

### Bridging Sustainability and Inclusion: Financial Access in the Environmental, Social, and Governance Landscape

Drago, Carlo and Costantiello, Alberto and Arnone, Massimo and Leogrande, Angelo

Unicusano University, LUM University Giuseppe Degennaro, University of Catania, LUM University Giuseppe Degennaro

 $23~\mathrm{May}~2025$ 

Online at https://mpra.ub.uni-muenchen.de/124827/ MPRA Paper No. 124827, posted 23 May 2025 14:45 UTC

# **Bridging Sustainability and Inclusion: Financial Access in the Environmental, Social, and Governance Landscape**

Carlo Drago, Unicusano University, Rome, <u>carlo.drago@unicusano.it</u>

Alberto Costantiello, LUM University Giuseppe Degennaro, Casamassima, <u>Costantiello@lum.it</u> Massimo Arnone, University of Catania, Catania, <u>massimo.arnone@unict.it</u> Angelo Leogrande, LUM University Giuseppe Degennaro, Casamassima, <u>leogrande.cultore@lum.it</u>

#### Abstract

This paper explores the correlation between financial inclusion and the Environment, Social, and Governance (ESG) aspects of sustainable development for a big panel of 103 developing nations over 12 years. Financial inclusion as a measure is taken through the Account Age variable capturing adults having access to formal financial institutions as a percentage. The analysis revolves around the three main ESG pillars each through panel data regressions complemented by instrumental variable (IV) approaches in addressing endogeneity concerns. In the Environment (E) dimension, we find conventional agricultural forms (e.g., extensive agricultural land areas and agriculture value added) as having a negative effect on financial inclusion, but the environmental modernization proxies-renewable energy utilization, food production, climate resilience, and areas under protection—exhibit positive and significant correlations. In the Social (S) dimension, development indicator variables like spending on education, internet penetration, life years at birth, sanitation, and gender equity emerge as strong predictors of higher financial inclusion, and labor market participation is found to have a negative effect, possibly due to the dynamics of employment in the informal sector. The Governance (G) analysis shows positive correlation with controlling corruption and innovation production (applications for patents) as arguments for increased financial access improving institutional transparency and economic ingenuity and a negative correlation with regulatory quality as a concern for capacity gaps in rapidly digitizing economies. Through the means of ESG-matched environmental instruments, this paper presents a unique crossdimensional approach to sustainable finance and shows through counterfactual analysis under both average and counterfactual distributions that policies supporting financial inclusion can be a path to multiple benefits on the environmental sustainability, social equity, and governance effectiveness axeskey requirements for the success of the Sustainable Development Goals (SDGs) in the Global South.

Keywords: Financial Inclusion, ESG Framework, Developing Countries, Instrumental Variables, Sustainable Development.

JEL CODE: G21, O16, Q56, I38, H55, O44, C33.

#### 1. Introduction

In the past decades, financial inclusion has become the center stage in the world development agenda as a principal force in the way people are expected to reduce the level of poverty, achieve economic stability and stability in growth and sustainable development. Financial inclusion is widely defined as access and use of financial services provided through the organized financial system by the entire populace-more so the poor and the vulnerable—has taken position in international policymaking documents such as the Sustainable Development Goals specifically SDG 8.10. Formidable attention and academic enthusiasm addressed to financial inclusion notwithstanding, large chasms still lay in the understanding of the deeper systemic and structural determinants and mechanisms propelling and constraining financial inclusion in developing economies (Zulkhibri, 2016). Of great concern is the nexus between financial access and the Environmental, Social, and Governance (ESG) aspects of development, and this still remains a poorly covered area in extant scholarship. This study bridges this significant research gap by developing and testing a new framework to analyze financial inclusion in developing countries using disaggregated ESG measures. Contrarily to the mainstream use of the aggregate or unified concept of ESG in the study of sustainability, this study breaks up ESG and its components and estimates their respective contributions to access to finance. More recent scholarship also brings attention to the moderating role played by poverty in the ESG-financial inclusion linkage and the complexities entangled with such interactions (Jain et al., 2024). Drawing samples from 103 developing countries over a 12-year time frame using comprehensive longitudinal panels from the World Bank and other international institutions' related datasets, we explore the respective contributions made individually to access to finance by environmental sustainability, social infrastructure, and good governance quality.

The research guiding framework used in this research is

In what way do the Environmental, Social, and Governance components of the ESG framework individually contribute to financial inclusion in developing economies, and how can each such effect reliably be quantified through econometric and machine-learning methods?

We resolve this issue through the adoption of a mixed methodological approach based upon the application of conventional panel econometrics supplemented with instrumental variables methods and advanced machine learning algorithms in order to apply them to regression and cluster analysis purposes. The econometric model allows us to manage the issues related to endogeneity and causality, while the machine learning models both supervised type (e.g., Random Forest regression) and unsupervised type (e.g., Fuzzy C-Means clustering) reveal profound non-linear relationships and allow countries' classification according to policy-related categories (Park et al., 2024).

Our lead financial inclusion indicator is Account Age, the share of adults in a nation who say they have an account at a bank or with a mobile money service provider. This measure has achieved broad consensus as a quantitative proxy for access to the formal financial system and reflects both the penetration of established financial system infrastructure and new digital finance innovations. By considering the effect each of the ESG pillars has separately on Account Age, we will endeavor to build a more nuanced view of the determinants driving financial inclusion in developing economies.

Environmentally, we look at the part played by ecological modernization through proxies such as the use of renewable energy, agricultural production levels and biodiversity conservation, and sensitivity to climatic stress. The concept being that green infrastructure (like solar energy off the grid) can actually promote financial access via facilitation of digital payments in remote areas while increasing climate risk can increase savings or insurance product take-up (Xie, 2024). While at the same time agrarian institution resilience—reflected in high proportions of agricultural land—may indicate structural impediment to financial access in off-grid rural economies.

Along with this, ESG-based analyses have also helped identify risk-mitigated and credible behaviour in individuals and institutions and lending legitimacy to the application of such variables to economic models (Aghaei et al., 2024). Blending this with our model helps to promote the analytical depth of our machine-learning models and predictive ability in the estimation of access dynamics in the Global South. On the social side, we apply human development indicators such as education expenditure, internet penetration rates, life expectancy at birth, access to sanitation, and female labor market participation. They are the social capacities upon which financial inclusion depends—i.e., literacy, access to digital technologies, health and well-being, and gender equality. The more advanced the human development in a nation is, the more it is expected to achieve better financial inclusion performance through improved public awareness, institutional trust, and improved ability to make use of financial products and services (Khalil & Siddiqui, 2020). Social infrastructure is a trigger to access to formal financial services if the digital expansion remains in place.

The governance factor assesses the extent to which institutional quality facilitates or deters access to financial services. Control of corruption, regulatory framework quality, rule of law, and voice and accountability are the measures used to test the hypothesis that good governance creates a facilitative environment for the financial sector to flourish. Governance is at the core of the role played in deciding if citizens are able to access identification systems, if contracts are enforceable or otherwise, and if protection mechanisms work or do not work for consumers—each being a need for a sustainable and inclusive financial system (Maket, 2024).

To estimate the empirical relations between aforementioned ESG pillars and financial inclusion, we apply panel data regressions complemented with instrumental variables (IV) in order to counter concerns of simultaneity, measurement errors, and omitted variables. This is particularly imperative in the governance-financial inclusion nexus where causality might run both ways. Our instruments are drawn from a universe of exogenous ESG-linked variables—such as climate indicators, demographic measures, and environmental stressors—that are correlated with the endogenous regressors but not the error term in the inclusion equation.

We estimate two econometric models: Generalized Two-Stage Least Squares (G2SLS) with random effects and Two-Stage Least Squares with fixed effects. Both are used over a balanced panel of more than 1,200 observations to enable robust inference across countries. The need to account for the interdependencies between the variables is further emphasized in recent evidence revealing that income inequality, institution quality, and human development co-determine financial depth and breadth of inclusion in developing countries like Africa (Kebede et al., 2023).

The econometric evidence is supplemented with a parallel alternative analytical pathway: machine learning modeling predicting and classifying financial inclusion levels based on ESG attributes. In this case, we apply a combination of supervised regressions including Random Forest, Neural Networks, Support Vector Regression, and the use of the Boosting techniques to model predictive performance using the common measures including Mean Squared Error (MSE), R<sup>2</sup> and Mean Absolute Percentage Error (MAPE). Of the models, Random Forest Regression performs the best and extracts leading non-linear relations and key importances across the variables. More importantly still, the strongest predictors include agricultural productivity (AGVA), renewable energy use (RENE), and the share of areas protected (PROT)—a result eminently derivable from the recent work based on advanced use of machine learning to uncover ESG-finance interlinkages (Li, 2025). In parallel, we apply Fuzzy C-Means clustering—a type of the unsupervised use of machine learning—enabling countries to be grouped based on ESG and financial inclusion variables according to soft clusters. This method enables flexible segmentation and the identification of typologies such as: "green but excluded," "socially advanced and inclusive," and "institutionally weak and financially marginalized." Such nuanced designations enable both interpretative depth and policy design since countries in the same cluster can in turn be addressed

with interventions proportionate to their context. The clustering framework also extends the use of the classical econometric analyses since the complex interdependences and latent patterns characterising the cross-country development context can be accommodated. In yet a further improvement, the use of stacked generalisation techniques using the ensemble methods to aggregate predictive models collectively and achieve greater aggregate robustness is also proving to be beneficial when the aim is to uncover nuanced ESG-finance interaction across different national contexts (Xu et al., 2024). This ensemble method extends the utility in the use of AI-informed insights within financial inclusion analysis, in particular where policy differentiation according to ESG type is the objective

We recognize each ESG factor to have a distinct and measurable impact upon financial inclusion performance. Environmental modernisation—captured through investment in renewables and protection of biodiversity—has a positive relation with financial access, and agrarian economies have a negative relation. Socially, digital access and gender equality measures are the robust positive predictors capturing the reality that digital connectedness and gender-inclusive labor markets are robust enablers of access to the financial system (Jain et al., 2024). Governance and financial inclusion have a positive relation and access to regulating in new forms where weak regulation remains the inhibitor.

These findings are consistent both across econometric and machine learning methods and are thus more widely applicable across a broad array of developing-country situations. More current innovation in predictive modeling based upon data, such as the inclusion of ESG attributes in financial outcome projections, further legitimizes our approach (Park et al., 2024).

This research makes three key contributions. First, we provide the first big-data, disaggregate financial inclusion analysis over developing countries based on panel data econometrics and sophisticated machine learning techniques. Second, this study bridges methodological paradigms via the integration of causal inference and predictive modeling to make more nuanced and rich insights available. Third, this study provides recommendations to direct the work of development practitioners, central banks, and multilateral institutions towards the development of context-specific and ESG-driven governance practices to promote financial inclusion. With sustainability, resilience, and inclusion increasingly recognized as interconnected pillars reinforcing one another, it is more and more important to understand the interlinkages between environmental-social-governance and financial access. By incorporating a structured and evidence-based perspective, this study encourages a more holistic understanding of development policy—financial inclusion addressed and considered as part of the very core of sustainable development systems rather than as a distinct outcome in itself (Li, 2025).

The article continues as follows: the second section presents the literature review, the third section contains the methodology, the fourth section presents the relationship between financial inclusion and the E-Environmental component within the ESG model, the fifth section investigates the relationship between S-Social component and the ESG model, the sixth section shows the relationship between financial inclusion and the G-Governance component within the ESG model, the SSG model, the sections analyses the policy implications, the eight section presents conclusions. The Appendixes contain further materials, data, summary statistics, hyper-parameters and abbreviations.

### 2. Literature Review

The nexus between ESG and financial inclusion in developing countries is a complex and dynamic relationship and acts as a vehicle for sustainable and inclusive development and sustainable finance. Originally conceived as a social goal, financial inclusion as a concept emerged as a cross-cutting force behind all the ESG elements—Environmental, Social and Governance—where structural disparities and

access to finance are weak. According to Buckley et al. (2021), FinTech innovations can ford the financial gaps effectively and map financial access with the UN SDGs, while Chafai et al. (2024) demonstrate empirically through their evidence that institutional quality through audit quality enhances the ESG impact of financial access in the MENA region. Bibliometric analysis of scholarly fascination with this point of meeting has been presented by Ahmad et al. (2024), but they do not deeply embed the Global South-specific issues concerned. The emergence of digital technologies accelerated inclusive financing through the confluence of various digital innovations as noted by Austin and Rawal (2023) and Dhanabhakyam and Suresh (2024), but innovation creates exclusion and algorithms risks and raise rollout and dissemination ethics concerns as well as fairness concern over outreach.

Meanwhile, researchers such as Al-Baraki (2022) and Anwar et al. (2023) stress impact measurement and targeted capital flows and persistent bottlenecks such as gender disparity and lack of formality. Gonzalez et al. (2025) emphasize ambitious and comprehensive action at the nexus between climate change, social responsibility, and e-finance access, but their normative perspective might overlook lowcapacity bottlenecks. Work such as Jain et al. (2024) illustrates the potential through empirical research in ESG and financial inclusion to promote sustainable growth, if the systemic moderator poverty is addressed. Li and Pang (2023) and Khalid et al. (2024) support the corporate aspect with evidence of digital inclusion mitigating ESG-related conflicts—though Li et al. (2024) also find the more pronounced role of greenwashing persists, lowering the efficacy of ESG disclosers. Halim (2024) expands the coverage to crowdfunding as a financing conduit and finds disparity across benefits' distributions. Macrolevel research such as Hassani et al. (2024) and Kiran et al. (2025) are correlated with financial development and good ESG integration and Khan et al. (2025) explore the enabling role of blockchain and green tech, albeit depending upon international financial integration and equitable infrastructure.

Further operational takeaways are provided by Liu and Naser (2024), noting increased bank performance—vital in weak economies—as a result of more inclusive finance, reflecting the views of Malik et al. (2022), recognizing social sustainability as a stabilizer in the case of Asia. Governance capacity mentioned by Lu and Cheng (2024) acts as a mediator between digital finance ESG effect and Liu et al. (2024) and Liu (2025) explore the use of AI by China in order to broaden ESG output, with governance ethics still unresolved as a contradiction. Portfolio-level applications are mentioned by Lindquist et al. (2022), with the focus being the incorporation of ESG information within developing economies and Mirza et al. (2025) establishing the connexion between ESG-based lending and techno-investments and stability in banks within BRICS economies. Mehdiyev (2023) and Lu et al. (2022) recognize the role played by digital inclusion in increasing ESG transparency in the case of China and the EU but warn those benefits are policy contingent.

On the firm or corporate level, Pan (2025) warns at the need for SME-oriented financial tools in digital environments and Qureshi (2025) points out digital finance as the key propellant towards ESG performance and long-term sustainability. Roy and Vasa (2025), using bibliometric analysis, document the trend towards FinTech-powered ESG financing in resource-constrained environments. Rajunčius and Miečinskienė (2024) outline a more comprehensive framework blending payment innovation, socially focused equity and ESG evaluation, and Shala and Berisha (2024), Ravichandran and Rao (2023) both highlight FinTech's pivotal role towards ESG progress but also the potential digital illiteracy risk. Legal foundations are weak institutionally, Schwarcz and Leonhardt (2021) remind us and and therefore policy coordination across regions is difficult. Tekin (2025) finds Islamic finance to be a culturally resonant model of inclusive ESG in the OIC economies. More advanced ways forward, such as those from Shah and colleagues (2024), require the incorporation of well-being goals such as happiness in ESG risk models and require a broader reimagining from the social impacts. Empirical work from Suresha and colleagues (2022) report financial evidence from Indian markets and illustrate the way in which ESG-

matched firms have better financial performance and liquidity, and this supports the materiality of inclusive ESG strategies.

Recent research supports the "S" of ESG—so woefully underrated—is key to sensitive and inclusive financial transformation. Ubeda et al. (2023) make this case in the context of the Global South and promote socially-responsible financial models. In the Indian context, Yadav et al. (2024) and Yadav, Premalatha & Patil (n.d.) discuss the ability of FinTech innovations to diffuse ESG-compatible services to remote zones and lead operations towards sustainability. Empirical evidence from Xu (2024) and Xue et al. (2023) ascertains this alteration in China too, where ESG performance shifts through the use of AI-guided financial instruments and risk in domains of digitization and accessibility. Rework by Thomas (2023) takes critical analysis on the greenwashing drift in ESG investment and reiterated by Li et al. (2024), warn more ESG action will prove pointless if left without accountability mechanisms. Cumulative analysis through bibliometrics by Trotta et al. (2024) and Upadhya et al. (2024) chart a fractured yet fast-growing body of work wherein FinTech emerges as a key enabler in re-locationing the arena of ESG debate. Though momentum gathers, Yadav et al. (2024) warn without coordination through policy and enabling infrastructures and importantly in the remote areas, the transformational potential of digital financial expansion through ESG models remains unrealized.

A synthesis of the literature review is synthetized in the following Table 1.

			Impact on Financial Inclusion-ESG
Macro-Theme	Key Insights	Representative Studies	Relationship
	Digital financial tools (e.g.,		
	mobile banking, AI,		
	blockchain) enhance access to	Buckley et al. (2021); Austin &	Enhances access and efficiency in ESG
FinTech &	ESG-aligned finance but raise	Rawal (2023); Liu et al. (2024);	implementation but requires
Digital	issues of digital inequality and	Xu (2024); Khan et al. (2025);	safeguards against algorithmic bias and
Innovation	governance.	Xue et al. (2023)	exclusion.
	Institutional capacity (e.g.,		
	audit quality, executive		Strong governance amplifies ESG
Governance &	expertise, regulation) mediates	Chafai et al. (2024); Lu & Cheng	outcomes from inclusion; weak or
Institutional	the ESG impact of financial	(2024); Lindquist et al. (2022);	fragmented systems can distort or
Quality	inclusion efforts.	Schwarcz & Leonhardt (2021)	dilute impact.
	Inclusion strategies aligned		
Social Equity &	with ESG have greater effect	Jain et al. (2024); Malik et al.	Reinforces the "S" in ESG; inclusion
Poverty	when addressing gender,	(2022); Ubeda et al. (2023); Tekin	leads to measurable social gains when
Alleviation	informality, and poverty.	(2025); Shah et al. (2024)	contextual factors are integrated.
	Increased digital ESG		
ESG	disclosures do not guarantee		
Measurement,	authenticity; issues of	Thomas (2023); Li et al. (2024);	Highlights need for standardized ESG
Integrity &	greenwashing and inconsistent	Roy & Vasa (2025); Ahmad et al.	metrics and impact validation to ensure
Greenwashing	metrics are prominent.	(2024)	inclusion leads to real sustainability.

Table 1. Synthesis of the literature by macro-themes.

### 3. Methodology

This study employs a robust hybrid methodological framework to critically assess the relationship between financial inclusion and the Environmental, Social, and Governance (ESG) components across 103 developing countries over a 12-year period. The originality and technical-scientific efficiency of this approach lie in its dual use of panel data econometrics with instrumental variables (IV) and machine

learning (ML) techniques, which together allow for both causal inference and predictive classification of inclusion-ESG dynamics.

Panel Data with Instrumental Variables. To examine the unique effect of each ESG pillar on financial inclusion—assessed through the Account Age variable (percentage of adults having access to formal financial services)-the study uses panel regression models on both Generalized Two-Stage Least Squares (G2SLS) and Fixed Effects Two-Stage Least Squares (TSLS) estimators. The two specifications facilitate checks on the robustness of estimates under alternative error specifications and unobserved differences between groups. Instrumental variables were chosen very carefully from exogenous sources, mainly from environmental and demographic variables (e.g., climate stress, access to energy, sanitation facilities), for dealing with the problem of endogeneity, measurement error, and simultaneity bias in the estimates. The IV approach improves the internal validity of the estimated coefficients and guarantees causal direction of ESG variables on financial inclusion and not vice versa (Jain et al., 2024). In all ESG dimensions, the IV panel models report analogous statistical significance and goodness of fit firmly established for all models except the governance model. In the environmental model, for instance, the R<sup>2</sup> values reach 35%, and the Wald  $\chi^2$  tests report very significant levels (p < 0.001), signaling the robustness of the models and the quality of the instruments used (Chininga et al., 2023). Analogous robustness also arises in the models of the social and governance variables, confirming the technical soundness of the econometric framework (Gidage & Bhide, 2024).

Machine Learning Models: Regression and Clustering. As a complement to the econometric analysis and in order to capture non-linear behavior and latent heterogeneity between countries, the analysis incorporates a set of supervised and unsupervised machine learning algorithms. In the regression analysis, a set of machine learning models-Random Forest, Neural Networks, Support Vector Regression, and Boosting-were compared on a set of normalized performance criteria like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R<sup>2</sup>. Random Forest Regression proved to be the most efficient, having the highest explanatory power ( $R^2 = 0.689$ ) and lowest prediction errors on all the indicators. Its ensemble-based nature is well suited for capturing non-linear and hierarchical ESG variable interactions, which is important in the structurally complex and heterogeneous developing countries (Li, 2025). In the clustering analysis, a series of algorithms like Hierarchical Clustering, DBSCAN, and Model-Based Clustering were tried out, but Fuzzy C-Means (FCM) performed better according to a set of cluster validity indices like Calinski-Harabasz index, Entropy, and Silhouette scores. FCM's soft clustering approach is well suited in the Global South region where countries tend to share overlapping ESG and financial inclusion traits. FCM's partial membership in multiple clusters captures the fuzzy and often hybrid development patterns witnessed in those settings and thus has great utility in policy segmentation as well as targeted policies (Shak et al., 2024). These machine learning processes not only corroborated the econometric findings but also facilitated interpretable segmentation of countries into clear ESGinclusion archetypes and higher utility in policy design for practical application. The methodological framework used here proves very efficient for three interconnected reasons. First, it accommodates multidimensionality by breaking the ESG framework down into their respective pillars-Environmental, Social, and Governance—so as to better analyze their unique impacts on financial inclusion. It prevents the oversimplification found in composite ESG indices and gives a better insight into what drives the relationship between a set of ESG variables and their outcome variables as well as the relationship between ESG variables and financial inclusion outcomes. Second, it balances prediction and causality through the integration of instrumental variable (IV) panel data econometrics and machine algorithms. Whereas the IV models provide strong causal inference, machine learning methods-particularly

clustering and regression—detect non-linear and complex dynamics and improve predictive performance. Third, the approach provides strong policy utility by facilitating country segmentation via clustering and thus pinpointing unique development profiles like "environmentally progressive but financially excluded" or "socially advanced but institutionally weak." These findings underpin the development of more specific and context-driven financial inclusion policy approaches. Generally speaking, combining rigorous econometric modeling and AI-driven tools produces a methodologically sound, science-driven, and policy-worthy approach for studying the relationship between financial inclusion and ESG in developing nations.

Data. The data employed in this analysis were obtained mainly from the World Bank, an internationally well-established organization for development data of a high quality and completeness level. The variable Account Age was used as a measure of financial inclusion and represents the proportion of respondents who report having an account (resp., jointly or individually) at a bank or other type of financial institution or who used a mobile money service themselves in the last year. This measure comes from the World Bank's Global Findex Database and has become accepted as a good proxy for having formal access to financial services-digital included-in the academic community and in particular in developing economies where the traditional banking system may be underdeveloped (Hicham & Hicham, 2020). The variables corresponding to the ESG (Environmental, Social, and Governance) framework's three pillars were obtained from the Foreign ESG Database of the World Bank, which carries a harmonized and internationally comparable set of indicators. These encompass, for instance, the consumption of renewables, female labor force participation rate, regulatory quality, and other variables related to each ESG dimension. The employment of disaggregated variables on ESG allows for a finer analysis of their unique impact on financial inclusion (Eldomiaty et al., 2020). To guarantee the completeness and consistency of the dataset and prevent imbalances in the sample's distribution when applying econometric and machine-learning algorithms, a missing data interpolation method based on mean substitution has been applied. Despite being a very basic method, it has proven efficient in panel data settings where missing data become sparse and not regularly concentrated at specific units. Mean substitution keeps the marginal distribution of variables and prevents the incorporation of bias in econometric and machine-learning models. The method ensured the sample's statistical relevance, enhanced the stability of the models, and increased the general reliability of the empirical results (Waliszewski, 2023).

### 4. Sustainable Finance and Environmental Resilience: Modelling Financial Inclusion through the "E" in ESG

This empirical analysis investigates the relationship between environmental sustainability and financial participation in developing nations through the proxy Account Age as a measure of participation in financial services. As financial access is rapidly emerging as a central driver of inclusive growth, it has become imperative to investigate the manner in which the environment and environmental considerations shape such access. Under the ESG (Environmental, Social, Governance) framework, the current study focuses on the "Environmental" pillar and examines the effect of six significant environmental variables on a multi-period panel dataset. To address the possible issue of simultaneity or endogeneity, both G2SLS (Generalized Two-Stage Least Squares) and TSLS (Two-Stage Least Squares) specifications are used in estimation and yield statistically reliable results. Prior research supports the use of environmental indicators such as renewable energy use and land quality in studying the financial inclusion-environment nexus (Jain et al., 2024). The overall goal is to provide empirical insight helpful for sustainable policy

making and show that well-designed environmental policies can also enhance financial participation. Recent findings show that digital financial inclusion can inhibit greenwashing while strengthening authentic ESG performance, supporting the idea that financial access and environmental responsibility can reinforce each other (Li et al., 2024). Moreover, inclusive finance is shown to improve environmental quality through enhanced service accessibility and reduced household vulnerability, especially when financial systems are aligned with ecological priorities (Wang et al., 2022). The findings establish positive synergies between economic access and climate resilience and suggest that environmental stewardship will not be a restraint but rather a facilitator of sustainable and inclusive development (Table 2).

Table 2.	Environmental	Determinants	of	Financial	Inclusion:	G2SLS	and	TSLS	Estimates	for
Developir	ng Economies									

Y			Accou	ınt Age				
Instruments	CFTC ELEC ED	UE NETU LABR L	LIFE LITR U5MR N	MIGR WATS SANS POP6 PDEN OVWT UNDT FLMR				
	SEPR SEC	GP UNEM CCOR I	ESRP GDPG GOVI	E PSTB WPAR REG	Q RULE JART SL	RI VOAC		
	G2SLS randor	n effects, using 123	5 observations	Fixed-effects TSLS, using 1235 observations				
	Coefficient	std. error	Z	coefficient	std. error	Z		
Constat	30.9837***	9.35984	3.310	15.2505	10.4782	1.455		
AGRL	-0.593386***	0.140859	-4.213	-0.697766***	0.144949	-4.814		
AGVA	-1.97514***	0.308340	-6.406	-0.735469*	0.389733	-1.887		
FOOD	0.411509***	0.0721429	5.704	0.444373***	0.0612962	7.250		
HIDX	0.375768***	0.128120	2.933	0.481345***	0.135033	3.565		
RENE	0.139831***	0.0447439	3.125	0.120447***	0.0412597	2.919		
PROT	1.50538***	0.203737	7.389	1.37595 ***	0.179812	7.652		
Statistics		SSR = 1414.82			SSR = 122192			
	sigma-	-hat = 1.07338 (df =	1228)	sigma-hat = 10.4172 (df = 1126)				
	R-squared	$l = corr(y, yhat)^2 =$	0.345050	R-squared	$= corr(y, yhat)^2 =$	0.347297		
	]	Included units $= 103$	3	I	ncluded units $= 103$	3		
	Time-serie	es length: $min = 11$ ,	$\max = 12$	Time-serie	es length: $min = 11$ ,	$\max = 12$		
	Wald chi-	square(6) = 457.630	6 [0.0000]	Wald chi-	square(6) = 491.936	6 [0.0000]		
	sigma	-hat(within) = $10.4$	17219	]				
	sigma-	hat(between) = $19.3$	360108					

The principal goal of this empirical research involves investigating the connection between financial inclusion, as captured by the measure used here for this research—namely, the Account Age variable— defined as the proportion of those who reported having an account (in their own right or jointly) at a bank or other type of financial organization or who used a mobile money service themselves in the previous year—against a set of environmental variables collected together as representing the "Environmental" part of the ESG approach (Environmental, Social and Governance). The analysis will be confined exclusively to developing countries since their countries present specific differences in their structural, socio-economic and institutional contexts in relation to high-income economies and since their levels of financial inclusion and success or prominence of environmental policies largely depend on them.

The empirical approach employs six explanatory variables with close links to the Environmental pillar of ESG: (1) AGRL – agricultural land as a percentage of total area, (2) AGVA – value added in agriculture, forestry, and fisheries as a percentage of GDP, (3) FOOD – food production index, (4) HIDX – heat index of the number of days on which the apparent temperature rises above  $35^{\circ}$ C, (5) RENE – share of consumption of renewable energy as a percentage of total energy consumption, and (6) PROT – protected areas on the land and seabed as a percentage of total territorial area. The variables were

employed since they capture key environmental dynamics for developing countries and should impact financial inclusion directly or indirectly.

The empirical specification employs two instrumental variable forms in order to take account of potential endogeneity: a G2SLS (Generalized Two-Stage Least Squares) random effects and a TSLS (Two-Stage Least Squares) fixed effects specification. The two specifications share a balanced panel of 1235 observations covering 103 countries over a time series ranging from 11 years to 12 years in duration. The use of instrumental variables serves as a necessity in order to address issues related to simultaneity-based endogeneity concerns, omitted variable bias, or measurement error. The instruments used are diverse and potent and include human development proxies (e.g., access to electrification and the internet, literacy rate, school enrollments), public health indicators (e.g., access to sanitation and clean water and under-5 mortality rate), demographic variables (e.g., crude population density and proportion of the age group 65 and above), and governance-quality variables (e.g., control of corruption sub-indices and sub-indices of rule of law, regulatory quality and voice and accountability).

Both models yield the same results and are statistically significant and consistently depict clear understanding of the effects of environmental variables on financial inclusion in developing countries. The AGRL (land area used for agriculture) variable has a negative and highly significant relationship to financial inclusion in both models (G2SLS: -0.593, z = -4.213 and TSLS: -0.698, z = -4.814). It reflects countries with large agricultural land areas having lower financial inclusion. It may be a reflection of domination by agrarian economies and informal economies where financial institutions remain underdeveloped and formal financial services and instruments remain limited. The same holds true for the AGVA variable—value added through agriculture, forestry, and fish—referring a negative and significant relationship (G2SLS: -1.975, z = -6.406 and TSLS: -0.735, z = -1.887), supporting the argument that agrarian economies without integration or modernization do not integrate or become included in financial systems.

On the other hand, other environmental variables show a positive and significant relationship. The FOOD variable (index of food production) has a positive and significant relationship both in the two models above (G2SLS: 0.412, z = 5.704; TSLS: 0.444, z = 7.250). It reflects the fact that in countries where levels of agricultural productivity tend to be stronger, the inhabitants tend to be in contact with financial institutions more often. This can be justified by the reasoning that a higher food production level may foster well-organized value chains, contract farming, agricultural insurance and microcredit availability, and thus encourage demand for financial products and services as well as both demand- and supply-push motives for financial institutions. Likewise, the HIDX variable (heat index) also shows a positive and significant relationship (G2SLS: 0.376, z = 2.933; TSLS: 0.481, z = 3.565). This may be less intuitive and may be a sign of adaptation behavior: in countries with high frequencies of extreme weather conditions, investment in resilience may be enhanced and thus maybe the take-up of digital financial means such as climate insurance or environmental shock coping savings accounts.

The other significant finding relates to RENE (consumption of renewable energy), which has a positive and statistically significant coefficient on both specifications (G2SLS: 0.140, z = 3.125; TSLS: 0.120, z = 2.919). The correlation might be driven by the increasing penetration of renewable energy facilities, such as off-grid mini-grids and photovoltaic panels, in remote villages and small towns. These systems would be based on digital payment systems (e.g., pay-as-you-go solar) and funded by microfinance or leasing schemes, which in turn create an incentive for opening a financial account or the use of mobile money services on the part of the consumers (Ababio et al., 2023).

The PROT variable, which accounts for the proportion of protected areas, has the strongest positive correlation with financial inclusion (G2SLS: 1.505, z = 7.389; TSLS: 1.376, z = 7.652). The finding can be explained through the prism of inclusive environmental governance: if active conservation policies take effect, they tend to be grounded on community participation, formalized funding channels, and

coordination between NGOs and external institutions, which may stimulate access to financial services in local communities (Paliienko & Diachenko, 2024).

Both estimations from a model performance perspective hold good explanatory power. The R<sup>2</sup> measure of determination approximates 34.5% for the G2SLS and approximately 34.7% for the TSLS fixed effects method, hence explaining a good percentage of variance in the Account Age indicator. Additionally, the Wald test for the collective explanatory significance of the explanatory variable produces highly significant values ( $\chi^2 = 457.6$  for G2SLS and  $\chi^2 = 491.9$  for TSLS; both p < 0.001), confirming the statistical validity of the models.

In short, this analysis offers systematic and strong empirical evidence that the environmental dimension of ESG has a crucial impact on the level of financial inclusion in developing countries. Precisely, those traditional and underdeveloped agricultural system-related variables tend to behave in a negative correlation manner vis-a-vis formal access to financial services. However, those environmental variables related to productivity, resistance, accessible sustainable energy, and policy enforcement of conservation efforts tend to be correlated in a positive manner vis-a-vis better levels of inclusion (Dovbiy, 2022).

These findings are most directly relevant for policymakers and development institutions as they reveal the complementarities between the access to finance and environmental sustainability. Encouragement of environmentally sustainable behavior may not only be beneficial from an environmental angle but also may entail economic and societal spillovers through enhanced coverage of financial services by previously under-banked groups.

With regard to the 2030 Agenda of Sustainable Development and in particular in relation to SDG 8 (Decent Work and Economic Growth) and SDG 13 (Climate Action), this study contributes to an emerging literature on the means through which environmental factors may be leveraged as drivers of inclusive economic growth. The cross-cutting ESG approach taken by this research calls for a holistic development comprehension whereby environmental stewardship is not a constraint but an accelerator of inclusive, resilient, and sustainable financial ecosystems in the Global South. The determined correlation between the "E" of ESG and financial inclusion means policies must be couched so they may address simultaneously both climate resilience and economic access and consequently achieve synergies that promote both ecological integrity and social equity.

### 4.1 Financial Access in the Climate Era: A Cluster-Based Exploration of Environmental-Economic Interactions

This section presents advanced clustering analysis aimed at revealing dynamics between environmental and financial variables in developing countries through the use of the Fuzzy C-Means method. The countries were grouped by similarity in the six environmental indicators and the Account Age measurement as a proxy measure of access to financial services. The nine-cluster solution based on best statistical fit as well as interpretability was finalized. The cluster analysis presents divergent profiles and illustrates ways in which environmental participation, agricultural organization, and proportion of renewables intersect with participation stages in the economy. The results offer actionable insights in the Complex and sometimes patchwork relationship between economic participation and sustainability in the Global South (Table 3).

Table 3. Clustering Model Comparison for Environmental-Financial Profiles: Fuzzy C-Means and Alternatives

	Density	Fuzzy C-		Model	Neighborhood-	Random
	Based	Means	Hierarchical	Based	Based	Forest
	Clustering	Clustering	Clustering	Clustering	Clustering	Clustering
Maximum diameter	0.0	1.0	0.0	0.0	0.0	0.0
Minimum						
separation	1.0	0.03	0.27	0.0	0.04	0.16
Pearson's Î <sup>3</sup>	0.0	0.54	1.0	0.05	0.7	0.09
Dunn index	1.0	0.04	0.4	0.0	0.1	0.16
Entropy	0.0	1.0	0.0	0.0	0.0	0.0
Calinski-Harabasz						
index	0.0	1.0	0.0	0.0	0.0	0.0
Clusters	0.0	1.0	1.0	0.88	1.0	0.88
Ν	0.0	0.0	0.0	0.0	0.0	0.0
R <sup>2</sup>	0.0	0.84	0.53	0.61	1.0	0.6
AIC	1.0	0.17	0.49	0.37	0.0	0.42
BIC	1.0	0.18	0.53	0.39	0.0	0.44
Silhouette	1.0	0.56	0.68	0.0	0.84	0.16

Among the cluster algorithms we experimented with—Density-Based Clustering, Hierarchical Clustering, Model-Based Clustering, Neighborhood-Based Clustering, Random Forest Clustering, and Fuzzy C-Means Clustering—the latter's Fuzzy C-Means method emerged as the best-placed and efficient method based on a multi-metric performance assessment (Kaushal et al., 2024). By normalizing and comparing twelve performance measures—compactness measures and cluster validity and separation measures alike—FCM produced the highest mean performance score, pointing towards FCM as the method typically superior in this use case.

One of the great strengths of FCM as a technique is its ability to accommodate data points being part of several groups to any extents of membership. This soft type of clustering method is very useful with complex data where different groups may lack distinct boundaries. In contrast with hard clustering methods such as k-means, where each point is assigned exclusively to a particular group, FCM captures more subtle patterns in the data and provides more interpretability and flexibility (Zhang et al., 2020). FCM's flexibility is evident in the very high performance across all significant parameters such as the R<sup>2</sup> value (0.84), reflecting the variance accounted for in the case model, and the silhouette score (0.56), reflecting good cluster separation and cohesion.

FCM achieved a 1.0 normalized value in our comparison experiment in terms of the Calinski-Harabasz index, Entropy, and Clusters. The measure of the Calinski-Harabasz index measures the ratio of betweencluster variance and the within-cluster variance and a high value reflects well-separated and compact clusters. Achieving a 1.0 value in this case provides a clue that FCM developed compact and wellseparated clusters. Entropy measure in clustering can reflect randomness or uniformity in cluster assignment distribution and also achieved a maximum score in FCM case. A value of 1.0 means FCM achieved high consistency in the clustering structure. That it also scored 1.0 in Clusters means the number of achieved clusters lay at the optimum or very near the optimum value to be expected in the dataset (Shen et al., 2021).

In addition, FCM performed comparably on other key measures too. On the measure for the minimum between-clusters distance and the maximum within-clusters diameter in the form of the Dunn index, FCM value at 0.04 can also be considered low but comparable since the maximum value 1.0 could only

be achieved by the Density-Based Clustering method alone. FCM scored 0.54 at Pearson's  $\gamma$ , the measure of the distance between points and the degree to which the points belong to the same cluster or different ones, being behind Hierarchical Clustering alone at 1.0 when it comes to its score in this measure. This reflects good geometric form preservation of the data through FCM. Even in the case of Minimum Separation and Maximum Diameter measuring the distance between the clusters and the size of the largest cluster respectively, FCM retained a good measure and proved to have good compactness and well-separatedness simultaneously.

Both Neighborhood-Based Clustering and Hierarchical Clustering did well on both measures—a silhouette score of 0.68 and 0.84, respectively—but were far from as reliably high-performing as FCM. The levels of R<sup>2</sup> were lower both times (0.53 and 1.0 respectively), but both those algorithms scored 0.0 on important measures such as the Calinski-Harabasz index, AIC, and BIC, indicative of their potential inefficiencies in model selection and complexity penalty. FCM's mean overall score of 0.58 is indicative of a well-balanced performance without falling low on any of the measures provided (Paliienko & Diachenko, 2024).

The key point also resides in a good interpretability in high-dimensional spaces, and this makes FCM a good choice for exploratory analysis in bioinformatics, segmentation of the market, and image classification. Its use of fuzzy logic also aligns with the reality of the overlapping classes and imprecise groupings in the real world. In practical terms, this can lead to actionableness over strict partitioning algorithms.

Fuzzy C-Means Clustering outperforms in performance based on good cluster compactness and separability and model validity as well as being more flexible and interpretative. Its highest mean score of the algorithms considered speaks volumes about the flexibility and potency of the algorithm. No cluster method is its best across all measures, yet since FCM performs well across a range of measures over and over again, it is the optimum method for this dataset. This conclusion based upon empirical performance is further enriched through the theoretical advantages of the algorithm's capacity to accommodate fuzzy boundaries and preserve the internal organization of the data. (Table 4).

Cluster	1	2	3	4	5	6	7	8	9
Size	90	211	109	126	326	108	179	62	25
Explaine	0.111	0.250	0.041	0.082	0.237	0.098	0.121	0.037	0.024
d									
proportio									
n within-									
cluster									
heterogen									
eity									
Within	571308	128352	212323	423023	121571	501548	619337	187747	121334
sum of	,00	9,00	,00	,00	9,00	,00	,00	,00	,00
squares									
Silhouett	0.078	-0.065	0.222	0.198	-0.055	0.103	-0.030	0.128	0.159
e score									

Table 4. Comparative Evaluation of Clustering Algorithms: Fuzzy C-Means as the Optimal Method for ESG-Finance Segmentation

Center	-1.512	0.450	-0.546	-0.667	-0.185	-0.848	-0.819	0.302	0.317		
Account											
Age											
Center	-0.007	0.339	-0.393	1.417	-0.442	-1.577	-0.113	-0.253	0.755		
AGRL											
Center	0.995	-0.813	-0.451	1.592	-0.569	0.392	0.143	-0.803	-0.076		
AGVA											
Center	-0.795	0.589	0.072	-0.105	0.596	-0.120	-0.467	-1.943	-1.404		
FOOD											
Center	0.779	-0.507	-0.541	-0.557	-0.539	-0.554	-0.470	-0.332	-0.563		
HIDX											
Center	-1.412	-0.903	0.207	0.644	-0.370	1.914	0.255	-0.435	-1.115		
RENE											
Center	-1.146	1.100	-0.591	-1.181	-0.364	0.317	-0.365	0.680	1.593		
PROT											
Note. The	Note. The Between Sum of Squares of the 9 cluster model is 4089.09										
Note. The	Total Sun	n of Square	es of the 9	cluster mo	del is 9224	4.96					

Fuzzy C-Means clustering outputs reveal a sharp segmentation of countries (or regions) based on the interaction between financial inclusion in terms of Account Age and a range of environmental variables within the E (Environment) corner of the ESG framework. Cluster analysis, within the developing country context, shows the nexus between financial infrastructure and environmental engagement, resilience, and sustainability (Zioło et al., 2022). The variables employed are AGRL (land used for agriculture), AGVA (agriculture, forestry and fishery value added), FOOD (% food production index), HIDX (% heat index 35), RENE (% renewable energy consumption), and PROT (% terrestrial and marine protected areas). The overall goodness of fit of the clustering model as a ratio of BSS/TSS at circa 44.3% reflects significant but relative capture of the heterogeneity and supports significant overlapping country profiles (López-Oriona et al., 2022). Financial inclusion perspective-wise, Cluster 2 (with the highest proportion explained at 25%) possesses a moderately high Account Age (0.450) and high environmental responsibility through PROT (1.100), but low values in RENE (-0.903) and AGVA (-0.813). This might reflect the regions with good institutions and conservation system but poorly developed renewables and agri-economies. Cluster 4 with its high values in both AGRL (1.417) and AGVA (1.592) overtly reflects regions with a high agrarian core. The Account Age in this case surprisingly emerges as negative (-0.667), and this reflects even agri-advanced areas and financial service bottlenecks might arise. This contrast indicates economic activity-financial access mismatch and a key development bottleneck. Cluster 9 comprises the smallest number of companies (only 25) but is striking in the high level of environmental responsibility: PROT (1.593), RENE (-1.115), and positive Account Age (0.317). Its relatively low renewable energy consumption is inconsistent with high levels of biodiversity conservation levels. The fit may be the result of conservation policies advocated and pursued by the state within those contexts where green technologies are still inaccessible. Group 1 possesses a clear profile of weaknesses: very low Account Age (-1.512), lower usage of renewable energy (-1.412), and minimal protection of environments (-1.146). It has moderately high AGVA (0.995) but extremely low financial cover and heat stress (HIDX = 0.779), corresponding to climate-vulnerable and institutionally undercovered locales. Cluster 6 is a very interesting paradox. It has the largest RENE (1.914) and positive PROT value (0.317), yet very low Account Age (-0.848) and AGRL (-1.577). This corresponds to green transformation at the nascent stage—possibly driven by donors or technologically enabled—in a context

of limited classical agriculture and minimal access to finance. This might reflect the case of countries turning to renewables while no overall financial safety net exists for their populations (Azkeskin & Aladağ, 2025). Cluster 3 shares the middle Account Age (-0.546) and relatively evenly distributed environmental values, specifically RENE (0.207), and therefore can be deemed middle-ground segment with lower variablity since it also witnessed lower explained heterogeneity (4.1%) and highest silhouette value (0.222), being a dense and homogeneous segment. Cluster 5, being the largest cluster with 326 units, contains mixed attributes: Account Age (-0.185), AGVA (-0.569), and barely better-than-average FOOD (0.596), yet both negative RENE (-0.370) and PROT (-0.364). These contradictory bits of information suggest a type of profile of the "developing majority"-economies with modest food producing capacity but low ecological and financial inclusion levels. Cluster 8 possesses low FOOD (-1.943) and Account Age (0.302) and below-average readings in all other measures. Its small size (N = 62) and low value of the silhouette (0.128) indicate it might include outliers or the economies in the process of transforming towards the outside conditions. Finally, Cluster 7, whose values are close to the mean, specifically in RENE and PROT measures (0.255 and -0.365), might serve as a "benchmark" profile—neither too underdeveloped and yet high performances are below its level yet also far from the highest level and thus it still has room to develop and improve both the fiscal and the environmental fronts. On the whole, the clustering shows financial access does not have high inter correlation across environmentally varied segments in developing countries. Conditions in agriculture and the environment do not directly correlate with high financial access and vice versa. The framework pushes the importance of the integration planning financial access and the environment in the first place in areas exposed to climate risk or sustainability transitions. The findings support the unified ESG strategy where social infrastructure (i.e., access to banks) should never be decoupled from green investment, particularly in emerging and vulnerable economies (Table 5).

	Account Age	AGRL	AGVA	FOOD	HIDX	RENE	PROT
Cluster 1	0.315	0.476	-0.797	-0.458	2.169	-0.540	-0.506
Cluster 2	0.355	-0.435	0.795	0.693	-0.060	0.990	-0.347
Cluster 3	-0.158	-0.575	-0.258	-0.176	-0.331	-0.485	0.187
Cluster 4	1.185	1.409	-0.653	0.064	-0.233	-0.438	1.126
Cluster 5	-0.353	-0.576	0.441	0.334	-0.153	-0.410	-0.611
Cluster 6	-1.255	0.939	-0.592	-0.069	-0.307	0.611	1.481
Cluster 7	0.218	0.412	-0.763	-0.577	-0.104	-0.144	0.382
Cluster 8	-0.519	-0.697	0.904	-1.071	-0.137	-0.138	-0.830
Cluster 9	0.333	-0.399	0.603	-1.020	-0.286	2.005	-0.843

Table 5. Profiling the Environment-Finance Nexus: Cluster Characteristics from ESG-Based Fuzzy C-Means Analysis

The outcomes from Fuzzy C-Means clustering present a refined breakdown of countries (regions) according to the interaction of financial inclusion, as quantified by Account Age, and a set of environmental variables taken from the E (Environment) pillar of the ESG framework. The clustering analysis in the developing countries context shows the relationship between environmental engagement, resilience, and sustainability and financial infrastructure (Latifah, 2022). The variables used comprise AGRL (agricultural area), AGVA (agriculture, forestry, and fishing value added), FOOD (index of food production), HIDX (heat index 35), RENE (consumption of renewable energy), and PROT (land and ocean and island protected areas). The global fit of the clustering model, wherein the ratio of the Between Sum of Squares (BSS) and the Total Sum of Squares (TSS) is about 44.3%, shows that the nine-cluster solution accounts for a considerable but not full extent of the heterogeneity and therefore indicates meaningful although overlapping country profiles (Sasmita et al., 2023).

From a financial inclusion viewpoint, Cluster 2 (the largest proportion explained at 25%) has a moderately high Account Age (0.450) and high environmental concern through PROT (1.100), but low scores on RENE (-0.903) and AGVA (-0.813). This may indicate those regions with good institutions and conservation policies but not well-developed renewable infrastructures and agricultural economies. Cluster 4, having high scores in AGRL (1.417) and AGVA (1.592), evidently locates regions having a well-established agricultural base. The Account Age here remains surprisingly negative (-0.667), however, and may indicate even agriculturally developed regions will be encumbered by access to financial services. This gap reflects a disconnect between economic activity and access to finance, a key development bottleneck.

Cluster 9 has the lowest number (only 25) but has a strong environmental orientation: PROT (1.593), RENE (-1.115), and positive Account Age (0.317). Paradoxically, although it has minimal use of renewables, it has excellent biodiversity preservation. The cluster might be a reflection of government-run policies for conservation in areas where green technologies remain inaccessible.

On the other hand, Cluster 1 has a definite outline of vulnerability profile: extremely low Account Age (-1.512), low levels of renewable energy consumption (-1.412), and low levels of environmental protection (-1.146). It has a moderately high AGVA (0.995) but low levels of financial inclusion and heat stress (HIDX = 0.779), reflecting climate-vulnerable, institutionally underserved areas.

Cluster 6 presents a very interesting contradiction. It has the largest RENE (1.914) and a positive PROT (0.317) score but very low Account Age (-0.848) and AGRL (-1.577). This indicates an early-stage green transition—potentially technologically enabled or donor-driven—wherein there is low traditional agriculture and limited financial coverage. This may be representative of those countries adding or embracing renewables without a general financial cushion of safety in their populations (Bondarenko et al., 2025).

Cluster 3 has a medium Account Age (-0.546) and fairly well-balanced environmental values, primarily RENE (0.207), and thus appears a middle-of-the-way group with minimal variability, as also reflected through low explained heterogeneity (4.1%) as well as maximum silhouette score (0.222), qualifying it as a close-knit, stable segment.

Cluster 5 is the largest cluster at 326 units but has mixed features: Account Age (-0.185), AGVA (-0.569), and modestly above-average FOOD (0.596), but negative RENE (-0.370) and PROT (-0.364). These contradictory signs indicate a profile of a "developing majority"—modest food productivity economies but not ecologically and financially inclusive economies.

Cluster 8 stands out for having a low FOOD (-1.943) and Account Age (0.302) along with below-average scores on all other indicators. Its modest size (62 units) and low silhouette measure (0.128) suggests it may be comprised of outliers or transit economies.

Lastly, Cluster 7, at levels close to the mean in RENE (0.255) and PROT (-0.365), can be called a "baseline" profile—neither highly underdeveloped nor high performers but having space for advancement both on the economic and environmental sides.

By and large, the clustering indicates uneven distribution of financial inclusion over environmentally differentiated regions in developing nations. Strong agriculture and environmental performance by regions do not necessarily translate to high financial access and vice versa. The model emphasizes the necessity of aligning financial inclusion policies with environmental policy, particularly in those regions under environmental risk or undergoing sustainability transitions. These observations justify a holistic ESG approach in which social infrastructure (such as access to banks) must not be differentiated from environmental investment, especially in vulnerable and emerging economies (Duan & Sun, 2020), (Figure 1).

Figure 1. Fuzzy C-Means Clustering of Developing Economies: Optimal Cluster Determination and Profile Mapping



The figure displays the outcome of Fuzzy C-Means cluster analysis through combination of statistical evaluation of the clustering model and visualization through cluster assignments inspection. On the left figure in the illustration, the model performance using cluster numbers two through ten is shown through taking the evaluation of the three important measures, namely Within Sum of Squares (WSS), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). These measures direct the choice of the ideal number of clusters through a compromise between fit and model complexity. Due to heavier penalization against complexity than AIC, BIC plays a pivotal role in the choice between competing models. Here, the red marker point on the curve is put at the point where the corresponding solution is a nine-cluster value; this means nine clusters since it provides the best compromise between maintaining the structure of the data and overfitting the data. The declining WSS and flattening of BIC and AIC support this choice as it shows a sign of approaching the point where marginal improvement in the quality of the clustering in returning more than nine clusters will be extremely minimal. Having nine units thus remains statistically warranted and parsimonious compared to the complexity of the data (Kaushal et al., 2024).

On the right-hand side of the figure we have a two-dimensional visualization of the clustering result. The two-dimensional representation, presumably produced through the use of dimension reduction techniques such as t-SNE or PCA, diminishes high-dimensional data to a format permitting visual

interpretation. Each marker marks a single record and coloring illustrates cluster membership according to the Fuzzy C-Means function. The clustering appears typically coherent and dense aggregations of points and discrete color-coded sets of clusters occupy the plot area. There is undoubtedly some overlap—especially in the middle of the plot—due to the fuzzy approach to the method where each and every marker will belong in part to two or more collections of clusters (Shen et al., 2021). However, visual patterning assures success in segment capture by the function in the data. Large clusters labeled 2, 3, 4, and 6 dominate the visual area and correspond to the size range of the clusters the researcher had been working with earlier. Small cluster 8 and 9 cluster at the edge and this will possibly reflect those occupying niche or specialist portions of the population.

On average, the image demonstrates good supportive evidence of the suitability of a nine-cluster solution for this analysis. Statistical assessment assures nine clusters offer the optimal compromise between model fit and model complexity and the map illustrates the fact the clusters are not just sensible but also are quite well-separated from each other (Sunori et al., 2023). While potentially some fuzziness across boundaries is apparent—a typical concomitant feature of the Fuzzy C-Means algorithm—clusters are revealed with internal coherence and external differentiation. It validates the ability of the model to reveal complex and overlapping patterns in the data set, and in the majority of real-world datasets usually reflects a situation when categories do not tend to demarcate each other with sharp boundaries. The combination of the model diagnostics and visual inspection reflects the robustness and explanatory potential of the nine-cluster solution and renders it a good point of commencement for further interpretation, segmentation, or decision-making over the revealed groups (Figure 2).

Figure 2. Exploring ESG Interactions in Developing Economies: Cluster-Based Pair plot of Account Age and Environmental Metrics



The figure displays a pairplot matrix examining the interaction between the level of financial inclusion as captured through Account Age—the proportion of individuals who own a bank account or a mobile money service—against a group of variables representative of the E (Environment) component of the ESG framework for developing economies. The included environmental variables used in the time series analysis were Agricultural Land (AGRL), Agriculture, Forestry, and Fishing Value Added (AGVA), Food Production Index (FOOD), Heat Index 35 (HIDX), Renewable Energy Consumption (RENE), and Terrestrial and Marine Protected Areas (PROT). The observations are color-coded by cluster membership from the Fuzzy C-Means clustering since each cluster defines a segment or profile of observations. As part of this ESG framework emphasizing development, Account Age behaves as a surrogate measure

As part of this ESG framework emphasizing development, Account Age behaves as a surrogate measure of financial participation and access—essential elements of inclusive growth. A central question here is whether higher financial inclusion also correlates with better environmental performance or engagement. The distribution of Account Age is highly varied between clusters. Clusters 2, 4, and 9 show relatively high financial inclusion levels, and others like 6 and 8 cluster at the lower levels. These tendencies indicate that in developing nations, financial inclusion is not evenly spread and possibly correlated with dissimilar levels of environmental capacity and policy engagement (Essel-Gaisey & Chiang, 2022).

Examining AGRL and AGVA in particular, which capture land use and economic productivity in agriculture—key sectors in many developing economies—higher Account Age clusters tend to have medium and high values. Illustratively, Cluster 4 has a high proportion of account holders and high agricultural participation. This could be taken as a reinforcing relationship: higher financial access may be generating agricultural productivity through loans, insurance coverage, and investments in technology or vice versa, whereby higher productivity areas may be drawing better financial infrastructure (Supartoyo, 2023).

The Food Production Index (FOOD) also offers added depth. Clusters 2 and 4 show a crossover between better financial inclusion and food production and thus could mean food security and access to finance complement each other. Conversely, both low FOOD values and low Account Age in Cluster 8 might reflect underserved areas having both production challenges and a lack of financial access—typical of more vulnerable segments in developing nations.

Renewable Energy Consumption (RENE) emerges as a very insightful variable. Cluster 9 stands out having both a high scoring for both renewable energy consumption and financial inclusion, which indicates a correlation between environmentally innovative policies or infrastructure and improved access to financial services (Ababio et al., 2023). This trend confirms the assumption that green transition investment—e.g., off-grid solar or micro-financing for clean energy—may be concurrent with or promote financial inclusion.

PROT, a measure of the percentage of protected environmental areas, also displays a somewhat parallel trend. More affluent Account Age clusters like 4 and 6 also report above-average PROT levels, possibly as a result of policies at the local or national level combining financial access and environmental conservation efforts (Said, 2024). In contrast, low-protection clusters like cluster 8 also report limited access to finance, supporting a trend of developmental weakness.

The function of HIDX (Heat Index 35), as an indicator of climatic stress, seems less specific, although cluster 1, which indicates high heat exposure also has medium Account Age values. This could be a sign that a portion of populations experiencing climatic adversity may not always enjoy proportionate financial assistance and thus present equity concerns in climatic resilience planning.

As a whole, the pairplot verifies that financial inclusion—here measured through Account Age—crosses meaningfully with environmental variables at the heart of the E dimension of ESG. In developing nations, increased financial inclusion seems to co-occur alongside higher agricultural productivity, enhanced food security, the consumption of renewables, and the preservation of the environment. Conversely, areas of limited access to finance tend to coincide with economic and environmental deprivation. These findings reinstate the essence of aligning financial inclusion policy in the contexts of environmental sustainability policy, especially in emerging and under-resourced areas where ESG alignment presents both developmental and environmental returns on investment.

## 4.2 From Land to Finance: Machine Learning Insights on Environmental Determinants of Inclusion

We used eight regression models: Boosting Regression, Decision Tree Regression, K-Nearest Neighbors Regression, Linear Regression, Neural Network Regression, Random Forest Regression, Regularized Linear Regression, and Support Vector Machine Regression. These models were compared based on an extensive set of performance metrics: Mean Squared Error (MSE), Scaled MSE, Root Mean Squared Error (RMSE), Mean Absolute Error / Median Absolute Deviation (MAE/MAD), Mean Absolute Percentage Error (MAPE), and R<sup>2</sup> (coefficient of determination). To make the results comparable, results

from all models were normalized based on Min-Max scaling, enabling us to map a 0-1 value range to each metric on all models (Table 6).

		MSE	DIGE			Da
Algorithms	MSE	(scaled)	RMSE	MAE / MAD	MAPE	R <sup>2</sup>
Boosting Regression	0.66	0.64	0.72	0.73	0.56	0.28
Decision Tree Regression	0.25	0.24	0.3	0.22	0.23	0.7
K-Nearest Neighbors Regression	0.24	0.16	0.29	0.0	0.0	0.79
Linear Regression	0.76	0.83	0.8	0.8	0.64	0.12
Neural Network Regression	0.82	1.0	0.85	0.84	1.0	0.0
Random Forest Regression	0.0	0.0	0.0	0.08	0.21	1.0
Regularized Linear	1.0	0.78	1.0	1.0	0.86	0.16
Support Vector Machine	0.79	0.83	0.83	0.83	0.67	0.12

Table 6. Model Evaluation of Predictive Algorithms on Financial Inclusion Using Environmental Indicators

After normalization and aggregation of performance scores, Random Forest Regression came out on top as the best-performing model with a normalized average score of only 0.048, far below any other method. It also boasted the lowest MSE (162.916), lowest RMSE (12.764), and highest R<sup>2</sup> (0.689), which clearly placed it as the best at explaining variance in financial inclusion on the basis of environmental predictors. Its performance stood far above conventional methods like Linear Regression (R<sup>2</sup> = 0.226) or even sophisticated techniques like Neural Networks (R<sup>2</sup> = 0.164), which although possessing the theoretical advantage of being able to model non-linear patterns, did poorly on error scores.

These outcomes are theoretically in line what would be expected from Random Forests. As a decision tree ensemble method based on bagging decision trees, Random Forest lends itself well to capturing nonlinear, hierarchical, and interactive relationships between variables—properties found in environmental systems. The interaction between climate vulnerability (HIDX) and natural capital (AGRL, PROT), e.g., may be poorly captured through linear models. Likewise, the relationship between renewable energy take-up (RENE) and access to finance may be dependent on intervening conditions such as agricultural productivity or land use intensity, upon which a model like Random Forest can learn without specification in advance.

Furthermore, Random Forest has strong resistance against overfitting, particularly when compared to single decision trees or boosting models. This has a critical application in modeling environmental indicators between countries or regions that can differ significantly in size, resources endowed, as well as in institutional capacity. The capacity of the model in managing missing data, ranking feature importances, as well as the facility to accommodate noise, also make it highly useful when dealing with imperfect real-world environmental data.

Conversely, other models in the comparison suffered from considerable weaknesses. As an example, K-Nearest Neighbors performed competitively on a few of the metrics but were sensitive to the presence of outliers and unsuited for the large environmental dimensionality. Boosting and Neural Networks, as strong performers generally, exhibited high error rates—likely as a result of overfitting or poorly optimized hyper parameters. Linear and Regularized Regression models, as easy-to-interpret models, performed poorly as a result of their failure to identify complex non-linearities inherent in the

environmental space. The Support Vector Machine model also performed middling on all of the metrics and needs heavy tuning in order to perform at a peak level—less ideal for exploratory policy modeling. From a policy and interpretability viewpoint, Random Forest also has other benefits. It allows for the estimation of the importance scores of features, giving insight on which environmental variables most significantly affect financial inclusion outcomes. As a case in point, initial model results indicate protected areas (PROT) and the consumption of renewably generated energy (RENE) carry high predictive weight—identifying the possible effect of conservation policy and green infrastructure on participation in the financial system. Such insight has particular utility in development planning, where the allocation of resources and environmental reform can be aligned with socio-economic inclusion on a strategic level.

The consequences of what we've found are significant. By showing that environmental factors can be used to predict financial inclusion at high levels of accuracy, and particularly when a Random Forest framework is used as the approach, we're helping create a better-integrated view of sustainability. Financial inclusion is not simply a social or economic issue—it's strongly rooted in the environmental situation of a country. In developing countries, where access to banks and the spread of mobile money increases at a very high rate, the environmental conditions of a region—whether it's a high heatwave exposure, the take-up of renewable energy sources, or degree of ecologic preservation—will either facilitate or obstruct those changes.

Also, this research highlights the necessity of connecting ESG areas and not addressing them as disparate variables. The E component, so often dealt with as discrete from societal and governance institutions, here proves highly correlated with economic access and financial infrastructure. In the wider debate on sustainable growth, this finding lends credibility to the policy argument that efforts to make the environment more sustainable—such as conservation of resources or investments in green energy—ought not be merely regarded as climate actions but also as interventions having good spillovers in financial access and inclusion.

Ultimately, our thorough examination validates that Random Forest Regression is the most reliable and efficient approach for modeling the relationship between environmental variables and financial inclusivity in developing nations. Its theoretical flexibility, statistical performance, and empirical interpretability make it exceptionally suitable for ESG modeling—especially in rich data but structurally complex areas like the E component. As the worldwide community encourages enhanced integrated and facts-driven ESG disclosure, Random Forest presents a grounded and actionable solution path toward modeling and predicting environmental-socioeconomic linkages critical to ESG decision-making.

The application of Random Forest Regression as a means of explaining the link between environmental and financial inclusion, as reflected through the Account Age measure, offers insightful information on the interface between sustainability and financial access in developing economies. Account Age as a proportion of individuals who report having a financial institution or receiving money services in the last year can be taken as a good proxy for financial access. In this regard, a set of explanatory variables representing the most critical elements of the Environmental (E) pillar of the ESG approach were subjected to analysis. Such variables included Agricultural Land (AGRL), Agriculture, Forestry, and Fishing Value Added (AGVA), Food Production Index (FOOD), Heat Index 35 (HIDX), Renewable Energy Consumption (RENE), and Terrestrial and Marine Protected Areas (PROT). By explaining the impact of those environmental variables on financial access, the model provides a precise insight into development dynamics bridging eco-resources and infrastructures and socio-economic accessibility (Figure 3).





Among the environmental predictors, the most important contributor to the predictive performance of the model came from AGVA. This was seen on all three importance metrics: mean reduction in accuracy, total boost in node purity, and mean dropout loss. The dominance of AGVA in the model indicates that areas of stronger economic production in agriculture and agribusiness sectors experience increased levels of financial inclusion. The relationship may be explained by the fact that agricultural value chains tend to demand financial services like credit, savings, and insurance as well as production inputs and intermediation services, where smallholder farmers and local enterprises are involved and engaged in production for part or full cash payments. The relationship between economic production in agriculture and access to financial systems looks fairly strong and may indicate that investment in productive agriculture can simultaneously advance financial inclusion on a larger scale.

The second most powerful variable, renewable energy consumption, emphasizes the facilitating role of energy infrastructure in supporting financial participation. Renewable energy, and especially decentralized sources of it such as wind and solar, has proven a key facilitator of digital finance and mobile money in areas without traditional access to electricity. Studies have shown that financial development and inclusion can significantly enhance renewable energy adoption, especially in developing countries (Shahbaz et al., 2021). The framework's identification of RENE as a key driver confirms the proposition that environmental sustainability and financial inclusion do not act at cross-purposes and work best together. By making energy available, renewables indirectly facilitate the utilization of financial technologies and promote inclusive development (Feng et al., 2022). Agricultural land, another critical predictor, points toward the role of land availability and utilization in determining economic conditions affecting financial behavior. Areas with larger agricultural land areas may enjoy higher involvement in agriculture and related industries, prompting interaction with financial institutions. Although less impactful than AGVA or RENE, AGRL's significance indicates physical land resources continue to be central to development pathways, especially in agrarian economies. The performance of

protected areas also deserves mention. PROT performed similarly to AGRL and RENE on the accuracy and dropout loss of the models. This might mean that environmental conservation policies may be linked to higher levels of financial inclusion, possibly due to stronger institutional arrangements or locally based approaches to natural resource management involving participation in formal systems. Or protected areas at higher levels may encourage sustainable investment or development financing containing financial services as a part of wider socio-ecological resilience.

Food production index and heat index, although useful, were less impactful on the strength of the model's predictions. FOOD, a measure of agricultural production, indicated modest significance possibly because subsistence and food security may be supported by food production but isn't necessarily directly converted to financial participation unless linked to wider economic activity as captured in a more direct manner by AGVA. The lower significance of HIDX as a measurement of climate stress indicates that heat-associated impacts of the climate may be crucial for long-run sustainability but less immediately impactful on financial participation. The implication here isn't that climate variables don't matter, but their impacts might be intermediated through other channels not captured in this model, e.g., migration, productivity losses, or effects on health.

As a whole, the Random Forest model has yielded a rich breakdown of the drivers of financial inclusion from an environmental perspective. Its capacity for dealing with non-linear relations and interaction between variables makes it well suited for this form of multi-dimensional data where relations between land use, climatic exposure, resource management, and economic activity involve complex interlinkages. The variable importance scores of the model give a clear indication of where environmental policy and financial inclusion policy may intersect. By pointing to AGVA, RENE, and PROT as the central levers, the analysis emphasizes sustainable environmental practices and efficient land use not only as critical for sustaining ecology but also as directly linked with access and empowerment related to finance. This aligns with emerging cross-country evidence that links renewable energy use and financial inclusion with inclusive growth outcomes (Cui et al., 2022). This points toward the necessity for integrated policy responses not viewing environmental sustainability and socio-economic inclusion as discrete agendas but as reinforcing aspects of development policy.

### **4.3 Forecasting Finance through the Environment: A Case-Based Additive Explanation Approach**

The "Base" value of 41.987 in every case indicates what the model's expected prediction would be without any particular feature inputs included. The difference between the predicted value and base would be the net effect of a positive or negative relationship each variable has on the final prediction. For Case 1, the final forecast value (31.054) drops significantly below the base largely due to huge negative contributions from AGVA (-7.839) and PROT (-3.647) dominating over positive contributions of RENE (+2.495) and HIDX (+0.401). It shows low agriculture value and little protected lands significantly depress predicted financial inclusion even when renewably powered energy consumption has considerable levels (Table 7).

Table 7. Additive Feature Contributions to Financial Inclusion Predictions: Case-Level Analysis from Random Forest Model

Case	Predicted	Base	AGRL	AGVA	FOOD	HIDX	RENE	PROT
1	31.054	41.987	-0.995	-7.839	-1.348	0.401	2.495	-3.647
2	30.988	41.987	-0.268	-8.530	-1.061	0.312	2.732	-4.185

3	32.340	41.987	-0.890	-8.462	0.120	1.023	2.718	-4.156
4	38.849	41.987	-2.461	-8.273	-1.825	0.463	1.640	7.318
5	54.252	41.987	-1.446	8.798	-0.332	1.630	5.772	-2.157

*Note.* Displayed values represent feature contributions to the predicted value without features (column 'Base') for the test set.

Case 2 follows the same trend with a predicted value of 30.988—once more much lower than the base level. AGVA also exerts a very strong negative impact (-8.530), and PROT has a negative effect (-4.185), but RENE has a positive effect (+2.732). These ongoing patterns reinforce the argument concerning levels of agricultural productivity and levels of protection determining the model and pulling projections down when levels are low even when compensatory action from the consumption of renewable energy sources exists.

The Case 3 forecast for the model stands at 32.340, weighed down by AGVA (-8.462) as well as PROT (-4.156), but boosted by RENE (+2.718) and HIDX (+1.023). Significantly in this instance, however, the FOOD variable enters a positive value (+0.120) for the first time and hence minimally alleviates the overall decline, suggesting some relief on the part of food production capacity in contributing toward upward momentum in projections of membership.

Case 4 presents a remarkably uncommon trend where the forecasted value (38.849) converges toward the base value. Although negative contributions from AGVA (-8.273) and AGRL (-2.461) persist, PROT strongly improves the forecast (+7.318), offsetting deficits. Case 4 illuminates the worth of protected areas as a potentially game-changer in the improvement of financial inclusion projections through signaling improved institutional arrangements or development assistance (Figure 4).



Figure 4. Model Fit and Convergence of Random Forest Predictions on Financial Inclusion from Environmental Variables

Lastly, Case 5 has the strongest forecasting value (54.252), far above the base. The boost comes from high positive impacts from AGVA (+8.798), RENE (+5.772), and HIDX (+1.630), which dominate the small negative impacts from AGRL (-1.446) and PROT (-2.157). This case demonstrates the model's

ability to capture synergy between environmental and economic productivity, energy access, and good climatic conditions in improving financial inclusion outcomes.

On the whole, both instances identify AGVA and PROT as salient variables that tend to impose strong directional effects—negative when below average and positive when above average. The positive contributions from RENE consistently affirm the status of a key facilitator. The additive accounts well capture the ways in which sets of environmental variables determine the model's outcome and provide clear insight into the mechanism by which environmental conditions impinge on financial inclusion in developing settings.

### 5. Social Foundations of Financial Inclusion: Empirical Evidence from Developing Economies

This research evaluates the effect of critical social variables on financial participation in developing countries on the "S" part of ESG. Literacy levels, educational attainment levels, internet penetration levels, labor market participation levels, life expectancy at birth levels, availability of sanitation facilities, and employment sex parity make up the proxy variables employed for financial participation in this research. Using both the G2SLS and TSLS models over a 12-year time panel, the study employs environmental instruments for addressing the challenge of endogeneity. The findings determine educational levels, internet coverage, good health levels, and employment sex equality as prominent drivers of financial participation and chart their usability for sustainable policy and development programs (Table 8).

Y		Account Age									
Instruments	NRSD NFSD A	AGRL AGVA AF	WW CO2P CLD	D ENIN FOOD FI	RST HIDX HDDY	LSTP WSTR					
		CH4P N2	2OP PM25 RENE	RELE SPEI PRC	OT TCLH						
Models	G2SLS random	G2SLS random effects, using 1236 observations Fixed-effects TSLS, using 1236 obs									
	coefficient	std. error	Z	coefficient	std. error	Z					
Constant	-199.526***	74.1612	-2.690	187.744***	61.2268	-3.066					
FERT	12.1565***	2.90271	4.188	12.2159***	2.43462	5.018					
EDUE	3.16223***	1.14172	2.770	2.48709**	0.979344	2.540					
NETU	0.575839***	0.0714890	8.055	0.534013***	0.0689459	7.745					
LABR	-3.39819***	1.19358	-2.847	-3.55157***	1.05204	-3.376					
LIFE	2.30672***	0.622509	3.706	2.22245***	0.478428	4.645					
SANS	1.13794***	0.365177	3.116	1.28825***	0.435367	2.959					
FLMR	2.05370***	0.631862	3.250	2.18771***	0.755028	2.898					
Statistics	SSR = 1644.79			SSR = 159850							
	sigma-hat = 1.15	5733 (df = 1228)		sigma-hat = 11	.9148 (df = 1126)						
	R-squared = corr	$(y, yhat)^2 = 0.2^2$	75947	R-squared = co	$rr(y, yhat)^2 = 0.3$	305085					
	Included units =	103		Included units	= 103						
	Time-series leng	th: $\min = 12$ , $\max$	x = 12	Time-series length: $min = 12$ , $max = 12$							
	Wald chi-square	$(7) = 288.308 \ [0.0]$	0000]	Wald chi-squar	re(7) = 374.518 [0]	.0000]					
	sigma-hat(within	n) = 11.914822									
	sigma-hat(betwe	en) = 19.830205									

Table 8. Social Determinants of Financial Inclusion: G2SLS and TSLS Regression Results from Developing Economies

This empirical study sought to investigate the nexus between financial inclusion and the "S" (Social) pillar of the ESG framework and developing countries specifically. Financial inclusion via the Account Age measure—calculated as the share of people who have a financial account or used a mobile money service in the preceding year—is a principal determinant for the equal access to economic engagement and formal financial system. In the ESG framework, the social pillar covers a subset of human development measures such as education level, health, gender equality, access to digital technologies, and labor participation. Such measures have a direct impact on people being able to use financial services and thus research the impact they have upon financial inclusion is both time and policy-oriented—especially in those countries where a significant proportion of the people are unbanked or underbanked (Shahid & Asghar, 2024).

We examine this relationship through the use of a panel-data instrumental variable approach with both G2SLS random effects and TSLS fixed effects estimations. There are 1236 observations from 103 developing countries over 12 years. There are seven significant explanatory parameters all capturing the Social pillar of ESG: (1) Fertility rate (FERT), (2) Education expenditure through the government (EDUE), (3) Internet usage (NETU), (4) Labour force participation ratio (LABR), (5) Life expectancy at birth (LIFE), (6) Access to safely managed services and facilities to sanitation (SANS), and (7) Female-to-male labour force ratio (FLMR). These parameters capture the social infrastructure facilitating or limiting financial inclusion (Jain et al., 2024).

The empirical findings based on both the G2SLS and the TSLS models are consistent and reinforce the evidence and validity of the findings to a great extent. In particular, the female-to-male labor force ratio (FLMR) is a positive determinant of financial inclusion and vindicates the role assigned to gender equality in sustainable development (Wani & Khanday, 2024). Further, the use of the internet (NETU) is also the highest predictor of financial inclusion and verifies the role of the digital divide in economic participation (Shahid & Asghar, 2024).

Government expenditure on education (EDUE) also shows strong correlation with evidence referenced in research to highlight financial literacy and education as key enablers of access to services (Zubair et al., 2023).

Finally, the approach aligns with the broader context of the integration of ESG factors, especially taking the socio-political complexities of policymaking towards sustainability in the Global South into account. Poor governance, misinformation, and ethics gaps can destabilize the financial system if left unresolved (Alibašić, 2024).

## 5.1 Clustering Sustainability: Evaluating Algorithmic Performance on ESG-Environmental and Social Indicators

In the context of clustering algorithm evaluation on the E (Environmental) dimension of the ESG (Environmental, Social, Governance) framework, a comparison of five approaches - Random Forest Clustering, Neighborhood-Based Clustering, Model-Based Clustering, Hierarchical Clustering, and Fuzzy C-Means Clustering - was performed. The aim was to determine the superior-performing algorithm based on a set of normalization evaluation criteria: Maximum Diameter, Minimum Separation, Pearson's  $\gamma$ , Dunn Index, Entropy, Calinski-Harabasz Index, R<sup>2</sup>, AIC, BIC, and Silhouette. These evaluation criteria measure various aspects of clustering quality - compactness, separation, cohesion, and statistical fit of the model - and the evaluation was conducted based on the average of the normalized

scores. The best-performing approach proved to be Hierarchical Clustering on average across the evaluation criteria used (Table 9).

	Maximum	Minimum	Pearson'	Dunn	Entrop	Calinski-	R <sup>2</sup>	AI	BI	Silhouett
	diameter	separation	s Î <sup>3</sup>	index	у	Harabasz		C	C	e
						index				
Random	0.0	0.25	0.16	0.12	0.0	0.33	0.67	1.0	1.0	0.43
Forest										
Neighborhood	0.0	0.1	0.47	0.12	0.0	1.0	0.44	0.0	0.0	1.0
Based										
Model Based	1.0	0.08	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
TT' 1' 1	0.0	1.0	1.0	1.0	0.0	0.21	0.44	0.0	0.0	1.0
Hierarchical	0.0	1.0	1.0	1.0	0.0	0.21	0.44	0.0	0.0	1.0
Fuzzy c-	0.0	0.0	0.39	0.0	0.0	0.27	1.0	0.0	0.0	0.57
Means										

Table 9. Evaluation of Clustering Algorithms on ESG-Environmental Data: Hierarchical Clustering as the Optimal Approach

Hierarchical Clustering performed superbly on a variety of measurements. Specifically, it attained the highest possible score (1.0) on Minimum Separation, Pearson's  $\gamma$ , Dunn Index, and Silhouette. This indicates very good clusters from a separation viewpoint, high internal correlation between clusters, an optimum intra- and inter-cluster dispersion ratio, and very good internal consistency. These are all key features when trying to make meaningful and interpretable distinctions between clusters, particularly in the case of environmental data where explicit groupings tend to reflect unique ecologies or policy regimes (Rusu et al., 2023). The method also reported a relatively good Calinski-Harabasz Index (0.21), supporting the indication of structurally stable and statistically consistent clusters.

Hierarchical Clustering did not result in the optimum AIC or BIC values—both of which were established here at 0.0—but did provide a stable measure of  $R^2$  at 0.44. This indicates a reasonably good varianceexplaining ability. Such a compromise can be worthwhile when interpretability and cluster stability become more important than statistical optimum in and of itself. In sustainability and ESG applications, the ability to interpret and explain results in cluster analysis clearly will often be of more value than small gains in fit (Saltık, 2024).

The second key strength of Hierarchical Clustering lies in the manner it presents the structure in visual form. The approach has the ability to create a dendrogram, a graph in the form of a tree representing the union progress and relationship between points or clusters visually. In ESG analysis—particularly environmental aspects—hierarchical visualization would be very helpful here (Ishizaka et al., 2021). It not only gives the final cluster classification but also the constructive relationship between regions, countries, or observations. Its flexibility and clarity add a level of depth and may be used as a guide when making a decision on an optimum number of clusters based on policy or research at hand.

On the other hand, Random Forest Clustering, whose R<sup>2</sup> (0.67) and AIC and BIC values were the best, fell behind on structural metrics such as Silhouette (0.43), Pearson's  $\gamma$  (0.16), and Dunn Index (0.12). That means although it is statistically efficient, it may form less differentiated or compact groups, making interpretation and qualitative understanding less easy. Its predictive strength may be in being able to over-fit the patterns, and hence reduce the reliability of capturing true group structure.

The Neighborhood-Based Clustering emerged well on Calinski-Harabasz Index (1.0) and on the Silhouette (1.0) as the best of the two since it indicates well-defined clusters in some respects. It did poorly on Minimum Separation (0.10) and R<sup>2</sup> (0.44) and functioned somewhat better on general

interpretability-related functionalities than Hierarchical Clustering. It forms compact clusters but a lower level of separation and explainability diminishes its usability in analysis where both statistical and policy relevance would be required.

Fuzzy C-Means Clustering attained the highest possible  $R^2$  value (1.0), thus theoretically highly capable of explaining variance. It failed on structure-related scores: Minimum Separation (0.0), Dunn Index (0.0), and Silhouette (0.57). While the soft clustering approach might be useful in those situations where observations truly belong to several categories, it can complicate the analysis when certain group labels must be established, particularly in policy settings where clear classification takes precedence.

The poorest performer on every measure was Model-Based Clustering, producing lowest R<sup>2</sup>, Pearson's  $\gamma$ , Dunn Index, Silhouette, and Calinski-Harabasz Index values. While model-based techniques give statistical rigor and probabilistic assumptions as their foundations, the results here portend failure on the part of the algorithm in extracting meaningful structure from the data. It has little interpretive insight and predictive utility and so is a poor candidate for this ESG-focused analysis.

Briefly, the preference for Hierarchical Clustering is warranted by a balanced and stable performance on those ESG context-relevant metrics of chief concern: clarity, consistency, and structure. Its better performance in separation, internal cohesion, and visual interpretability harmonizes well with properties of environmental clustering where complexity of data and policy decision-making call for models not just statistically sound but also interpretable from a stakeholders' viewpoint (Morelli et al., 2025). Hierarchical Clustering possesses a perfect balance between analytical accuracy and communicative capacity, which is essential in applications demanding clear and actionable information. Accordingly, it stands out as the best option from the algorithms evaluated as it offers the best possible tradeoff between statistical performance and usability in environmental ESG analysis (Figure 5).

Figure 5. Cluster Validation and Visualization of ESG-Environmental Segmentation: Hierarchical Clustering Solution (k = 10)



The Hierarchical Clustering output presents a meaningful division of the performance and shape of the model over the ten clusters found through the analysis. The sizes and consistency of the clusters differ and their corresponding contributions toward explaining the variability in the dataset as a whole offer some interpretation over the extent to which the partitioning of data by the algorithm occurred in relation to group similarity and between-group spread.

The cluster group sizes vary from very small ones like Cluster 5 (11 units), Cluster 9 (5 units), and Cluster 10 (12 units) to very big ones like Cluster 2 (486 units) and Cluster 4 (279 units). It would not be abnormal if hierarchical models were of differing size because they cluster based on the merit of distance measurement and not size uniformity. Interestingly enough, the single cluster of Cluster 2 accounts for over 42% of the total of this within-cluster heterogeneity because it contains a big entry in the column on "explained proportion of within-cluster heterogeneity" (0.422). Similarly, Cluster 4 contains 25.6% of the within-cluster heterogeneity. These two clusters together explain most internal variability and may contain very diverse or loosely related observations and possibly may call for a finer refinement or sub-clustering (Table 10).

Table 10. Structural Characteristics of Hierarchical Clusters: Size, Variance, and Cohesion Across ESG Dimensions

Cluster	1	2	3	4	5	6	7	8	9	10
Size	136	486	83	279	11	64	136	24	5	12
Explained proportion within-cluster heterogeneity	0.091	0.422	0.062	0.256	0.002	0.048	0.102	0.011	0.002	0.004
Within sum of squares	408.221	1.888	276.617	1.143	8.014	215.893	455.912	47.400	9.955	17.539
Silhouette score	0.229	0.066	0.185	0.067	0.656	0.183	0.228	0.516	0.492	0.493

Notes. The Between Sum of Squares of the 10-cluster model is 5408.53. The Total Sum of Squares of the 10-cluster model is 9880

The WSS for size differences and heterogeneity reflect those differences as well. The largest cluster, Cluster 2, has by far the greatest WSS (1,888,726), reflecting the high internal variance of the cluster. The smallest clusters (e.g., Cluster 5, Cluster 9, Cluster 10) also hold very small WSS values, reflecting their small and compact aggregations of very comparable observations. The small clusters may be capturing niche or extreme profiles and could be very informative in the identification of outliers or high-interest groups, even if they capture a negligible amount of the dataset's heterogeneity.

Silhouette scores give us another look at the quality of the clustering. By and large, the silhouette scores report a mixed story. Clusters 5, 8, 9, and 10 all contain very good scores (between 0.492 and 0.656), which means well-separated and internally homogeneous clusters. The single best-scoring cluster (0.656) is cluster 5, and this cluster also has the best internal cohesion and separation from other clusters in our measure. The two largest and most heterogeneous groups, Clusters 2 and 4, both contain the worst silhouette scores (0.066 and 0.067), which means they may contain overlap points from other clusters and urge us to be cautious in our interpretation. The fact they both contain high internal variance and low silhouette scores means they may be aggregating heterogeneous or margin points and would be better split or refined upon later analysis.

At a global level, the ratio between between sum of squares (BSS) and total sum of squares (TSS) stands at 54.7. This makes about 54.7% of total variance explained through the clustering model—a very good percentage given the complexity and heterogeneity typically present in socio-environmental datasets. This reflects the goodness of the model in partitioning the data in a meaningful way although adjustment will be necessary in the largest clusters.

Overall, the hierarchical clustering solution has interpretable and well-balanced partitioning of the data set. It well discovers small and well-separated clusters (e.g., Clusters 5, 8, 9, 10) and large general groupings (Clusters 2 and 4). The high silhouette scores of the smaller clusters attest to the appropriateness of the method in detecting niche patterns, and the explained variance (54.7%) validates good general performance. The findings justify the applicability of hierarchical clustering for ESG-environmental analyses where interpretability and discrimination between patterns are most critical. Further analytical procedures may entail re-clustering big clusters or applying other models so as to test and complement those results.

Cluster means. The clusters reveal evident patterns of correspondence between financial access as measured by Account Age and a set of the concomitant social variables representing the S-Social column of the ESG framework in developing countries. Notably, Cluster 8 features the greatest standardized value of Account Age (2.046) and correlates highly with high life expectancies (LIFE = 1.339), social solidarity (SANS = 1.869), and female labor market participation (FLMR = 2.296). The implication here would be a clear positive correlation between financial access and social development-specifically institutional solidarity and gender participation. Conversely, Cluster 1 presents the lowest value of Account Age (-0.955) and encompasses lower education enrollment (EDUE = 0.438), labor market participation (LABR = -1.715), and female labor market participation (FLMR = 0.123), signifying a correspondence between poor societal institutions and lower financial access. Clusters 6 and 2 also reflect moderately positive Account Age values and couple them with sets of moderately positive labor dynamics and social solidarity, confirming multi-dimensional support as essential. At the opposite extreme, clusters 5, 7, 9, and 10 report negative values of Account Age as they record their respective levels of education (EDUE) and fertility (FERT), signifying the latter may not be every assurance of financial access without supporting socio-economic institutions. The complete set of patterns thus validates close interdependence between financial access and societal aspects most importantly gender participation, quality of life and welfare provision and calls once again for coordinated societal policies promoting inclusive financial ecosystems in developing countries (Table 11).

	Account	FERT	EDUE	NETU	LABR	LIFE	SANS	FLMR
	Age							
Cluster 1	-0.955	-0.389	0.438	-1.351	-1.715	-0.287	-0.501	0.123
Cluster 2	0.434	0.080	-0.706	0.119	0.378	0.643	0.602	0.154
Cluster 3	0.192	0.245	-0.409	-1.861	-1.379	0.812	0.797	1.214
Cluster 4	-0.259	-0.162	0.631	0.909	0.433	-0.899	-0.700	-0.037
Cluster 5	-0.581	-1.624	-0.978	1.083	0.523	0.438	1.460	1.021
Cluster 6	0.860	-0.797	-0.716	-1.189	-0.704	0.761	0.064	-0.166
Cluster 7	-0.838	0.431	1.559	0.685	0.666	-1.262	-1.123	-1.591

Table 11. Cluster-Level Means of Financial Inclusion and Social Indicators: Hierarchical Segmentation on ESG-Social Variables

Cluster 8	2.046	0.469	-1.097	-0.036	0.401	1.339	1.869	2.296
Cluster 9	-0.791	4.101	0.297	0.847	-0.521	-0.572	-0.624	-1.098
Cluster 10	-0.390	1.436	0.898	-0.472	-1.253	-0.101	-0.386	-1.325

#### 5.2 Education, Equity, and the Economy: A Social ESG Perspective on Financial Participation

Of the eight algorithms under review—Boosting Regression, Decision Tree, K-Nearest Neighbors, Linear Regression, Neural Network, Random Forest, Regularized Linear Regression, and Support Vector Machine—the comparison of performance metrics on a uniform scale has a definitive winner. The comparison of the performance criteria is done using six standardized criteria: MSE, MSE(scaled), RMSE, MAE/MAD, MAPE, and R<sup>2</sup>. These scores, when normalized, yield a common scale for comparison free from the variances brought about by differences in measurement units and magnitude (Table 12)

Table 12. Comparative Evaluation of Machine Learning Regressors: Scaled Error Metrics and Exp	plained
Variance	

		MSE		MAE /		
	MSE	(scaled)	RMSE	MAD	MAPE	R <sup>2</sup>
Boosting Regression	0.64	0.44	0.67	0.82	0.71	0.53
Decision Tree	1.0	0.87	1.0	0.69	0.0	0.11
K-Nearest Neighbors	0.0	0.03	0.0	0.0	0.1	0.96
Linear Regression	0.89	1.0	0.9	1.0	0.59	0.0
Neural Network	0.81	0.89	0.84	0.98	0.97	0.1
Random Forest	0.01	0.0	0.02	0.11	0.18	1.0
Regularized Linear Regression	0.9	0.98	0.91	0.98	1.0	0.02
Support Vector Machine	0.82	0.97	0.85	0.92	0.6	0.02

K-Nearest Neighbors (KNN) is the top performer. It possesses the lowest values of MSE and RMSE at 0.00 each and also the lowest value of MAE at 0.00, figures which reflect little predictive error on the part of the algorithm. It also possesses the greatest R<sup>2</sup> value at 0.96 and hence accounts for nearly all the variance of the target variable. These figures firmly secure the top ranking in the rankings for this regression task on behalf of KNN. It would be a sign that the organization of the data suits well the instance-based and non-parametric nature of KNN, which excels at settings where locally salient patterns exist and are clearly defined (Srisuradetchai & Suksrikran, 2024). All success notwithstanding, however, the weaknesses of KNN must be realized. It possesses weak scalability on higher datasets since it makes heavy reliance on lazy learning, and it's highly sensitive to scaling and noise on features. Its performance on larger or dirtier datasets may be severely hampered unless pre-processing is extremely conscientious.

Boosting offers a good alternative. It does not lead on any single measure but scores well on all of them. Its normalized scores place it firmly in the middle or above, even at a modest tradeoff between variance and bias. At an R<sup>2</sup> of 0.53, it accounts for well over half the variance, a good showing considering the generally high variance-bias trade-off in ensemble algorithms. It also has an edge over KNN when it comes to complex data distributions via iterative error correction. This makes it generally better for real-world data where data will be noisy and may well embody complex inter-relations between variables (Ibarz et al., 2022).

Their performance here however is not as good as theory and they perform poorly as well. The values here are very high in error and they are particularly high in MAPE and MAE and also indicative of high predictive imprecision. Their measure of  $R^2$  is also low and indicative of poor explanatory value as well. It indicates overfitting or data size and complexity of features aren't sufficient for a neural architecture here. Without deep tuning and enormous data sets, their complexity here isn't warranted.

Linear and regularized Linear Regression perform worst here. Both share the assumptions of linearity and independence of predictors, both obviously violated here. Both models produce the largest error rates and lowest  $R^2$  scores, and Linear Regression has a normalized  $R^2$  of 0.00. These results indicate the limits of linear models when working with those datasets having non-linear data or interactive complexities between variables (Jin, 2022).

Decision Trees and Support Vector Machines also fall in between. Decision Trees perform erratically enough on a normalized MAPE of 0.00 (a probable normalization scale anomaly rather than performance), and an R<sup>2</sup> of 0.11—which is far from adequate. Their overfitting tendencies are well established and here as well. Support Vector Machines fare a little better on consistency but do not perform well on any of the metrics and require very high computational resources without giving any superior accuracy.

From a generalizability and stability viewpoint, Boosting Regression could be better since it's an ensemble method which smoothes out the variances between models. Nevertheless, from a strictly empirical viewpoint, the best on this dataset has been posted by the KNN, both in reducing each metric of error and optimizing explained variance on several performance metrics.

Overall, the top-performing model for this test is KNN. It beats out on key metrics and possesses high accuracy and explanatory power. If scalability, generalization, or immunity from noisy inputs were paramount concerns in a given situation, then a more adaptable and equally impressive solution option exists in the form of Boosting Regression. Currently, under these conditions and on these metrics, KNN stands as the statistically superior option.

In the context of modeling financial inclusion (Y), the K-Nearest Neighbors (KNN) algorithm outputs variable importance via *mean dropout loss*, which represents the increase in prediction error (specifically RMSE) when a given predictor is permuted or excluded from the model. This metric quantifies the marginal contribution of each variable to predictive accuracy, allowing for an assessment of their relative influence within the model's structure (Table 13).

Table 13. Variable Importance in Predicting Financial Inclusion: Mean Dropout Loss from K-Nearest Neighbors Model

Variables	Mean dropout loss
NETU	18.755
SANS	15.465
FERT	15.351

FLMR	15.312
LABR	14.863
EDUE	14.447
LIFE	13.670

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

From the S (Social) corner of the ESG (Environmental, Social, and Governance) framework and from the perspective of developing economies, evaluation based on the deployment using the K-Nearest Neighbors (KNN) model to estimate financial inclusion holds key implications (Das & Nayak, 2020). Defined in this context as the share of the population who report holding a financial institution or mobile money account, financial inclusion is a core economic and social inclusion measure—cornerstones of the ESG social pillar.

The result based on mean dropout loss—calculated as the mean increase in root mean squared error (RMSE) subsequent to each predictor variable being permuted in 50 simulations—provides a measure of the relative role played by the different socioeconomic variables in the determination of financial inclusion in low- and middle-income economies.

The largest variable is NETU (internet-using individuals), having the highest dropout loss at 18.755. This presupposes the key role played by digital access in facilitating financial inclusion in developing countries. Internet usage is more likely to serve as the starting portal point to digital bank services, cellular money apps, and other fintech services. The importance of the variable indicates the direction towards the improvement in social inclusion and access to finance (Lee et al., 2023).

Next is SANS (access to safely managed sanitation services), with a dropout loss of 15.465. SANS is a structural quality-of-life and access-to-public-services measure. In developing countries, access to sanitation safety is related to urbanization, stages of development, and institutional coverage—and all are related to access to the organized networks, including the financial system. SANS' predictive power reflects the relation between economic inclusion and the infrastructure of basic public services.

The FERT (Fertility rate), measured as -15.351, provides a demographic background. Fertility is high in areas with lower education and access to medical care, conditions usually indicated through lower levels of financial inclusion. Fertility in this case acts as a proxy measure for overall conditions of human development (Balde et al., 2022).

FLMR (female-to-male labor force) ratio decreases 15.312 and also holds a significant position in this regard. Inclusion of women in the labor force is hand in hand with financial inclusion: when women work, the probability of their maintaining bank accounts or utilizing financial instruments enhances. This illustrates the significance of gender equality in the social part of ESG measurement.

LABR (Labor force participation rate), 14.863 in value, reflects the significance of overall economic activity. Access to the labor market—even insecure (informal) access to the labor market—encourages the use of financial services in the management of income, savings, and expenditure in developing economies.

On the lower rung in the list are the EDUE (Government expenditure on education) and the LIFE (Life expectancy at birth), both of which suffer losses to the tune of 14.447 and 13.670 respectively. Though they are significant gauges of people's development in their own right, their implications with regard to financial inclusion can tend to be more lagged or indirect. The expenditure on education can impact financial literacy but the tendency is towards being a long-term rather than immediate effect. The life expectancy has more of a contextual rather than direct impact upon financial behaviour as a gauge of well-being in the community.

Overall, when looking at financial inclusion as the social ESG aspect in emerging economies, the findings clearly prioritize digital access, infrastructure fundamentals, and labor market participation—especially female participation. The highest-quality variables in terms of highest loss in value are the categories suggesting structural and societal barriers to inclusion. Consistent with this logic then are initiatives to promote internet penetration, the improvement in sanitation delivery, and women's employment as high-impact interventions towards increasing financial inclusion and improving the social ESG ground in emerging economies

#### 5.2.1 Forecasting Financial Access: Interpretable AI Insights from KNN in ESG-Social Contexts

The table presents additive descriptions of the financial inclusion forecasts calculated from the K-Nearest Neighbors (KNN) model utilized in the context of the Social (S) component of the ESG framework for developing countries. The goal variable utilized here defines financial inclusion as the percentage of individuals who own a financial account or make use of (mobile) money services. The forecasted value for every case has been provided from a base forecast of 42.032 and each feature then adds or detracts from it based on the strength of the effect (Table 14).

Table 14. Additive Feature Contributions to Financial Inclusion Predictions: KNN-Based Case-Level Explanations

Case	Predicted	Base	FERT	EDUE	NETU	LABR	LIFE	SANS	FLMR
1	10.238	42.032	-2.094	-6.935	-9.933	-3.346	-0.476	-0.164	-8.846
2	12.288	42.032	-9.403	-6.884	-8.276	-4.699	-0.541	-0.062	0.120
3	18.242	42.032	-18.092	-0.563	-6.775	-2.549	1.220	1.546	1.424
4	31.890	42.032	0.465	-2.837	-0.017	-6.171	-3.849	4.762	-2.495
5	36.722	42.032	3.775	-16.512	4.358	0.337	6.618	-5.693	1.808

For Case 1, the forecasted value of financial inclusion stands at 10.238, much below the base level. The decline here is mainly due to negative contributions from internet penetration, female-to-male labor force ratio, and government spending on education. These indicate poor access to digitalization, uneven employment opportunities between genders, and possibly wasteful spending on education as drivers of financial exclusion. Case 2 forecasts a value of 12.288 but remains well below the base level. High birth rates, weak internet penetration, and low spending on education contribute adversely once more, but a small positive effect from the female labor force entry ratio is observed. Case 3 has a forecasted value of 18.242 where negative contributions from a high birth rate are compensated for in part by positive contributions from life expectancy at birth, access to sanitation facilities, and participation of females in labor forces. The case indicates how enhanced basic infrastructure and gender equality can counteract negative effects from demographic stress (Figure 6).

Figure 6. KNN Regression Model Evaluation: Forecast Accuracy and Neighbour Optimization in Financial Inclusion Modelling


The predicted value rises in Case 4 to 31.890. Better sanitation access and a modestly positive fertility rate reinforce the prediction here despite adverse contributions from labor force participation and life expectancy. This indicates even when labor participation and health outcomes are less positive, good infrastructure in sanitation can remain a driving lever for financial inclusion. Case 5 experiences the greatest predicted financial inclusion at 36.722. This is driven primarily by high life expectancy and good internet use and a positive fertility rate, but heavy negative contributions from expenditure on education and access to sanitation reveal possible inefficiencies or inequalities in the distribution of services. The extreme negative effect of education expenditure may indicate misalignment between expenditure and actual educational programs related to financial behavior. Generally speaking, the model results emphasize the stronger predictors of financial inclusion in the ESG's social dimension as being internet access, fertility rates and gender labor equity alongside wider human development indicators such as sanitation and life expectancy. Whilst education has conventionally been regarded as a driver of financial participation, their repeated negative effect in these instances means investment itself will not be enough and quality and effectiveness of education programs must be the deciding factors. This points out the multi-sided nature of financial participation as guided not just by access directly to services but by the wider socio-economic context as well.

#### 6. Governance Meets Inclusion: Empirical Insights from ESG Analysis in Developing Economies

This analysis looks at financial inclusion's relationship with the "G" (Governance) pillar of the ESG framework in developing economies. As a financial inclusion measure, Account Age guides the analysis of its correlation with the three governance proxies: control of corruption (CCOR), patent activity (PATR), and regulatory quality (REGQ). Both G2SLS and TSLS models utilizing environmental instruments guide the analysis as it investigates whether access to financial services at a wider level fosters institutional development and accountability. The research confirms both opportunities—e.g., increased transparency and innovation—and challenges—e.g., potential regulatory loopholes—emphasizing the multifaceted dynamics between governance and financial inclusion (Table 15).

Table 15. Governance Determinants of Financial Inclusion: G2SLS and TSLS Estimates for Developing Economies

Y			Accou	nt Age		
Instruments	NRSD NFSD	AGRL AGVA	AFWW CO2I	CLDD ENIN FOOD FRST HIDX HDDY		
	LST	P WSTR CH4	P N2OP PM25	RENE RELE S	SPEI PROT TC	CLH
Models	G2SLS rar	ndom effects, u	sing 1236	Fixed-eff	fects TSLS, usi	ng 1236
		observations			observations	
	coefficient	std. error	Z	coefficient	std. error	Z
Constant	70.0248***	6.37110	10.99	39.3086***	9.75481	4.030
CCOR	82.7466***	16.0177	5.166	76.1752 ***	17.8926	4.257
PATR	0.000289***	6.29571e-	4.598	0.00022***	7.21168e-	3.100
		05			05	
REGQ	-36.4103***	12.1262	-3.003	-93.98***	17.4674	-5.381
Statistics		SSR = 1625.98		SSR = 390242		
	sigma-ha	t = 1.14882 (df	f = 1232)	sigma-hat = 18.5835 (df = 1130)		
	R-squared =	corr(y, yhat)^2	2 = 0.011619	R-squared =	corr(y, yhat)^2	= 0.020903
	Included units = 103			Inc	luded units = 1	03
	Time-series length: $min = 12$ , $max = 12$			Time-series	length: $min = 1$	2, max = $12$
	Wald chi-square(3) = 49.7846 [0.0000]			Wald chi-squ	uare(3) = 42.56	38 [0.0000]
	sigma-ha	t(within) = 18	5.583517			
	sigma-hat	t(between) = 19	9.202094			

This empirical study aims to explore the nexus of financial inclusion and the "G" (Governance) factor of the ESG (Environmental, Social, Governance) model in the context of developing countries. The principal assumption in our case is the fact that greater financial inclusion—measured through the Account Age variable, reflecting the proportion of the adult population with a bank account or using a mobile money service to make payments—might actually define the key governance performance measures. Financial inclusion is more than a finance marker: it is a force behind institutional growth and a citizens' and governments' accountability lever (Zeqiraj et al., 2022). In this work, we are exploring whether more financially included societies are also more robust in their governance measures and specifically when focusing upon the control over corruption, regulatory quality, and innovation output respectively captured through variables CCOR, REGQ, and PATR.

As a counterfactual to account for concerns over endogeneity (e.g., since improved governance might also lead to greater financial inclusion), we apply an instrumental variables strategy using a G2SLS (Generalized Two-Stage Least Squares) random-effects and a TSLS (Two-Stage Least Squares) fixed-effects strategy using a panel dataset of 1,236 observations from 103 developing economies over 12 years. Identification is achieved through a rich and carefully constructed set of environmental instruments—a methodological innovation spanning the environmental and the governance sides of the ESG framework (Borgi et al., 2023).

Results from the two models both reflect statistically significant and economically significant relationships albeit somewhat qualified in the implications. Most importantly, however, Control of Corruption (CCOR) is revealed to have a very strong and highly significant positive relationship with financial inclusion in the G2SLS and the TSLS models. This result verifies theoretical expectations and the literature: the greater the proportion of the population gravitates towards access to the formal financial

system, the more difficult for the government to perpetuate corruption (Ben Khelifa, 2023). Financial inclusion fosters transparency and traceability of activities and formal cashing out of transactions, both of which restrict the scope for instances of corrupt behavior.

The other institutional capacity and domestic innovation proxy, the second governance indicator, Residents' Patent Applications (PATR), also exerts a positive and statistically significant impact on financial inclusion. Increasing financial inclusion should lead to enhanced innovation ecosystems— because greater access to funding enables more people and small firms to make research investments and engage in entrepreneurship activities (Sanderson et al., 2018).

But the results for Regulatory Quality (REGQ) are more complex. While statistically significant, the coefficient is negative and shows that financial access extension can surpass institutional ability in the regulation of such services in some developing nations (Kawor, 2023). This disconnect signals a key problem where fintech growth and penetration of mobile banking accelerate beyond the establishment of corresponding regulatory environments.

Overall, this study contributes to the research at the nexus of financial inclusion, governance, and sustainability. This study evidences robust empirical support for a positive linkage between financial inclusion and both anti-corruption and innovation performance and also identifies governance weaknesses concerning regulatory quality. The incorporation of green instruments in the estimation further more firmly supports causal inference and is consistent with the ESG logic of systemic interdependence.

## 6.1 Clustering Governance: Evaluating Algorithmic Performance on ESG-G Data

The comparison of clustering algorithms—Fuzzy C-Means, Hierarchical, Model-Based, Neighborhood Based, and Random Forest—across the normalized performance criteria gives a holistic view of their relative performance in the unsupervised classification task. The evaluation parameters used here comprise R<sup>2</sup>, AIC, BIC, silhouette score, maximum diameter, minimum separation, Pearson's  $\gamma$ , Dunn index, and the Calinski-Harabasz index. All the criteria have been min-max normalized between 0 and 1 so that comparisons can be made directly between them. The aim here is to identify the best-performing algorithm in general in the aspects of cluster quality, cohesion, separation, and interpretability (Table 16).

	Fuzzy C-	Hierarchica			
	Means	1	Model-Based	Neighborhood Based	Random Forest
R <sup>2</sup>	0.16	0.73	0.18	1.0	0.0
AIC	0.0	0.0	0.0	1.0	0.0
BIC	0.75	0.29	0.82	0.0	1.0
Silhouette	0.49	0.9	0.21	1.0	0.0
Maximum diameter	1.0	0.0	0.86	0.28	0.98
Minimum separation	0.02	1.0	0.0	0.08	0.14

Table 16. Comparative Performance of Clustering Algorithms on ESG-Social Data: Normalized Multi-Metric Evaluation

Pearson's Î <sup>3</sup>	0.0	1.0	0.02	0.45	0.37
Dunn index	0.0	1.0	0.0	0.11	0.03
Entropy	0.98	0.0	0.92	1.0	0.12
Calinski-Harabasz index	0.04	0.39	0.01	1.0	0.0

From R<sup>2</sup>, which indicates the extent of the clustering structure explaining the variance in the data, Neighborhood Based has a perfect 1.0, showing high explanatory value. Hierarchical comes next at a good 0.73 rating, and the rest lag behind and score only 0.0—meaning little or no clustering structure is present. This already pushes Neighborhood Based and Hierarchical approaches ahead since variance explanation has a lot of relevance as a measure of clustering quality when clusters are anticipated to capture clear-cut informative groupings (Blasilli et al., 2024).

The Akaike Information Criterion (AIC) is generally applied to measure model parsimony and lower values are desirable. As AIC has been inversely normalised, Neighborhood Based has a clear 1.0 value representing the optimum fit for complexity. The remainder of the models record a value of 0.0 and therefore perform apparently less well in relation to simplicity of the model versus fit. This must be taken as read, however, as AIC proves very lenient on complex models if they're not compared against other measures of separation. The Bayesian Information Criterion (BIC) offers a stronger penalty on complexity than AIC. In this case, the Model-Based method has the best performance (0.82), showing good explanatory strength without overtraining, followed by Fuzzy C-Means (0.75) (Ambarsari et al., 2023). Neighborhood Based, although doing well at AIC, has a score of 0.0 at BIC, which may indicate that it has an excessive complexity when measured through a stricter penalty framework. Random Forest has a perfect 1.0 here, in a seeming contradiction, although this may be the effect of the normalization scale rather than true interpretability, especially since it fares poorly at other scores (Sarmas et al., 2024). Silhouette score, which indicates the similarity of an object to its own cluster versus other clusters, benefits the Neighborhood Based approach again at a top score of 1.0 once more. Hierarchical also scores well at 0.9 and suggests well-separated and well-clustered groups. Fuzzy C-Means (0.49) and Model Based (0.21) lag behind and Random Forest (0.0) falls at the bottom once more. Since the silhouette score directly relates cluster cohesion and cluster separation, the findings strongly affirm the consistency of Neighborhood Based and Hierarchical approaches (Putra & Abdulloh, 2024).

With respect to Maximum Diameter, smaller scores describe more dense clusters. Hierarchical clustering has the best (0.0), with very dense cluster formations, and Fuzzy C-Means (1.0) and Random Forest (0.98) making the loosest clusters. Model-Based and Neighborhood Based take a middle ground, although compactness will not be enough on its own in order to measure clustering quality since extreme dispersion, as in Fuzzy C-Means and Random Forest, usually goes together with poorly defined clusters (Azkeskin & Aladağ, 2025).

On the other hand, Minimum Separation—a larger value corresponding to a larger inter-cluster distance—is greatest in Hierarchical clustering (1.0), reflecting superior discrimination, followed by Random Forest (0.14) and Neighborhood Based (0.08). Fuzzy C-Means (0.02) and Model-Based (0.0) indicate extreme cluster overlap. This confirms Hierarchical clustering's performance in creating clear groupings, a key requirement for explainability in applications such as customer segmentation or ESG scoring.

The Pearson's  $\gamma$  between cluster assignments and true distances is maximum for Hierarchical clustering (1.0), yet another reflection of the clarity of its structure. Neighborhood Based (0.45) and Random Forest (0.37) also perform well in terms of structure, although Fuzzy C-Means and Model-Based perform very little. It seems therefore that Hierarchical clustering has the best preservation of the natural order of the data in the partitioning. The intersection of compactness and separation occurs at the peak of the Dunn

index (1.0), validating its ability to produce dense and well-separated clusters. Neighborhood Based has a minimal value (0.11), and all other models approach zero, demonstrating poor trade-offs between intracluster cohesion and inter-cluster separation (Vikrant & Bhattacharjee, 2024).

Entropy, conventionally utilized as a measure of the uniformity of cluster distribution (lower being preferred), has here a reverse sense: high values denote increased disorderliness. Neighborhood Based (1.0), and Fuzzy C-Means (0.98) both share the greatest entropy and hence irregular cluster size or mixed distribution of classes—unwanted in most application contexts. Hierarchical has the lowest entropy (0.0), as expected from a method having a proclivity for producing purer and evenly split clusters. Finally, the Calinski-Harabasz index, which favors clusters having high between-cluster dispersion and minimum intra-cluster variance, strongly favors Neighborhood Based (1.0), followed by Hierarchical (0.39). Fuzzy C-Means (0.04) and the rest lag behind, Random Forest also scoring 0.0. This final measure reiterates the perception that Neighborhood Based performs best in forming statistically stable and well-defined clusters.

Summing over all aspects, the Neighborhood Based method always ranks at or close to the best in the most important criteria:  $R^2$ , AIC, silhouette score, and Calinski-Harabasz index. Nevertheless, a low BIC and high entropy indicate possible overfit or cluster size imbalance risk. The Hierarchical method presents a good alternative as a stable approach, ranking highly in silhouette, minimum separation, Pearson's  $\gamma$ , Dunn index, and entropy—attributes leading to well-defined and interpretable clusters with preserved structure. Model-Based and Fuzzy C-Means clustering algorithms show some single-point strengths like good BIC (Model-Based) or entropy alignment (Fuzzy), without consistently performing well on the entire set of measurement techniques. Random Forest emerges as the worst clustering approach in this situation and performs poorly in almost all categories except a good BIC score, which itself is not enough to counteract the poor performance on silhouette,  $R^2$ , and Calinski-Harabasz index.

Finally, Neighborhood Based clustering proves the best performer in general on this dataset and evaluation approach. Its superiority in multiple key measurements renders it the ideal solution for realizing significant patterns from highly internally coherent and well-separated data. Hierarchical clustering comes in as a close runner-up, producing better-balanced and interpretable clusters and being well placed for applications demanding clear hierarchical organization and small-cluster size. Depending on application necessity—whether prioritizing precision, simplicity, or interpretability—one or the other may be the best solution and Neighborhood Based preferred when performance on the model is the top priority and Hierarchical when clarity of structure matters most.

The Neighborhood clustering results present a differentiated interpretation of financial inclusion in developing nations from the Governance (G) component of the ESG framework. The estimation of financial inclusion (measured through Account Age) employs the Control of Corruption (CCOR), Patent Applications by Residents (PATR), and Regulatory Quality (REGQ) as the three key governance predictors. The clustering analysis sheds light on governance-related profiles correlated with enhanced or reduced levels of financial inclusion in various settings (Table 17).

Table 17. Governance-Based Clusters of Financial Inclusion: Neighborhood Clustering Analysis Across 10 Profiles

Cluster	Size	Explained proportion within-cluster heterogeneity	Within sum of squares	Silhouette score	Center Account Age	Center CCOR	Center PATR	Center REGQ
1	77	0.123	101.089	0.266	1.116	1.994	0.066	1.812
2	34	0.045	37.137	0.350	1.409	-0.911	-0.075	-1.834

3	260	0.144	118.379	0.311	-0.925	-0.318	-0.082	-0.287
4	132	0.075	61.903	0.308	-0.073	0.034	-0.162	1.033
5	9	0.065	53.295	0.694	1.717	0.670	11.038	0.328
6	17	0.022	18.218	0.400	-0.326	3.114	-0.116	0.088
7	212	0.136	111.382	0.248	0.190	-0.367	-0.109	-0.394
8	178	0.174	142.581	0.380	1.457	0.573	-0.085	0.736
9	139	0.087	71.018	0.269	-0.063	0.805	-0.078	0.334
10	178	0.129	105.827	0.376	-1.036	-1.344	-0.046	-1.333

Ten clusters were detected, each corresponding to a unique governance-financial inclusion profile. The cluster size ranges surprisingly from as few as 9 observations (Cluster 5) to as many as 260 (Cluster 3), reflecting varied distributions of governance features in the data. Explained proportion of within-cluster heterogeneity and within sum of squares (WSS) measure compactness; lower WSS and higher explained proportions tend to reflect well-coherent clusters. As an example, small but statistically significant (34) Cluster 2 has a relatively modest WSS (37.137) and high silhouette score (0.350), reflecting well-defined internal consistency. The largest cluster in the sense of explained within-cluster heterogeneity (0.174)has a good silhouette score (0.380), reflecting it as both cohesive and differentiated from surrounding clusters (Cluster 8). The cluster silhouette scores between 0.248 (Cluster 7) and 0.694 (Cluster 5) also measure cluster quality. The highest silhouette score accompanies Cluster 5, which denotes unique distinguishability from other clusters and indicates high uniqueness in governance-financial inclusion tendencies. In contrast, although having a sizeable share (212), Cluster 7 has a rather low silhouette score of 0.248 and indicates overlapping with adjacent clusters or intra-cluster heterogeneity. These insights are consistent with recent findings highlighting the power of clustering in revealing latent structure in financial and governance data (Bester & Rosman, 2024); (Bhaskaran, 2023). Examining the cluster centers of Account Age, we can start deciphering the relationship between governance characteristics and financial inclusion. Clusters at positive Account Age centers-e.g., Clusters 1 (1.116), 2 (1.409), 5 (1.717), 8 (1.457)—reflect higher levels of financial inclusion. These are especially interesting when combined with good governance scores. Cluster 1 has high financial inclusion coupled with good governance indexes (CCOR = 1.994, REGQ = 1.812), reflecting a situation where low corruption and good regulatory quality facilitate access to financial institutions. Cluster 5, although small in size, has the highest Account Age center (1.717) and very high patent applications (PATR = 11.038), possibly a sign of a subgroup more innovative and economically active and thus possessing good institutional features (CCOR = 0.670, REGQ = 0.328). On the other hand, negative Account Age centers—most notably Clusters 3 (-0.925), 4 (-0.073), 10 (-1.036)—indicate lower financial inclusion. Cluster 10 stands out as most noteworthy: not only has it the lowest center for Account Age but also highly negative governance scores (CCOR = -1.344, REGQ = -1.333), which are consistent with a framework of poor institutions, low trust levels, and poor regulation and thus probable obstacles to financial inclusion. Likewise, Cluster 2 indicates strong financial inclusion coupled with negative governance scores (CCOR = -0.911, REGQ = -1.834), a counterintuitive finding which may indicate local policies or local financial arrangements independent of governance quality (Майорова et al., 2023). Another interesting example is Cluster 6, with the greatest CCOR score (3.114) and a mildly negative Account Age center (-0.326). Although possessing superior corruption control, financial inclusion here is not as robust—presumably because other missing elements of the necessary structures, e.g., education or access to the digital platform, are not reflected here given the governance-centric approach of the model. The activity of PATR in clusters tends to be subdued generally, as a majority of clusters reveal near-zero or negative

standardized values, and only Cluster 5 appears as a clear outlier here. This indicates that patent activity might not be a robust direct driver of financial inclusion in most instances, although the dominance of Cluster 5 suggests a niche where innovation and financial access may be correlated (Figure 7).



Number of Clusters

Figure 7. Optimal Cluster Selection and Mapping of Governance-Financial Inclusion Profiles in Developing Countries



The Neighborhood Based clustering gives a useful overview of the interplay between governance variables and financial inclusion in developing economies. The aim of the analysis is estimation of financial inclusion—defined as Account Age, or the availability of access to institutions or mobile money accounts—according to three governance variables: Control of Corruption (CCOR), Patent Applications by Residents (PATR), and Regulatory Quality (REGQ). The clustering creates ten profiles in the dataset differentiated by their average values (mean) in the four variables (Table 18).

• 10

	Account Age	CCOR	PATR	REGQ
Cluster 1	1.116	1.994	0.066	1.812
Cluster 2	1.409	-0.911	-0.075	-1.834
Cluster 3	-0.925	-0.318	-0.082	-0.287
Cluster 4	-0.073	0.034	-0.162	1.033
Cluster 5	1.717	0.670	11.038	0.328
Cluster 6	-0.326	3.114	-0.116	0.088
Cluster 7	0.190	-0.367	-0.109	-0.394
Cluster 8	1.457	0.573	-0.085	0.736
Cluster 9	-0.063	0.805	-0.078	0.334
Cluster 10	-1.036	-1.344	-0.046	-1.333

Table 18. Governance-Driven Clustering of Financial Inclusion: A Ten-Group Profile Analysis

At the beginning, the mean Account Age also varies extensively by cluster from a high of 1.717 for Cluster 5 through a low of -1.036 for Cluster 10. The standardized values indicate differences in financial access or utilization. High cluster values for Account Age might be understood as reflecting populations with increased access or participation in formal financial institutions and lower values reflecting lower financial access or utilization (Ahmad et al., 2022).

Cluster 5 is a telling example. It possesses not only the largest Account Age (1.717), but also a very high patent application value (PATR = 11.038). Both of them suggest an ecosystem dominated by innovation and financial engagement. The governance indices on this profile, namely, the CCOR (0.670) and the REGQ (0.328), are fairly positive, i.e., the quality of governance is good but not excellent. This profile would define a niche of nascent innovation economies where good quality governance coexists with high creative productivity and good-quality financial facilities (Kaplan et al., 2024).

The lowest Account Age at -1.036 also comes from cluster 10 and the worst governance scores: -1.344 for CCOR and -1.333 for REGQ. These reflect a pervasive corruption and weak regulatory capacity context. The negative PATR at -0.046 here also suggests minimal innovation activity. This cluster would likely be for failing or fragile governance regimes where both institutional trust and innovation would be low and therefore where financial inclusion would be confronted strongly by obstacles (Zeqiraj et al., 2022).

Cluster 1 also features high financial inclusivity (1.116 Account Age) underpinned by high levels of CCOR (1.994) and REGQ (1.812), and a low level of PATR (0.066). This cluster would then capture a good governance and stable context featuring acceptable levels of innovation and good regulatory quality. This configuration aligns with development models in which institutions' strength has a direct impact on financial integration (Acheampong & Said, 2024).

Cluster 2, on the other hand, has a contradictory narrative as well. It has negative governance scores— CCOR at -0.911 and REGQ at -1.834—but the second highest Account Age at 1.409. It may be the case that financial inclusion can progress regardless of—or even despite—bad governance in some contexts. It may be because of the impacts of non-governmental financial systems, e.g., microfinance or mobile money platforms, which may be able to persist even in low levels of institutional quality. Or this cluster may be demonstrating states in transition where legacy-infrastructure continues to provide financial access as governance drives through chaos (Jungo et al., 2022).

Cluster 8 presents another performing scenario when it comes to financial inclusion (Account Age = 1.457), coupled with good governance scores: CCOR (0.573) and REGQ (0.736). The low and negative

PATR score of -0.085 merely suggests a scenario where a stable and well-regulated system provides financial services even without high levels of innovation. This validates the hypothesis that regulatory quality plays a central role in achieving financial inclusion through removal of entry and operation obstacles, consumer protection, and fostering competitive financial sectors.

Clusters 3 and 4 offer middle-ground examples. Cluster 3 offers low Account Age (-0.925), poor governance (CCOR = -0.318, REGQ = -0.287), and minimal patent activity. The cluster likely contains contexts of stagnation where both governance and innovation are lacking and thus diminish financial access. Cluster 4 offers a nearly neutral Account Age (-0.073) and a moderately positive REGQ (1.033), which suggests formal strength in regulation may not be enough to stimulate financial inclusion in the absence of trust in governance or innovation. This adds strength to the contention that regulatory systems must be positioned effectively and supplemented by other system strengths if they are to be impactful on inclusion (Figure 8).



Figure 8. Account Age and Institutional Integrity: Cluster Patterns in ESG-Governance Modelling

Cluster 6 stands out through a very high level of CCOR (3.114), the highest in the sample, and a negative Account Age (-0.326). Although having a highly controlled corruption, low financial inclusion indicates that controlling corruption by itself may not be enough to promote access to the financial system (Zeqiraj et al., 2022). Since the moderate REGQ (0.088) and low PATR (-0.116) both show a lack of regulatory movement and innovative economic activity, the cluster may be indicative of post-reform settings in which corruption has already been tamed but other institutional or economic development has not caught up yet.

Cluster 7 has average scores on all indices (Account Age = 0.190, CCOR = -0.367 and REGQ = -0.394) and depicts a middle-to-low level of governance and limited financial intermediation. The countries in

this cluster must be at a nascent phase of their institutional reform or be unevenly developed in sectors (Ben Khelifa, 2023).

Cluster 9, which has a near-zero Account Age of -0.063, weakly positive governance (CCOR = 0.805 and REGQ = 0.334), and little innovation activity, is another interesting case in point. It has fairly good governance scores but not necessarily corresponding financial inclusion strength. This might mean governance effects take time to show themselves in concrete socio-economic means such as financial access or possibly other forces—such as financial literacy, infrastructure, or cultural attitudes—intervening as inhibitors (Okereke et al., 2023).

Throughout the clusters, certain patterns emerge. Strong governance scores tend to be associated generally with higher financial inclusion, as in Clusters 1 and 8. This would be in line with the ESG approach's assumption that good governance is the bedrock upon which inclusive economic systems operate. Secondly, outliers such as Clusters 2 and 6 contradict this story and indicate that financial inclusion may take place even in low-governance environments or be non-existent even in high governance ones. These aberrations illustrate the governance-inclusion relationship's complexity and the necessity of mediating variables such as technology, education, and the financial systems present in the informal sectors.

Thirdly, innovation activity (PATR) seems a less significant determinant. Except in cluster 5, which has very high innovation and corresponds also to the most financial inclusion, in all other clusters, PATR stays close to or even below zero. This means innovation can be a catalyst for financial systems but is not a panacea in governance in the sense of inclusivity.

Finally, the Neighborhood Based approach effectively pinpoints the complexity of developing country governance-financial inclusion dynamics. It not only indicates anticipated correlation patterns but also key exceptions in between, providing a data-driven platform for policy interventions aimed at specific areas of action. Its implication for policymakers is that governance has to be enhanced indeed, but complemented by efforts in infrastructure development, financial education, and rollout of technology. This multi-dimensional understanding harmonizes well with the ESG worldview where governance as a standalone measurement but a supporting pillar for wider development objectives.

## 6.2 Predicting Inclusion: A Machine Learning Benchmark on Governance Predictors

The performance of regression algorithms must be evaluated in a multi-dimensional manner balancing both error minimization and explanatory strength. Eight algorithms (Boosting Regression, Decision Tree Regression, K-Nearest Neighbors (KNN) Regression, Linear Regression, Neural Networks, Random Forest, Regularized Linear Regression, and Support Vector Machine (SVM)) are compared here on a set of six normalized performance scores: Mean Squared Error (MSE), Scaled MSE, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE/MAD), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R<sup>2</sup>). All performance scores are minmax-normalized on the scale 0–1 so they can be directly compared, where smaller values for error scores (MSE, RMSE, MAE, and MAPE) denote better performance and larger values for R<sup>2</sup> indicate better model fit (Table 19).

Table 19. Balancing Accuracy and Fit: A Normalized Performance Benchmark of Eight Regression Models

				MAE /		
	MSE	MSE(scaled)	RMSE	MAD	MAPE	R <sup>2</sup>
Boosting Regression	0.12	0.01	0.13	0.17	0.0	1.0
Decision Tree Regression	0.5	0.45	0.52	0.59	0.58	0.52
K-Nearest Neighbors Regression	0.05	0.0	0.06	0.09	0.09	1.0
Linear Regression	0.31	0.77	0.34	0.46	0.15	0.2
Neural Networks	1.0	1.0	1.0	1.0	0.65	0.0
Random Forest	0.0	0.0	0.0	0.0	0.14	1.0
Regularized Linear	0.48	0.67	0.51	0.67	0.53	0.3
Support Vector Machine	0.78	0.84	0.8	0.79	1.0	0.14

The first model to consider would be Random Forest since it performs almost perfectly across all measures of error: MSE (0.0), RMSE (0.0), MAE (0.0), and scaled MSE (0.0). It also registers the highest achievable  $R^2$  (1.0), since it captures the entire variance of the dependent variable (Du et al., 2022). The relatively low value of MAPE (0.14) indicates this model strikes a good accuracy vs. generalizability balance. The zeroes placed in the error measures on a scaled space mean that out of the models considered, Random Forest registered lowest raw errors. Pairing this with the highest  $R^2$ , a strong case would be made for Random Forest to emerge the highest-performing model in this assessment (Zhang et al., 2020).

Following close behind is also K-Nearest Neighbors Regression (KNN), also with excellent performance in a normalized MSE (0.05), RMSE (0.06), MAE (0.09), and scaled MSE (0.0) all measuring minimal error in predictions. KNN also gets a perfect R<sup>2</sup> value of 1.0, indicative of excellent explanatory ability. Its slightly higher MAPE score of 0.09 versus 0.14 from Random Forest indicates slightly improved performance in percent error. Although KNN does have excellent performance, however, it is sensitive to noisy data and computationally ineffective in big datasets and so potentially limited in real-world application. Nevertheless, based strictly upon performance in terms of results, the close similarity between KNN and Random Forest means the slight trade-off in interpretability and scalability might be worth it with small or cleansed datasets.

Boosting Regression ranks third and also performs well on MSE (0.12), RMSE (0.13), MAE (0.17), and MAPE (0.0), all excellent performance metrics. The high point to note is it also returns the high  $R^2$  value of 1.0. The extremely low MAPE value of 0.0 might reflect very accurate computation of percentage error and Boosting Regression can be very beneficial where proportion errors play a big role, say in finance or marketing. The extremely low figure of error against other domains in comparison with Random Forest and KNN means while Boosting can be very potent, it might not be as consistent with broader performance metrics (Chauhan, 2024).

Decision Tree Regression performs poorly in all the measures: MSE (0.5), RMSE (0.52), MAE (0.59), and scaled MSE (0.45) indicate significantly higher levels of error relative to the top three models. Its  $R^2$  value (0.52) also talks about poor explanatory ability in explaining less than half the variation within the data. The MAPE (0.58) is also higher and reflects lower reliability in making percent predictions. The interpretability and simplicity of the Decision Tree are well-documented, but this overfitting and sensitivity over varied datasets can be observed from the above result.

Linear Regression, being simple to interpret and conceptual in character, performs poorly across all the measures. Its error levels are average: MSE (0.31), RMSE (0.34), MAE (0.46), and scaled MSE (0.77). Its R<sup>2</sup> value is just 0.20, reflecting the fact that it can explain very little variance in the data. While the MAPE is respectable (0.15), showing improved proportion forecasting, as a performance measure

overall, it shows linear regression hasn't got the ability to understand the complexity in the relation between the variables in the data.

Regularized Linear Regression also seeks to better the general linear model in its aim to reduce overfitting but provides inconsistent predictions. While the MSE value (0.48) and the RMSE value (0.51) are better than those of the Decision Trees, both its MAE value (0.67) and MAPE value (0.53) are high. Its  $R^2$  value (0.3) beats the  $R^2$  value of Linear Regression but lags behind the rest of the non-linear algorithms. This shows while the use of regularization improves the linear model, it cannot entirely compensate for the flaws in the linear model in its ability to capture the complex patterns in the dataset.

Support Vector Machine Regression performs the worst in several key areas. It produces extremely high normalized MSE (0.78), RMSE (0.8), MAE (0.79), and worst MAPE (1.0) for purposes of percent error, and measures very low  $R^2$  (0.14), implying low explanatory value. SVMs do potentially become very strong if a suitable kernel and tuning are employed, but their extremely high complexity and sensitivity to their parameters likely stop them from doing very well here.

In the bottom position in the performance rankings is Neural Networks, worst-ranked across all the important measures of error: MSE (1.0), RMSE (1.0), MAE (1.0), scaled MSE (1.0), and worst-after-the-second MAPE (0.65). Its  $R^2$  (0.0) further means they have no predictive ability whatsoever. Even though they are very valuable in complex, higher-order problems, poor performance here shows as overfitting, weak training, or network-incompatibility with the data set. Their need for large data sets and computing brawn drains their utility here as well.

Out of all the eight algorithms, Random Forest is the highest performer with the minimum errors and highest model fit ( $R^2 = 1.0$ ), the highest degree of generalizability and reliability across a range of applications. Its ensemble property gives it the ability to define complex patterns without overfitting, a theoretical requirement in real predictive use cases. KNN performance is similar but can result in lower scalability and more demanding data preprocessing needs. While very accurate in relative error, Boosting does have slightly higher absolute error and may be susceptible to overfitting if left tuned improperly.

Mid-range models like Regression and Decision Trees are somewhat useful in terms of interpretability and convenience but are weak in performance. SVM and Neural Networks also don't do too well, and this illustrates the wisdom in the selection of a model based not only on theoretical potency but also empirical fit to the available data and context.

Lastly, based on this multi-metric ranking, the Random Forest is the best general-purpose regression predictor for predicting financial inclusion or similar socio-economic indicators in this sample. It displays the best combination of predictive accuracy, low error, and explanatory power and thereby emerges as a general-purpose machine learner suitably suited to many different sorts of regression predictions.

The initial algorithm to look at would be Random Forest, as it has nearly perfect performance on all error measures: MSE (0.0), RMSE (0.0), MAE (0.0), and scaled MSE (0.0). It also has the best possible R<sup>2</sup> (1.0), as it accounts for the full variance of the dependent variable. The modest value of MAPE (0.14) suggests this model has a good balance between accuracy and generalizability. These findings are consistent with studies such as those by Remegio (2024), who demonstrated the efficacy of Random Forest in predicting student performance with impressive results (Remegio, 2024), and Sood et al. (2023), who showed its predictive strength in real estate analytics during COVID-19 (Sood et al., 2023).

Coming close behind it is K-Nearest Neighbors Regression (KNN), which also has stellar performance, with minimal prediction error and a perfect R<sup>2</sup> of 1.0. Teguh et al. (2024) confirmed this effectiveness when applying KNN in environmental data classification alongside Random Forest, noting its reliability when working with clean datasets (Teguh et al., 2024).

Boosting Regression also scores high, especially in scenarios where proportional errors are significant. Meanwhile, Decision Tree Regression and Linear Regression perform moderately to poorly across several metrics. These conclusions are echoed in the comparative evaluations by Kaliappan et al. (2021),

who tested various regression models on COVID-19 reproduction data and found ensemble methods like Random Forest to significantly outperform linear models (Kaliappan et al., 2021).

Interestingly, while Support Vector Machines (SVM) and Neural Networks have theoretical strengths, their performance in this evaluation was weak, reinforcing findings from sports analytics by Sanjaykumar et al. (2024), where Random Forest clearly outperformed other models in predicting outcomes in cricket matches (Sanjaykumar et al., 2024).

Ultimately, this evaluation confirms that Random Forest offers the best balance of predictive accuracy, minimal error, and model robustness across a wide range of applications, aligning with the most recent literature (Figure 9).

Figure 9. What Drives Financial Inclusion? Feature Importance of Governance Indicators in Random Forest Models



The data below depict the application of a Random Forest Regression to estimate financial inclusion in developing nations. The target variable employed here is Account Age, defined as the proportion of individuals having a financial institution account or using mobile money accounts. It will act as a proxy for financial inclusion—a key driver for socioeconomic development and a pillar of the Social (S) part of the ESG framework. Yet in this model, we will be estimating financial inclusion based on variables representing the Governance (G) part of ESG: Control of Corruption (CCOR), Patent Applications by Residents (PATR), and Regulatory Quality (REGQ) (Table 20).

Table 20. Governance in Action: Additive Effects of Corruption, Regulation, and Innovation on Financial Inclusion Forecasts

Case	Predicted	Base	CCOR	PATR	REGQ
1	26.780	42.390	-7.830	-0.471	-7.309
2	40.801	42.390	-5.596	-0.915	4.923
3	45.063	42.390	-4.189	7.373	-0.511
4	45.580	42.390	-3.379	1.221	5.348
5	44.243	42.390	-3.260	2.469	2.644

The case's predicted value is then moved from the base by the additive effect of CCOR, PATR, and REGQ. The additive accounts allow us to view the effect of a change in each of the governance variables on the estimation of financial inclusion and thus insight into each of the governance components' relative effect on financial accessibility. The forecast value in Case 1 is 26.780, which falls below the base level significantly. The fall here is dominated by extremely negative contributions from both REGQ (-7.309) and CCOR (-7.830), and a lesser effect from PATR (-0.471). The magnitude of the effects suggests financial inclusion gets heavily repressed in the contexts where both corruption levels and regulatory quality are high. The result is not surprising: higher corruption levels tend to destroy trust and deter financial formalization in both government and the finance sector, and poor-quality regulatory institutions impose disincentives on access to finance. The minor effect of patent applications here suggests innovation itself cannot act as a proxy for poor institutions in promoting financial inclusion (Figure 10).

Figure 10. Governance and Inclusion: Random Forest Insights on the Institutional Drivers of Financial Access



Case 2 shows a much closer estimate to the base (40.801) under mixed governance contributions. Here, both CCOR (-5.596) and PATR (-0.915) both remain negative but less intense than Case 1 and are almost offset by a strong positive effect from REGQ (+4.923). This trend suggests that even if corruption persists as a challenge, good governance has a compensatory effect whereby it permits access to a greater number of financial services through regulated channels (Ben Khelifa, 2023). The example: well-regulated

mobile money services, consumer protection regulations or easy financial onboarding may drive inclusion even in contexts where trust in institutions remains low.

Case 3 has a forecasted Account Age value above the baseline (45.063) thanks mainly to a very large positive contributor from PATR (+7.373). Both CCOR (-4.189) and REGQ (-0.511) happen to be small and negative but not very significant quantitatively. This indicates a unique set of conditions where innovation—through fintech, digital platforms, or entrepreneurial ecosystems—bridges financial exclusion despite governance frailties (Okereke et al., 2023). In developing nations, private-sector innovations in financial tech have opened alternative channels of inclusion irrespective of, or even in the face of, government ineffectiveness.

Case 4's even stronger forecast (45.580) benefits from a high REGQ effect (+5.348) and positive PATR (+1.221), and a modest drag from CCOR (-3.379). This input set illustrates synergy between good regulatory conditions and modest innovation in supporting financial access. If institutions can establish inclusive financial policies—e.g., identity checks as a requirement for access, e-banking systems, or financial education drives—conditions become ripe for large-scale financial participation (Zeqiraj et al., 2022).

Case 5 also has a similar profile to Case 4 and features a projected value (44.243) above the base as well. The positive contributions from REGQ (+2.644) and PATR (+2.469) and negative impact of CCOR (-3.260) reinforce the previous conclusion: regulatory capacity and innovation together form a formidable force behind financial inclusion and may neutralize the adverse impact of corruption to some extent.

Across all cases, the Control of Corruption (CCOR) variable shows a consistent negative impact on financial inclusion, supporting the argument that corruption has a systematic role in reducing access. It deters participation in the formal system, induces institutional mistrust, and marginalizes vulnerable groups. On the other hand, Regulatory Quality (REGQ) plays a largely positive role, providing a supportive environment for inclusion via financial service rules, consumer protections, and outreach incentives. Patent applications (PATR), while impactful in some cases like Case 3, do not substitute for institutional strength and are most effective when paired with solid governance.

Overall, this Random Forest Regression illustrates that in developing countries, financial inclusion is strongly influenced by governance characteristics, where anti-corruption, regulatory quality, and innovation must work in tandem.

### 7. Financial Inclusion as a Governance Lever: Policy Insights from ESG Analytics

The results of this research present strong evidence-based findings that highlight important policy design implications for developing nations. Empirically quantifying the effect of financial inclusion on ESG performance—specifically in the areas of governance, innovation and social—this research provides tangible advice on how policymakers, development institutions, and financial institutions can encourage inclusive and sustainable development. Fundamentally, the research presents the proposition that financial inclusion is not simply a socioeconomic aspiration but also a policy tool able to reform governance institutions, catalyze innovation, and extend social strength [(Jain et al., 2024)].

The statistically significant and positive correlation between financial inclusion and governance of corruption means extending access to financial services can be a central mechanism for enhancing governance [(Luo et al., 2024)]. Specifically, through enhanced formal economic participation, financial

inclusion minimizes the room for corruption and informality. Policymakers must appreciate that inclusive financial systems not only advance economic equity targets but also complement wider governance change by improving transparency, facilitating digital trackability, and fostering trust in institutions. To make this insight a reality, anti-corruption agencies, central banks, and financial regulators must support each other in digitalizing financial services, instituting e-KYC (know-your-customer) procedures, and promoting mobile payment platforms in both private and government operations [(Li et al., 2024)].

Though the findings reveal a statistically significant relationship between regulatory quality and financial inclusion, the negative sign indicates the risk of rapid financial growth in developing economies outpacing institutions at times. The finding demands a systemic approach: financial innovation and access must be coupled with investments in commensurate regulatory capacity. If not, the system can become prone to fraud, predatory lending, and consumer injury. The policy response must involve the establishment of quick-response regulatory sandboxes where new financial technologies may be tried out under regulated conditions [(Zaroni et al., 2025)]. Regulators must also enhance their digital oversight capacity and invest in workforce capacity building on fintech dynamics and engage in cooperation with providers of technology on developing compliance procedures together. Public education campaigns on digital financial literacy must be part of the mix in order not to unintentionally exclude vulnerable groups from innovative financial services.

Patent filings—taken as a proxy for innovation—exhibited a high and positive correlation with financial inclusion. It indicates that fintech and digital entrepreneurship-centered innovation ecosystems can be inclusive engines if properly supported [(Jain et al., 2024)]. Yet their effect hinges on a certain quality threshold of governance. The policymakers must thus nurture innovation ecosystems through strategic incentives for fintech startups, investments in digital infrastructure, and public-private collaborations expanding coverage for inclusive financial services. Particular concern must be given to supporting innovations specific to rural, low-income, or informal sectors, including micro-savings apps, mobile credit scoring systems, and blockchain remittances.

Application of Fuzzy C-Means clustering in this research identified countries as having overlapping ESG and financial inclusion profiles, e.g., as "environmentally advanced but financially excluded" or "institutional weak but digitally innovative" [(Creţu et al., 2024)]. These diverse profiles indicate that a single one-size-fits-all approach will not be adequate. Accordingly, financial inclusion policies will be designed for specific country archetypes. For instance, regulatory capacity-strong but low-innovation countries will be enhanced by programs promoting fintech start-ups. Conversely, weak governance but high-digitization countries will need investments in institutional accountability prior to embarking on financial access activity. The cluster outputs may form the building blocks for differentiated country roadmaps and be supported by multilateral institutions and global donors.

The environmental factors employed as instrumental variables in the study—e.g., land use patterns, penetration of renewable energy, and productivity of agriculture—were seen to have a significant impact on financial inclusion outcomes [(Zaroni et al., 2025)]. This finding suggests that financial and environmental agendas ought not be siloed but rather be seen as complementary. To take a few examples, availing more sources of clean energy or promoting sustainable agriculture may indirectly create conditions for financial inclusion through making available new business models and expanding demands for financial services. Development planners must therefore incorporate financial services as part of wider climate and sustainable programs.

Climate adaptation funds, for instance, can be part of microcredit lines to smallholder farmers. Likewise, energy transition programs can be coupled with financing packages for household solar systems or sustainable irrigation systems. The study's social indicators—education levels, internet penetration, and gender equality—thus become key drivers of financial inclusion. More highly performing countries on these fronts tend toward greater levels of financial access [(Luo et al., 2024)]. This re-emphasizes the

value of integrating financial inclusion as part of a wider social development agenda. Policies must comprise digital skilling of underserved communities, broadband rollout, and gender-specific financial products. Data subsidy plans from telecos and the disaggregation of financial data by gender must be supported by governments as pre-requisites for wider progress.

The ministries of education must also look at integrating financial literacy as part of secondary school education as a starting point for lifetime financial engagement.

One of the methodological innovations of this research is the application of environmental variables as instruments for social and governance variables [(Li et al., 2024)]. Not only is this method a solution for addressing endogeneity, but it also demonstrates the interlinked nature of ESG pillars. From a policymaker's perspective, this corroborates the perception that financial inclusion, institutional development, and sustainability cannot be pursued separately from each other. Rather, holistic approaches—like "Green Financial Inclusion" programs—are likely to yield synergistic effects. Multilateral development banks, for example, should create cross-sectoral programs connecting climate adaptation financing with inclusive banking instruments. Governments might encourage banks to open in underserved locations by tying access to green financing facilities with inclusivity targets.

Last but not least, a key conclusion of the study is that institutional quality both precedes and results from financial inclusion. This two-way causality indicates that investments in institutional building—namely judiciary effectiveness, anti-corruption agencies, and regulatory autonomy—will be necessary in order to maintain financial inclusion progress [(Jain et al., 2024)]. Recommendations for policies here would be to enhance country-level financial inclusion policies through well-defined governance reform elements, enhance the central banks' role in consumer protection in financial services, and establish independent bodies for ESG monitoring by financial institutions.

## 8. Conclusions

This paper has studied the multi-dimensional link between ESG aspects of sustainable development and financial inclusion over a 12-year period in 103 developing nations. The research used panel data econometric models and instrumental variables and sophisticated machine-learning algorithms— Random Forest regression and Fuzzy C-Means clustering—to both establish causal and predictive effects of ESG dynamics on inclusive financial systems.

Empirical results report financial inclusion strongly and affirmatively related to multiple ESG drivers although their intensity and direction differ between the two pillars. At the environmental level, ecomodernization indicators—renewable energy consumption, agricultural production, and protected areas—appear as prominent facilitators of financial access, while agrarian traditional forms hamper inclusion. The social pillar presents as similarly strong effects: education, access to the digital economy, life expectancy, sanitation, and gender equality all emerge as strong predictors of formal financial participation. In contrast, higher labor force participation—with a strong link to the informal economy— negatively relates to inclusion and points toward structural restraints undermining formalization. At the governance level, financial inclusion relates affirmatively to corruption control and innovation capacity and indicates inclusive institutions encourage wider financial access. Nevertheless, the negative link between regulatory quality cautions about a possible interplay mismatch between rapid financial innovation and preparedness.

From a methodological perspective, both IV-based econometrics and machine learning proved a powerful combination for untangling the complex and sometimes non-linear inter-relationships between variables. Random Forest regression from machine learning provided good predictive performance and allowed us to pinpoint the most impactful ESG variables, and clustering identified the heterogeneity

between countries—qualifying that ESG advance and financial inclusion do not converge across all countries in a uniform manner. The country categorization into ESG-inclusion archetypes offers a useful guide for policy intervention on a targeted basis.

Altogether, this report adds to mounting empirical evidence that financial inclusion not only represents a matter of social necessity but also a strategic part of wider agendas for sustainability. Alignment of ESG priorities in financial inclusion can create cross-cutting benefits ranging from strengthening institutions to building an ability to cope with climate change and promoting gender equality. Policymakers and institutions at a global level and financial actors in the Global South must therefore tackle inclusion not separately from other priorities but as an inherent outcome of environmental responsibility, social progress, and governance reform. By so doing, they can tap the synergy effect of ESG approaches in developing fairer, stronger, and sustainable economic systems.

### 9. References

Ababio, J. O. M., Yiadom, E. B., Mawutor, J. K., Tuffour, J. K., & Attah-Botchwey, E. (2023). Sustainable energy for all: the link between financial inclusion, renewable energy and environmental sustainability in developing economies. International Journal of Energy Sector Management, 18(5), 1088-1108.

Acheampong, A. O., & Said, R. (2024). Financial inclusion and the global net-zero emissions target: Does governance matter? Environmental Science and Pollution Research, 31, 14318–14336.

Aghaei, V., Seiedkhani, R., Mohammadipour, R., Feizollahi, S., & Moradpour, M. (2024). Validation of Loan Applicants Using Environmental, Social, and Governance Indicators. Management Strategies and Engineering Sciences, 6(5), 92-104.

Ahmad, F., Cheong, C. W., Hussain, R. T., Salman, A., Hussain, K., Shahid, M. K., & Naseem, M. A. (2024). Financial inclusion and sustainability: A contemporary and future perspective from bibliometric analysis. In Financial inclusion, sustainability, and the influence of religion and technology (pp. 105-132). IGI Global.

Ahmad, M., Ahmed, Z., Kousar, R., & Adebayo, T. S. (2022). Financial inclusion, technological innovations, and sustainable development: A panel data analysis of BRICS economies. Technological Forecasting and Social Change, 178, 121604.

Al-Baraki, Mohammed. (2022). Investing in financial inclusion. Master Thesis. Copenhagen Business School M.Sc. Applied Economics & Finance

Alibašić, H. (2024). The Post-factual Polity: Ethical, Governance, Administrative, and Policy Crises in the Disinformation Era. IAP.

Ambarsari, E. W., Dwitiyanti, N., Selvia, N., Cholifah, W. N., & Mardika, P. D. (2023). Comparison Approaches of the Fuzzy C-Means and Gaussian Mixture Model in Clustering the Welfare of the Indonesian People. KnE Social Sciences, 16-22.

Anwar, S., Rusanti, E., & Mochlasin. (2023). Contemporary issues in sustainable finance: Exploring performance, impact measurement and financial inclusion: edited by Mario La Torre and Sabrina Leo, Cham, Switzerland, Palgrave Macmillan, 2023, XXII+ 294 pp., EUR 149.99 (hardback), ISBN 9783031225383.

Austin, T., & Rawal, B. S. (2023). Model Retraining: Predicting the Likelihood of Financial Inclusion in Kiva's Peer-to-Peer Lending to Promote Social Impact. Algorithms, 16(8), 363.

Azkeskin, S. A., & Aladağ, Z. (2025). Evaluating regional sustainable energy potential through hierarchical clustering and machine learning. Environmental Research Communications, 7(1), 015002.

Balde, A., Adhikari, S., Singh, A., Mabuela, N. P., & Choudhary, A. (2022, September). An Efficient Approach for Predicting the state of Financial Inclusion by Addressing Class Imbalance using the Oversampling Technique. In 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. 1-5). IEEE.

Barboza, F., & Altman, E. (2024). Predicting financial distress in Latin American companies: A comparative analysis of logistic regression and random forest models. The North American Journal of Economics and Finance, 72, 102158.

Ben Khelifa, S. (2023). Governance quality, social and macro-economic conditions: implications for financial inclusion. Journal of Sustainable Finance & Investment, 13(1), 463-476.

Bester, T., & Rosman, B. (2024). Towards financially inclusive credit products through financial time series clustering. arXiv preprint arXiv:2402.11066.

Bhaskaran, E. (2023). Financial inclusion in the industrial cluster. Journal of The Institution of Engineers (India): Series A, 98(1-2), 115-120.

Blasilli, G., Kerrigan, D., Bertini, E., & Santucci, G. (2024, October). Towards a Visual Perception-Based Analysis of Clustering Quality Metrics. In 2024 IEEE Visualization in Data Science (VDS) (pp. 15-24). IEEE.

Bondarenko, S., Kalaman, O., & Danilova, O. (2025). DIGITAL TRANSFORMATION OF THE TOURISM BUSINESS IN AN OPEN ECONOMY: A COMPREHENSIVE APPROACH TO TRAVEL PLANNING AND PARTNERSHIPS. Social Development: Economic and Legal Issues, (4).

Borgi, H., Mabrouk, F., Bousrih, J., & Mekni, M. M. (2023). Environmental change and inclusive finance: does governance quality matter for african countries?. Sustainability, 15(4), 3533.

Buckley, R. P., Zetzsche, D. A., Arner, D. W., & Veidt, R. (2021). FinTech, financial inclusion and the UN Sustainable Development Goals. In Routledge Handbook of Financial Technology and Law (pp. 247-272). Routledge.

Chafai, A., Khémiri, W., Tobar, R., Attia, E. F., & Abozeid, H. O. (2024). The moderating effect of audit quality on the relationship between financial inclusion and corporate investment: new evidence from the Middle East and North Africa region. Cogent Economics & Finance, 12(1), 2378121.

Chauhan, S. (2024, November). Machine Learning Models for Loan Default Forecasting: Accuracy Comparison. In 2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES) (pp. 1-5). IEEE.

Chininga, E., Alhassan, A. L., & Zeka, B. (2023). ESG ratings and corporate financial performance in South Africa. Journal of Accounting in Emerging Economies, 14(3), 692-713.

Crețu, R. F., Banța, V. C., Oancea-Negescu, M. D., Anica-Popa, A., & Crecană, C. D. Challenges in the Food Market and the Agenda of Future Research Directions in the Context of Sustainability: A Bibliometric Exploration.

Cui, L., Weng, S., & Song, M. (2022). Financial inclusion, renewable energy consumption, and inclusive growth: Cross-country evidence. Energy Efficiency, 15(6), 43.

Das, S. G., & Nayak, B. (2020). Financial inclusion: An application of machine learning in collaborative filtering recommender systems. Int. J. Recent Technol. Engg, 8(6).

de Oliveira, M. A. (2023). ENVIRONMENTAL, SOCIAL AND GOVERNANCE (ESG) PRACTICES AND ORGANIZATIONAL RESILIENCE IN CREDIT UNIONS. Revista Alcance (Online), 30(2), 13-27.

Dhanabhakyam, M., & Suresh, S. P. (2024). Exploring the role of fintech innovations in driving sustainability: A study on user adoption and ESG integration in financial services. IJAR, 10(12), 65-69.

Dovbiy, I. P. (2022). Financial and economic conditions of the energy transition for the national economy. Finansovyj žhurnal—Financial Journal, (5), 25-42.

Du, S., Hao, D., & Li, X. (2022, October). Research on stock forecasting based on random forest. In 2022 IEEE 2nd International Conference on Data Science and Computer Application (ICDSCA) (pp. 301-305). IEEE.

Duan, J., & Sun, H. (2020, December). Comprehensive financial analysis of a company relying on fuzzy c-mean (FCM) clustering algorithm. In 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE) (pp. 1904-1907). IEEE.

Eldomiaty, T., Hammam, R., & El Bakry, R. (2020). Institutional determinants of financial inclusion: evidence from world economies. International journal of development issues, 19(2), 217-228.

Essel-Gaisey, F., & Chiang, T. F. (2022). Turning the tide on environmental poverty in Ghana: does financial inclusion matter?. Sustainable Production and consumption, 33, 88-100.

Fau, J. F., Wau, M., Laia, F., & Wau, F. T. (2025). Fintech for ESG and the circular economy: by Ernesto DR Santibanez Gonzalez, Vinay Kandpal, Peterson K. Ozili and Prasenjit Chatterjee (Eds.), Chile, India, Nigeria, India, Wiley-Scrivener, 2024, 320 pp., 156.00(E-book), 100.40 (hardcover), ISBN-10: 1394233965, ISBN-13: 9781394233960.

Feng, J., Sun, Q., & Sohail, S. (2022). Financial inclusion and its influence on renewable energy consumption-environmental performance: the role of ICTs in China. Environmental Science and Pollution Research, 29(35), 52724-52731.

Galeone, G., Ranaldo, S., & Fusco, A. (2024). ESG and FinTech: Are they connected?. Research in International Business and Finance, 69, 102225.

Gidage, M., & Bhide, S. (2024). ESG and economic growth: Catalysts for achieving sustainable development goals in developing economies. Sustainable Development.

Gonzalez, E. D., Sijariya, R., Singh, A. K., & Garg, V. (2025). Integrating Climate Change, Social Responsibility and Electronic Financial Inclusion: A Pathway to Sustainable Development. In Climate Change and Social Responsibility (pp. 147-164). Emerald Publishing Limited.

Halim, M. A. (2024). Does crowdfunding contribute to digital financial inclusion?. Research in Globalization, 100238.

Hassani, B. K., Mushtaq, R., & Bahini, Y. (2024). ESG (Environment, Social, and Governance) development: What is the role of financial development at the global level?. Bankers, Markets & Investors, 177(2), 14-29.

Hicham, A., & Hicham, O. (2020). Modelling financial inclusion using a panel macroeconomic approach: evidence from african countries. European Journal of Economic and Financial Research, 4(3).

Ibarz, B., Kurin, V., Papamakarios, G., Nikiforou, K., Bennani, M., Csordás, R., ... & Veličković, P. (2022, December). A generalist neural algorithmic learner. In Learning on graphs conference (pp. 2-1). PMLR.

Ishizaka, A., Lokman, B., & Tasiou, M. (2021). A stochastic multi-criteria divisive hierarchical clustering algorithm. Omega, 103, 102370.

Jain, P., Kanoujiya, J., Gautam, R. S., & Rastogi, S. (2024). Impact of Financial Inclusion and ESG on the Sustainable Growth of the Nation: Moderating Role of Poverty. Vision, 09722629241271509.

Jin, Y. (2022). On \$ p \$-adic Measures for Quaternionic Modular Forms. arXiv preprint arXiv:2209.11822.

Jin, Y. (2025). Distinctive impacts of esg pillars on corporate financial performance: A random forest analysis of korean listed firms. Finance Research Letters, 71, 106395.

Johnson, E. A., Inyangetoh, J. A., Rahmon, H. A., Jimoh, T. G., Dan, E. E., & Esang, M. O. (2024). An intelligent analytic framework for predicting students academic performance using multiple linear regression and random forest. European Journal of Computer Science and Information Technology, 12(3), 56-70.

Jungo, D., Madaleno, M., & Botelho, A. (2022). The relationship between social inclusion, financial innovation, and economic growth in the Euro area. Journal of Policy Modeling, 44(2), 367–385.

Kaliappan, J., Srinivasan, K., Mian Qaisar, S., Sundararajan, K., & Chang, C. Y. (2021). Performance evaluation of regression models for the prediction of the COVID-19 reproduction rate. Frontiers in Public Health, 9, 729795.

Kaplan, H., Rahman, M. M., Khan, S., & Vergil, H. (2024). The triple impact of innovation, financial inclusion, and governance on sustainable development in emerging markets. Sustainability, 16(6), 3011.

Kaushal, M., Danish Lohani, Q. M., & Castillo, O. (2024). Weighted intuitionistic fuzzy c-means clustering algorithms. International Journal of Fuzzy Systems, 26(3), 943-977.

Kawor, S. (2023). Effects Of Financial Development And Digital Financial Exclusion On Elements Of Welfare In Sub-Saharan Africa (Doctoral dissertation, University of Cape Coast).

Kebede, J., Naranpanawa, A., & Selvanathan, S. (2023). Financial inclusion and income inequality nexus: A case of Africa. Economic Analysis and Policy, 77, 539-557.

Khalid, F., Irfan, M., & Srivastava, M. (2024). The impact of digital inclusive finance on ESG disputes: Evidence from Chinese non-financial listed companies. Technological Forecasting and Social Change, 204, 123415.

Khalil, I., & Siddiqui, D. A. (2020). Effect of Economic Growth, Human Development, and Governance Factors on Financial Inclusion: A Cross Country Comparative Analysis. Human Development, and Governance Factors on Financial Inclusion: A Cross Country Comparative Analysis.(December 31, 2020).

Khan, S., Ullah, A., Liu, Y., & Kashif, M. (2025). Examining the blockchain and green innovation technologies on sustainability (ESG): the moderating role of global financial integration. Journal of Sustainable Finance & Investment, 15(1), 145-181.

Kiran, M., Naeem, M. A., & Rabbani, M. R. (2025). Evolving dynamics of ecological sustainability: investigating ESG investments and governance quality in G8 nations. Journal of the Knowledge Economy, 1-29.

Latifah, U. W. (2022). K-means and fuzzy c-means algorithm comparison on regency/city grouping in Central Java Province. Desimal: Jurnal Matematika, 5(2), 155-168.

Lee, H. S., Kuang, K. S., Liew, P. X., Har, W. M., & Low, C. W. (2023, September). Supporting financial inclusion with digital financial services: comparative analysis of machine learning models. In 2023 IEEE

International Conference on Artificial Intelligence in Engineering and Technology (IICAIET) (pp. 283-288). IEEE.

Li, K. (2025). Big Data and Machine Learning in ESG Research. Asia-Pacific Journal of Financial Studies, 54(1), 6-21.

Li, W., & Pang, W. (2023). The impact of digital inclusive finance on corporate ESG performance: based on the perspective of corporate green technology innovation. Environmental Science and Pollution Research, 30(24), 65314-65327.

Li, W., Shi, C., Xiao, Z., & Zhang, X. (2024). Bridging the green gap: How digital financial inclusion affects corporate ESG greenwashing. Finance Research Letters, 69, 106018.

Li, Weiping, Chang Shi, Zhongyi Xiao, and Xuezhi Zhang. "Bridging the green gap: How digital financial inclusion affects corporate ESG greenwashing." Finance Research Letters 69 (2024): 106018.

Lindquist, W. B., Rachev, S. T., Hu, Y., & Shirvani, A. (2022). Inclusion of ESG ratings in optimization. In Advanced REIT Portfolio Optimization: Innovative Tools for Risk Management (pp. 227-245). Cham: Springer International Publishing.

Liu, J., & Naser, M. A. (2024). Investigating the interplay of financial inclusion with bank performance.

Liu, R. (2025, February). The Social Effects of Corporate ESG on Digital Financial. In Proceedings of the International Workshop on Navigating the Digital Business Frontier for Sustainable Financial Innovation (ICDEBA 2024) (Vol. 315, p. 56). Springer Nature.

Liu, X., Ma, C., & Ren, Y. S. (2024). How AI powers ESG performance in China's digital frontier?. Finance Research Letters, 70, 106324.

López-Oriona, Á., D'Urso, P., Vilar, J. A., & Lafuente-Rego, B. (2022). Quantile-based fuzzy C-means clustering of multivariate time series: Robust techniques. International Journal of Approximate Reasoning, 150, 55-82.

Lu, H., & Cheng, Z. (2024). Digital inclusive finance and corporate ESG performance: The moderating role of executives with financial backgrounds. Finance Research Letters, 60, 104858.

Lu, Hongyu, and Zhao Cheng. "Digital inclusive finance and corporate ESG performance: The moderating role of executives with financial backgrounds." Finance Research Letters 60 (2024): 104858.

Lu, Y., Wang, L., & Zhang, Y. (2022). Does digital financial inclusion matter for firms' ESG disclosure? Evidence from China. Frontiers in Environmental Science, 10, 1029975.

Luo, Z., Li, Y., Nguyen, L. T., Jo, I., & Zhao, J. (2024). The moderating role of country governance in the link between ESG and financial performance: a study of listed companies in 58 countries. Sustainability, 16(13), 5410.

Maket, I. (2024). Rethinking energy poverty alleviation through financial inclusion: Do institutional quality and climate change risk matter?. Utilities Policy, 91, 101820.

Malik, A. H., Jais, M. B., Md Isa, A. H., & Rehman, A. U. (2022). Role of social sustainability for financial inclusion and stability among Asian countries. International Journal of Social Economics, 49(9), 1324-1348.

Mehdiyev, A. (2023). Development of the ESG Integration in Financial Inclusions in EU: A Comparative Analysis of Financial Performance and Environmental Impact.

Mirza, N., Umar, M., Lobont, O. R., & Safi, A. (2025). ESG lending, technology investment and banking performance in BRICS: navigating sustainability and financial stability. China Finance Review International.

Morelli, C., Boccaletti, S., Maranzano, P., & Otto, P. (2025). Multidimensional spatiotemporal clustering–an application to environmental sustainability scores in europe. Environmetrics, 36(2), e2893.

Mushagalusa, C. A., Fandohan, A. B., & Glèlè Kakaï, R. (2022). Random forests in count data modelling: An analysis of the influence of data features and overdispersion on regression performance. Journal of Probability and Statistics, 2022(1), 2833537.

Okereke, S. F., Kyarem, R., & Iyoboyi, M. (2023). interaction effect of financial inclusion and institutional quality on living standard in Sub-Saharan Africa. Lapai Journal of Economics, 7(2), 1-17.

Paliienko, T., & Diachenko, V. (2024). Green Indices and Their Role in the Implementation of Esg in the Context of Sustainable Economic Development. Економіка розвитку систем, 6(2), 107-116.

Pan, J. (2025, June). and Medium-Sized Enterprises in the Context of the Digital Economy and Financial Inclusion. In Proceedings of the 2025 4th International Conference on Big Data Economy and Digital Management (BDEDM 2025) (Vol. 191, p. 51). Springer Nature.FINANCIAL PRACTICES. International Journal of Business Ethics in Developing Economies, 13(1).

Park, J., Na, H. J., & Kim, H. (2024). Development of a Success Prediction Model for Crowdfunding Based on Machine Learning Reflecting ESG Information. IEEE Access.

Putra, D. M., & Abdulloh, F. F. (2024). Comparison of Clustering Algorithms: Fuzzy C-Means, K-Means, and DBSCAN for House Classification Based on Specifications and Price. Journal of Applied Informatics and Computing, 8(2), 509-515.

Qureshi, S. H. (2025). Nexus of Digital Inclusive Finance and Corporate ESG Performance to Achieve Financial Sustainability.

Rajunčius, M., & Miečinskienė, A. (2024). Measuring the Impact of Payment Innovations in Sustainable Finance: A Refined Framework for Evaluating ESG, Social Equity, Financial Inclusion, and Efficiency.

Ravichandran, M. R., & Rao, L. (2023). FINTECH IMPORTANCE, ADOPTION AND INNOVATIONS HAPPENING IN THE SECTOR–FINANCIAL INCLUSION AND LINKS TO SUSTAINABLE ECONOMIC GROWTH-A PERSPECTIVE. In NATIONAL CONFERENCE–2023.

Remegio, F. M. C. (2024, March). Predicting Student Performance to Boost Educational Outcomes: The Efficacy of a Random Forest Approach. In 2024 13th International Conference on Educational and Information Technology (ICEIT) (pp. 253-260). IEEE.

Roy, J. K., & Vasa, L. (2025). Financial technology and environmental, social and governance in sustainable finance: a bibliometric and thematic content analysis. Discover Sustainability, 6(1), 1-22.

Rusu, Ş., Boloş, M. I., & Leordeanu, M. (2023). K-means and agglomerative hierarchical clustering analysis of esg scores, yearly variations, and stock returns: insights from the energy sector in Europe and the United States. Journal of Financial Studies, 8(Special-June\_2023), 166-180.

Said, R. (2024). Financial inclusion and environmental pollution in sub-Saharan Africa: Moderating effects of economic growth and renewable energy. Environmental Science and Pollution Research, 31(43), 55346-55360.

Saltık, Ö. (2024). Algorithmic Approaches to Unraveling ESGC Impact on Financial Per-formance Based on the Ethical Investing in the Energy Sector. International Journal of Pure and Applied Sciences, 10(2), 459-517.

Sanderson, A., Mutandwa, L., & Le Roux, P. (2018). A review of determinants of financial inclusion. International Journal of Economics and Financial Issues, 8(3), 1.

Sanjaykumar, S., Natarajan, S., Lakshmi, P. Y., & Boby, F. A. (2024). Predicting team success in the Indian Premier League Cricket 2024 season using random forest analysis. Physical Education Theory and Methodology, 24(2), 304-309.

Sarmas, E., Fragkiadaki, A., & Marinakis, V. (2024). Explainable AI-Based Ensemble Clustering for Load Profiling and Demand Response. Energies, 17(22), 5559.

Sasmita, N. R., Khairul, M., Sofyan, H., Kruba, R., Mardalena, S., Dahlawy, A., ... & Maula, A. W. (2023). A Statistical Clustering Approach: Mapping Population Indicators Through Probabilistic Analysis in Aceh Province, Indonesia. Infolitika Journal of Data Science, 1(2), 63-72.

Schwarcz, S. L., & Leonhardt, T. L. (2021). Scoping and defining financial inclusion, access to credit, and sustainable finance. Law & Contemp. Probs., 84, 1.

Shah, S. F., Damra, Y., Yasin, S., & Albaity, M. (2024). Deciphering the Impact of Happiness, Financial Inclusion, ESG Metrics, and Economic Policy Uncertainty on Bank-Level Risk Dynamics.

Shahbaz, M., Topcu, B. A., Sarıgül, S. S., & Vo, X. V. (2021). The effect of financial development on renewable energy demand: The case of developing countries. Renewable Energy, 178, 1370-1380.

Shahid, S., & Asghar, N. (2024). The Role of Financial Inclusion to Enhance Social Progress in Developed and Developing Countries. Pakistan Journal of Humanities and Social Sciences, 3383-3393.

Shak, M. S., Uddin, A., Rahman, M. H., Anjum, N., Al Bony, M. N. V., Alam, M., ... & Pervin, T. (2024). INNOVATIVE MACHINE LEARNING APPROACHES TO FOSTER FINANCIAL INCLUSION IN MICROFINANCE. International Interdisciplinary Business Economics Advancement Journal, 5(11), 6-20.

Shala, A., & Berisha, V. (2024). The impact of Fintech on the achievement of environmental, social, and governance (ESG) goals. In Sustainable Development Goals: The impact of sustainability measures on wellbeing (pp. 55-77). Emerald Publishing Limited.

Shen, Y., E, H., Chen, T., Xiao, Z., Liu, B., & Chen, Y. (2021, June). High-dimensional data clustering with fuzzy C-Means: problem, reason, and solution. In International Work-Conference on Artificial Neural Networks (pp. 89-100). Cham: Springer International Publishing.

Sood, K., Shang, S., & Nijad, N. (2023, June). Predictive Analytics for Real Estate During COVID-19 Pandemic. In 2023 IEEE World AI IoT Congress (AIIoT) (pp. 0338-0344). IEEE.

Srisuradetchai, P., & Suksrikran, K. (2024). Random kernel k-nearest neighbors regression. Frontiers in big Data, 7, 1402384.

Sunori, S. K., Mittal, A., Singh, N., Manu, M., & Juneja, P. (2023, June). Fuzzy c-means clustering of soil data samples. In 2023 3rd International Conference on Intelligent Technologies (CONIT) (pp. 1-4). IEEE.

Supartoyo, Y. H. (2023). The state of financial inclusion for agriculture in Indonesia. Jurnal Ekonomi Pertanian dan Agribisnis, 7(2), 926-932.

Suresha, B., Srinidhi, V. R., Verma, D., Manu, K. S., & Krishna, T. A. (2022). The impact of ESG inclusion on price, liquidity and financial performance of Indian stocks: Evidence from stocks listed in BSE and NSE ESG indices. Investment Management & Financial Innovations, 19(4), 40.

Teguh, S., Muhammad Rafli, A., Haldi, B., M Rezqy, N. R., Usman, S., & Noor, A. (2024). Comparison of Logistic Regression, Random Forest, SVM, KNN Algorithm for Water Quality Classification Based on Contaminant Parameters. Journal of Data Science, 2024(48), 1-7.

Tekin, H. (2025). Financial inclusion and sustainable development of banks: the place of organization of Islamic cooperation (OIC) in emerging markets. Journal of Economic Studies, 52(1), 161-178.

Thomas, S. (2023). The problem of greenwashing in ESG investments in Finance. BQ Prime.

Trotta, A., Rania, F., & Strano, E. (2024). Exploring the linkages between FinTech and ESG: a bibliometric perspective. Research in International Business and Finance, 69, 102200.

Ubeda, F., Mendez, A., & Forcadell, F. J. (2024). Sustainable banking and trust in the global South. Global Policy, 15, 34-44.

Ubeda, Fernando, Alvaro Méndez, and Francisco Javier Forcadell. "Sustainable banking & financial inclusion in the global South: the "S" in ESG." Available at SSRN 4384975 (2023).

Upadhya, V., Bendigeri, D., & Nadiger, S. (2024). EXPLORING THE IMPACT OF FINTECH ON ESG (ENVIRONMENTAL, SOCIAL AND GOVERNANCE) INVESTMENTS: AN EMPIRICAL ANALYSIS OF TECHNOLOGICAL INNOVATIONS SHAPING SUSTAINABLE

Vikrant, V., & Bhattacharjee, K. (2024). G\" odel Number based Clustering Algorithm with Decimal First Degree Cellular Automata. arXiv preprint arXiv:2405.04881.

Waliszewski, K. (2023). Socio-economic determinants of financial inclusion in Central and Eastern Europe during the COVID-19 pandemic. Przegląd Wschodnioeuropejski, 14(1), 25-41.

Wang, Y., Fahad, S., Wei, L., Luo, B., & Luo, J. (2022). Assessing the role of financial development and financial inclusion to enhance environmental sustainability: do financial inclusion and eco-innovation promote sustainable development?. Frontiers in Environmental Science, 10, 1056478.

Wani, I. U., & Khanday, I. N. (2024). Dynamic interaction between financial inclusion and gender equality in determining economic growth: A study of developing countries. Journal of the Knowledge Economy, 1-26.

Xie, Y. (2024). The impact of ESG performance on corporate sustainable growth from the perspective of carbon sentiment. Journal of Environmental Management, 367, 121913.

Xu, J. (2024). AI in ESG for financial institutions: an industrial survey. arXiv preprint arXiv:2403.05541.

Xu, K., Gan, Y., & Wilson, A. (2024). Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances. Innovations in Applied Engineering and Technology, 1-12.

Xue, L., Dong, J., & Zha, Y. (2023). How does digital finance affect firm environmental, social and governance (ESG) performance?—Evidence from Chinese listed firms. Heliyon, 9(10).

Yadav, M., Jadon, P., Agarwal, B., & Rastogi, S. (2024, December). Improving Bank Performance in Rural India Through Financial Inclusion, Sustainability, and Fintech. In 2024 2nd International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES) (pp. 1-5). IEEE. Yadav, R. A., Premalatha, K. P., & Patil, S. ADVANCING SUSTAINABLE BANKING AND FINANCIAL INCLUSION IN INDIA THROUGH ESG INTEGRATION IN TECHNOLOGICAL DISRUPTIONS.

Yadav, Rashmi Akshay, K. P. Premalatha, and Shrinivas Patil. "ADVANCING SUSTAINABLE BANKING AND FINANCIAL INCLUSION IN INDIA THROUGH ESG INTEGRATION IN TECHNOLOGICAL DISRUPTIONS."

Zaroni, H. W. P., de Carvalho Miranda, R., & de Pinho, A. F. (2025). Enhancing sustainable efficiency: applications of data envelopment analysis in ESG performance measurement. International Journal of Productivity and Performance Management.

Zeqiraj, V., Sohag, K., & Hammoudeh, S. (2022). Financial inclusion in developing countries: Do quality institutions matter?. Journal of International Financial Markets, Institutions and Money, 81, 101677.

Zhang, H., Bi, Y., Jiang, W., Luo, C., Cao, S., Guo, P., & Zhang, J. (2020). Application of random forest classifier in loan default forecast. In Artificial Intelligence and Security: 6th International Conference, ICAIS 2020, Hohhot, China, July 17–20, 2020, Proceedings, Part III 6 (pp. 410-420). Springer Singapore.

Zhang, L., Luo, M., Liu, J., Li, Z., & Zheng, Q. (2020). Diverse fuzzy c-means for image clustering. Pattern Recognition Letters, 130, 275-283.

Zioło, M., Spoz, A., Bąk, I., & Oesterreich, M. (2022). Fuzzy model of sustainable development with the inclusion of financial variables. Technological and Economic Development of Economy, 28(5), 1368-1391.

Zubair, N. F., Saipul, F., Shafii, Z., & Alwi, K. (2023). ENHANCING ESG INTEGRATION IN TAKAFUL THROUGH FINANCIAL EDUCATION: CASE STUDY OF FINANCIAL LITERACY PROGRAMME IN A TAKAFUL OPERATOR. I-iECONS e-proceedings, 954-959. Zulkhibri, M. (2016). Financial inclusion, governance and institution: Evidence from developing economies. Governance and Institution: Evidence from Developing Economies (January 22, 2016).

Майорова, Т., Тимошенко, I., Урванцева, С., Черняк, Р., & Луців, П. (2023). Cluster approach in assessing the pre-war level of financial inclusion of population from different regions of Ukraine in the market of bank loan. Financial and credit activity problems of theory and practice, 1(48), 64-76.

Туре	Variable	Acronym
Е	Adjusted savings: natural resources depletion	NRSD
Е	Adjusted savings: net forest depletion	NFSD
Е	Agricultural land	AGRL
Е	Agriculture, forestry, and fishing, value added	AGVA
Е	Annual freshwater withdrawals	AFWW
Е	CO <sub>2</sub> emissions (per capita)	CO2P
Е	Coastal protection	CSTP
Е	Cooling Degree Days	CLDD
Е	Electricity from coal sources	ECOA
Е	Energy imports, net	ENIM
Е	Energy intensity	ENIN
Е	Energy use (per capita)	ENUS
Е	Food production index	FOOD
Е	Forest area (% of land area)	FRST
Е	Fossil fuel energy consumption	FOSC
Е	GHG net emissions/removals by LUCF	GHGL
Е	Heat Index 35	HIDX
Е	Heating Degree Days	HDDY
Е	Land Surface Temperature	LSTP
Е	Level of water stress	WSTR
Е	Mammal species, threatened	MAMT
Е	Methane emissions (per capita)	CH4P
Е	Nitrous oxide emissions (per capita)	N2OP
Е	PM2.5 air pollution exposure	PM25

Appendix A. List of Selected Variables in the ESG framework

Е	Water bodies with good quality	WBQL
Е	Renewable electricity output	RELE
Е	Renewable energy consumption	RENE
Е	Standardized Precipitation-Evapotranspiration Index	SPEI
Е	Terrestrial and marine protected areas (% of total)	PROT
Е	Tree Cover Loss (hectares)	TCLH
S	Access to clean fuels and technologies for cooking	CFTC
S	Access to electricity	ELEC
S	Growth rate in per capita real income	IGRW
S	Cause of death: communicable, maternal, etc.	CDTH
S	Children in employment	CHLD
S	Fertility rate, total (births per woman)	FERT
S	Gini index	GINI
S	Government expenditure on education	EDUE
S	Hospital beds (per 1,000 people)	HOSP
S	Income share held by lowest 20%	L20I
S	Individuals using the Internet	NETU
S	Labor force participation rate	LABR
S	Life expectancy at birth	LIFE
S	Literacy rate, adult total	LITR
S	Mortality rate, under-5	U5MR
S	Net migration	MIGR
S	Safely managed drinking water (% of population)	WATS
S	Safely managed sanitation (% of population)	SANS
S	Population ages 65+	POP6
S	Population density	PDEN
S	Poverty headcount ratio	POVT
S	Prevalence of overweight (% of adults)	OVWT
S	Prevalence of undernourishment	UNDT
S	Female-to-male labor force ratio	FLMR
S	School enrollment, primary (% gross)	SEPR
S	School enrollment, primary & secondary, GPI	SEGP
S	Unemployment, total (% of labor force)	UNEM
S	Unmet need for contraception (% of women 15–49)	UNMC
G	Control of Corruption: Estimate	CCOR
G	Economic and Social Rights Performance Score	ESRP
G	GDP growth (annual %)	GDPG
G	Government Effectiveness: Estimate	GOVE
G	Patent applications, residents	PATR
G	Political Stability & Absence of Violence	PSTB
G	Women in national parliaments	WPAR
G	Regulatory Quality: Estimate	REGQ

G	R&D expenditure (% of GDP)	RDSP
G	Rule of Law: Estimate	RULE
G	Scientific and technical journal articles	JART
G	Strength of legal rights index (0–12)	SLRI
G	Voice and Accountability: Estimate	VOAC

# **APPENDIX B E-Environmental ESG**

Accounting Age

Y	Account Age	The percentage of respondents who report having an account (by themselves or together with someone else) at a bank or another type of financial institution (see the definition for "financial institution account") or report personally using a mobile money service in
		or report personally using a mobile money service in the past year (see the definition for "mobile money
		account").

Х	Agricultural land	AGRL
	Agriculture, forestry, and fishing, value added	AGVA
	Food production index	FOOD
	Heat Index 35	HIDX
	Renewable energy consumption	RENE
	Terrestrial and marine protected areas (% of total)	PROT

Instruments	Access to clean fuels and technologies for cooking	CFTC
	Access to electricity	ELEC
	Government expenditure on education	EDUE
	Individuals using the Internet	NETU
	Labor force participation rate	LABR
	Life expectancy at birth	LIFE
	Literacy rate, adult total	LITR
	Mortality rate, under-5	U5MR
	Net migration	MIGR
	Safely managed drinking water (% of population)	WATS
	Safely managed sanitation (% of population)	SANS
	Population ages 65+	POP6
	Population density	PDEN
	Prevalence of overweight (% of adults)	OVWT
	Prevalence of undernourishment	UNDT
	Female-to-male labor force ratio	FLMR

School enrollment, primary (% gross)	SEPR
School enrollment, primary & secondary, GPI	SEGP
Unemployment, total (% of labor force)	UNEM
Control of Corruption: Estimate	CCOR
Economic and Social Rights Performance Score	ESRP
GDP growth (annual %)	GDPG
Government Effectiveness: Estimate	GOVE
Political Stability & Absence of Violence	PSTB
Women in national parliaments	WPAR
Regulatory Quality: Estimate	REGQ
Rule of Law: Estimate	RULE
Scientific and technical journal articles	JART
Strength of legal rights index (0–12)	SLRI
Voice and Accountability: Estimate	VOAC

	Account Age	AGRL	AGVA	FOOD	HIDX	RENE	PROT
Valid	1236	1236	1236	1236	1236	1236	1236
Missing	0	0	0	0	0	0	0
Mode	12.270	44.483	14.664	99.600	0.000	42.768	6.970
Median	39.980	44.483	11.770	101.800	2.265	42.768	6.970
Mean	42.572	44.483	14.664	103.787	8.543	42.768	10.039
Std. Error of Mean	0.642	0.531	0.293	0.401	0.419	0.791	0.229
95% CI Mean Upper	43.832	45.526	15.238	104.574	9.366	44.319	10.489
95% CI Mean Lower	41.313	43.441	14.089	103.000	7.720	41.217	9.589
Std. Deviation	22.567	18.674	10.294	14.100	14.747	27.797	8.067
95% CI Std. Dev. Upper	23.493	19.440	10.716	14.679	15.353	28.938	8.398
95% CI Std. Dev. Lower	21.711	17.965	9.903	13.566	14.188	26.743	7.761
Coefficient of variation	0.530	0.420	0.702	0.136	1.726	0.650	0.804
MAD	15.400	12.506	6.501	6.585	2.265	23.520	1.965
MAD robust	22.832	18.542	9.639	9.763	3.358	34.871	2.914
IQR	30.690	24.268	14.559	13.608	8.523	47.508	5.461
Variance	509.263	348.707	105.961	198.822	217.486	772.672	65.075
95% CI Variance Upper	551.944	377.931	114.841	215.486	235.713	837.429	70.529
95% CI Variance Lower	471.367	322.758	98.076	184.027	201.302	715.175	60.233
Skewness	0.396	-0.022	1.152	1.242	2.939	0.197	1.930
Std. Error of Skewness	0.070	0.070	0.070	0.070	0.070	0.070	0.070
Kurtosis	-0.637	-0.512	1.676	4.394	10.438	-1.070	4.304
Std. Error of Kurtosis	0.139	0.139	0.139	0.139	0.139	0.139	0.139
Kurtosis Std. Error of Kurtosis	-0.637 0.139	-0.512 0.139	1.676 0.139	4.394 0.139	10.438 0.139	-1.070 0.139	
Shapiro-Wilk	0.970	0.978	0.910	0.920	0.616	0.946	0.777
-----------------------------	--------	--------	--------	---------	---------	--------	--------
P-value of Shapiro- Wilk	< .001	< .001	< .001	< .001	< .001	< .001	< .001
Range	97.560	78.256	60.153	126.700	107.430	97.020	49.565
Minimum	1.520	3.637	0.457	56.750	0.000	0.010	0.105
Maximum	99.080	81.893	60.610	183.450	107.430	97.030	49.669
25th percentile	25.780	33.714	6.675	96.058	0.020	19.447	6.970
50th percentile	39.980	44.483	11.770	101.800	2.265	42.768	6.970
75th percentile	56.470	57.982	21.234	109.665	8.543	66.955	12.432
25th percentile	25.780	33.714	6.675	96.058	0.020	19.447	6.970
50th percentile	39.980	44.483	11.770	101.800	2.265	42.768	6.970
75th percentile	56.470	57.982	21.234	109.665	8.543	66.955	12.432
Sum	52.619	54.981	18.124	128.280	10.558	52.860	12.408

### Covariance

	Account Age	AGRL	AGVA	FOOD	HIDX	RENE	PROT
Account Age	509.263	-24.734	-107.633	30.823	-29.299	-215.118	20.796
AGRL	-24.734	348.707	18.220	24.161	30.148	8.889	-26.760
AGVA	-107.633	18.220	105.961	6.429	13.882	165.018	1.367
FOOD	30.823	24.161	6.429	198.822	21.869	17.814	23.565
HIDX	-29.299	30.148	13.882	21.869	217.486	-20.033	-6.698
RENE	-215.118	8.889	165.018	17.814	-20.033	772.672	43.691
PROT	20.796	-26.760	1.367	23.565	-6.698	43.691	65.075

# Correlations

	Account Age	AGRL	AGVA	FOOD	HIDX	RENE	PROT
Account Age	1.000	-0.059	-0.463	0.097	-0.088	-0.343	0.114
AGRL	-0.059	1.000	0.095	0.092	0.109	0.017	-0.178
AGVA	-0.463	0.095	1.000	0.044	0.091	0.577	0.016
FOOD	0.097	0.092	0.044	1.000	0.105	0.045	0.207
HIDX	-0.088	0.109	0.091	0.105	1.000	-0.049	-0.056
RENE	-0.343	0.017	0.577	0.045	-0.049	1.000	0.195
PROT	0.114	-0.178	0.016	0.207	-0.056	0.195	1.000









# APPENDIX C S-SOCIAL ESG

Y	Account Age	The percentage of respondents who report having an account (by
		themselves or together with someone else) at a bank or another type of
		financial institution (see the definition for "financial institution account")
		or report personally using a mobile money service in the past year (see the
		definition for "mobile money account").

Х	Fertility rate, total (births per woman)	FERT
	Government expenditure on education	EDUE
	Individuals using the Internet	NETU
	Labor force participation rate	LABR
	Life expectancy at birth	LIFE
	Safely managed sanitation (% of population)	SANS

Female-to-male labor force ratio	FLMR

Ζ	Adjusted savings: natural resources depletion	NRSD
	Adjusted savings: net forest depletion	NFSD
	Agricultural land	AGRL
	Agriculture, forestry, and fishing, value added	AGVA
	Annual freshwater withdrawals	AFWW
	CO <sub>2</sub> emissions (per capita)	CO2P
	Cooling Degree Days	CLDD
	Energy intensity	ENIN
	Food production index	FOOD
	Forest area (% of land area)	FRST
	Heat Index 35	HIDX
	Heating Degree Days	HDDY
	Land Surface Temperature	LSTP
	Level of water stress	WSTR
	Methane emissions (per capita)	CH4P
	Nitrous oxide emissions (per capita)	N2OP
	PM2.5 air pollution exposure	PM25
	Renewable energy consumption	RENE
	Renewable electricity output	RELE
	Standardized Precipitation-Evapotranspiration Index	SPEI
	Terrestrial and marine protected areas (% of total)	PROT
	Tree Cover Loss (hectares)	TCLH

	Account Age	FERT	EDUE	NETU	LABR	LIFE	SANS	FLMR
Valid	1236	1236	1236	1236	1236	1236	1236	1236
Missing	0	0	0	0	0	0	0	0
Mode	12.270	3.314	15.688	38.177	49.810	67.818	51.472	68.648
Median	39.980	3.314	15.688	38.177	65.837	67.818	51.472	73.129
Mean	42.572	3.314	15.688	38.177	63.657	67.818	43.166	68.648
Std. Error of Mean	0.642	0.038	0.126	0.705	0.318	0.192	0.578	0.607
95% CI Mean Upper	43.832	3.388	15.935	39.560	64.280	68.195	44.300	69.838
95% CI Mean Lower	41.313	3.239	15.441	36.794	63.033	67.441	42.031	67.457

Std. Deviation	22.567	1.338	4.426	24.783	11.174	6.762	20.338	21.336
95% CI Std. Dev. Upper	23.493	1.393	4.608	25.800	11.633	7.040	21.173	22.212
95% CI Std. Dev. Lower	21.711	1.287	4.258	23.843	10.750	6.506	19.567	20.526
Coefficient of variation	0.530	0.404	0.282	0.649	0.176	0.100	0.471	0.311
MAD	15.400	1.052	2.409	20.733	6.266	5.212	10.211	13.073
MAD robust	22.832	1.559	3.571	30.739	9.291	7.727	15.138	19.383
IQR	30.690	2.042	4.758	40.616	13.025	10.351	25.085	26.327
Variance	509.263	1.790	19.592	614.184	124.853	45.728	413.633	455.204
95% CI Variance Upper	551.944	1.940	21.234	665.658	135.317	49.560	448.299	493.354
95% CI Variance Lower	471.367	1.657	18.134	568.480	115.563	42.325	382.853	421.331
Skewness	0.396	0.627	0.330	0.342	-0.526	-0.392	-0.039	-0.817
Std. Error of Skewness	0.070	0.070	0.070	0.070	0.070	0.070	0.070	0.070
Kurtosis	-0.637	-0.234	1.172	-0.875	-0.066	-0.351	-0.256	-0.059
Std. Error of Kurtosis	0.139	0.139	0.139	0.139	0.139	0.139	0.139	0.139
Shapiro- Wilk	0.970	0.954	0.970	0.956	0.971	0.978	0.935	0.935
P-value of Shapiro-Wilk	<.001	< .001	< .001	< .001	< .001	<.001	< .001	<.001
Range	97.560	6.612	31.778	99.100	54.830	36.735	97.551	96.659
Minimum	1.520	0.837	3.228	0.900	32.742	46.692	1.928	7.345
Maximum	99.080	7.449	35.006	100.000	87.572	83.427	99.479	104.004
25th percentile	25.780	2.238	13.219	16.309	57.727	62.897	26.388	58.583
50th percentile	39.980	3.314	15.688	38.177	65.837	67.818	51.472	73.129

75th percentile	56.470	4.279	17.977	56.925	70.752	73.248	51.472	84.910
25th percentile	25.780	2.238	13.219	16.309	57.727	62.897	26.388	58.583
50th percentile	39.980	3.314	15.688	38.177	65.837	67.818	51.472	73.129
75th percentile	56.470	4.279	17.977	56.925	70.752	73.248	51.472	84.910
Sum	52.619	4.095	19.390	47.187	78.679	83.823	53.352	84.848

### Covariance

	Account Age	FERT	EDUE	NETU	LABR	LIFE	SANS	FLMR
Account Age	509.263	-16.451	1.658	351.290	31.759	71.156	110.890	15.288
FERT	-16.451	1.790	0.143	-21.777	0.709	-7.516	-9.804	6.152
EDUE	1.658	0.143	19.592	-7.095	2.370	-2.296	-5.960	4.440
NETU	351.290	-21.777	-7.095	614.184	-6.385	109.620	198.844	-96.518
LABR	31.759	0.709	2.370	-6.385	124.853	-4.654	-39.327	182.862
LIFE	71.156	-7.516	-2.296	109.620	-4.654	45.728	45.850	-53.253
SANS	110.890	-9.804	-5.960	198.844	-39.327	45.850	413.633	-106.303
FLMR	15.288	6.152	4.440	-96.518	182.862	-53.253	-106.303	455.204

## Correlation

	Account Age	FERT	EDUE	NETU	LABR	LIFE	SANS	FLMR
Account Age	1.000	-0.545	0.017	0.628	0.126	0.466	0.242	0.032
FERT	-0.545	1.000	0.024	-0.657	0.047	-0.831	-0.360	0.216
EDUE	0.017	0.024	1.000	-0.065	0.048	-0.077	-0.066	0.047
NETU	0.628	-0.657	-0.065	1.000	-0.023	0.654	0.395	-0.183
LABR	0.126	0.047	0.048	-0.023	1.000	-0.062	-0.173	0.767
LIFE	0.466	-0.831	-0.077	0.654	-0.062	1.000	0.333	-0.369
SANS	0.242	-0.360	-0.066	0.395	-0.173	0.333	1.000	-0.245
FLMR	0.032	0.216	0.047	-0.183	0.767	-0.369	-0.245	1.000







#### Appendix D G-GOVERNANCE ESG

Y	Account Age	The percentage of respondents who report having an account (by
		themselves or together with someone else) at a bank or another type of
		financial institution (see the definition for "financial institution account")
		or report personally using a mobile money service in the past year (see the
		definition for "mobile money account").

X	Control of Corruption: Estimate	CCOR
	Patent applications, residents	PATR
Regulatory Quality: Estimate		REGQ

Z	Adjusted savings: natural resources depletion	NRSD
	Adjusted savings: net forest depletion	NFSD
	Agricultural land	AGRL
	Agriculture, forestry, and fishing, value added	AGVA
	Annual freshwater withdrawals	AFWW
	CO <sub>2</sub> emissions (per capita)	CO2P
	Cooling Degree Days	CLDD
	Energy intensity	ENIN
	Food production index	FOOD
	Forest area (% of land area)	FRST
	Heat Index 35	HIDX
	Heating Degree Days	HDDY
	Land Surface Temperature	LSTP
	Level of water stress	WSTR

Methane emissions (per capita)	CH4P
Nitrous oxide emissions (per capita)	N2OP
PM2.5 air pollution exposure	PM25
Renewable energy consumption	RENE
Renewable electricity output	RELE
Standardized Precipitation-Evapotranspiration Index	SPEI
Terrestrial and marine protected areas (% of total)	PROT

	Account Age	CCOR	PATR	REGQ
Valid	1236	1236	1236	1236
Missing	0	0	0	0
Mode	12.270	-0.903	20.201	-0.797
Median	39.980	-0.642	4.680	-0.479
Mean	42.572	-0.616	20.201	-0.486
Std. Error of Mean	0.642	0.016	2.906	0.017
95% CI Mean Upper	43.832	-0.585	25.904	-0.452
95% CI Mean Lower	41.313	-0.647	14.499	-0.520
Std. Deviation	22.567	0.557	102.195	0.609
95% CI Std. Dev. Upper	23.493	0.579	106.391	0.634
95% CI Std. Dev. Lower	21.711	0.536	98.319	0.586
Coefficient of variation	0.530	-0.903	5.059	-1.253
MAD	15.400	0.341	4.679	0.423
MAD robust	22.832	0.505	6.937	0.628
IQR	30.690	0.694	20.125	0.851
Variance	509.263	0.310	$1.044 \times 10^{+10}$	0.371
95% CI Variance Upper	551.944	0.336	1.132×10 <sup>+10</sup>	0.403
95% CI Variance Lower	471.367	0.287	9.667×10 <sup>+9</sup>	0.344
Skewness	0.396	0.661	11.329	-0.024
Std. Error of Skewness	0.070	0.070	0.070	0.070
Kurtosis	-0.637	0.912	135.122	-0.119
Std. Error of Kurtosis	0.139	0.139	0.139	0.139
Shapiro-Wilk	0.970	0.974	0.123	0.998
P-value of Shapiro- Wilk	< .001	< .001	< .001	0.170
Range	97.560	3.555	$1.427 \times 10^{+6}$	3.532
Minimum	1.520	-1.937	1.000	-2.387
Maximum	99.080	1.618	1.427×10 <sup>+6</sup>	1.145
25th percentile	25.780	-1.010	76.500	-0.899
50th percentile	39.980	-0.642	4.680	-0.479

75th percentile	56.470	-0.316	20.201	-0.047
25th percentile	25.780	-1.010	76.500	-0.899
50th percentile	39.980	-0.642	4.680	-0.479
75th percentile	56.470	-0.316	20.201	-0.047
Sum	52.619	-761.869	2.497×10 <sup>+7</sup>	-601.000

#### Covariance

	Account Age	CCOR	PATR	REGQ
Account Age	509.263	5.663	373.045	6.036
CCOR	5.663	0.310	3.912	0.241
PATR	373.045	3.912	$1.044 \times 10^{+10}$	2.106
REGQ	6.036	0.241	2.106	0.371

#### Correlation

	Account Age	CCOR	PATR	REGQ
Account Age	1.000	0.451	0.162	0.439
CCOR	0.451	1.000	0.069	0.710
PATR	0.162	0.069	1.000	0.034
REGQ	0.439	0.710	0.034	1.000





Scatter Plots V



