V.; Bogdan, O.; Bunget, O.C.; Dumitrescu, A.C.; Imbrescu, C.M. Financial Implications of Supply Chains Transition to ESG Models. *Explor. ESG Chall. Oppor. Navig. Towards A Better Future* **2024**, *116*, 127–143.
52. Kurniawan, F.; Musa, S.N.; Nurfauzi, B.; Ferdian, R.; Khair, F. Container Terminal Performance: System Dynamic Approach with Port Capacity Constraints and ESG Integration. *Jordan J. Mech. Ind. Eng.* **2024**, *18*, 59–73.
53. Niu, B.; Dong, J.; Wang, H. Smart port vs. port integration to mitigate congestion: ESG performance and data validation. *Transp. Res. Part E Logist. Transp. Rev.* **2024**, *191*, 103741.

54. Li, W.; Wang, Y. A procurement Advantage in Disruptive times: New Perspectives on ESG Strategy and Firm Performance. SSRN 4817562. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4817562 (accessed on 13 January 2025).
55. Fatimah, Y.A.; Kannan, D.; Govindan, K.; Hasibuan, Z.A. Circular economy e-business model portfolio development for e-business applications: Impacts on ESG and sustainability performance. *J. Clean. Prod.* **2023**, *415*, 137528.
56. Gürler, H.E.; Özçalıcı, M.; Pamucar, D. Determining criteria weights with genetic algorithms for multi-criteria decision making methods: The case of logistics performance index rankings of European Union countries. *Socio-Econ. Plan. Sci.* **2024**, *91*, 101758.
57. Taskin, D.; Sariyer, G.; Acar, E.; Cagli, E.C. Do past ESG scores efficiently predict future ESG performance? *Res. Int. Bus. Financ.* **2025**, *74*, 102706.
58. Costantiello, A.; Leogrande, A. The regulatory quality in the light of environmental, social and governance framework at world level. *Discov. Glob. Soc.* **2024**, *2*, 1.
59. Magazzino, C.; Alola, A.A.; Schneider, N. The trilemma of innovation, logistics performance, and environmental quality in 25 topmost logistics countries: A quantile regression evidence. *J. Clean. Prod.* **2021**, *322*, 129050.
60. Hayyat, U.; Qian, L.; Saeed, M.; Nawaz, W. Modeling the Growth Dynamics of Logistics Performance: Industrialization, Environmental Technology, and Economic Transformation in Manufacturing Economies. *Systems* **2025**, *13*, 375.
61. Cheong, T.S.; Liu, S.; Ma, N.; Han, T. The Impact of Public Environmental Concern on Corporate ESG Performance. *J. Risk Financ. Manag.* **2025**, *18*, 82.
62. Zheng, C.; Khan, M.A.M.; Islam, R.; Chowdhury, M.M. Exploring the relationship between ESG performance and firm value in Chinese and US banks: The moderating impact of environmental uncertainty and competitive advantage. *Int. J. Res. Bus. Soc. Sci.* **2025**, *14*, 1–16.
63. Wan, B.; Wan, W.; Hanif, N.; Ahmed, Z. Logistics performance and environmental sustainability: Do green innovation, renewable energy, and economic globalization matter? *Front. Environ. Sci.* **2022**, *10*, 996341.
64. Gholami, H.; Mohammadifar, A.; Bui, D.T.; Collins, A.L. Mapping wind erosion hazard with regression-based machine learning algorithms. *Sci. Rep.* **2020**, *10*, 20494.
65. Wang, T.; Qin, L.; Dai, C.; Wang, Z.; Gong, C. Heterogeneous Learning of Functional Clustering Regression and Application to Chinese Air Pollution Data. *Int. J. Environ. Res. Public Health* **2023**, *20*, 4155.
66. Xiao, I.; Jaller, M.; Phong, D.; Zhu, H. Spatial analysis of the 2018 logistics performance index using multivariate kernel function to improve geographically weighted regression models. *Transp. Res. Rec.* **2022**, *2676*, 44–58.
67. Xuan, T.T.T.; Quach, P.H.; Van Thinh, N.; Hoa, T.T.; Tu, N.T. The efficiency and the performance of the logistics global supply chain activities to Vietnam exportation: An empirical case study. *Int. J. Prof. Bus. Rev.* **2023**, *8*, 48.
68. Conștângioară, A.; Florian, G.L. Is Logistics Mediating the Relationship Between Pollution and Economic Complexity? In Proceedings of the 16th International Management Conference, Bucharest, Romania, 3–4 November 2022; Volume 16, pp. 312–321.
69. Karaduman, H.A.; Karaman-Akgül, A.; Çağlar, M.; Akbaş, H.E. The relationship between logistics performance and carbon emissions: An empirical investigation on Balkan countries. *Int. J. Clim. Change Strateg. Manag.* **2020**, *12*, 449–461.
70. Akram, M.W.; Hafeez, M.; Yang, S.; Sethi, N.; Mahar, S.; Salahodjaev, R. Asian logistics industry efficiency under low carbon environment: Policy implications for sustainable development. *Environ. Sci. Pollut. Res.* **2023**, *30*, 59793–59801.
71. Zhao, L.; Yu, Q.; Li, M.; Wang, Y.; Li, G.; Sun, S.; Jia, F.; Liu, Y. A review of the innovative application of phase change materials to cold-chain logistics for agricultural product storage. *J. Mol. Liq.* **2022**, *365*, 120088.
72. Liang, Y.; Ge, X.; Jin, Y.; Zheng, Z.; Zhang, Y.; Jiang, Y. Economic optimization of fresh logistics pick-up routing problems with time windows based on gray prediction. *J. Intell. Fuzzy Syst.* **2024**, *46*, 10813–10832.
73. Filassi, M.; Oliveira, A.L.R.D.; Elias, A.A.; Braga Marsola, K. Analyzing complexities in the Brazilian soybean supply chain: A systems thinking and modeling approach. *RAUSP Manag. J.* **2022**, *57*, 280–297.
74. Okanda, T.L.; Zhang, J.; Sarfo, P.A.; Amankwah, O. Exploring the Nexus between Debt Financing and Firm Performance: A Robustness Analysis Using Instrumental Variables. *Int. J. Adv. Eng. Res. Sci.* **2025**, *12*, 8–22.
75. Li, X.; Sohail, S.; Majeed, M.T.; Ahmad, W. Green logistics, economic growth, and environmental quality: Evidence from one belt and road initiative economies. *Environ. Sci. Pollut. Res.* **2021**, *28*, 30664–30674.
76. Nawurunnage, K.R.; Prasadika, A.P.K.J.; Wijayanayake, A.N. TQM and Green Supply Chain Management Practices on Supply Chain Performance of Third-Party Logistics Services in Sri Lanka: A Systematic Review of Literature. In Proceedings of the 2023 3rd International Conference on Advanced Research in Computing (ICARC), Belihuloya, Sri Lanka, 23–24 February 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 274–279.
77. Onukwulu, E.C.; Agho, M.O.; Eyo-Udo, N.L. Advances in green logistics integration for sustainability in energy supply chains. *World J. Adv. Sci. Technol.* **2022**, *2*, 47–68.

78. Nagy, G.; Szentesi, S. Green logistics: Transforming supply chains for a sustainable future. *Adv. Logist. Syst.-Theory Pract.* **2024**, *18*, 29–42.
79. Sun, Y.; Li, Y.; Jia, Y.; Yang, J.; Peng, Y.; Guo, X. A Random Forest-based Model for Cargo Volume Prediction and Personnel Scheduling in Logistics Sorting Centers. In Proceedings of the 2024 3rd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI), Zakopane, Poland, 18–20 October 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 289–294.
80. Thummala, G.S.R.; Baskar, R. Prediction of Heart Disease using Random Forest in Comparison with Logistic Regression to Measure Accuracy. In Proceedings of the 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 25–26 May 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5.
81. Jomthanachai, S.; Wong, W.P.; Khaw, K.W. An application of machine learning regression to feature selection: A study of logistics performance and economic attribute. *Neural Comput. Appl.* **2022**, *34*, 15781–15805.
82. Kocabaş, M.B.; Tashan, W.; Shaye, I.; Alibek, M. Comparative Analysis of One-Dimensional Regression Techniques. In Proceedings of the 2024 IEEE 16th International Conference on Computational Intelligence and Communication Networks (CICN), Indore, India, 22–23 December 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1365–1370.
83. Al Bony, M.N.V.; Das, P.; Pervin, T.; Shak, M.S.; Akter, S.; Anjum, N.; Alam, M.; Akter, S.; Rahman, M.K. Comparative Performance Analysis of Machine Learning Algorithms for Business Intelligence: A Study on Classification And Regression Models. *Frontline Mark. Manag. Econ. J.* **2024**, *4*, 72–92.
84. Samy, S.; Jaini, K.; Preheim, S. A Novel Machine Learning-Driven Approach for Predicting Nitrous Oxide Flux in Precision Managed Agricultural Systems. SSRN 4976901. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4976901 (accessed on 13 January 2025).
85. Maier, R.; Hörtnagl, L.; Buchmann, N. Greenhouse gas fluxes (CO₂, N₂O and CH₄) of pea and maize during two cropping seasons: Drivers, budgets, and emission factors for nitrous oxide. *Sci. Total Environ.* **2022**, *849*, 157541.
86. Shang, Y.J.; Mao, Y.H.; Liao, H.; Hu, J.L.; Zou, Z.Y. Response of PM 2.5 and O₃ to Emission Reductions in Nanjing Based on Random Forest Algorithm. *Huan Jing ke Xue= Huanjing Kexue* **2023**, *44*, 4250–4261.
87. Novian, T.R.; Hardi, I.; Zahriah, Z.; Sofyan, R.; Sasmita, N.R.; Hilal, I.S.; Idroes, G.M. Environmental and economic clustering of Indonesian provinces: Insights from K-Means analysis. *Leuser J. Environ. Stud.* **2024**, *2*, 41–51.
88. Zhu, C.; Liu, Z. Semi-supervised clustering of PM2. 5 pollution. In Proceedings of the International Conference on Computer Application and Information Security (ICCAIS 2023), Wuhan, China, 27–29 November 2024; SPIE: Bellingham, WA, USA, 2023; Volume 13090, pp. 427–432.
89. Nakhjiri, A.; Kakroodi, A.A. Air pollution in industrial clusters: A comprehensive analysis and prediction using multi-source data. *Ecol. Inform.* **2024**, *80*, 102504.
90. Gebreyesus, Y.; Dalton, D.; Nixon, S.; De Chiara, D.; Chinnici, M. Machine learning for data center optimizations: Feature selection using Shapley additive exPlanation (SHAP). *Future Internet* **2023**, *15*, 88.
91. Mohanty, P.K.; Francis, S.A.J.; Barik, R.K.; Roy, D.S.; Saikia, M.J. Leveraging Shapley Additive Explanations for Feature Selection in Ensemble Models for Diabetes Prediction. *Bioengineering* **2024**, *11*, 1215.
92. Varkiani, S.M.; Pattarin, F.; Fabbri, T.; Fantoni, G. Predicting employee attrition and explaining its determinants. *Expert Syst. Appl.* **2025**, *272*, 126575.
93. Syed, M.N. Neighborhood density information in clustering. *Ann. Math. Artif. Intell.* **2022**, *90*, 855–872.
94. Fu, X.; Feng, L.; Zhang, L. Data-driven estimation of TBM performance in soft soils using density-based spatial clustering and random forest. *Appl. Soft Comput.* **2022**, *120*, 108686.
95. Bicego, M.; Escolano, F. On learning random forests for random forest-clustering. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 3451–3458.
96. Yıldırım, M. Cluster Analysis on Supply Chain Management-Related Indicators. *İnsan Ve Toplum Bilim. Araştırmaları Derg.* **2023**, *12*, 2499–2520.
97. Kara, K. Clustering of Developing Countries in Terms of Logistics Market Development with Fuzzy Clustering and Discriminant Analysis. *Yaşar Üniversitesi E-Derg.* **2023**, *18*, 19–40.
98. Effiong, U.E.; Udofia, L.E.; Garba, I.H. Governance and economic development in West Africa: Linking governance with economic misery. *Path Sci.* **2023**, *9*, 2009–2025.
99. Pinjaman, S.; Thani, M.A.M.; Bakar, M.; Hadi, S. The Nexus between Governance Quality and Economic Growth of Malaysia: Short-And Long-Run Analyses. *Int. J. Res. Innov. Soc. Sci.* **2025**, *9*, 115–129.
100. Sadriu, M.; Balaj, D. Assessing the Role of Governance Indicators on Foreign Direct Investment: Insights from Southeastern European Countries. *J. Gov. Regul./Vol.* **2024**, *13*, 316–321.

101. Baciu, L.E. The impact of governance upon sustainable development. Empirical evidence. *Stud. Univ. Babes Bolyai-Oeconomica* **2023**, *68*, 73–86.
102. Rawat, D.S. Political Governance and Stock Market Performance: An Autoregressive Distributed Lag Analysis of the Nepalese Market. *KMC J.* **2025**, *7*, 272–294.
103. Long, J.P.; Zhu, H.; Do, K.A.; Ha, M.J. Estimating causal effects with hidden confounding using instrumental variables and environments. *Electron. J. Stat.* **2023**, *17*, 2849.
104. Qu, Z.; Kwon, Y. Distributionally Robust Instrumental Variables Estimation. *arXiv* **2024**, arXiv:2410.15634.
105. Troncoso, J.A.; Quijije, Á.T.; Oviedo, B.; Zambrano-Vega, C. Solar Radiation Prediction in the UTEQ based on Machine Learning Models. *arXiv* **2023**, arXiv:2312.17659.
106. Jenifel, M.G.; Jasmine, R.A.; Umanandhini, D. Bitcoin Price Predictive Dynamics Using Machine Learning Models. In Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 24–28 June 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6.
107. Maheshwari, A.; Malhotra, A.; Hada, B.S.; Ranka, M.; Basha, M.S.A. Towards an Improved Model for Stability Score Prediction: Harnessing Machine Learning in National Stability Forecasting. In Proceedings of the 2024 IEEE North Karnataka Subsection Flagship International Conference (NKCon), Bagalkote, India, 21–22 September 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1–7.
108. Mukhtar, M. Unravelling Structural Underdevelopment: Is Governance Quality the Key? Ph.D. Thesis, Minhaj University Lahore, Lahore, Pakistan, 2023.
109. Peiman, F.; Khalilzadeh, M.; Shahsavari-Pour, N.; Ravanshadnia, M. Estimation of building project completion duration using a natural gradient boosting ensemble model and legal and institutional variables. *Eng. Constr. Archit. Manag.* **2023**, *32*, 2069–2104.
110. Boukrouh, I.; Tayalati, F.; Azmani, A. Comparative SHAP Analysis on SVM and K-NN: Impacts of Hyperparameter Tuning on Model Explainability. In Proceedings of the 2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAET), Kota Kinabalu, Malaysia, 26–28 August 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 194–198.
111. Ilyas, M. Unveiling the education paradox: Conflict, pandemic and schooling in Kashmir. *Int. Rev. Educ.* **2024**, *70*, 869–891.
112. Guo, J.; Dong, R.; Zhang, R.; Yang, F.; Wang, Y.; Miao, W. Interpretable machine learning model for predicting the prognosis of antibody positive autoimmune encephalitis patients. *J. Affect. Disord.* **2025**, *369*, 352–363.
113. Tao, Y.; Wang, S.; Wu, J.; Zhao, M.; Yang, Z. Logistic network construction and economic linkage development in the Guangdong-Hong Kong-Macao Greater Bay Area: An analysis based on spatial perspective. *Sustainability* **2022**, *14*, 15652.
114. Gagolewski, M.; Bartoszek, M.; Cena, A. Are cluster validity measures (in) valid? *Inf. Sci.* **2021**, *581*, 620–636.
115. Sarma, E.; Fragkiadaki, A.; Marinakis, V. Explainable AI-Based Ensemble Clustering for Load Profiling and Demand Response. *Energies* **2024**, *17*, 5559.
116. Hossen, M.B.; Auwal, M.R. Comparative study of K-means, partitioning around medoids, agglomerative hierarchical, and DI-ANA clustering algorithms by using cancer datasets. *Biomed. Stat. Inform.* **2020**, *5*, 20–25.
117. Slezák, J. Relations between Development of E-Government and Government Effectiveness, Control of Corruption and Rule of Law in 2010–2020: A Cluster Analysis. *Acta VŠFS-Ekon. Stud. A Analýzy* **2023**, *17*, 161–187.
118. Pehlivan, P.; Aslan, A.I.; David, S.; Bacalum, S. Determination of Logistics Performance of G20 Countries Using Quantitative Decision-Making Techniques. *Sustainability* **2024**, *16*, 1852.
119. Ulkhaq, M.M. Clustering countries according to the logistics performance index. *JATISI (J. Tek. Inform. Dan Sist. Inf.)* **2023**, *10*, 1010–1018.
120. Sharawi, H.; Alsaadi, L.; Alsagri, M. The impact of LPIs' indicators on the global logistics performance index: Global perspective. *Multidiscip. Sci. J.* **2025**, *7*, 2025361.
121. Göçer, A.; Özpeynirci, Ö.; Semiz, M. Logistics performance index-driven policy development: An application to Turkey. *Transp. Policy* **2022**, *124*, 20–32.
122. Rodrigues, N.V.F.; Fiorini, P.D.C.; Piato, É.L. Logistics 4.0 and corporate sustainability: An organizational theory perspective. *Gepros Gestão Da Produção Operações E Sist.* **2022**, *17*, 108.
123. Yontar, E. Assessment of the logistics activities with a structural model on the basis of improvement of sustainability performance. *Environ. Sci. Pollut. Res.* **2022**, *29*, 68904–68922.
124. Yoo, B.C. Exploring Stakeholder and Organizational Influences on ESG Management in the Logistics Sector. *Sustainability* **2025**, *17*, 4243.

125. Karountzos, P.; Sakas, D.P.; Nasiopoulos, D.K.; Toudas, K. Redefining Development Through Logistics Performance and ESG Metrics. *Account. Audit.* **2025**, *1*, 11.
126. Truant, E.; Borlatto, E.; Crocco, E.; Sahore, N. Environmental, social and governance issues in supply chains. A systematic review for strategic performance. *J. Clean. Prod.* **2024**, *434*, 140024.
127. Popescu, C.A.; Ifrim, A.M.; Silvestru, C.I.; Dobrescu, T.G.; Petcu, C. An evaluation of the environmental impact of logistics activities: A case study of a logistics centre. *Sustainability* **2024**, *16*, 4061.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.