Empirical Project

Determinants of Adoption of Digital Payment Systems in Emerging Economies

Wael Ben Hamouda
Linnea Dalvi

1 Introduction

Digital payment systems (DPS) have revolutionized financial transactions in emerging economies and drastically improved financial inclusion. While DPS were initially developed to make retail payments more efficient, they have since evolved to provide essential financial services to previously unbanked and underbanked populations. For example, Fabregas and Yokossi (2022) show that M-Pesa, a Kenyan DPS, complements traditional banking by reducing transaction costs and facilitating capital allocation, thereby fostering economic growth, particularly in urban areas. The increasing popularity of DPS among individuals and businesses have made them a core part of the economy. Frost et al. (2024) show that the uptake of fast payment systems has been rapid, particularly in emerging economies like Brazil (Pix) and India (UPI): this trend reflects the digitalization of financial services and is only expected to accelerate.

The existing literature on DPS has largely focused on descriptive analyses of system design and adoption trends. Frost et al. (2024) and the World Bank document various approaches to DPS design and measure their success primarily through transaction volume per capita.[1]

Liu et al. (2024) and Aurazo and Gasmi (2024) develop theoretical frameworks based on existing payment systems to identify key components that drive their success. Aurazo and Gasmi (2024) analyze four prominent payment systems in emerging economies—M-Pesa (Kenya), UPI (India), Pix (Brazil), and Yape (Peru)—and propose a comparative framework centered on three key dimensions: the role of payment service providers (PSP), regulators and central banks, and interoperability. They highlight the role of mobile money in promoting financial access and economic growth, something also shown by Gasmi et al. (2024), Aguilar et al. (2024), Qiu (2022), and Fabregas and Yokossi (2022). Institutional and regulatory factors play a crucial role in DPS adoption. Aurazo and Gasmi (2024) and Liu et al. (2024) emphasize the importance of central bank involvement, participation of non-bank financial institutions, and cross-border interoperability in driving fast payment adoption. Cámara and Tuesta (2014) propose a multidimensional financial inclusion index, which has been instrumental in assessing DPS performance.[2]

Despite these contributions, several gaps remain. First, while existing studies document key design choices in DPS, there is limited research on the structural determinants of their adoption across multiple economies. The role of governance, infrastructure, and PSP involvement in shaping DPS diffusion requires further empirical investigation. Second, comparative analyses have largely focused on individual case studies rather than systematically assessing multiple economies using a unified framework. Third, due to their relatively recent uptake and challenges in obtaining data, there are still gaps in empirically analyzing the structural determinants of their success. There is a need for a data-driven approach that quantitatively measures the performance of DPS and its determinants, building upon conceptual frameworks like that of Aurazo and Gasmi (2024) or Liu et al. (2024).

In this paper, we address these gaps through an analysis of the structural determinants influencing DPS adoption in emerging economies. We use transactions per capita as a proxy

[1] The World Bank's case studies on retail and wholesale payment systems are available on their fast payments resources webpage.

[2] Later in the paper we adapt this methodology to create indices to measure regulation and system interoperability.

for adoption, following the work of Frost et al. (2024) and Liu et al. (2024). Building on the three-dimensional framework proposed by Aurazo and Gasmi (2024), we compile a novel dataset covering 20 DPS in 17 countries across 4 years and quantitatively assess the influence of regulation, interoperability, and PSP involvement on DPS adoption. We estimate the effects of each dimension − role of payment service providers (PSP), interoperability, regulators and central banks − using econometric and machine learning techniques. We use Principal Component Analysis (PCA) to construct composite indices to proxy two of the three dimensions, and employ an interactive fixed effects model to estimate their effects on adoption. To test whether our choice of covariates is correct, we use a random forest model and Shapley Additive Explanation values (SHAP) to identify the most important covariates and confirm the relative impact of the three dimensions proposed by Aurazo and Gasmi (2024).

Section 2 discusses data collection, preparation, and imputation. Section 3 explains how we construct relevant indices using PCA. Section 4 outlines our empirical strategy and machine learning framework. Section 5 presents the results, and Section 6 concludes.

2 Data

2.1 Data Sources and Descriptions

We collect data from a variety of sources, which we describe below according to thematic categories. Our panel data consists of 20 DPS across 17 countries from 2020 to 2023, and the data is annual. Table 1 presents the list of countries and payment systems included in our study.

DPS Data We gathered DPS data from annual reports and official websites. Due to the diverse nature of these sources — some countries only published transaction volumes and values in annual reports (several African DPS), while others made system statistics publicly available through online visualizations (e.g., PIX[3]) — we were unable to employ web scraping techniques.

For African countries, we obtained data from the AfricaNenda Foundation. Since 2021, the foundation has worked to improve financial inclusion in Africa by making financial services accessible to all African adults by 2030. AfricaNenda publishes an annual report titled *The State of Inclusive Instant Payment Systems in Africa* (SIIPS)[4], which analyzes the evolution of Instant Payment Systems in the region. These reports outline improvements, challenges, and opportunities African countries face in their digitalization efforts. Additionally, the foundation offers technical support, shares best practices, and contributes to public policy development aimed at promoting Instant Payment Systems.

[3] We gathered data from the PIX statistics website (Brazilian Central Bank). For previous years, we used the Wayback Machine, a popular Internet initiative that allows users to visit old snapshots of websites.

[4] See the following annual reports: Africa Nenda Foundation (2022), Africa Nenda Foundation (2023), Africa Nenda Foundation (2024). They document transaction values and volumes from the year prior in the appendices of the reports.

World Bank Open Data We collected population[5], GDP[6], and Internet access[7] data from the World Bank's Open Data portal.

Banking Data Banking variables have been extracted from the Financial Access Survey (FAS), supervised by the International Monetary Fund (2023). We used the following set of variables: Institutions of commercial banks, Automated Teller Machines (ATMs), Number of ATMs per 1,000 km^2, Number of ATMs per 100,000 adults, Branches of commercial banks, Outstanding loans from commercial banks, Number of commercial bank branches per 1,000 km^2, Number of commercial bank branches per 100,000 adults, Outstanding loans from commercial banks (% of GDP), Outstanding deposits with commercial banks, and Outstanding deposits with commercial banks (% of GDP).

Regulation Data For the regulation dimension of DPS, we used data from the Worldwide Governance Indicators (WGI) database. The WGI database, developed by the World Bank (2024), measures the governance levels of countries worldwide. It provides indicators on Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Rule of Law, Control of Corruption, and Regulatory Quality. Following the two-step principal component analysis (PCA) method from Cámara and Tuesta (2014), we construct a composite index of regulation. We explain our approach in Section 3.

Role of Payment System Providers (PSP) Data assessing the role of PSP was collected from the Global System for Mobile Communication Association (GSMA) website.[8] GSMA, founded in 1995, represents the interests of over one thousand mobile sector actors and promotes the Global System for Mobile Communication (GSM) standard. We used GSMA's *Prevalence Index*, which measures the activity and accessibility of mobile money services at the country level. The index is on a scale from 0 to 100, where a higher score reflects better accessibility to mobile money services.

Interoperability Data To measure interoperability, we used GSMA's *Regulatory Index*. It contains variables such as Interoperability Solutions, International Remittances and Cross-Border Data Flow, which we use to construct an index by performing a one-stage PCA. We explain our approach in Section 3.

2.2 Factor Model for Missing Values Estimation

We were unable to gather transaction data for all years for all DPS, as some systems had not yet reported all 2022 and 2023 values at the time of writing. To avoid dropping DPS with some missing variables, we use a factor model proposed by Stock and Watson (2002) to impute them. The factor model treats the observed data as being generated by a smaller set of underlying common factors that drive variation across all DPS. The model extrapolates the missing data by using the variation of the transactions over time and across observations.

[5]See the World Bank (2025c) Population Indicator.
[6]See the World Bank (2025b) GDP Indicator in current USD.
[7]See the World Bank (2025a) for the percentage of the population with Internet access.
[8]See Global System for Mobile Communication Association (2024).

Country	**Digital Payment System (DPS)**
Brazil	Pix
Ghana	Ghana Mobile Money Interoperability (MMI)
Ghana	GhIPSS Instant Pay (GIP)
India	UPI
Kenya	Kenya mobile money
Kenya	PesaLink
Morocco	MarocPay
Madagascar	Madagascar mobile money
Mexico	CoDi
Mozambique	Sociedade Interbancaria De Mocambique (SIMO)
Mauritius	Mauritius Central Automated Switch (MauCAS)
Malawi	Natswitch
Nigeria	NIBSS Instant Payment (NIP)
Nigeria	Nigeria mobile money
Thailand	PromptPay
Tanzania	Taifa Moja
Uganda	Uganda mobile money
South Africa	Real Time Clearing (RTC)
Zambia	National Financial Switch (NFS)
Zimbabwe	Zimswitch Instant Payment Interchange Technology (ZIPIT)

Table 1: List of Selected Digital Payment Systems (DPS)

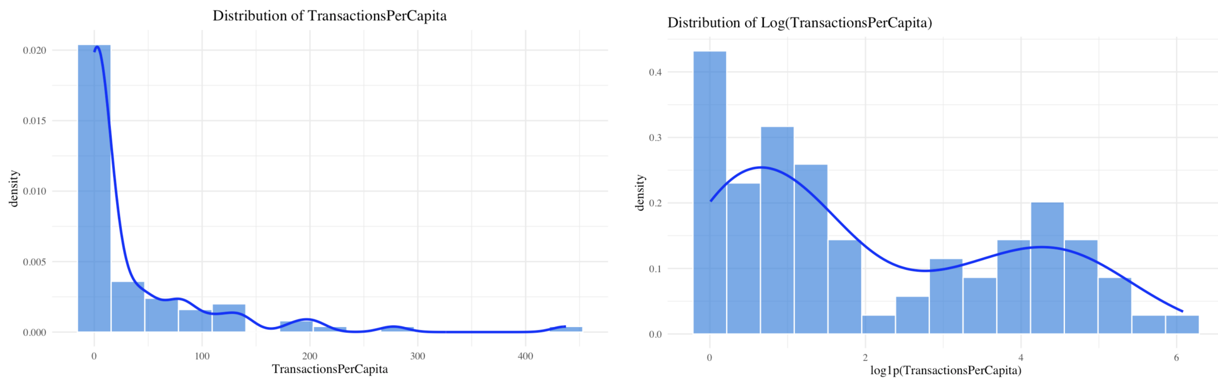
Step 1: Constructing a Data Matrix We start by creating a version of the data matrix X where missing data are replaced by zeros. We use an indicator matrix D, where each element D_{ij} is 0 if we have a missing value and 1 if not. The element-wise product $X \circ D$ keeps the observed values and replaces missing entries with zeros. Next, we apply a weighting step to ensure that the estimation focuses on the reliable parts of the data. Let W be a diagonal matrix where each diagonal element w_i represents the proportion of observed values for DPS i. We then scale the data matrix by the inverse W:

$$Z = W^{-1}(X \circ D)$$

This transformation gives more weight to DPS with less observed data, which ensures that the imputed values are informed by patterns in the all of the data, rather than risk being overwhelmed by the zeros, biasing the factor estimates toward zero. By scaling the observed entries by the inverse of the proportion of available data, we give them more influence in the factor model.

Step 2: Factor Extraction Using PCA We apply the factor model to the weighted matrix Z to extract the underlying individual-specific and time-specific trends that drive the observed variation across DPS. We do this by computing the variance-covariance matrix of Z:

Figure 1: Distribution of Transactions per Capita



$$\hat{\Omega} = \frac{1}{T} Z'Z$$

We then apply PCA to $\hat{\Sigma}$ to extract a low-dimensional representation of the data. This gives:

- Loading Factors A: An $N \times P$ matrix containing the first P eigenvectors, representing the influence of latent factors on each DPS.
- Common Factors F: A $P \times T$ matrix of time-varying factors, computed as $F_t = A'Z_t$.

Step 3: Imputing Missing Values The missing entries in X are estimated as the product of loading and common factors:

$$\hat{x}_{it}^{\text{missing}} = \hat{A}_i \hat{F}_t$$

Finally, these imputed values replace the missing elements in X, allowing us to retain all DPS in the dataset without compromising the validity of our analysis.

2.3 Exploration

We are interested in understanding the key drivers behind a DPS's transaction per capita, which proxies adoption. Figure 1 shows the distribution of transactions per capita (left) and its log-transformed counterpart (right). We can quickly see in the left-hand plot that the distribution is heavily right-skewed. Taking the logarithm of this variable results in a bimodal distribution, as shown in the right-hand plot. Therefore, there is substantial variation in the adoption of our DPS from which we aim to infer factors of success.

We can also use a simple linear regression to get an idea about potential patterns in our data. Table 2 suggests that the role of PSP and the ATMs per 100,000 adults play a statistically significant positive role in determining DPS adoption.

Many coefficients in Table 2 are statistically insignificant, which could be a result of our small sample size, the assumption of linearity, or multicollinearity among covariates that may bias the estimates. Additionally, since we apply a simple OLS regression to panel data without accounting for the structure of the data, our errors are likely correlated which may lead to biased standard errors and coefficient estimates. We describe our methodology for

Table 2: Simple Linear Regression Results

	Dependent variable:
	Transactions per Capita
Role of PSP Index	0.897***
	(0.221)
Regulation Index	−1.450
	(1.300)
Interoperability Index	−0.326
	(0.421)
ATMs per 100,000 Adults	1.852**
	(0.856)
Bank Branches per 1,000 km^2	−0.248
	(1.072)
Bank Deposits (% of GDP)	0.769
	(0.571)
Internet Access	−0.772
	(0.497)
Africa (Dummy)	−25.961
	(44.481)
GDP per Capita	−0.007
	(0.005)
Years Since Launch	−0.631
	(2.363)
Constant	76.026
	(48.568)
Observations	80
R^2	0.375
Adjusted R^2	0.285
Residual Std. Error	61.096 (df = 69)
F Statistic	4.142*** (df = 10; 69)
Note:	*p<0.1; **p<0.05; ***p<0.01

constructing composite indices and selecting appropriate empirical models in Sections 3 and 4, respectively, below.

3 Index Construction

3.1 Regulation Index Calculation: Two-Stage PCA

As described in Section 2, we use the WGI database developed by the World Bank (2024) to construct a composite index as a measure of regulation. Following Cámara and Tuesta (2014), we build a multidimensional index by splitting the relevant variables into two distinct subsets: Stability and Governance. We explain our variable grouping and each of the two stages, and conclude the subsection by explaining our rationale for using a two-stage PCA over a one-stage PCA.

Indicator Grouping We begin by categorizing six WGI indicators into two conceptually distinct dimensions (World Bank, 2024):

- **Stability:**
  - **Political stability and absence of violence/terrorism:** the likelihood of political instability and/or politically-motivated violence, including terrorism.
  - **Voice and accountability:** the extent to which a country's citizens are able to participate in selecting their government, as well as freedoms of expression, association, and the press.
- **Governance:**
  - **Government effectiveness:** the quality of public services and policy implementation, including the independence and competence of the civil service.
  - **Regulatory quality:** the government's ability to formulate and implement sound policies and regulations that promote private sector development.
  - **Rule of law:** the degree of confidence in and adherence to society's rules, including contract enforcement, property rights, and legal institutions.
  - **Control of corruption:** the extent to which public power is used for private gain, including both petty and grand corruption.

This grouping ensures that the resulting sub-indices, $Stability_{it}$ and $Governance_{it}$, capture distinct dimensions of regulation that we can interpret.

Stage 1: Constructing the Sub-Indices We apply PCA separately to each group to extract all latent factors (principal components) that represents each dimension. The resulting sub-indices are computed as:

$$\begin{aligned} \text{Stability}_{it} &= \beta_1 \cdot \text{PoliticalStability}_{it} + \beta_2 \cdot \text{VoiceAccountability}_{it} \\ \text{Governance}_{it} &= \gamma_1 \cdot \text{ControlCorruption}_{it} + \gamma_2 \cdot \text{RuleLaw}_{it} + \gamma_3 \cdot \text{RegulQuality}_{it} + \gamma_4 \cdot \text{GovEffectiveness}_{it} \end{aligned}$$

Here, β_j and γ_j are the component loadings from the first principal component for each group, capturing the contribution of each variable to the underlying dimension.

Stage 2: Aggregating the Sub-Indices To form the overall regulation index, we perform PCA again on the two sub-indices, Stability_{it} and Governance_{it}. The final *Regulation Index* is a weighted average of the two sub-indices:

$$\text{Regulation}_{it} = \psi_1 \cdot \text{Stability}_{it} + \psi_2 \cdot \text{Governance}_{it}$$

The weights ψ_1 and ψ_2 reflect the relative contribution of each dimension to the overall variation in regulation (i.e., the eigenvalues). Formally:

$$\text{Regulation}_{it} = \frac{\lambda_1 \cdot \text{PC}_{1,it} + \lambda_2 \cdot \text{PC}_{2,it}}{\lambda_1 + \lambda_2}$$

where λ_j is the eigenvalue corresponding to the j-th principal component of the second-stage PCA, and $\text{PC}_{j,it}$ is the score of observation i on component j.

Why Two-Stage PCA? We use a two-stage PCA rather than a one-stage approach because it better reflects the structure of our data. In a one-stage PCA, all variables are analyzed together, which can obscure the underlying dimensions and lead to biased results, especially when variables reflect conceptually distinct aspects of regulation. The two-stage PCA solves this by first grouping variables into distinct categories based on how closely related they are (correlation), and then performing PCA on those groups. Doing so allows us to understand the relative impact of each dimension.

The resulting index is on a scale from 0 to 100, where a lower score indicates less regulation in a given country in a given year.

3.2 Interoperability Index Calculation: One-Stage PCA

We use three primary variables from Global System for Mobile Communication Association (2024) to construct the Interoperability Index: interoperability solutions, international remittances, and cross-border data flow. Given that we only have three components (compared to the six variables used to construct the regulation index above) which reflect similar measures, a one-step PCA is sufficient.

$$\text{Interoperability}_i = \frac{\sum_{j=1}^{3} \lambda_j^{(Int)} P_{ji}^{(Int)}}{\sum_{j=1}^{3} \lambda_j^{(Int)}} \tag{1}$$

Here, $\lambda_j^{(Int)}$ denotes the eigenvalue corresponding to the j-th principal component, and $P_{ji}^{(Int)}$ is the component score for country i on component j. This index, which is on a scale from 0 to 100, captures the state of system interoperability for each country, where a higher score means more interoperability.[9]

[9] Due to GSMA's restricted access and the high cost associated with obtaining additional years, we rely on data from 2024 only. We therefore assume that interoperability is time-invariant. While this is a strong assumption, we believe it is justified by the fact that interoperability is typically embedded in the infrastructure at the time the launch of a DPS and does not change significantly year to year. We discuss the impact of this assumption on our estimation in Section 4.1.

4 Methodology

4.1 Estimation Using the Interactive Fixed Effects Model

To ensure we account for any unobserved heterogeneity across individuals and over time, we apply an interactive fixed effects model as proposed by Bai (2009). The model is given by:

$$Y_{it} = A_i F_t + X'_{it}\beta + \varepsilon_{it} \tag{2}$$

where:

- A_i are individual-specific loading factors (i.e., coefficients showing how much each DPS is influenced by country-specific trends),
- F_t are common time-varying factors,
- X_{it} are observed variables, and
- ε_{it} is the error term.

Estimation Procedure

1. **Initial Estimation:** Estimate $\beta^{(0)}$ using pooled OLS:
$$Y_{it} = X'_{it}\beta^{(0)} + \hat{\varepsilon}_{it}$$
2. **PCA on Residuals:** Compute residuals $u_{it} = Y_{it} - X'_{it}\beta^{(0)}$ and apply PCA to extract loading factors $\hat{A}_i^{(0)}$ and common factors $\hat{F}_t^{(0)}$ from the residual matrix:
$$u_{it} = \hat{A}_i^{(0)}\hat{F}_t^{(0)}$$
3. **Update β:** Compute adjusted dependent variable:
$$\hat{v}_{it} = Y_{it} - \hat{A}_i^{(0)}\hat{F}_t^{(0)}$$
and re-estimate:
$$\hat{v}_{it} = X'_{it}\beta^{(1)}$$
4. **Iterate:** Repeat the estimation of β, A_i, and F_t until the estimates converge. They converge when the relative change between iterations falls below a tolerance threshold:
$$\max\left(100 \times \left|\frac{A_i^{(k)}F_t^{(k)} - A_i^{(k-1)}F_t^{(k-1)}}{A_i^{(k-1)}F_t^{(k-1)}}\right|, 100 \times \left|\frac{\beta^{(k)} - \beta^{(k-1)}}{\beta^{(k-1)}}\right|\right) < \text{Tolerance Threshold}$$

Unlike standard fixed effects models which erase time-invariant variables, the interactive fixed effects model allows us to estimate such variables. This is because the model accounts for unobserved heterogeneity using latent time-varying factors interacted with unit-specific loadings rather than demeaning (i.e., performing a within transformation) as in a traditional fixed effects model. As a result, the impact of time-invariant variables – like interoperability, in our case – can still be estimated, given that they are not perfectly collinear with the structure of the unobserved factors. This makes the interactive fixed effects model ideal for our analysis.

4.2 Machine Learning Methods

PCA is a commonly used technique in machine learning, but it is primarily a dimensionality reduction technique that helps identify underlying patterns in the data. However, PCA has some limitations: it assumes linear relationships between covariates and adoption, and it does not capture interactions between variables. Consequently, to better understand variable interactions and nonlinear effects, we complement our econometric model with a random forest model. We interpret the results using SHAP (SHapley Additive exPlanations) values and present our findings in Section 5.

In this section, we explain our machine learning framework. We first compare various classification algorithms to determine which performs the best in the classification task. Next, we use the built-in feature importances of our chosen model − random forest − to get a brief overview of the relative importance of the three dimensions proposed by Aurazo and Gasmi (2024). The feature importances serve as a launch pad into our results, which we present in Section 5.

4.2.1 Model Selection

We define success by adoption, which we measure by transactions per capita. We use two key approaches: first, predicting the level of DPS adoption's prediction based on a set of input features; and second, feature importance, which helps identify the relative impact of different variables (e.g., regulation, interoperability, and the role of PSP) on DPS adoption. In doing so, we can pinpoint the key structural determinants of DPS success and use our findings to complement our econometric approach.

To simplify, we define ”success” as having an adoption level above the median, thereby transforming the transactions per capita variable into a binary variable (successful or not). This reframes our research question as a classification task, allowing us to use classification algorithms to assess the impact of various factors. Many classification algorithms, such as random forest, adaptive boosting, and gradient boosting, have built-in feature importance extraction in R. From this we can get a better understanding of which structural determinants have the most impact on DPS adoption, thus providing some guidance on which areas policymakers can focus on in order to ensure the success of their DPS.

Table 3: Predictive Accuracy of Different Classification Algorithms

Model	Accuracy
Naive Bayes	0.700
Random Forest	0.800
Adaptive Boosting	0.700
KNN (Floor)	0.650
KNN (Ceiling)	0.650
SVM	0.750
Neural Network	0.800

We compare seven different classification algorithms: Naive Bayes, Random Forest, Adaptive

Boosting, K-Nearest Neighbours (KNN) with floor and ceiling models[10], Support Vector Machines (SVM), and Neural Networks.[11] We find that the random forest and neural network perform the best of all the models (see Table 3). We choose random forest for several reasons. First, neural networks operate more as a "black box" model. The amount of layers and non-linear transformations involved in training make it difficult to trace exactly how individual inputs influence the final prediction, whereas random forests provide more direct interpretability through measures like feature importance. Second, neural networks are more likely to overfit in small samples, while random forest is more robust, making it better suited for a dataset as small as ours. Lastly, neural networks are more complex and require hyperparameter tuning, whereas random forests are simpler and more straightforward to use.

4.2.2 Random Forest

In addition to its predictive accuracy compared to other models, there are several other benefits to using random forest instead of other classification algorithms. Random forest models can handle large datasets with many features, which is critical in our study, where we have 38 variable columns. Although we narrow down our set of features to 10 to narrow our focus, the flexibility of random forest eliminates the need for manual feature selection, because the model can automatically rank the importance of variables based on their contribution to the prediction. Furthermore, random forest is more robust to collinearity, which is important given the presence of categorical variables like continent dummies and indices for the role of PSP.

Our primary motivation for using random forest is its ability to rank the importance of each variable (Breiman (2001)). In a random forest for classification, at each point in the decision tree, the algorithm chooses a feature on which to split the data. The goal of the decision tree is to pick the feature that best separates the data into different classes at each split, thereby making it as pure as possible. The feature importance tells us which features are most helpful in making those splits. The more a feature helps in reducing the impurity in the data, the more important it is. The measure of impurity is given by the Gini index. A Gini index of 0 indicates that all data points in a node are of the same class, making it "pure"; likewise, 0.5 indicates an impure node (in binary classification models). Impurity measures the likelihood of a new data point being mislabeled. The feature importance seeks to identify which features decrease the impurity of the nodes. We are interested in understanding the influence of the three dimensions proposed by Aurazo and Gasmi (2024) − role of PSP, regulation, and interoperability − on adoption, both relative to one another and overall. The results from our random forest model and its associated feature importances can tell us which features are the most important in correctly determining whether or not a DPS is successful.

To better understand the feature importances identified by the random forest, we use SHAP (SHapley Additive exPlanations), a framework proposed by Lundberg and Lee (2017) to better understand the role of each variable in the individual and overall predictions of the model (Louhichi et al. (2023)). Each observation gets its own set of SHAP values that

[10]The floor and ceiling models refer to using either the floor or ceiling functions to round the k-value (i.e., the number of nearest neighbours) to the nearest integer.

[11]Gradient boosting was considered but was infeasible given our relatively small dataset.

reflect both the magnitude and direction of a feature's influence on the prediction[12]. The directionality of each feature's SHAP value indicates the role of each feature in drawing the prediction away from the mean of transactions per capita. To use SHAP effectively, we train a second random forest model on the continuous version of our transactions per capita instead of the binary classification version. This serves as a robustness check to assess whether the same structural determinants emerge as most influential when predicting DPS success on a continuous scale.

5 Results

This section presents the findings from our two main analyses. In Section 5.1 we compare three different econometric models and present the results of the most compelling specification. In Section 5.2 we present the results of our random forest model, drawing on both built-in feature importances and SHAP values to corroborate the findings from our econometric model and understand how different variables influence adoption when a functional form is not specified. The key takeaways from both approaches are summarized in the conclusion (Section 6).

5.1 Econometric Analysis

We estimate three different models using different approaches:

- **Model 1:** Includes banking variables and each of the three dimensions proposed by Aurazo and Gasmi (2024). We also interact each dimension with the Africa dummy variable.
- **Model 2:** We build on Model 1 by including Internet accessibility and interacting it with the Africa dummy variable. [13]
- **Model 3:** We build on Model 2 by replacing individual banking variables with their first principal components extracted from PCA. As in Model 2, we include interaction terms between the three structural dimensions (Aurazo and Gasmi, 2024) and Internet access with the Africa dummy variable.

There are some assumptions that we must make when estimating the three models above, in no particular order of importance. First, we assume that interoperability is time-invariant; this was briefly mentioned in Sections 3 and 4.1. Second, we allow for autocorrelation within each DPS over time by clustering standard errors at the DPS level and using a heteroskedasticity-consistent HC1 variance-covariance matrix. Third, we assume that all explanatory variables are exogenous; that is, we assume there is no reverse causality or omitted variable bias affecting the relationship between the covariates and the dependent variable. Finally, we assume that the missing values in our original dataset are random (that is, the missingness; not the actual values) and therefore do not bias our results. Due to missing observations in certain years and countries, particularly for Internet access in 2023, we work

[12]For example, regulation (measured by our regulation index) increases adoption overall, but for DPS x, regulation decreased adoption.

[13]Unfortunately, incorporating this variable requires us to drop the 2023 observations due to poor data quality and estimation of Internet access rates in that year.

with an unbalanced panel. To address this we imputed the missing values using a factor model, as described in Section 2.2).

5.1.1 Model 1

$$\begin{aligned}\frac{Transactions}{Population}_{it} &= A_iF_t + \beta_0 + \beta_1 RolePSP_{it} + \beta_2 Regulation_{it} + \beta_3 Interoperability_i \\ &+ \beta_4 RolePSP_{it} \times Africa_i + \beta_5 Regulation_{it} \times Africa_i \\ &+ \beta_6 Interoperability_i \times Africa_i \\ &+ \gamma X_{it} + \alpha Africa_i + \epsilon_{it}\end{aligned}$$

In this specification, X_{it} denotes a matrix of selected banking-related controls, including the number of ATMs per 100,000 adults (capturing access to cash withdrawal infrastructure) and the number of commercial bank branches per 100,000 adults (reflecting general banking accessibility). The model also includes interaction terms between the three structural dimensions from Aurazo and Gasmi (2024) and a dummy variable indicating whether a country is in Africa, thereby capturing potential regional heterogeneity in the effectiveness of these determinants.

This model serves as a preliminary specification to inform our later, more robust econometric models. Here, our goal is to explore the direct influence of selected banking variables while identifying potential differences in how structural factors affect adoption across African and non-African contexts.

Table 4: Model 1.

Variable	**Estimate**	**Std. Err.**	**P-value**	**Signif.**
(Intercept)	-1026.1798	72.7304	$< 2.2 \times 10^{-16}$	***
Role of PSP Index	1.8334	1.5390	0.2376	
Role of PSP Index:Africa	-1.0065	1.4672	0.4950	
Regulation Index	14.9574	3.7037	1.36×10^{-4}	***
Regulation Index:Africa	-17.1458	4.0111	5.94×10^{-5}	***
Interoperability Index	10.3995	11.2936	0.3603	
Interoperability Index:Africa	-11.0275	11.2790	0.3316	
ATMs per 100k adults	1.2814	0.6669	0.0588	.
Bank branches per 100k adults	0.6014	2.7896	0.8299	
Africa	1106.5928	104.9747	4.26×10^{-16}	***
Info.	**DPS:** 20, **Periods:** 4, $\mathbf{R^2}$**:** 0.4060 (**adj.** 0.3296)			

In Table 4, we can clearly see that the impact of each variable depends on whether or not the country is African. However, given that we only have 80 observations, we should interpret the results cautiously.

The regulation index is the only statistically significant variable of the three key dimensions proposed by Aurazo and Gasmi (2024). In non-African countries, the effect is strongly positive (14.96); in African countries, it is strongly negative (-17.15). These effects are anything but trivial. Given that the average transactions per capita is approximately 28.44, the magnitude of the impact is substantial. This implies that regulation is harmful for African countries. One possible explanation is that regulatory frameworks pose administrative burdens requirements that hinder adoption in an African context, whereas in non-African countries, a strong regulatory environment may on the contrary serve to build trust and foster more DPS uptake. The interoperability and role of PSP indices are insignificant for countries in- and outside of Africa. Note that the signs of the estimates are opposite for each dimension: what helps non-African countries seems to hurt African ones. The adjusted R^2 tells us that with this specification we capture about 33% of the variation in adoption.

Model 1 builds a foundation for understanding structural differences across regions. The results suggest that there is substantial regional heterogeneity to the point where key determinants of DPS adoption may have differing or even opposite effects depending on whether or not the system is African. We take these insights into account in subsequent models, where we combine regional heterogeneity with access to Internet.

5.1.2 Model 2

In this second model, we evaluate how incorporating Internet access improves the estimates and the overall explanatory power of the model. This new model improves the robustness and quality of estimation compared to Model 1.

$$
\begin{aligned}
\frac{Transactions}{Population}_{it} = {} & A_iF_t + \beta_0 + \beta_1 RolePSP_{it} + \beta_2 Regulation_{it} + \beta_3 Interoperability_i + \beta_4 Internet \\
& + \beta_5 RolePSP_{it} \times Africa_i + \beta_6 Regulation_{it} \times Africa_i \\
& + \beta_7 Interoperability_i \times Africa_i + \beta_8 Internet \times Africa \\
& + \gamma X_{it} + \alpha Africa_i + \epsilon_{it}
\end{aligned}
$$

Here, X includes the same banking control variables as in Model 1.

In Model 1, we chose to omit Internet access to avoid reducing the sample size, as it would have required dropping all 2023 values due to poor imputations. Since our factor model for missing data relies on sufficient variation across individuals and periods, including a weakly estimated variable for 2023 would bias the results. This presented a trade-off between preserving statistical power and avoiding omitted variable bias. However, as shown in Table 5, we see that incorporating Internet access − even when dropping an entire year of observations − significantly improves the model's fit. We see the adjusted R^2 jump from 32.96% in Model 1 to 50.79%. Omitting Internet poses a much greater cost than including it, even at the expense of dropping one year.

Table 5: Model 2.

Variable	Estimate	Std. Error	P-value	Signif.
(Intercept)	-1068.5888	112.5713	1.35×10^{-12}	***
Role of PSP Index	2.4389	0.4724	4.63×10^{-6}	***
Role of PSP Index:Africa	-1.8324	0.4308	9.68×10^{-5}	***
Regulation Index	12.6504	1.8043	7.07×10^{-9}	***
Regulation Index:Africa	-13.1516	1.8607	5.78×10^{-9}	***
Interoperability Index	9.3851	2.5968	0.0007	***
Interoperability Index:Africa	-9.5932	2.7712	0.0011	**
Internet Access	2.2255	0.8055	0.0081	**
Internet Access:Africa	-3.3939	0.7709	5.96×10^{-5}	***
ATMs per 100k adults	0.6067	0.2991	0.0481	*
Bank branches per 100k adults	3.4862	1.5349	0.0276	*
Africa	1104.1661	120.1878	3.77×10^{-12}	***
Info.	**DPS:** 20, **Periods:** 3, $\mathbf{R^2}$**:** 0.5997 (**adj.** 0.5079)			

Table 5 shows that with the inclusion of Internet access, all variables are statistically significant to some degree. All three dimensions proposed by Aurazo and Gasmi (2024) are highly statistically significant, with the exception of interoperability in African countries. As in Model 1, we see the same opposite effects for each variable in African and non-African countries. In fact, we see that the Africa dummy variable on its own is extremely statistically significant, with a p-value below 3.77×10^{-12}.

5.1.3 Model 3

In the third and final model, we capture more variation in banking variables without overfitting through PCA, which reduces the dimensionality. We construct principal components PC1 and PC2 using four variables:

- ATMs per 100,000 adults
- ATMs per 1,000 km^2
- Commercial bank branches per 100,000 adults
- Commercial bank branches per 1,000 km^2

We include PC1 and PC2 alongside our three core dimensions and Internet access, and their interactions with the Africa dummy:

$$\begin{aligned}\frac{Transactions}{Population}_{it} &= A_iF_t + \beta_0 + \beta_1 RolePSP_{it} + \beta_2 Regulation_{it} + \beta_3 Interoperability_i + \beta_4 Internet \\ &+ \beta_5 RolePSP_{it} \times Africa_i + \beta_6 Regulation_{it} \times Africa_i \\ &+ \beta_7 Interoperability_i \times Africa_i + \beta_8 Internet \times Africa \\ &+ \gamma_1 PC1_{it} + \gamma_2 PC2_{it} + \alpha Africa_i + \epsilon_{it}\end{aligned}$$

Model 3 aims to incorporate banking variables without overfitting. Initially, we considered conducting a PCA on all 11 banking variables (please see the Appendix for a list of all variables in our dataset) and testing different combinations of principal components. However, we ultimately opted for a more conservative approach using only the four core infrastructure variables listed above for simplicity.

Table 6 shows the results of the estimation. While Model 3 did slightly improve the predictive power (adjusted R^2 increased by about 1%), one coefficient lost its statistical significance. For example, the Internet access variable for non-African countries lost its significance at even the 5% level. This result suggests that indiscriminately incorporating banking variables into PCA may dilute explanatory power if irrelevant indicators are included.

Table 6: Model 3.

Variable	**Estimate**	**Std. Error**	**P-value**	**Signif.**
(Intercept)	-961.2214	113.4949	4.34×10^{-11}	***
Role of PSP Index	2.6658	0.4460	2.72×10^{-7}	***
Role of PSP Index:Africa	-2.0365	0.4484	3.76×10^{-5}	***
Regulation Index	13.0019	1.7142	9.40×10^{-10}	***
Regulation Index:Africa	-13.5126	1.8587	2.84×10^{-9}	***
Interoperability Index	6.9779	2.3373	0.0044	**
Interoperability Index:Africa	-7.1732	2.1991	0.0020	**
Internet Access	1.5702	0.9209	0.0946	.
Internet Access:Africa	-2.8141	0.7112	0.0002	***
PC1	20.5957	7.6476	0.0097	**
PC2	-27.1039	10.0229	0.0094	**
Africa	1044.7134	102.7011	1.44×10^{-13}	***
Info.	**DPS:** 20, **Periods:** 3, $\mathbf{R^2}$: 0.6055 (**adj.** 0.5151)			

As shown in Table 6, all coefficients are statistically significant at the 5% level or better—except for the main Internet access term— which is marginally significant at the 10% level. Compared to Model 2, most coefficients increased in magnitude by roughly 0.2 points. The Internet-related coefficients increased by 0.7 points, despite the lost significance for the non-interacted term. Overall, the inclusion of PC1 and PC2 did not lead to a major improvement in fit.

5.1.4 Choosing the Best Econometric Model

The three models above experimented with how to incorporate banking variables and Internet access into an interactive fixed effects model, including through dimensionality reduction via PCA. Model 1 provided a baseline and Model 3 offered the broadest specification; we ultimately choose Model 2 as the most suitable for economic interpretation and actionable policy recommendations.

Model 2 includes two key banking infrastructure variables (ATMs per 100,000 adults and commercial bank branches per 100,000 adults), as well as Internet access and our three main institutional dimensions (Aurazo and Gasmi, 2024). We interacted each of these variables with the Africa dummy variable to capture any regional heterogeneity.

Of the three key dimensions proposed by Aurazo and Gasmi (2024), Model 2 ranks regulation as the most influential and the role of PSP as the least.[14] For each dimension, a higher score helps non-African countries but hurts African ones with small magnitude.

[14]Nevertheless, the role of PSP has a greater statistical significance than interoperability. Since all three dimensions are statistically significant at the traditional 5% level, we rank their influence by the magnitude of their impact.

- **Role of PSP:** In non-African countries, increased involvement of PSP is associated with moderate increase in adoption. A 1-point increase in the PSP index (which is on a scale of 0 to 100) raises DPS usage by approximately 2.44 transactions per capita annually. However in African countries, the effect is largely mitigated: a 1-point increase in PSP involvement increases annual transactions per capita by 0.61.

- **Regulation:** Regulation is estimated to be the most influential dimension. In non-African countries, a 1-point increase in the regulation index corresponds to an increase of 12.65 transactions per capita. However, in African countries, a 1-point increase in regulation results in a decrease of 0.50 transactions per capita annually. This suggests that while a strong regulatory environment supports adoption in non-African countries, in Africa this may come as a burden that hinders growth, potentially due to weaker institutions or inconsistent enforcement.

- **Interoperability:** In non-African countries, a 1-point improvement in interoperability leads to an increase of 9.39 transactions per capita. In African countries, however, the effect is slightly negative at –0.3. This implies that while interoperability enhances system efficiency and user convenience in more developed markets, its impact is limited—or even counterproductive—when digital infrastructure is weak or fragmented.

- **Internet Access:** While it is not one of the three dimensions, Internet access plays a significant role. For non-African countries, more widespread Internet access increases usage by 2.23 transactions per capita, whereas in African countries, the effect is about −1.16. This is a bit of a puzzling result, as we would expect better access to Internet to improve the take-up of a *digital* payment system.

Model 2 highlights the opposite effects of each dimension when a system operates inside of outside of Africa. We see clearly that the same drivers of adoption in one country may have completely different implications in another. In the next subsection, we use a machine learning model to complement our econometric model and see whether alternative methods also give the same results.

5.2 Machine Learning Analysis

5.2.1 Built-in Feature Importance from Random Forest

We train our random forest model using all available data from 2020 to 2022 to fully leverage the information at hand and understand how each variable influences predictions, even though this approach does not allow for out-of-sample prediction.[15] The feature importances serve as a descriptive tool to enhance our understanding of the key determinants of DPS adoption.

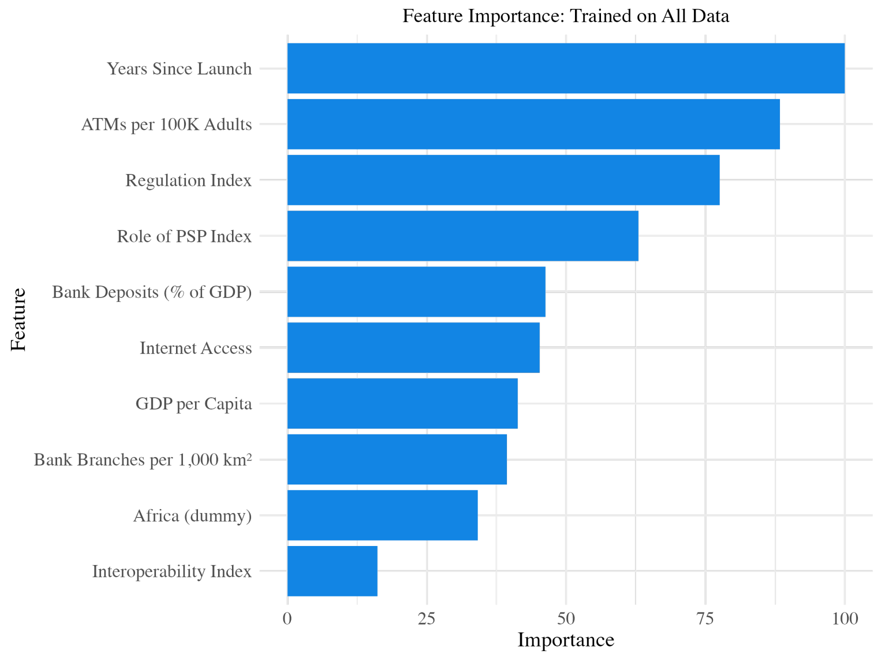
Figure 2 shows the importance of each feature ranked from most to least influential.[16] The most significant determinant of DPS success seems to be the years since launch, closely followed by the number of ATMs per 100,000 adults and the regulation index.[17] The results are somewhat consistent

[15] Because we include Internet as a variable, we omit the 2023 observations as discussed in the specification of Model 2 (Section 5.1.2).

[16] The numbers represent the relative importance of each feature in predicting the success of DPS, with 100 being the most influential feature. All other values are expressed as a percentage of this maximum importance, indicating how much each feature contributes in comparison to the most significant one. For example, a feature with an importance score of 50 means it is 50% as important as the feature with the highest score. The importances are not directional.

[17] Years since launch is not a particularly useful determinant in a policy sense, as it is not something that can act as a policy lever when designing a payment system. The direction of the impact is also not clear: longer-established systems may be successful due to having gained traction over years, or they may be less successful because of outdated technology.

Figure 2: Most Influential Features in Random Forest



with our hypotheses during the exploratory linear regression analysis (see Table 2), where the role of PSP and the ATMs per 100,000 adults emerged as the only statistically significant variables. The ranking is also consistent with the results of our econometric model, with regulation having the most influence over the success of a DPS. The role of PSP ranks second (in relative importance to the three key dimensions) while the interoperability index ranks as the least important. While this may seem contradictory, recall that Table 5 showed that of the three dimensions of interest, the interoperability index had the lowest statistical significance. The most striking difference between the machine learning approach and the econometric approach is the relative irrelevance of the Africa dummy.[18]

Again, the direction of the influences suggested by Figure 2 is not known: it could be a positive effect, in the sense that e.g. better regulation creates a better service; or, it could be a negative effect, where excessive regulations may hinder DPS growth. To understand whether these features have a positive or negative impact on DPS success, we use SHAP values.

5.2.2 SHAP Values for Feature Importance

While feature importances tell us which variables have the most impact, SHAP values tell us how each variable influences the individual predictions, specifically as a deviation from the mean transactions per capita. SHAP values show the marginal contribution of each feature to a specific prediction. A positive SHAP value means that increasing that variable will increase the prediction, while a negative SHAP value would mean decreasing the prediction. Each observation (a DPS for a given year) is assigned a set of SHAP values, which allows us to analyze the direction and magnitude of each feature's influence in that specific prediction.

For this analysis, we train a random forest model on all available data to predict transactions

[18]We are not sure of the reasons for this. The machine learning model may be doing a better job at handling potential collinearity between variables compared to the econometric model, or it may be producing odd estimates due to being fed a small dataset.

per capita as a continuous outcome, rather than using the discretized version employed in the earlier classification models (see Section 4.2.1). This allows us to use SHAP values to examine the contribution of each feature to the exact level of DPS success, making it more comparable to the results of our econometric model presented in Section 5.1.

Figure 3 shows the mean absolute SHAP values and the mean SHAP values. While SHAP values for individual observations show how much each feature influences a specific prediction, mean absolute SHAP values tell us how important a feature is on average across all predictions. In this sense, the SHAP analysis serves as a robustness check against the built-in feature importances from the random forest classification model, because it allows us us to assess whether the same features are important when predicting DPS success as a continuous variable.

To get an understanding of the direction of each feature's influence, we can look at mean SHAP values, which are the average (signed) SHAP contributions of each feature across all data points. This allows us to get a sense of whether a variable helps or hurts DPS adoption.

From Figure 3, we see that the role of PSP is the most influential feature, closely followed by the regulation index and the number of ATMs per 100,000 adults. The results are slightly different from that of the discretized random forest model, but there are many complementary findings. Regulatory quality and ATMs still play a critical role, and the interoperability index and Africa dummy play a relatively small role. The role of PSP emerges as the most influential in the continuous model.

The bottom panel in Figure 3 indicates the average directionality across all observations. Here, we see that the role of PSP index has a very strong positive impact on DPS adoption, whereas regulation index has a weakly negative effect and the interoperability index a near-to-negligible positive effect. ATMs per 100,000 adults also indicates a strong positive impact, suggesting that physical banking presence may increase the success of a DPS. The regulation index may be small here since, in calculating the mean SHAP values, different signs of importance may cancel each other out. This suggests that both too little and too much regulation can hurt DPS adoption, meaning that increasing the regulation index may in some instances increase transactions per capita, and lower it in other cases. Therefore, when calculating the mean, some of these values may cancel out and result in a mean SHAP value with a relatively small magnitude.[19] Overall the findings compliment the built-in random forest model importances.

6 Conclusion

This paper set out to identify the structural determinants of DPS adoption in emerging economies using a data-driven approach. We constructed a novel dataset of 20 DPS across 17 emerging economies over a four-year period. Building on existing theoretical frameworks and empirical case studies, our goal was to quantitatively assess how factors such as regulation, interoperability, and PSP involvement shape the success of these systems (three key dimensions identified by Aurazo and Gasmi (2024)). We used an interactive fixed effects model and found that across all key dimensions (e.g., the role of PSPs), higher scores tended to help DPS adoption in non-African countries but hurt or mitigated it in African countries. We found that regulation was the most influential of the three dimensions, and the role of PSP the least. We also found that incorporating Internet access as a covariate significantly improved the model's predictive power, even at the expense of dropping the last year of observations. To corroborate our findings, we used a random forest classification

[19]This may indicate why the Africa dummy has such a small impact in the bottom panel of Figure 3; the existence of some extremely successful DPS like Kenya mobile money (135 transactions per capita) existing alongside a very underutilized DPS like MarocPay (0 transactions per capita) may lead to an overall negligible effect for whether or not a DPS is African on whether or not it succeeds.

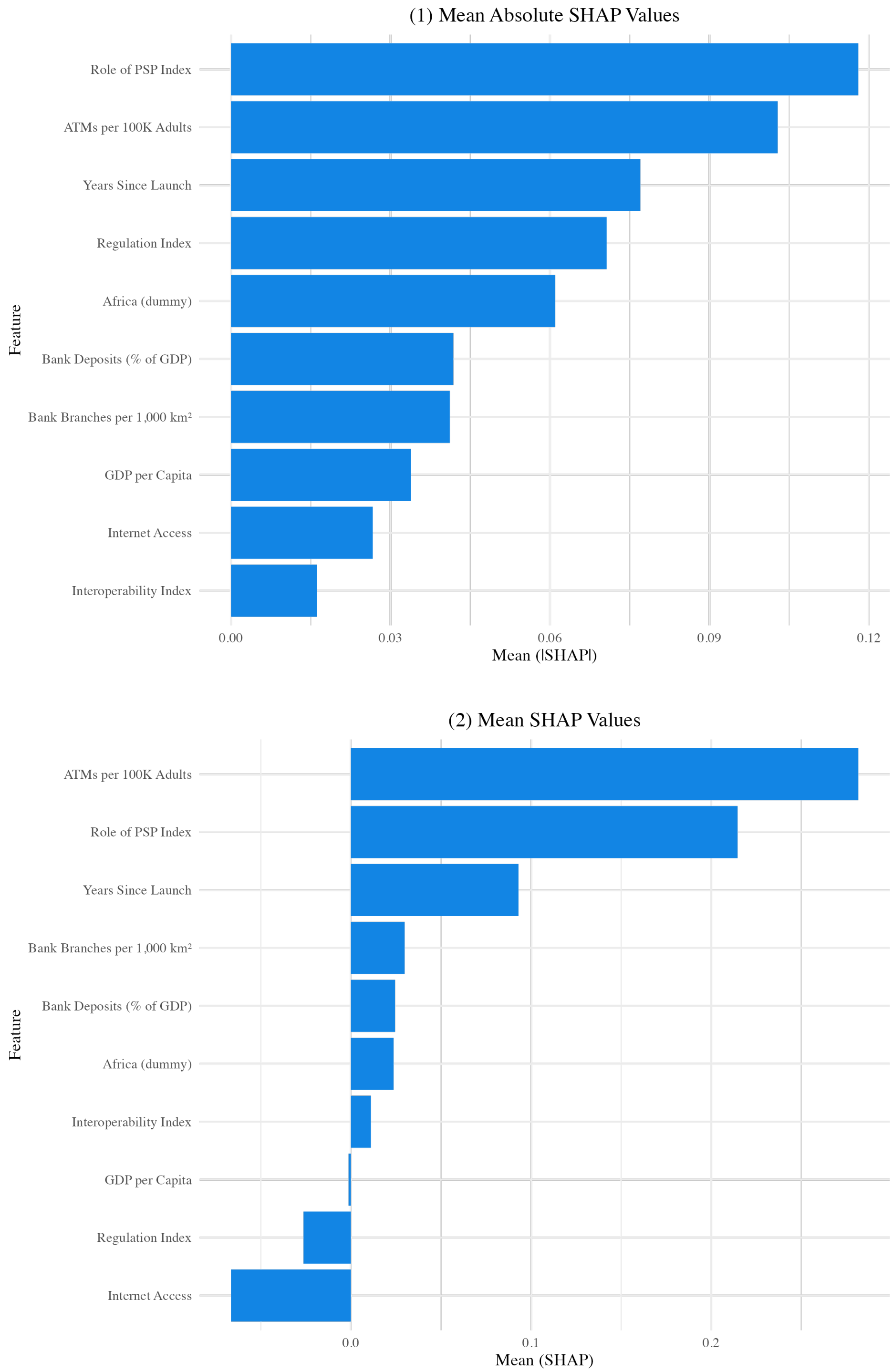


Figure 3: SHAP Values for Random Forest (Continuous Outcome)

model to extract feature importances and conducted a robustness check with SHAP values from a continuous prediction model to validate and interpret the direction and magnitude of each variable's influence.

It is striking that the lack of infrastructure, funding, and coordination among the entities responsible for promoting digital financial inclusion is not having the expected impact on DPS adoption. This paper aims to highlight a clear governance gap between African and non-African countries that must be narrowed in the coming decades if governments wish to expand access to banking services, stimulating their economies and ensuring long-term development.

Our findings highlight several consistent patterns. Regulatory quality, the role of payment service providers, and Internet access are key factors that influence the success of a DPS. The econometric model revealed that whether or not a DPS was African played a highly significant role in the level of DPS adoption. The random forest model and SHAP values found that being in- or outside Africa was negligible for the adoption of a DPS, but confirmed the importance of the regulation and role of PSP indices. Some factors, such as regulation, had a mixed impact across observations: in some cases, better regulation was linked with higher DPS adoption, while in others it was linked with low adoption. Higher regulatory quality may hurt or help DPS adoption depending on the context; this was reflected in the econometric model, where there was substantial regional heterogeneity in the effects of each dimension. The SHAP values added an important dimension to the analysis, since it is only through them that we were able to isolate the direction of the feature impacts using machine learning. In a policy context, understanding how each feature will influence the success of a DPS is necessary to foster an environment where DPS can thrive.

More broadly, this paper provides a flexible framework for future research into DPS adoption, especially as more data becomes available. One potential extension is to incorporate data from developed economies, which could be used to better train a model and provide clearer insights into the factors that distinguish high-performing DPS in emerging economies to those in developing economies. More data could be gathered on financial inclusion-specific metrics, such as the share of the population that is unbanked or the proportion with access to credit. More complete data would improve the precision of both the econometric and random forest models and thus allow us to better understand the institutional factors that hinder DPS success.

As digital financial systems continue to evolve, identifying the factors that support or hinder their success will be essential for improving financial inclusion, especially in emerging economies.

References

Africa Nenda Foundation. State of inclusive instant payment systems in Africa: 2022 report. Annual report, https://www.africanenda.org/en/siips2022, 2022.

Africa Nenda Foundation. State of inclusive instant payment systems in Africa: 2023 report. Annual report, https://www.africanenda.org/en/siips2023, 2023.

Africa Nenda Foundation. State of inclusive instant payment systems in Africa: 2024 report. Annual report, https://www.africanenda.org/en/siips2024, 2024.

Ana Aguilar, Jon Frost, Rafael Guerra, Steven Kamin, and Alexandre Tombini. Digital payments, informality and economic growth. *BIS Working Papers*, 1196, 2024.

Jose Aurazo and Farid Gasmi. Digital payment systems in emerging economies: Lessons from Kenya, India, Brazil, and Peru. *Information Economics and Policy*, 69, 2024.

Jushan Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, July 2009.

Brazilian Central Bank. Pix statistics. https://www.bcb.gov.br/en/financialstability/pixstatistics.

Leo Breiman. Random Forests. *Machine Learning*, 45:5–32, 2001.

Noelia Cámara and David Alfredo Tuesta. Measuring Financial Inclusion: A Multidimensional Index. *BBVA Research Paper*, 14(26), 2014. URL https://ssrn.com/abstract=2634616.

Raissa Fabregas and Tite Yokossi. Mobile Money and Economic Activity: Evidence from Kenya. *The World Bank Economic Review*, 36(3):734–756, 2022.

Jon Frost, Priscilla Koo Wilkens, Anneke Kosse, Vatsala Shreeti, and Carolina Velásquez. Fast payments: design and adoption. *BIS Quarterly Review*, March 2024:31–44, 2024.

Farid Gasmi, Dorgyles C.M. Kouakou, Paul Noumba Um, and Pedro Rojas Milla. An empirical analysis of the social contract in the Middle East and North Africa region and the role of digitalization in its transformation. *Economics of Transition and Institutional Change*, pages 1–51, 2024.

Global System for Mobile Communication Association. Mobile money metrics. https://www.gsma.com/mobile-money-metrics/#, 2024.

International Monetary Fund. Financial Access Survey (FAS), 2023. URL https://data.imf.org/fr-FR/Data-Explorer. Accessed: 2025-04-07.

Youming Liu, Francisco Rivadeneyra, Edona Reshidi, Oleksandr Shcherbakov, and André Stenzel. Ecosystem Models for a Central Bank Digital Currency: Analysis Framework and Models. *Bank of Canada Staff Discussion Paper*, 13, 2024.

Mouad Louhichi, Redwane Nesmaoui, Marwan Mbarek, and Mohamed Lazaar. Shapley Values for Explaining the Black Box Nature of Machine Learning Model Clustering. *Procedia Computer Science*, 220:806–811, 2023.

Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

Christina M. Qiu. Regionalized liquidity: A cross-country analysis of mobile money deployment and inflation in developing economies. *World Development*, 152, 2022.

J. H. Stock and M. W. Watson. Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002.

World Bank. Worldwide governance indicators, 2024 update. https://www.worldbank.org/en/publication/worldwide-governance-indicators, 2024.

World Bank. Individuals using the internet (% of population). https://data.worldbank.org/indicator/IT.NET.USER.ZS, 2025a.

World Bank. GDP (current US$). https://data.worldbank.org/indicator/NY.GDP.MKTP.CD, 2025b.

World Bank. Population, total. https://data.worldbank.org/indicator/SP.POP.TOTL, 2025c.