Predicting Formal Financial Account Ownership Using Machine Learning: A Development Perspective

Theresa R Henry

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Rachel M Heath, Chair

Seth Garz

Melissa Knox

Program Authorized to Offer Degree:

Economics

University of Washington

Abstract

Predicting Formal Financial Account Ownership Using Machine Learning: A Development Perspective

Theresa R Henry

Chair of the Supervisory Committee:

Rachel M Heath

Department of Economics

Financial inclusion has rapidly increased globally, yet large gaps persist between developed and developing countries in the ownership of accounts at formal financial institutions. Although several cross-country studies exist, few have specifically studied financial inclusion in a development context. Even so, demand-side studies have been constrained to demographic and socioeconomic variables and have included measures of key enablers to financial inclusion using country-level indicators. This limits our understanding of important enablers and barriers to financial inclusion. Our study overcomes these constraints by exploiting rich surveys from Financial Inclusion Insights. Using machine learning, we take a data-driven approach to identifying variables that vitally contribute to the predictive performance of our Random Forest

and LASSO classifiers. We then compare our findings to hypotheses about perceived enablers and barriers to financial inclusion. Our analysis reveals that contrary to prior studies, various demographic and socioeconomic characteristics such as literacy, living in an urban area, gender, and poverty status are not consistently predictive of financial inclusion. Instead, we find that the most important predictors are distance to financial points of service, trust in financial service providers, and consistent sources of income, such as regular government-to-person and peer-to-peer transfers. Despite important variation in predictors across the markets examined, each of these indicators is the first or second most important predictor of financial inclusion in every country in our sample. This suggests data captured in individual-level questionnaires is meaningful for understanding priority enablers and barriers to financial account access. These results also provide policymakers and industry practitioners with evidence of actionable ways to increase financial inclusion.

This research is joint work with Seth Garz.

Dedication

To Granddaddy and Papa, whose love lives on within me, always.
To Aunt Carrie and Aunt Eunice, who instilled in me a love for education.

Acknowledgments

I am deeply indebted to all of my committee members, past and present, but in particular, to Rachel Heath and Seth Garz. Rachel, thank you for your continuous support both inside and outside of research. Seth, thank you for the opportunity to work for you and with you, for your guidance (tough love included), and for never giving up on me.

I am also deeply indebted to the Bonderman Travel Fellowship Team. Because of this life-changing experience of exploring and wandering into the unfamiliar, I could find my way back to my passion for care, community, and advocacy for marginalized and under-resourced groups and engage these values in my work.

Finally, I want to thank my Mom, Dad, Stepdad, Stepmom, and Aunt Tanya; my sisters Rayven, Anaya, and Amber; my closest friends Ashleigh, Elba, Fariha, Liz, Matthew, Reva, and Suwilanji; my economics family Alejandro, Mishita, Pushpak, and Yurim, and of course, my dog, Oreo. I am so grateful for your encouragement and support throughout this long, trying journey. Without all of you, none of this would be possible. Last but not least, thank you to my ancestors. I love and miss you all dearly.

Table of Contents

List of Figures

List of Tables

1 Introduction

Many studies have researched the relationship between the development of a country's financial system and its impact on economic growth[1]; however, with the recent availability of data at more granular levels, researchers have sought to understand the effects of financial inclusion on firms, households, and individuals (Ayyagari et al. 2017; Beck et al. 2009; Allen et al. 2016). At a micro-level, financial inclusion is defined as ownership of a formal financial account or product, such as transactions, savings, credit, and insurance. Between 2011 and 2017, the number of adults who opened an account at a financial institution increased from 51 percent to 69 percent. Yet, large gaps still exist between countries at different levels of financial development and for specific population segments (Demirgüç-Kunt et al. 2018). In particular, according to the World Bank's 2017 Global Findex Database, nearly half of the world's unbanked live in seven countries[2] (Barajas et al. 2020). The United Nations (UN) has identified financial inclusion as a critical enabler to several of its Sustainable Development Goals (SDGs)[3] (Klapper et al. 2016), as the body of literature has pointed to many potential benefits of the use of financial services. These include improving adults' earning potential, facilitating financial risk management, lowering transaction costs, helping accumulate savings, reducing corruption, and improving efficiency (Demirgüç-Kunt et al. 2018).

Despite this evidence, only some studies focus on understanding the underpinnings of financial inclusion across countries with lower shares of financial inclusion or at lower stages of

---

[1] See Levine (2005) and Popov (2018) for reviews of this literature.
[2] These seven countries are Bangladesh, China, India, Indonesia, Mexico, Nigeria, and Pakistan. Our study includes four of these countries, Bangladesh, Indonesia, Nigeria, and Pakistan.
[3] eliminating extreme poverty, reducing hunger and promoting food security, and promoting gender equality (Klapper et al. 2016)

development[4]. In this study, we use a rich dataset of individual-level financial inclusion indicators to understand the underpinnings of financial inclusion in seven developing countries: Bangladesh, Indonesia, Kenya, Nigeria, Pakistan, Tanzania, and Uganda. We build on this area of research by focusing on the development context, taking an agnostic, data-driven approach, and considering a comprehensive financial inclusion measure. We use machine learning algorithms to predict financial inclusion and determine which variables are most important for prediction. Additionally, we use a measure of financial inclusion that includes three formal financial services: banks, mobile money, and non-bank financial services such as microfinance.

Current research typically focuses on a single financial service, such as bank payments and deposits, mobile money, savings, or credit. To our knowledge, few studies consider multiple types of financial services in a microeconomic setting[5], and no study has yet considered more than two types of financial services when trying to understand the underpinnings of financial inclusion. Additionally, cross-country financial inclusion research typically focuses on a breadth of countries across different development stages instead of concentrating on countries at similar economic and formal financial institution development levels. The constraints mentioned above are likely due to either data limitations or limitations imposed by linear modeling. Much of the available data that is appropriate for studying financial inclusion has limited demographic and socioeconomic information, limited or non-existent financial behavior information, or limited financial service indicators. Even in datasets less constrained by these limitations, linear models restrict the number of variables researchers can include relative to the number of observations available.

[4] Honohan and King (2012)
[5] Barajas et al. (2020) present studies that create composite indices of financial inclusion that assess macroeconomic impacts.

We overcome these constraints by applying machine learning to Kantar/Intermedia's Financial Inclusion Insights Surveys (FII). To our knowledge, FII contains information from one of the richest surveys of harmonized individual-level financial inclusion indicators specifically for lower-middle-income countries. Along with Kenya, Tanzania, and Uganda, FII includes five out of seven countries where nearly half of the world's unbanked live: Bangladesh, India, Indonesia, Nigeria, and Pakistan. Focusing on this subsample of developing countries is important as vulnerability, or a lack of resources required for resiliency, remains a crucial concern for many people in these settings. Although FII is limited to these eight countries, the sheer number of demographic, socioeconomic, digital engagement, infrastructure, and economic empowerment variables sets it apart from other financial inclusion datasets. FII surveys also include data following the rapid expansion of digital financial services. Thus, the FII survey's available sample of countries, harmonization, depth of information, and coverage of a time of technological and financial innovation make it a compelling dataset to understand the underpinnings of financial inclusion using predictive modeling.

By utilizing machine learning, we take advantage of nearly the entire span of variables available in the data. These additional variables can provide new information that can be used for prediction, which may be more costly if included in linear models commonly used in the literature. Specifically, we employ Random Forest (RF) (Breiman 2001) and the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani 1996) to understand which variables contribute the most information to predicting financial account ownership.

While policymakers and practitioners have much to celebrate about expanding formal financial services, they continue to seek additional understanding and evidence on barriers and opportunities to financial inclusion (World Bank 2021). Machine learning can reveal important predictors that are different from our existing knowledge. Policymakers and development economists can use these predictors to inform data collection exercises, as proxies for hard or expensive-to-measure variables, to generate new theories, and, more generally, to inform research in the financial inclusion literature. Additionally, transfer learning, an application of machine learning, allows policymakers and other stakeholders to gain insights from comparable countries when there is a limited supply or non-existent availability of financial inclusion data (Weiss et al. 2016). By approaching financial inclusion from a predictive perspective, we generate insights that contribute to a richer understanding of the factors associated with financial inclusion than any single approach can on its own (e.g., randomized controlled trials (RCTs) or linear observational studies). We also provide stakeholders with a potentially promising way to use predictive modeling even when data availability is limited.

To investigate the differences between our model and models used in the existing literature, we compare variables identified as important for the accuracy of the prediction algorithms to (1) priority correlates from prior analyses that used the Global Findex Database and FinMark Trust's FinScope Consumer Surveys and (2) variables commonly identified as critical "enablers" of financial inclusion by practitioners and other stakeholders, such as the UN, the World Bank, and their partners. These comparisons amount to the difference between theoretically unconstrained models produced for predictive accuracy (i.e., black box models) versus models based on informed priors designed to attribute causality.

Our results reveal that despite important variation in predictors of account ownership across the seven markets examined, our models perform remarkably well. Contrary to some prior studies, various demographic and socioeconomic characteristics, such as literacy, living in an urban area, gender, and poverty status, are not consistently predictive of account ownership across countries. Instead, we find that distance to financial points of service, trust in financial service providers, source of income, and identification are amongst the most important predictors across our sample of countries. These findings support supply-side findings in the literature and suggest that data captured in individual-level questionnaires are meaningful for understanding priority enablers and barriers to financial account ownership. Additionally, the differences between the predictors chosen by each machine learning algorithm are not only illustrative of the differences between their objective functions but also illustrative of which type of variables the algorithms favor.

The remainder of the dissertation proceeds as follows: Section 2 provides an overview of financial inclusion indicators and their relation to economic theory and the use of machine learning methods in economics. Section 3 presents high-level technical details of the RF and LASSSO algorithms, the metrics used for performance evaluation, and predictive variable importance values. Section 4 describes the data and our empirical approach. Section 5 presents the performance evaluation, variable importance, and classification error analysis results with discussion. Section 6 presents the results of the transfer learning exercise. Finally, section 7 concludes.

2 Literature Review

2.1 Why financial inclusion matters

Financial inclusion has been identified as a critical enabler to many of the United Nations' SDGs, including eliminating extreme poverty, reducing hunger and promoting food security, achieving good health and well-being, fostering quality education, and promoting gender equality (Klapper et al. 2016). This situates financial inclusion at a junction of impacting multiple channels. Observational and RCT research has shown financial inclusion to facilitate risk sharing and expand networks, allowing for an increased ability to cope with shocks (Jack and Suri 2014; Riley 2018), increase savings (Breza et al. 2020; Blumenstock et al. 2018), increase female labor supply (Field et al. 2021), reduce energy poverty (Koomson and Danquah 2021), and increase consumption (Munyegera and Matsumoto 2016; Lee et al. 2021).

Since the availability of better data on consumer access to financial services (i.e., the Global Findex Database and the FinScope Consumer Surveys), researchers have studied some form of the question, "What are the underpinnings of financial inclusion?" Our research builds upon studies that specifically used bank deposit account data and linear probit models in their analysis. We add to this literature by using new and richer data that includes indicators for banks, mobile money, and non-bank financial institutions and by applying data-driven methods to understand what predicts financial inclusion. What follows is a discussion of various financial inclusion indicators and datasets, financial inclusion's relation to economic theory, and a description of the use of machine learning methods in economics.

2.2 Financial inclusion data and cross-country analysis

In contrast with the current analysis, prior quantitative studies on the correlates of financial inclusion primarily exploit the Global Findex Database (Demirgüç-Kunt et al. 2018). For example, Allen et al. (2016) and Klapper et al. (2021) merge individual-level data from the Global Findex Database with country-level data from other sources to run probit models with fixed effects. These studies compared an average of 131 countries. Cross-country approaches have advantages in attempting to quantitatively capture variation in country-level economic parameters, such as regulatory regimes and macro-policy, but also suffer from several disadvantages, including sample size restrictions, canceling effects, and variation in financial institutions, products, and policies that would make individual country analysis more amenable.

The Global Findex Database's country sample sizes are often too small to precisely capture individual, household, and community-level barriers to adopting formal financial accounts among diverse national populations. For example, the modal national sample in the Global Findex Database is 1,000 observations. In contrast, most datasets examined herein range from 2,000 to 5,000 observations.

Additionally, at any given time, countries occupy different stages of retail financial market maturity and maintain different demographic profiles; therefore, pooling data across markets in cross-country regressions could cause informative variation in a highly developed market to cancel out informative variation in a poorly developed market. For example, as illustrated below, possession of formal identification is highly predictive of account access in two East African markets we examine, Kenya and Uganda. Still, it is not particularly informative in the other markets when considering the Random Forest classifier. This may reflect the lack of variation captured in markets where national identification campaigns have reduced the relevance of identification as a barrier to financial inclusion. Notably, 20 percent of unbanked adults report a lack of identity documentation as a reason for not having an account (Demirgüç-Kunt et al. 2018); however, when using a country-level indicator for know-your-customer (KYC) identification requirements, Allen et al. (2016) find that identification is not associated with financial inclusion for the general population nor the poorest quintile.

Lastly, financial institutions, products, and policy choices often vary across national boundaries. Limiting our models to single countries acknowledges that countries are quite different and renders the results more relevant to the administrative level at which market actors make decisions. While it is tempting to think of global analysis as more externally valid than national analysis, it is more likely that national analysis is relevant for well-defined policy challenges within clusters of countries that exhibit similar stages of development or geographic proximity. For example, our transfer learning exercise finds that training the RF classifier with Kenya's data and running Uganda's data through Kenya's classifier produces a better fit than running Bangladesh's data through Kenya's classifier. We attribute this result to similar distributions of financial indicators and geographical proximity.

2.3 Financial Inclusion Indicators

Development and poverty research often assumes that demographic and socioeconomic characteristics are strong predictors of human development outcomes. In Allen et al. (2016) and Klapper et al. (2021), virtually all demographic indicators reveal statistically significant associations with account ownership, though gender is notably only significant in Klapper et al. (2021). Furthermore, observational studies and randomized experiments frequently use these characteristics as controls and have proposed that the accuracy of proxy-means testing using demographics is a reasonable standard against which alternative targeting methods should be tested (Brown et al., 2018). In contrast, our data-driven approach finds that only a few demographic and socioeconomic indicators are informative for predicting account ownership in select countries. This may suggest that other covarying observable characteristics contain more meaningful information from a financial inclusion access perspective. Specifically, we find that sources of income, proximity to POS, trust in FSPs, and identification are some of the most important predictors of financial inclusion.

The 2017 Global Findex Database cites insufficient money as the most common explanation for the lack of a bank account, with two-thirds of respondents indicating this as a critical barrier (Demirgüç-Kunt et al. 2018). Unsurprisingly, Allen et al. (2016) and Klapper et al. (2021) both show positive associations with employment and negative associations with unemployment. Lack of money, and relatedly lack of employment, is a difficult barrier to applying policy actions geared towards promoting account access. Thus, interventions to expand financial access may benefit from focusing on more targeted segments and their specific sources of income, as several public policies designed for income support and growth already exist. In fact, as a means of expanding financial access, advocates have pointed to the opportunity to directly deposit public cash transfers (Government-to-Person or G2P) and public and private sector wages into formal accounts as the default means of payment (Field et al. 2021; Blumenstock et al. 2018). We find that the predictive importance of receiving G2P transfers is high in Indonesia, Bangladesh, and Kenya; however, regular support from family, friends, or spouses is the most common income source predicting financial inclusion. Regular Peer-to-Peer (P2P) support is a top predictor in Bangladesh, Indonesia, Kenya, Nigeria, and Tanzania. Our results show that consistent sources of income, such as P2P and G2P, are more predictive of financial inclusion than poverty or employment.

In addition to demographic and socioeconomic indicators, various evidence suggests that key aspects of what Allen et al. (2016) refer to as "the enabling environment[6]" are important mediators of account access (Camner 2012; Johnson and Arnold 2012; Economides and Jeziorski 2017). The retail finance distribution network is often considered one such enabler. Causal studies exploit variation in proximity to mobile money agents arising from the staggered rollout

[6] Sahay et al. (2020) list customer identification, digital infrastructure, financial literacy, and a supportive regulatory and legal environment as enabling factors for financial inclusion. Additionally, Staschen and Meagher (2018) list nonbank e-money issuance, use of agents, risk-based customer due diligence, and consumer protection as four basic enables of digital financial inclusion.

of agent infrastructure (Jack and Suri 2014; Suri and Jack 2016). Due to the growth of mobile money systems at that time, this rollout was a quasi-random driver of mobile money account adoption that was otherwise exogenous to household characteristics associated with financial behavior (Suri 2017; Suri et al. 2021). The strong association between proximity to retail finance infrastructure and account access is vital to these analyses. Indeed, 22 percent of unbanked respondents in the 2017 Global Findex Database reported that greater distances to financial institutions were a critical barrier to account access. Thus, expanding agent infrastructure in low- and middle-income countries is a crucial focus of leading financial inclusion advocacy institutions (Demirgüç-Kunt et al. 2018). Using the 2011 Global Findex Database, Allen et al. (2016) find that the average branch and automated teller machine (ATM) penetration per 1000 square kilometers (km) is meaningfully associated with account access among the general population, except in the rural or poorest quintile sub-samples. This lack of robust association may reflect that country-level indicators of average financial infrastructure penetration fail to capture between community variation that is key to the causal studies referenced above. Our analysis uses an individual-level measure of financial infrastructure that measures the approximate distance from the respondent’s home to a bank branch, an ATM, a mobile money agent, or a microfinance institution. We find that distance to point-of-service (POS) is the most or second most important predictor in all countries in our sample. This result holds regardless of a country’s mobile money penetration. It suggests that measuring financial infrastructure from a demand-side perspective may be more accurate and thus can capture the important microeconomic variation in distance to POS. Future research will assess whether this result is robust to the rural and poor segments of the population.

We also find evidence of consumer protection as an enabler of financial inclusion. FSPs are required to keep consumer data confidential, but approximately 20 percent of financially excluded adults indicated distrust of the financial system as a reason for not having an account (Demirgüç-Kunt et al. 2018). Honohan and King (2012) studied the association between bank account ownership and trust in banks and found a positive association. Our results support the evidence from the literature. The FII trust variable measures if an individual is concerned about FSPs keeping their information private, irrespective of financial inclusion status[7]. Although consumer trust is more difficult to target in an actionable way (Garz et al. 2021), evidence contributing to an understanding of the barriers to financial inclusion can inform regulatory bodies of the need to either develop or enforce consumer protection regulations or ensure users are well-informed of the policies and enforcements in place to build trust.

Access to personal identity documentation is also often considered an enabling feature of the financial ecosystem, given the regulatory requirement for financial service providers (FSPs) to comply with KYC requirements (World Bank 2019). In the 2017 Global Findex Database, 20 percent of unbanked respondents reported the lack of required documentation as a key barrier to account access. As mentioned, country-level indicators for KYC exceptions in cross-country regressions are not associated with account access for the general population or the poorest quintile (Allen et al. 2016). Again, these national indicators are likely weak proxies for capturing the role of documentation in KYC processes required to open accounts. In contrast, our methods exploit individual variation in access to documentation and confirm that access to identification

[7] Trust information is only available for financially excluded individuals in the Global Findex Database.

is highly informative for predicting account access in East African countries where identification is less universal than in countries like Bangladesh, Indonesia, or Pakistan.

Cross-country global analyses have yet to examine the relationship between phone access and account ownership. The causal literature has demonstrated impressive social welfare impacts from access to mobile money (Suri et al. 2021), which has spurred many development practitioners to pursue interventions to boost mobile phone access among low-income populations (Highet et al. 2021). Employing an experimental design in Tanzania, Roessler et al. (2021) demonstrate that distributing free basic phones and smartphones to women who lack phones increases the use of mobile money by 77 percent and 48 percent, respectively, after 13 months. The authors suggest that the larger effects of basic phones versus smartphones are likely due to the lack of advanced literacy required to use smartphones effectively. In contrast, our RF classifier identifies smartphones as an important predictor in Kenya and Nigeria; however, mobile phone ownership (owning any phone) is not an important predictor of financial inclusion.

2.4 Machine Learning

Another relevant strand of literature is prediction questions in economics. Machine learning algorithms excel at prediction relative to the linear models commonly used in economics. Machine learning also offers different value than RCTs and other causal methods; that is, it has the potential to generate different or complementary insights that differ from traditional approaches in economics. While RCTs and other causal studies provide value by minimizing selection biases, insights from these studies could benefit from studying similar questions at a larger scale, given their limitation of internal validity. Policymakers implement policies at scale and in imperfect conditions (relative to RCTs), and having results from predictive analyses can aid in potentially more effective targeting of development policies, programs, and interventions. For example, in development economics, researchers have employed machine learning to predict resilience (Knippenburg et al. 2019; Garbero and Letta 2022) and poverty (McBride and Nichols 2018; Jean et al. 2016; Blumenstock et al. 2015). These studies use traditional data but with many more features than linear models can accommodate, as well as newer types of data such as call records, high-frequency data, and satellite data, to name a few. Machine learning is also helpful in that it has the potential to uncover generalizable patterns across different settings, such as across developing countries, if any exist (Mullainathan and Spiess, 2017).

As described above, causal research on financial inclusion has contributed to the findings that provide evidence in favor of financial inclusion, yet as Morduch (2020) notes, "[RCT] studies reveal challenges in drawing inferences across RCTs." Studying models that use the same set of features but where the algorithm is separately run on each country allows us to make comparisons (although endogenous) across countries and provide value and insight for policymakers. In particular, compared to RCTs, which rely heavily on local context, design, and marginal customers whose characteristics, behaviors, and impacts are most likely different than those on the "inframargin" (Banerjee et al. 2015), focusing on prediction can provide useful targeting of both marginal and inframarginal customers for either direct policy implementation or experimentation, given that financial inclusion programs are well-suited for RCTs in developing countries.

2.5 Economic Theory: Risk-sharing and Resilience

Financial inclusion is not an end in and of itself but a means to an end and has the potential to increase resilience and facilitate risk-sharing. Still, ultimately, these changes may have an impact on existing risk-sharing networks if users self-insure or if users can rely on long-distance networks. Since Townsend (1994), researchers have studied risk-sharing and resilience to economic shocks in formal and informal insurance networks. Much of the more recent theoretical literature typically addresses access to mobile money and how it reduces individual transaction costs. Jack and Suri (2014) study insurance through an informal risk-sharing network by including explicit transaction costs in their framework. They find that non-mobile money users who live in villages with mobile money users experience a 7 percent reduction in consumption when a negative shock is experienced. This implies an increase in risk-sharing that they attribute to reduced transaction costs. More recently, researchers have examined if individuals are more inclined to self-insure and the spillover effects this has on other community members, locally and further away (Riley 2018, Dizon et al. 2020). Although the impact of formal financial services on existing informal institutions and risk-sharing networks has been challenging to pin down, our results support evidence that risk-sharing through P2P transfers is an important predictor of financial inclusion.

3 Methodology

We aim to build a classifier to predict financial inclusion based on some or all of the variables[8] present in the data. Given that our problem of interest is predictive rather than causal, we aim to minimize prediction error on data not used to train the classifier. We prefer that the classifier depends on a subset of variables to aid our understanding of the underlying structure of financial inclusion; in other words, in addition to building a well-performing classifier, we want to determine which characteristics are most important for predicting financial inclusion. To

[8] Used interchangeably with "predictor" and "feature."

extract interpretable results from our models, we use the Least Absolute Shrinkage and Selection Operator algorithm (LASSO) (Tibshirani 1996) and the Random Forest (RF) (Breiman 2001) algorithm and compare their results. These flexible and complex algorithms allow us to exploit the richness of the data while still providing interpretable results. Each algorithm predicts the probability of each observation being either financially included or excluded. Given a decision threshold, we can classify an observation into one of the two groups.

In addition to interpretability, these algorithms also allow us to compare results from both a parametric and nonparametric perspective. The LASSO is a linear, parametric algorithm that estimates coefficients most closely to traditional econometric methods; however, these coefficients are biased by design. This bias is reduced by rerunning logistic regression on the selected subset (more detail below); however, caution is still needed relative to how economists readily interpret coefficients of linear models. On the other hand, RF is a nonparametric algorithm that allows for more complex data relationships, such as determining complex interactions between variables during its prediction procedure. Because RF does not have directly interpretable output, we use a feature importance metric to determine which predictors were most influential during prediction.

3.1 Random Forest

Figure 1[9] describes the Random Forest (RF) algorithm (Breiman 2001). RF is a classifier consisting of a collection of decision tree classifiers that are decorrelated by bootstrap aggregation (aka bagging)[10]. Since decision trees are approximately unbiased but have high variance, the bagging procedure averages classifiers, thus reducing the variance.

[9] Tibshirani et al. (2009).

[10] See Breiman (1996) for details.

Figure 1 *Random Forest for Regression or Classification.*

1. For $b = 1$ to B:
   (a) Draw a bootstrap sample $\mathbf{Z}^*$ of size N from the training data.
   (b) Grow a random-forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached.
      i. Select m variables at random from the p variables.
      ii. Pick the best variable/split-point among the m.
      iii. Split the node into two daughter nodes.
2. Output the ensemble of trees $\{T_b\}_1^B$.

To make a prediction at a new point x:

Regression: $\hat{f}_{\text{rf}}^B(x) = \frac{1}{B}\sum_{b=1}^B T_b(x)$.

Classification: Let $\hat{C}_b(x)$ be the class prediction of the bth random-forest tree. Then $\hat{C}_{\text{rf}}^B(x) = \textit{majority vote}\ \{\hat{C}_b(x)\}_1^B$.

Notice that for a given tree, the algorithm only considers a subset of predictors at each node. This prevents the algorithm from growing trees with the same few strong predictors dominating the first splits (James et al. 2022). Even still, a few highly correlated variables can dominate the growth of the forest since there is a higher probability that a predictor from a set of highly correlated predictors is selected and considered for splitting a node. This is unfavorable when one of two highly correlated predictors strongly correlates with the response, but the other does not. This can occur if continuous predictors vary in scale or if categorical predictors vary in the number of categories, thus biases variable importance measures (Strobl et al. 2007). We discuss the impact of this on variable importance measures below.

3.2 Least Absolute Shrinkage and Selection Operator (LASSO)

Tibshirani (1996) developed the LASSO estimator. This estimator combines a least square lost function and an l_1-constraint, as shown below

$$minimize_{\beta_0,\beta}\left\{\frac{1}{2}\sum_{i=1}^{N}(y_i-\beta_0-\sum_{j=1}^{p}x_{ij}\beta_j)^2\right\} \text{ subject to } \sum_{j=1}^{p}|\beta_j| \le t,$$

where $i = 1,\ldots,N$ denotes the number of observations, $j = 1,\ldots,p$ denotes the number of feature variables, and t denotes the regularization size. Equivalently, the LASSO estimator is

$$\hat{\beta}^{lasso} = \min_{\beta_0,\beta} \left\{\frac{1}{2}\sum_{i=1}^{N}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|\right\}.$$

There is no theoretically optimal value of lambda, so lambda is chosen using cross-validation[11]. The lambda value that produces the most accurate model on the validation set is selected.

The LASSO optimization problem's constraint allows for sparse solution vectors, i.e., the solution identifies a relatively small number of feature coefficients that are important for the optimization problem and thus "recover[s] the underlying signal in a set of data." (Hastie et al. 2015, 1) This problem produces better prediction accuracy relative to the least-squares estimator by trading off low bias for low variance by shrinking the coefficient values, some of which become zero. When there are many features, this shrinkage also allows for interpretability by identifying the most important features of the prediction problem.

3.3 Performance Evaluation

Our classifiers will be evaluated using various metrics. The most straightforward metric, accuracy, is the percentage of correctly classified observations; however, it has been deemed a poor performance metric and should be considered in conjunction with other metrics (Provost et al. 1998). Additionally, we consider sensitivity, specificity, and the area under the receiver operating characteristic (ROC) curve. Sensitivity and specificity evaluate how well the classifier predicted each group. Sensitivity is defined as the proportion of actual observations in the

[11] Cross-validation consists of randomly splitting the training set into *k*-folds. The model is then trained using *k-1* folds. The remaining fold, the validation set, is used to compute the model's error (note that the validation set differs from the test set). This is repeated until each fold has served as the validation set. Averaging each model's error as a function of λ produces an error curve, and λ is chosen by minimizing this curve. See Hastie et al. (2009) for more details.

minority group correctly identified. In contrast, specificity is the proportion of actual observations in the majority group correctly identified. When each class of a binary problem has equal proportions, the minority group (or positive group) is assigned to the group of interest. In this study, the financially excluded group is the minority group.

The ROC curve plots the true positive rate (sensitivity) and the false positive rate (1-specificity) at each possible classification decision threshold and provides a visual representation of the classifier's performance (Swets 1988). The closer the curve to a boundary from (0,0) to (0,1) to (1,1), the better the performance. If the curve is a line from (0,0) to (1,1), the classifier performs no better than random chance. The ROC curve is summarized by taking the area under the curve (AUC) and ranges from one-half (0.5) to one (1.0). The AUC is applicable for evaluating the performance of a given classifier and the performance across classifiers trained on the same sample of data. This allows for comparing the performance of our model between the RF classifier and the LASSO classifier. Using the ROC curve, we can show that a classifier outperforms another classifier if the former has a higher true positive rate for each false positive rate (Hoens and Chawla 2013).

3.4 Variable Importance

Variable importance informs us about which predictors are most influential to the model's performance and provides insight into the underlying structure of the response. In Section 5, we will compare the predictors deemed important by our classifiers to the ones the literature identified as important.

Permutation feature importance (PI) measures the importance of predictors in the RF classifier. This metric is interpretable and comparable across countries. Higher values imply larger decreases in model accuracy and are thus deemed more important for predicting financial inclusion (Fisher et al. 2019). Figure 2[12] (left) illustrates the scheme. PI calculates the change in a tree's prediction error when one variable is randomly permuted. If the error increases following the predictor's permutation, this predictor is considered important. For each variable, the difference in error is averaged across all trees; this average is the PI. After this procedure is carried out for all predictors, the importance values are ranked. PI only tells us whether a predictor is important but does not provide information on the direction of association with the response.

Y	X_j	Z
y_1	$x_{\pi_j(1),j}$	z_1
$\vdots$	$\vdots$	$\vdots$
y_i	$x_{\pi_j(i),j}$	z_i
$\vdots$	$\vdots$	$\vdots$
y_n	$x_{\pi_j(n),j}$	z_n

Y	X_j	Z
y_1	$x_{\pi_{j\|Z=a}(1),j}$	$z_1 = a$
y_3	$x_{\pi_{j\|Z=a}(3),j}$	$z_3 = a$
y_{27}	$x_{\pi_{j\|Z=a}(27),j}$	$z_{27} = a$
y_6	$x_{\pi_{j\|Z=b}(6),j}$	$z_6 = b$
y_{14}	$x_{\pi_{j\|Z=b}(14),j}$	$z_{14} = b$
y_{21}	$x_{\pi_{j\|Z=b}(21),j}$	$z_{21} = b$
$\vdots$	$\vdots$	$\vdots$

Figure 2
Permutation scheme for the original marginal (left) and for the newly suggested conditional (right) permutation importance.

Variable importance measures produced by this scheme suffer from bias based on the RF split selection bias described above. Highly correlated variables will dominate the ranking even if some variables do not correlate directly with the response. Strobl et al. (2008) developed a conditional permutation feature importance (CPI) measure to reduce the bias for highly correlated predictors. Rather than permuting a given variable across all observations, this

[12] Strobl et al. (2008).

measure conditions the permutation of each predictor on a grid space partitioned by the split points induced by a given tree; in other words, the permutation occurs within subgroups of the observations. This scheme is illustrated on the right side of Figure 2.

The LASSO estimator's coefficients directly provide feature importance measures. The estimator is applied to 1,000 bootstrapped training data samples to identify the most important predictors. In each bootstrap sample, predictors with zero-valued coefficients are filtered, and logistic regression is run on the remaining predictors. Finally, each predictor's coefficients' mean and standard deviation are computed, and the absolute value is taken to determine its rank of importance. Notice that, unlike importance values used for RF, the LASSO importance values have information on the direction of the association with the response.

4 Dataset and Data Preprocessing

Our analysis uses Kantar/Intermedia's Financial Inclusion Insights (FII) surveys. Adult-aged, nationally representative surveys were conducted in eight low- and lower-middle-income countries, including Bangladesh, India, Pakistan, Indonesia, Kenya, Tanzania, Uganda, and Nigeria, from 2013 to 2018. FII's questionnaires collect nearly identical information across countries, with some exceptions for country-specific programs and some variation in questionnaires over time. Kantar created a harmonized version of data to overcome survey variation between countries and across time. The most complete set of variables available are from 2017-2018, as additional sections and questions were added or amended over time. Specifically, our analysis includes cross-sections from 2017 for Bangladesh, Kenya, Nigeria, Pakistan, Tanzania, and Uganda and 2018 for Indonesia. After cleaning the data, there are 24,738 observations, and each country's observations range from approximately 2,000 to 5,000.

Our variable of interest is formal account ownership. In the FII data, account ownership consists of three types of formal financial services: banks, mobile money, and non-bank financial institutions (NBFI)[13]. Individuals with accounts registered in their name or jointly are considered financially included. Six of the seven countries in our sample are classified as low-middle-income countries by the World Bank, while Uganda is classified as a low-income country (World Bank 2024). Even so, we see variation in financial inclusion across countries.

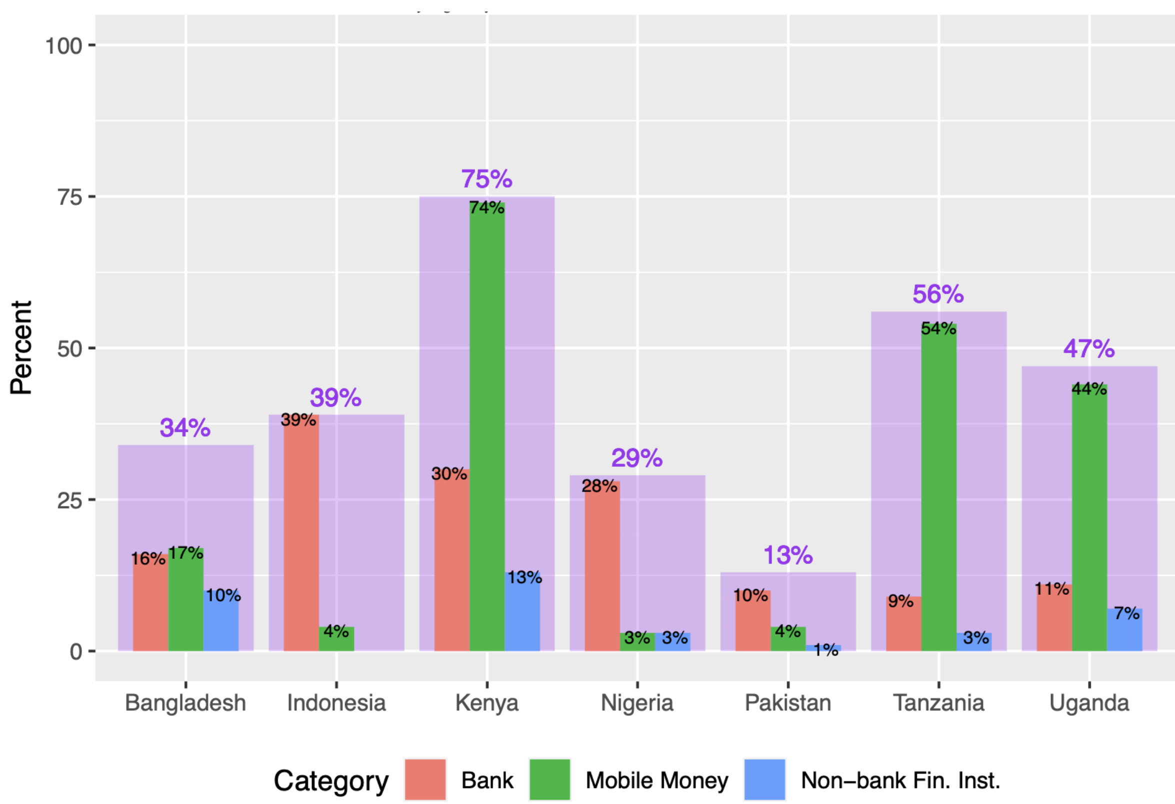


Figure 3: Financial Inclusion - Overall and by Institution

Figure 3 shows the percentage of respondents who are financially included overall (across all three financial service types) and by financial service type. Kenya, Tanzania, and

[13] Non-bank financial institutions include microfinance institutions, cooperatives, post offices, SACCOs, pawnshops, payment banks, and savings groups. Institutions differ by country. No NBFI data is available for Indonesia 2018.

Uganda have the highest financial inclusion rates, ranging from 53 percent to 78 percent, whereas Pakistan has the lowest rate in the sample of countries, at 14 percent. Financial inclusion varies by type of financial institution as well. For nearly all countries, a single financial service institution dominates. Bank account ownership drives inclusion for Indonesia, Nigeria, and Pakistan, whereas mobile money account ownership drives inclusion for Kenya, Tanzania, and Uganda. Notably, Indonesia, Nigeria, and Pakistan have less than 5 percent of individuals with a registered mobile money account. In Bangladesh, bank and mobile money accounts have similar shares of ownership (16 percent and 17 percent, respectively). Finally, NBFIs constitute the smallest share of account ownership in all countries. However, the difference between the share of NBFI account holders and the share of account holders in the second largest financial institution only ranges from 0-6 percentage points.

The data include information on financial service ownership, demographics, socioeconomic characteristics, aptitude for using technology (phone ownership and ability to perform actions), economic empowerment (e.g., involvement in household spending decisions), and financial infrastructure (travel distance to points of sale). A complete list of variables can be found in Appendix Table A1.

Before analysis, we take several actions to preprocess the data and tune the classifiers. First, due to the imbalanced distribution of financial inclusion in most countries, the data is upsampled to balance the share of included and excluded individuals[14][15]. We then split each country's data into training and testing sets to avoid overfitting the classifier. The training set comprises 80 percent of the original data, and the testing set comprises 20 percent. The testing set will remain untouched throughout the analysis. It will act as our unseen independent sample to test out-of-sample performance (i.e., data not used to train the algorithms). Splitting the data helps reduce fitting the classifier too closely to the data used to train the classifier, as we want the classifier to perform well on unseen data.

All variables in the data are categorical, so the categories of each variable are coded as binary. This results in 56 predictors included in the model[16]. We perform cross-validation to tune the RF classifier's parameter, which determines the number of predictors randomly selected at each node split, and the LASSO classifier's penalization parameter, λ. Finally, neither algorithm can account for missing data; thus, as an initial approach, observations with missing data are dropped. More sophisticated techniques will be implemented in the future. Finally, our model takes a black-box approach and leverages all the variables in Appendix Table A1.

---

[14] The data were balanced using up-sampling following an analysis demonstrating that setting the threshold to a more appropriate level did not improve performance. Thus, the data were rebalanced, and this greatly improved performance. The analyses using each country's original financial inclusion distributions were evaluated using two thresholds to assign probabilities to be included or excluded, 0.5 and the threshold that optimized the f-score (a metric that considers the trade-off between precision and recall, which is preferred to sensitivity and specificity with imbalanced data). Few, if any, of the classifiers performed well with the 0.5 threshold. Using the optimal f-score threshold improved recall but at the cost of precision, and thus, the classifiers still performed poorly. These practices are suggested by Provost (2008) and Brownlee (2020).

[15] Up-sampling entails randomly sampling the minority class to have the same number of observations as the majority class. Typically, down-sampling is performed; however, given that some countries have such imbalanced data, down-sampling would comprise the information available from the data (Chen et al. 2004). The analysis uses a naïve up-sampling procedure, but more sophisticated procedures, such as synthetic minority oversampling (SMOTE), will be considered in the future. Additionally, this sampling does not account for the survey design.

[16] This value excludes reference levels. Unlike linear regression, the choice of reference levels impacts the fit of the LASSO, but LASSO (and RF) can also fit models that include reference levels. To my knowledge, there is no documented common practice, so the least advantaged group was assigned as the reference group for categorical variables with three or more categories, as is common in economics.

In the next section, we compare the results of this approach to the results from theoretically motivated analyses of financial inclusion; however, to assess the validity of this comparison, we evaluate the performance of the RF and LASSO classifiers and compare their top-ranked important predictors, then we further assess performance by comparing select characteristics of correctly classified and misclassified individuals.

5 What predicts financial inclusion?

5.1 Overall Performance Evaluation

We evaluate the model's performance on each country's data across the RF and LASSO classifiers. That is, we determine the extent to which each algorithm correctly predicts each class and whether one classifier dominates the other.

Table 1 shows each country's accuracy, sensitivity, specificity, and AUC for RF (Panel A) and LASSO (Panel B). Each classifier performs well in terms of accuracy. In particular, all but one classifier has an accuracy greater than 70 percent. Pakistan's RF classifier is 94 percent accurate, 13 percentage points higher than the accuracy of Pakistan's LASSO classifier. In contrast, Nigeria's RF classifier is 84 percent accurate, only four percentage points more than Nigeria's LASSO classifier. Within RF, Kenya and Pakistan's models have the highest accuracy (94 percent), and within LASSO, Kenya's model is the most accurate (84 percent).

In terms of each classifier's ability to correctly predict individual classes, each classifier tends to classify the financially included class with few exceptions successfully. This is seen by comparing the sensitivity to specificity within each algorithm. In the RF algorithm, both sensitivity and specificity are approximately greater than 75 percent, implying that on average, across countries, 84 percent of the individuals classified as financially excluded are financially

Table 1: Performance Evaluation

Panel A: Random Forest

COUNTRY	Accuracy	Balanced Accuracy	Sensitivity	Specificity	AUC ROC
Kenya	0.94	0.94	0.98	0.90	0.99
Tanzania	0.90	0.90	0.91	0.89	0.94
Uganda	0.86	0.86	0.82	0.89	0.89
Bangladesh	0.78	0.78	0.74	0.81	0.87
Indonesia	0.76	0.76	0.74	0.79	0.84
Pakistan	0.94	0.94	0.88	0.99	0.99
Nigeria	0.84	0.84	0.81	0.87	0.92

Panel B: LASSO

COUNTRY	Accuracy	Balanced Accuracy	Sensitivity	Specificity	AUC ROC
Kenya	0.84	0.84	0.81	0.87	0.91
Tanzania	0.80	0.80	0.75	0.86	0.87
Uganda	0.81	0.81	0.73	0.88	0.86
Bangladesh	0.67	0.67	0.68	0.66	0.73
Indonesia	0.70	0.70	0.67	0.74	0.78
Pakistan	0.81	0.81	0.77	0.84	0.89
Nigeria	0.80	0.80	0.79	0.81	0.89

excluded (sensitivity), and 88 percent of individuals classified as financially included are financially included (specificity). Similarly, for LASSO, 74 percent of the individuals classified as financially excluded are financially excluded (sensitivity), and 81 percent are financially included (specificity). Lastly, notice that for both algorithms, specificity is greater than sensitivity, on average, across countries. This implies that the algorithms are better at predicting financially included individuals based on the included, observable predictors.

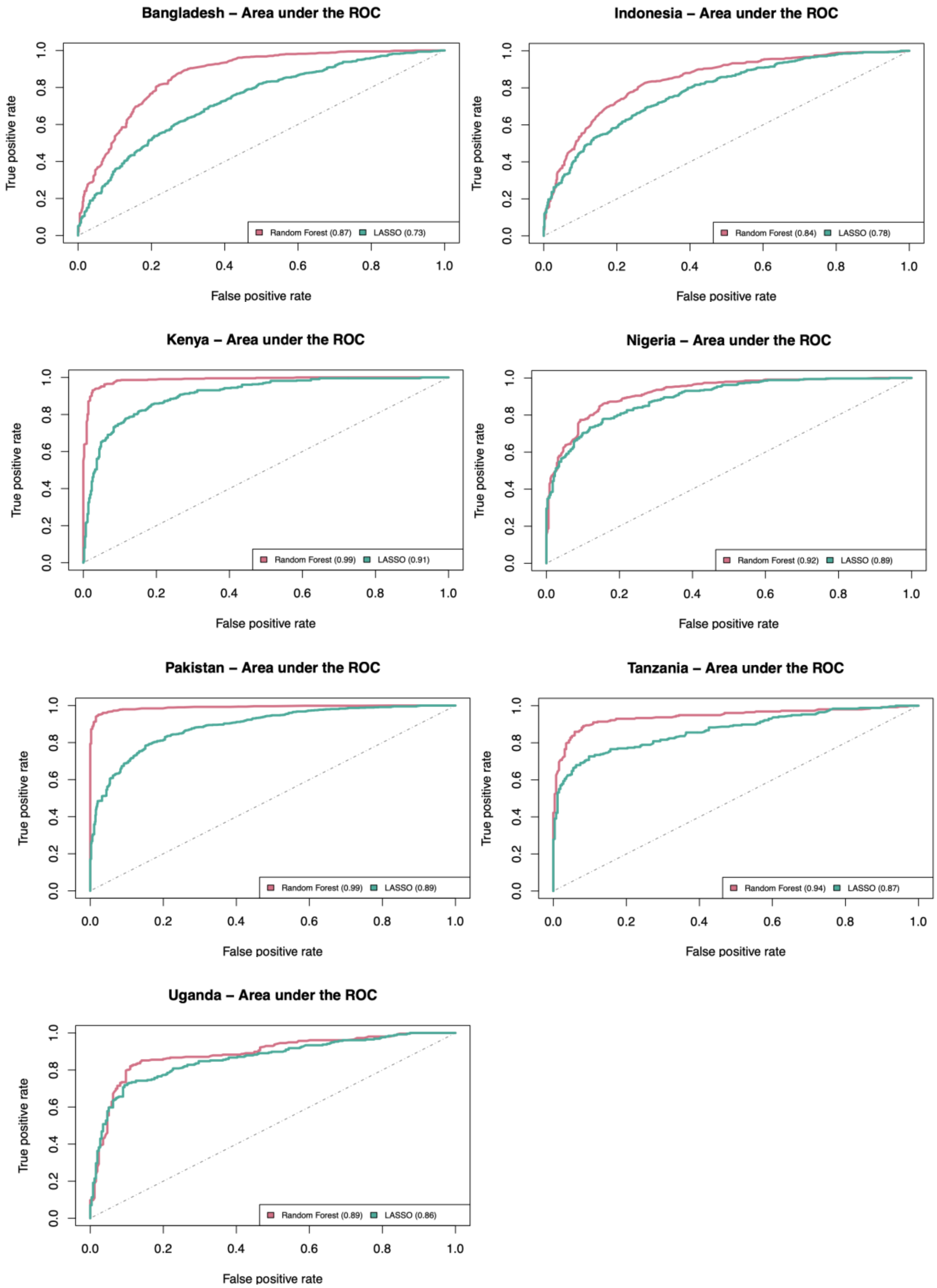
Finally, we compare each country's ROC curve for RF and LASSO. Figure 4 shows the plots for each country. Visually, we see that the RF and LASSO perform similarly in Nigeria and Uganda, but in all other countries, RF seems to dominate. We statistically test these differences using Robin et al. (2011), which tests for a difference in the AUC of the ROCs. The results indicate a statistically significant difference at the 95 percent level between the RF and LASSO for all countries. Thus, the RF classifier outperforms the LASSO classifier across all metrics.

RF is not an automatically better classifier than LASSO across all datasets simply because it is more flexible[17]. Our result suggests that using a richer dataset with a larger sample size than other financial inclusion datasets provides a better-predicting model that can provide insight into the complex interactions involved in financial inclusion. Given that the RF classifier is more accurate overall and outperforms LASSO when predicting individual classes, our primary focus will be describing importance values and ranking across countries for the RF classifier. Moreover, we prefer the conditional PI measure of variable importance as it considers the correlation structure between a given predictor and all the other predictor variables, which reflects a more accurate impact of each predictor (Strobl et al. 2008).

---

[17] For example, in Knippenberg et al. (2019), the performance of LASSO and Random Forest on their out-of-sample data was comparable, with LASSO slightly outperforming Random Forest.

Figure 4: Area Under the ROC Curve



5.2 Cross-country Comparison of Important Predictors

To gain insight into similarities across countries in predicting financial inclusion, we compare each country's top ten important predictors for a given algorithm. Figure 3 and Appendix Figure A1 show the results for RF and LASSO, respectively. The RF important predictor values are calculated using conditional permutation importance (conditional PI) (Strobl et al. 2008), and the LASSO important predictors were derived as the non-zero predictors re-run using logistic regression post-LASSO. The vertical axis shows the subset of top predictors that appear in at least two countries. These predictors are sorted by the number of times they appear in the ten most important predictors across all countries [18]. The circles' coloring indicates each variable's rank within each country. Darker colors have a higher rank than lighter colors. The size of the circle represents the mean decrease in accuracy when that predictor is conditionally permuted and indicates the extent of importance relative to other predictors both within and between countries. Lastly, the countries are sorted by the GSMA's Mobile Money Penetration Index (MMPI) (Andersson-Manjang 2021) in descending order from left to right[19].

For the RF classifiers, we find that the two most important predictors of financial inclusion are distance to point-of-service (POS) between 1 and 5 km and concerns about financial service providers (FSPs) keeping private information safe. Although these two predictors are in the top ten predictors for all countries, distance to POS is the top predictor in five of the seven countries and the second top predictor in the other two countries. In four of the seven countries, privacy is the most or second most important predictor. Additionally, in all

[18] This figure excludes top ten variables that are only present in a single country to conserve space.

[19] Appendix Table A2 shows the MMPI category of each country.

Figure 5: Random Forest Conditional Permutation Importance

countries except Uganda and Pakistan, a distance to POS of less than 1 km is also an important predictor.

To our knowledge, the financial inclusion literature has not explicitly incorporated an individual's distance (or time) to the point of service, particularly from a demand-side perspective, except for studies specifically considering mobile money. These studies have mixed results on proximity to mobile money agents. Riley (2018) finds that distance to the nearest agent does not impact the effectiveness of using mobile money to smooth shocks, while Jack and Suri (2014) find that proximity to mobile money agents positively impacts household risk-sharing. More general studies on banks or multiple formal accounts have included proxies, such as time to grocery (Honohan and King 2012) and bank and ATM penetration at the country level (Allen et al. 2016). Similar to the mobile money literature, these studies have mixed results. Honohan and King (2012) do not find that time to grocery is a significant variable associated with financial inclusion, while, in separate specifications, Allen et al. (2016) find that both bank and ATM penetration levels are positively and significantly associated with financial inclusion. Our results support the evidence that distance to the point of service matters. Although researchers have hypothesized this association, our study confirms this hypothesis from an individual-level, demand-side perspective. Additionally, this finding tells us that despite increased access to formal accounts via mobile money, financial infrastructure continues to be an important predictor of financial inclusion.

Even if proximity is not a barrier to inclusion, users may be more likely to voluntarily be excluded from formal financial services if they cannot trust their FSP. Without regulations and protections, consumers may be less likely to be financially included or be active in their account use. Consumer protection and financial sector regulations are priorities for policymakers. The literature has considered consumer protection indices and other measures from a supply-side perspective; however, if consumers lack knowledge of these protections or are risk-averse, they may be unwilling to engage with the formal financial system. Given that trust in FSPs is typically highly ranked amongst the top predictors, it suggests that whether or not the consumer trusts an FSP is an important predictor.

Further investigation on whether this result holds under variations in regulatory environments should be explored in future research. Preliminary exploration based on the GSMA's Mobile Money Regulatory Index (MMRI)[20] indicates that regulatory environments may not be highly correlated with formal account adoption, as the countries in our sample range from approximately 65 to 80 on an overall index range from approximately 30 to 90 (Bahia and Muthiora 2019). If a more careful analysis supports this, it would suggest that policymakers could implement policies that improve consumer trust, such as providing training on consumer protection regulations in the country (Garz et al. 2021; Galiani et al. 2022).

An individual's money source is another common top predictor across countries. In particular, money from family, friends, or spouses sent for regular support is among the top predictors in all countries except Pakistan and Uganda. This finding supports available evidence on risk sharing and the expansion of networks due to being financially included, more specifically, having a mobile money account (Jack and Suri 2014; Riley 2018). Additionally, neither Pakistan nor Uganda has any money source predictors as top predictors, but nearly all other countries have multiple money source predictors. Government assistance and ownership of

[20] See Appendix Table A2.

a business, property, or asset are top predictors in three and two countries, respectively, and employment by a private company is also a top predictor in Indonesia.

Although money source variables are not included in previous studies, income amounts are. Individuals with income in the lowest two quintiles are statistically and economically significant in Allen et al. (2016) and Klapper et al. (2021). Income is also significant in Honohan and King's (2012) study. Instead of a direct measure for income, the FII database has a poverty measure available. In the FII database, three poverty levels (extreme, poor, and above the poverty line) were derived using the Poverty Probability Index (PPI) based on ten questions about household characteristics and assets. Notably, the RF classifier only identifies being poor as a top predictor in Bangladesh and Tanzania, while, as previously noted, money sources are important variables in most countries.

Furthermore, employment status is also a significant variable in the literature; however, only "occasional" and "other" employment is important in just two countries. This suggests that an individual's source of income, particularly consistent sources of income, is a better predictor of account ownership than poverty (or income level) and employment status. Using these results, policymakers can continue to advocate for policies incentivizing or requiring specific income sources and government transfers to be deposited into formal accounts. Research has shown that mandatory deposits have additional benefits and provide access to formal financial services to the most vulnerable groups (Aker et al. 2016; Blumenstock et al. 2018; Field et al. 2021).

One final notable finding involves identification. This particular variable is notable because FII, until recently, was the only financial inclusion database that collected demand-side information for this important factor of financial inclusion. Without this variable, the literature has proxied identification using country-level variables such as know-your-customer (KYC) requirements and exemptions (Allen et al. 2016; Klapper et al. 2021) and has found that this predictor is not significantly associated with financial inclusion. In contrast, using our individual-level measure, identification appears as a top predictor for the RF classifier in two countries, Kenya and Uganda. Given that this essential enabler is an important predictor in just two countries, it speaks to the high levels of identification documentation due to the efforts of national identification campaigns.

5.3 Classification Error Evaluation

To understand the nature of errors that the black box model creates across countries, we compare the characteristics of correctly classified and misclassified individuals. We examine what type of individuals tend to be misclassified and whether there are any patterns of misclassification across countries. This exercise explores the misclassification of select variables.

First, let's consider individuals who own accounts. We consider if there are any systematic differences between those who are correctly classified and those who are incorrectly classified. Across the seven countries, there is little variation in identification. Over 80 percent of people in each country have some form of identification. Table 2 shows that the model performs similarly in correctly classifying and misclassifying account holders. One exception is in Kenya, where the model tends to classify account holders who do not have identification incorrectly. Whereas 97 percent of account holders who were classified correctly have identification, only 44 percent of those who were misclassified have identification. Alternatively, if we consider the classification performance within the non-account group, 90 percent of non-account holders that were incorrectly classified as account holders have identification. Other variables that do not have systematic differences between correct classification and incorrect classification are

Table 2: Random Forest Classification Analysis

Panel A: Account Holders (Differences Between True Included and False Excluded)									
	Female			Identification			Rural		
	True Included	False Excluded	Difference	True Included	False Excluded	Difference	True Included	False Excluded	Difference
Bangladesh	0.37 (0.02)	0.56 (0.05)	-0.18* (0.05)	0.99 (0.01)	0.99 (0.01)	0.00 (0.01)	0.68 (0.02)	0.70 (0.04)	-0.02 (0.05)
Indonesia	0.47 (0.02)	0.49 (0.05)	-0.02 (0.05)	0.99 (0.01)	0.97 (0.02)	0.01 (0.02)	0.36 (0.02)	0.44 (0.05)	-0.09 (0.05)
Kenya	0.44 (0.03)	0.51 (0.08)	-0.07 (0.09)	0.98 (0.01)	0.62 (0.08)	0.36* (0.08)	0.56 (0.03)	0.58 (0.08)	-0.01 (0.08)
Nigeria	0.39 (0.06)	0.28 (0.08)	0.12 (0.10)	0.97 (0.01)	0.74 (0.10)	0.23* (0.10)	0.55 (0.05)	0.68 (0.09)	-0.13 (0.10)
Pakistan	0.19 (0.01)	0.00 (0.00)	0.19* (0.01)	0.98 (0.01)	1.00 (0.00)	-0.02* (0.01)	0.60 (0.02)	0.19 (0.14)	0.41* (0.14)
Tanzania	0.42 (0.03)	0.48 (0.10)	-0.05 (0.10)	0.90 (0.02)	0.89 (0.07)	0.01 (0.07)	0.58 (0.03)	0.72 (0.09)	-0.14 (0.09)
Uganda	0.40 (0.04)	0.59 (0.12)	-0.19 (0.13)	0.95 (0.02)	0.86 (0.10)	0.09 (0.10)	0.65 (0.04)	0.80 (0.11)	-0.15 (0.12)
Panel B: Non-account Holders (Differences Between True Excluded and False Included)									
	Female			Identification			Rural		
	True Excluded	False Included	Difference	True Excluded	False Included	Difference	True Excluded	False Included	Difference
Bangladesh	0.54 (0.02)	0.34 (0.04)	0.20* (0.04)	0.96 (0.01)	0.99 (0.01)	-0.03* (0.01)	0.78 (0.02)	0.71 (0.04)	0.06 (0.04)
Indonesia	0.48 (0.03)	0.48 (0.04)	0.00 (0.05)	0.92 (0.02)	0.95 (0.02)	-0.03 (0.03)	0.53 (0.03)	0.35 (0.04)	0.18* (0.05)
Kenya	0.56 (0.03)	0.62 (0.21)	-0.06 (0.21)	0.49 (0.03)	0.82 (0.13)	-0.32* (0.13)	0.74 (0.02)	0.91 (0.09)	-0.17 (0.09)
Nigeria	0.53 (0.06)	0.41 (0.09)	0.12 (0.11)	0.74 (0.06)	0.95 (0.02)	-0.21* (0.06)	0.75 (0.06)	0.65 (0.08)	0.10 (0.10)
Pakistan	0.55 (0.02)	0.08 (0.03)	0.47* (0.03)	0.89 (0.01)	1.00 (0.00)	-0.11* (0.01)	0.71 (0.02)	0.54 (0.05)	0.17* (0.06)
Tanzania	0.59 (0.03)	0.45 (0.11)	0.13 (0.11)	0.63 (0.03)	0.92 (0.07)	-0.29* (0.08)	0.80 (0.03)	0.78 (0.09)	0.02 (0.10)
Uganda	0.53 (0.05)	0.38 (0.09)	0.15 (0.10)	0.74 (0.04)	0.96 (0.03)	-0.22* (0.05)	0.84 (0.03)	0.89 (0.07)	-0.05 (0.07)

Significance levels: *0.05

distance to the point of service, gender, and ownership of a mobile phone[21]. These results are consistent across learners. Conversely, account holders living in rural areas tend to be systematically misclassified across all countries except Pakistan.

Although the model performs well overall, understanding the systematic misclassifications can have important implications for individuals to become financially included. Relative to poverty and social welfare targeting, misclassification is less harmful but still significant in terms of costs.

6 Generalizability of learning models

The interest in harmonized data and analyzing across countries leads to considering how a country's fitted model performs on another country's data. In the machine learning literature, this is referred to as transfer learning. It allows for analyzing how generalizable and externally

[21] Results for these additional variables is available by request.

valid the results are. We examine if there are any patterns in countries where many are financially included or across countries that are geographically proximate or in proximity based on other characteristics.

The FII data is well-suited to carry out this exercise, given that the data collection was spearheaded by the same institution and harmonized. To study which results are generalizable, we consider the fitted models from Kenya, Uganda, and Bangladesh. Table 3 presents summary statistics. Most people in Kenya are financially included, so I compare Kenya and Uganda to make an intra-regional comparison. I also compared the fitted models with those in Bangladesh, which is geographically different. Ideally, the analysis will be extended to consider all combinations of countries in the data. Additionally, this analysis only uses the random forest algorithm[22].

Table 3: Transfer Learning Summary

	Training Set		Testing Set	
Country	Included	N	Included	N
Kenya	0.78	2219	0.78	556
Uganda	0.53	1924	0.53	483
Bangladesh	0.36	3876	0.36	970

Figure 6 demonstrates that the performance results vary across country pairs, and Table 4 provides additional evaluation metrics. Uganda's data applied to Kenya's fitted model performed well, and vice versa. These models performed well when evaluated using their own country's models. Although it may be expected that Kenya and Uganda's data perform well on each other's

[22] These analyses naively use the original sample sizes to compare fitted models. To improve the comparison, countries should be tested on models that were fitted with the same number of observations. Many of the datasets have very similar sample sizes, at least within regions, but following this procedure may provide more confidence in the comparisons.

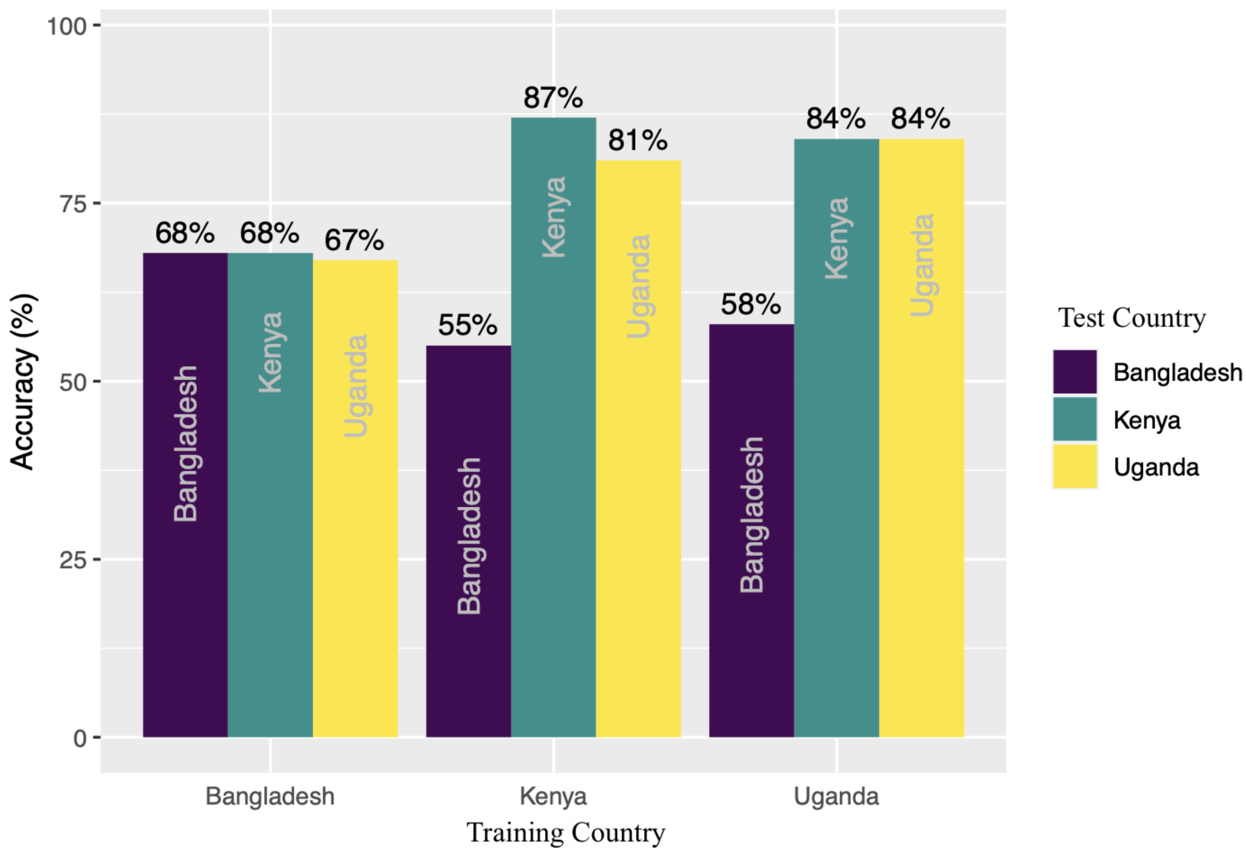


Figure 6: Random Forest Classifers Transferred Across Countries

Table 4: Transfer Learning Performance Evaluation

Model Fit Country	Model Test Country	Minority Class	Minority Class (%)	Sensitivity	Specificity
Kenya	Kenya	Excluded	0.22	0.61	0.94
Kenya	Uganda	Excluded	0.47	0.69	0.91
Kenya	Bangladesh	Included	0.36	0.87	0.37
Uganda	Uganda	Excluded	0.47	0.74	0.92
Uganda	Kenya	Excluded	0.22	0.59	0.92
Uganda	Bangladesh	Included	0.36	0.82	0.92
Bangladesh	Bangladesh	Included	0.36	0.40	0.84
Bangladesh	Kenya	Excluded	0.22	0.83	0.64
Bangladesh	Uganda	Excluded	0.47	0.91	0.46

models, further investigation should be considered, given that each model performs well with its own country's data.

When applied to Kenya and Uganda's models, Bangladesh's data performs relatively poorly in accuracy and specificity. Bangladesh's data is only 54 percent accurate on Kenya's model and 58 percent accurate relative to Uganda's model. When making this comparison, care is needed when evaluating sensitivity and specificity as Bangladesh is a country where financial account holders are a minority, whereas financial account holders are a majority in Kenya and Uganda. When evaluating Bangladesh's performance on Kenya and Tanzania's models, the models do a relatively good job correctly classifying non-account holders in Bangladesh (86 percent and 82 percent, respectively) but poorly classifying account holders correctly. Only 36 percent and 44 percent of Bangladesh's account holders are classified correctly. This seems like a poor fit; however, the performance of Bangladesh's data on its model is similar. Bangladesh's model only classifies 44 percent of its account holders correctly, relative to correctly classifying 87 percent of non-account holders. Finally, let's consider Kenya's and Uganda's performance when applied to Bangladesh’s model. Although both countries' accuracies are below 70 percent, their sensitivity and specificity are relatively high compared to all other transfer learning models. These results are presented for completeness, but this exercise may not be beneficial considering Kenya’s high financial inclusion rate.

7 Conclusion

As financial access becomes more widely available, researchers and practitioners seek more nuanced ways to reach individuals excluded from financial services and incentivize increased activity with financial account holders. In this study, we use a rich dataset of individual-level financial inclusion indicators to understand the underpinnings of financial inclusion in seven developing countries: Bangladesh, Indonesia, Kenya, Nigeria, Pakistan, Tanzania, and Uganda. By applying predictive machine learning algorithms, we find that distance to financial points of service, trust in financial service providers, and consistent sources of income, such as regular government-to-person and peer-to-peer transfers, are the most important predictors across our entire sample of countries. These findings contrast with prior analyses that find that demographic and socioeconomic characteristics consistently predict financial inclusion. This suggests data captured in individual-level questionnaires is meaningful for understanding priority enablers and barriers to financial account access. Finally, these results provide policymakers and industry practitioners with evidence of actionable ways to increase financial inclusion.

8 References

Aker, Jenny C., Rachid Boumnijel, Amanda McClelland, and Niall Tierney. 2016. "Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger." *Economic Development and Cultural Change* 65 (1): 1–37. https://doi.org/10.1086/687578.

Allen, Franklin, Asli Demirguc-Kunt, Leora Klapper, and Maria Soledad Martinez Peria. 2016. "The Foundations of Financial Inclusion: Understanding Ownership and Use of Formal Accounts." *Journal of Financial Intermediation* 27 (July):1–30. https://doi.org/10.1016/j.jfi.2015.12.003.

Andersson-Manjang, Simon K. 2021. "The Mobile Money Prevalence Index (MMPI): A Country-Level Indicator for Assessing the Adoption, Activity and Accessibility of Mobile Money." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3935919.

Ayyagari, Meghana, Asli Demirguc-Kunt, and Vojislav Maksimovic. 2017. "SME Finance." *World Bank Policy Research Working Paper,* 8241, . https://documents1.worldbank.org/curated/en/860711510585220714/pdf/WPS8241.pdf.

Bahia, Kalvin, and Brian Muthiora. 2019. "The Mobile Money Regulatory Index." GSMA. https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/wp-content/uploads/2019/03/The-Mobile-Money-Regulatory-Index-1.pdf.

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. 2015. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics* 7 (1): 1–21. https://doi.org/10.1257/app.20140287.

Barajas, Adolfo, Thorsten Beck, Mohammed Belhaj, and Samy Ben Naceur. 2020. "Financial Inclusion: What Have We Learned So Far? What Do We Have to Learn?" *IMF Working Papers* 2020 (157). https://doi.org/10.5089/9781513553009.001.A001.

Beck, Thorsten, Asli Demirgüç-Kunt, and Patrick Honohan. 2009. "Access to Financial Services." *World Bank Research Observer* 24 (1): 119–45. https://doi.org/10.1093/wbro/lkn008.

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. 2015. "Predicting Poverty and Wealth from Mobile Phone Metadata." *Science* 350 (6264): 1073–76. https://doi.org/10.1126/science.aac4420.

Blumenstock, Joshua, Michael Callen, and Tarek Ghani. 2018. "Why Do Defaults Affect Behavior? Experimental Evidence from Afghanistan." *American Economic Review* 108 (10): 2868–2901. https://doi.org/10.1257/aer.20171676.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. https://doi.org/10.1023/A:1010933404324.

Breza, Emily, Martin Kanz, and Leora Klapper. 2020. “Learning to Navigate a New Financial Technology: Evidence from Payroll Accounts.” w28249. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w28249.

Brown, Caitlin, Martin Ravallion, and Dominique Van De Walle. 2018. “A Poor Means Test? Econometric Targeting in Africa.” *Journal of Development Economics* 134 (September):109–24. https://doi.org/10.1016/j.jdeveco.2018.05.004.

Brownlee, Jason. 2020. “A Gentle Introduction to Threshold-Moving for Imbalanced Classification.” *MachineLearningMastery.Com* (blog). February 9, 2020. https://machinelearningmastery.com/threshold-moving-for-imbalanced-classification/.

Camner, Gunnar. 2012. “Mobile Money for the Unbanked Expanding the Ecosystem of Mobile Money: Considerations for Interoperability.” GSMA. https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/wp-content/uploads/2012/10/2012_MMU_Expanding-the-ecosystem-of-mobile-money.pdf.

Chen, Chao, Leo Breiman, and Andy Liaw. 2004. “Using Random Forest to Learn Imbalanced Data.” Technical Report. University of California, Berkeley.

Demirguc-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess. 2018. “The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution.” Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1259-0.

Dizon, Felipe, Erick Gong, and Kelly Jones. 2020. “The Effect of Promoting Savings on Informal Risk Sharing: Experimental Evidence from Vulnerable Women in Kenya.” *Journal of Human Resources* 55 (3): 963–98. https://doi.org/10.3368/jhr.55.3.0917-9077R2.

Economides, Nicholas, and Przemyslaw Jeziorski. 2017. “Mobile Money in Tanzania.” *Marketing Science*, November. https://doi.org/10.1287/mksc.2017.1027.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore. 2021. “On Her Own Account: How Strengthening Women’s Financial Control Impacts Labor Supply and Gender Norms.” *American Economic Review* 111 (7): 2342–75. https://doi.org/10.1257/aer.20200705.

Fisher, Aaron, Cynthia Rudin, and Francesca Dominici. 2019. “All Models Are Wrong, but Many Are Useful: Learning a Variable’s Importance by Studying an Entire Class of Prediction Models Simultaneously.” *Journal of Machine Learning Research : JMLR* 20:177.

Galiani, Sebastian, Paul Gertler, and Camila Navajas-Ahumada. 2022. “Trust and Saving in Financial Institutions by the Poor.”

Garbero, Alessandra, and Marco Letta. 2022. "Predicting Household Resilience with Machine Learning: Preliminary Cross-Country Tests." *Empirical Economics* 63 (4): 2057–70. https://doi.org/10.1007/s00181-022-02199-4.

Garz, Seth, Xavier Giné, Dean Karlan, Rafe Mazer, Caitlin Sanford, and Jonathan Zinman. 2021. "Consumer Protection for Financial Inclusion in Low- and Middle-Income Countries: Bridging Regulator and Academic Perspectives." *Annual Review of Financial Economics* 13 (1): 219–46. https://doi.org/10.1146/annurev-financial-071020-012008.

Hastie, Trevor, Robert Tibshirani, and Martin Wainwright. 2015. *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman & Hall/CRC. https://dl.acm.org/doi/10.5555/2834535.

Highet, Catherine, Nisha Singh, and Arisha Salman. 2021. "Can Free Phones Close the Digital Gender Divide?" February 25, 2021. https://www.cgap.org/blog/can-free-phones-close-digital-gender-divide, https://www.cgap.org/blog/can-free-phones-close-digital-gender-divide.

Hoens, T. Ryan, and Nitesh V. Chawla. 2013. "Imbalanced Datasets: From Sampling to Classifiers." In *Imbalanced Learning*, 43–59. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118646106.ch3.

Honohan, Patrick, and Michael King. 2012. "Cause and Effect of Financial Access: Cross-Country Evidence from the FinScope Surveys," December. https://doi.org/10.7551/mitpress/9517.003.0005.

Jack, William, and Tavneet Suri. 2014. "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution." *American Economic Review* 104 (1): 183–223. https://doi.org/10.1257/aer.104.1.183.

James, Gareth, Daniela Witten, Robert Hastie, and Robert Tibshirani. 2022. *An Introduction to Statistical Learning with Applications in R*. Vol. 6. Springer Science and Business Media. https://www.tandfonline.com/doi/full/10.1080/24754269.2021.1980261.

Jean, Neal, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell, and Stefano Ermon. 2016. "Combining Satellite Imagery and Machine Learning to Predict Poverty." *Science* 353 (6301): 790–94. https://doi.org/10.1126/science.aaf7894.

Johnson, Susan, and Steven Arnold. 2012. "Inclusive Financial Markets: Is Transformation Under Way in Kenya?" *Development Policy Review* 30 (6): 719–48. https://doi.org/10.1111/j.1467-7679.2012.00596.x.

Klapper, Leora, Jake Hess, Saniya Ansar, and Dorothe Singer. 2021. "Financial Inclusion and Financial Regulation." Washington, DC: Development Research Group, The World Bank. https://thedocs.worldbank.org/en/doc/2876cb5ebb0c3df2455ad372e3d89593-0050062022/original/FindexNote8-17.pdf.

Klapper, Leora, Jake Hess, and Mayada El-Zoghbi. 2016. “Achieveing the Sustainable Development Goals: The Role of Financial Inclusion.” CGAP. https://www.cgap.org/sites/default/files/researches/documents/Working-Paper-Achieving-Sustainable-Development-Goals-Apr-2016_0.pdf.

Knippenberg, Erwin, Nathaniel Jensen, and Mark Constas. 2019. “Quantifying Household Resilience with High Frequency Data: Temporal Dynamics and Methodological Options.” *World Development* 121 (September):1–15. https://doi.org/10.1016/j.worlddev.2019.04.010.

Koomson, Isaac, and Michael Danquah. 2021. “Financial Inclusion and Energy Poverty: Empirical Evidence from Ghana.” *Energy Economics* 94 (February):105085. https://doi.org/10.1016/j.eneco.2020.105085.

Lee, Jean N., Jonathan Morduch, Saravana Ravindran, Abu Shonchoy, and Hassan Zaman. 2021. “Poverty and Migration in the Digital Age: Experimental Evidence on Mobile Banking in Bangladesh.” *American Economic Journal: Applied Economics* 13 (1): 38–71. https://doi.org/10.1257/app.20190067.

Levine, Ross. 2005. “Chapter 12 Finance and Growth: Theory and Evidence.” In *Handbook of Economic Growth*, 1:865–934. Elsevier. https://doi.org/10.1016/S1574-0684(05)01012-9.

McBride, Linden, and Austin Nichols. 2018. “Retooling Poverty Targeting Using Out-of-Sample Validation and Machine Learning.” *The World Bank Economic Review* 32 (3): 531–50. https://doi.org/10.1093/wber/lhw056.

Morduch, Jonathan. 2020. “Why RCTs Failed to Answer the Biggest Questions about Microcredit Impact.” *World Development* 127 (March):104818. https://doi.org/10.1016/j.worlddev.2019.104818.

Mullainathan, Sendhil, and Jann Spiess. 2017. “Machine Learning: An Applied Econometric Approach.” *Journal of Economic Perspectives* 31 (2): 87–106. https://doi.org/10.1257/jep.31.2.87.

Munyegera, Ggombe Kasim, and Tomoya Matsumoto. 2016. “Mobile Money, Remittances, and Household Welfare: Panel Evidence from Rural Uganda.” *World Development* 79 (March):127–37. https://doi.org/10.1016/j.worlddev.2015.11.006.

Popov, Alexander. 2017. “Evidence on Finance and Economic Growth.” *European Central Bank*, Working Paper Series 2115.

Provost, Foster, Tom Fawcett, and Ron Kohavi. 1998. “The Case Against Accuracy Estimation for Comparing Induction Algorithms.” *ICML* 98:445–53.

Riley, Emma. 2018. “Mobile Money and Risk Sharing against Village Shocks.” *Journal of Development Economics* 135 (November):43–58. https://doi.org/10.1016/j.jdeveco.2018.06.015.

Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. 2011. "pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves." *BMC Bioinformatics* 12 (1): 77. https://doi.org/10.1186/1471-2105-12-77.

Roessler, Philip, Peter Carroll, Flora Myamba, Cornel Jahari, Blandina Kilama, and Daniel Nielson. 2021. "The Economic Impact of Mobile Phone Ownership: Results from a Randomized Controlled Trial in Tanzania." *Centre for the Study of African Economies*, CSAE Working Paper Series, , April.

Sahay, Ratna, Amina Lahreche, Sumiko Ogawa, Kimberly Beaton, Purva Khera, Majid Bazarbash, and Ulric Von Allmen. 2020. "The Promise of Fintech." *Departmental Papers* 2020 (009): 1. https://doi.org/10.5089/9781513512242.087.

Staschen, Stefan, and Patrick Meagher. 2018. *Basic Regulatory Enablers for Digital Financial Services*. World Bank, Washington, DC. https://doi.org/10.1596/30275.

Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. 2008. "Conditional Variable Importance for Random Forests." *BMC Bioinformatics* 9 (1): 307. https://doi.org/10.1186/1471-2105-9-307.

Strobl, Carolin, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. 2007. "Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution." *BMC Bioinformatics* 8 (January):25. https://doi.org/10.1186/1471-2105-8-25.

Suri, Tavneet. 2017. "Mobile Money." *Annual Review of Economics* 9 (1): 497–520. https://doi.org/10.1146/annurev-economics-063016-103638.

Suri, Tavneet, Prashant Bharadwaj, and William Jack. 2021. "Fintech and Household Resilience to Shocks: Evidence from Digital Loans in Kenya." *Journal of Development Economics* 153 (November):102697. https://doi.org/10.1016/j.jdeveco.2021.102697.

Suri, Tavneet, and William Jack. 2016. "The Long-Run Poverty and Gender Impacts of Mobile Money." *Science* 354 (6317): 1288–92. https://doi.org/10.1126/science.aah5309.

Swets, J. A. 1988. "Measuring the Accuracy of Diagnostic Systems." *Science (New York, N.Y.)* 240 (4857): 1285–93. https://doi.org/10.1126/science.3287615.

Tibshirani, Robert. 1996. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1): 267–88. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Townsend, Robert M. 1994. "Risk and Insurance in Village India." *Econometrica* 62 (3): 539–91. https://doi.org/10.2307/2951659.

Weiss, Karl, Taghi M. Khoshgoftaar, and DingDing Wang. 2016. "A Survey of Transfer Learning." *Journal of Big Data* 3 (1): 9. https://doi.org/10.1186/s40537-016-0043-6.

World Bank. 2019. "World Development Indicators: Female Labor Force Participation Rate." https://data.worldbank.org.

World Bank. 2021. "Lessons from Implementing a National Financial Inclusion Strategy." World Bank. https://doi.org/10.1596/36213.

World Bank. 2024. "World Bank Country and Lending Groups." 2024. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

9 Appendix

Table A1: List of Variables

Total Number of Variables: 65	
Financially Included	Phone Access: None
Age Group: 15 to 24	Phone Access: Borrows
Age Group: 25 to 34	Phone Access: Owns
Age Group: 35 to 44	Phone Smart
Age Group: 45 to 54	Phone Basic Last: Never
Age Group: 55 plus	Phone Basic Last: More than 90 days
Female Gender	Phone Basic Last: Past 30 days
Married	Phone Basic Last: Past 7 days
Edu: No formal education	Phone Adv Ever
Edu: Primary education	Able Basic: Call and SMS
Edu: Secondary education	Able Basic: Call only
Edu: Higher education	Able Basic: None
Edu: Other	Able Menu
Identification	Able Internet
Female Household Head	Distance to point of sale: Greater than 5 km
Household Head	Distance to point of sale: Between 1 and 5 km
Reading Literacy	Distance to point of sale: Less than 1 km
Rural	Female Empowerment: Household Spending
Poverty: Extreme poverty	Female Empowerment: Influence Some
Poverty: Poor	Female Empowerment: Self Advocacy
Poverty: Above poverty line	Female Empowerment: Own Spending
Employment: Unemployed	Female Empowerment: Financial service provider Privacy
Employment: Occasionally	Empowerment: Household Spending
Employment: Seasonally	Empowerment: Influence Some
Employment: Self employed	Empowerment: Self Advocacy
Employment: Employed with regular salary	Empowerment: Own Spending
Employment: Other	Empowerment: Financial service provider Privacy
Employment: Disabled or retired	
Employment: Stay at home	
Employment: Student	
Farm Owner	
Farm Worker	
Money Source: Ag	
Money Source: Govt Asst	
Money Source: Personal	
Money Source: Ownership	
Money Source: Govt Emp	
Money Source: Private Emp	

Table A2: Country-level Indicators

Country	WB Income Class	MMPI	MMRI
Indonesia	lower-middle	medium	-
Nigeria	lower-middle	medium	65.57
Pakistan	lower-middle	medium	80.65
Bangladesh	lower-middle	high	70.70
Kenya	lower-middle	very high	79.24
Tanzania	lower-middle	very high	78.17
Uganda	low	very high	72.50

World Bank Income Classification
MMPI - Mobile Money Penetration Index
MMRI - Mobile Money Regulatory Index

Figure A1: LASSO Variable Importance