# Assessing Incentives to Increase Digital Payment Acceptance and Usage: A Machine Learning Approach[*]

**Jeff Allen, Santiago Carbo-Valverde, Sujit Chakravorti, Francisco Rodriguez-Fernandez, and Oya Pinar Ardic[†]**

**January 18, 2022**

## Abstract

An important step to achieve greater financial inclusion is to increase the acceptance and usage of digital payments. Although consumer adoption of digital payments has improved dramatically globally, the acceptance and usage of digital payments for micro, small, and medium-sized retailers (MSMRs) remain challenging. Using random forest estimation, we identify 14 key predictors out of 190 variables with the largest predictive power for MSMR adoption and usage of digital payments. Using conditional inference trees, we study the importance of sequencing and interactions of various factors such as public policy initiatives, technological advancements, and private sector incentives. We find that in countries with low POS terminal adoption, killer applications such as mobile phone payment apps increase the likelihood of P2B digital transactions. We also find the likelihood of digital P2B payments at MSMRs increases when MSMRs pay their employees and suppliers digitally. The level of ownership of basic financial accounts by consumers and the size of the shadow economy are also important predictors of greater adoption and usage of digital payments. Using causal forest estimation, we find a positive and economically significant marginal effect for merchant and consumer fiscal incentives on POS terminal adoption on average. When countries implement financial inclusion initiatives, POS terminal adoption increases significantly and MSMRs' share of P2B digital payments also increases. Merchant and consumer fiscal incentives also increase MSMRs' share of P2B electronic payments.

JEL Classifications: D4, E4, G2, O3

Key Words: Payments, Incentives, Financial Inclusion, Regulation, Machine Learning, MSME,
Tax Policy, Shadow Economy

---

# 1    Introduction

In 2013, the World Bank Group President announced the Universal Financial Access (UFA) goal, which stresses that adults globally should have access to a transaction account to safely store money, send and receive payments as the basic building block to manage their financial lives.[1] Benoît Cœuré (2019) further elaborated on payments serving as a gateway to other financial services. As individuals make several payments daily to purchase food items and other basic needs, generally, the use of payment services is the first time an individual can potentially be introduced to the regulated financial sector. Based on their needs to mitigate risks and invest for the future, individuals can then start using other financial services such as savings, insurance and credit, typically tailored to their needs, ideally delivered responsibly and sustainably in an affordable way. Financial inclusion has been recognized as an important policy goal internationally for more than a decade.

Global Findex 2017 data, which is used to measure progress towards UFA, reports that account ownership at a regulated institution has increased: in the six-year period between 2011-2017, there were 1.2 billion new accountholders globally. Despite greater transaction account ownership, challenges remain as 1.7 billion adults lack access to a basic transaction account.[2] Importantly, nonbanks are increasingly playing a larger role in the provision of transaction accounts by increasing access to the unbanked.

---

[1] In 2016, Committee on Payments and Market Infrastructures (CPMI) and the World Bank (2016) provided a framework to achieve UFA.

[2] Progress has also been uneven: globally, 72 percent of men and 65 percent of women have an account—reflecting a gender gap that has remained unchanged since 2011 and 2014. Additionally, the gap between the richer 60 percent and the poorer 40 percent, in terms of account ownership, has remained the same throughout the period. The data shows that the unbanked are predominantly women, poor, not well-educated and unemployed.

Global Findex database also tracks how frequently transaction accounts are used to make digital payments.[3] The indicator "made or received digital payments" is calculated based on one payment made or received with the account owned. In 2017, 52 percent of adults globally made or received digital payments as opposed to 42 percent in 2014. Finally, the findings of Global Findex 2017 indicate that a significant proportion of accounts remain inactive for at least 12 months. Globally, one in five account owners had an inactive account during the past 12 months, while in India this figure was as high as one in two.

These findings suggest that policies that solely increase transaction account ownership are not sufficient to increase usage of digital payments. Account ownership must be coupled with greater opportunities for account owners to use their transaction accounts for making and receiving digital payments. Account owners need to have opportunities to use their digital payment instruments for everyday purchases, bill payments, online payments, and government payments. In addition, they would also benefit from digitally receiving their salaries and payments from government, businesses, and other individuals.

In this paper, we assess the effectiveness of various public and private sector initiatives including financial inclusion, better tax collection, and adoption of new technologies on digital payment acceptance by micro, small, and medium retailers (MSMRs) and the increase in the number of digital payments. In 2016, the World Bank commissioned a study to analyze the size of cash versus digital transactions made and received by MSMRs to better understand whether there are sufficient opportunities for consumers to use their transaction accounts to make payments at small, everyday merchants (World Bank Group, 2016). This issue was considered important because it is through these everyday merchants that retail payment solutions become valuable to

---

[3] Demirguc-Kunt *et al.*, (2017) provides empirical evidence regarding the benefits of digital payments including safety, quicker receipt of payments, and lower costs.

consumers and use of electronic payment instruments become habitual, generating an anchor for them in the regulated financial sector.[4] The study estimated that MSMRs globally made and accepted USD 34 trillion worth of payments annually, of which only 44 percent were made digitally suggesting that there is a USD 19 trillion opportunity.

We use machine learning techniques to evaluate the effectiveness of public and private sector initiatives along with technological advancements to increase the acceptance of digital payments by MSMRs and to increase the volume of digital payments at the point of sale. We identify key conditions and incentives that predict acceptance and usage of digital payments by MSMRs. Importantly, the order of implementation (or a specific combination) of conditions and the types of incentives matter. With large amounts of data and of covariates, we can only achieve this sequencing using machine learning. Furthermore, we do not limit our analysis to merchant incentives but also consider incentives given to other participants in the payment ecosystem along with improvements to the payment and telecommunication infrastructures that promote digital payment adoption and usage.

Our dependent variables are point of sale (POS) terminal adoption by MSMRs needed for payment card acceptance and the share of digital payments made to MSMRs by individuals.[5] Unfortunately, POS terminal adoption only captures the acceptance of payment cards. However, the volume of digital payments captures all forms of digital payments. While the volume of digital payments does not directly capture the adoption of acceptance infrastructure, it does indirectly capture the adoption of merchants of digital payments beyond card payments. Given the large

---

[4] While MSMRs' acceptance of digital payments is important for consumers, it is also important for these merchants to be able to make digital payments in the form of salaries or supplier payments. Therefore, the World Bank study also looked at business-to-business (B2B – in the form of immediate supplier payments) and business-to-person (B2P) payments. In other words, achieving greater digital liquidity is the goal of many policymakers.

[5] Amromin and Chakravorti (2009) use POS terminals to study the impact of payment cards on transactional cash demand in advanced economies. They find that POS terminal adoption does decrease transactional demand. However, in emerging and developing economies, bank account ownership is not always common.

number of covariates in our cross-country dataset, we use machine learning techniques to identify strong predictors without imposing a structure on the data as is the case for partial or structural standard econometric models. We analyze 81 country-level variables from 106 countries and 111 merchant-level variables for 576 merchants across seven countries. Using conditional inference trees, we identify combinations of predictors that increase the likelihood of increased acceptance and usage of digital payments. Finally, using causal forest estimation, we quantify the impact of different incentives on acceptance and usage of digital payments by comparing countries that have implemented to those that did not.

Our main results are as follows. Using random forest estimation techniques, we identify several factors that are strong predictors for greater POS terminal adoption and higher shares of person-to-business (P2B) electronic payments at MSMRs such as information and communication technologies (ICT) infrastructure, level of transaction account ownership in the economy, fiscal incentives, the size of the shadow economy, digitization of payment chain, and introduction of "killer applications."[6]

Using conditional inference trees, we are able to study these factors and predict a sequence of incentives or initiatives that enables greater acceptance and usage of digital payments. For example, if a country is above the median in ICT infrastructure and killer applications are implemented or the proportion of payment service agents are high, the share of P2B MSMR digital payments are predicted to be up to 60 percent higher.

Using causal forest estimation, we consider incentive implementation in different countries as a quasi-natural experiment. While some caveats should be considered as this is not a purely randomized experiment and, therefore, our results cannot be directly extrapolated to experiences

---

[6] In our sample, a killer application is defined as a successful mobile application that enables digital payments.

in jurisdictions not considered in our sample, the estimated treatment effects reinforce the idea that incentives to electronic payment acceptance (EPA) may have significant economic effects. In our setting, the treatment effects consisting in evaluating a number of country-level policies at merchant level. Some of these policies (e.g. cash limits) are applied in various countries while others are just applied in specific countries. Hence, our approach is connected to a strand of research that uses causal forest to evaluate cross-country public policies (see, for example, Tiffin, 2019), while the impact is evaluated at the individual level. In particular, we find that fiscal incentives, killer applications along with government policies to limit cash transactions or mandate digital payments for certain types of transactions are effective to increase merchant acceptance and usage of digital payments. In addition, we find that when killer applications are introduced in countries that do not have mass adoption and usage of payment cards, mobile payments are able to leapfrog payment cards as the dominant digital payment instrument at the point of sale (POS).

This article is structured as follows. In the next section, we discuss the current literature on payment merchant acceptance and usage. In section 3, we discuss the data and our empirical approach. In section 4, we discuss our results. In section 5, we discuss the policy implications of our findings. Finally, in section 6, we offer some conclusions.

## 2    Literature Review

Our paper contributes to various strands of the payments and non-payments literatures. Given that the market for payments has two distinct end-users—consumers and merchants— payment networks or platforms need to bring both sides on board for a transaction to occur. This aspect of payment services is often referred to as a two-sided market.[7] Generally, incentives on the

---

[7] For a discussion of two-sided markets, see Rochet and Tirole (2006) and Rysman (2009).

consumer side have been greater than the merchant side, however incentives to merchants continue to increase especially for MSMRs.

For the most part, this literature focused on payment cards. We extend the two-sided market literature by looking at much broader policy levers than regulating payment card fees, the structure of card networks, and the ability of merchants to pass on payment costs to consumers by imposing surcharges or discounts based on the payment instrument used.

As discussed, consumer access to transaction accounts and digital payment instruments continues to increase around the world. Several empirical studies focus on the how consumers choose to use digital payments instead of cash and checks (Klee, 2008, and Koulayev *et al*., 2016, and Rysman and Schuh, 2017). Cohen and Rysman (2013) find that transaction size and consumer demographic factors determine how consumers pay at grocery stores. In some countries, payment platforms bundle goods and services to encourage digital payments at the point of sale (POS), e.g. Alipay and WeChat Pay in China (World Bank, 2021).

Empirical research on the merchant side is more limited. As noted by Arango *et al.* (2015) most studies have ignored or only tangentially consider merchant acceptance. These authors stress the interactions between acceptance and demand. They highlight the challenges for payment platforms to attract critical mass of consumers and merchants. However, theoretical research on why merchants accept payment cards has been growing including the ability of merchants to increase sales to cash and credit-constrained customers and reduce the costs of safekeeping and transporting cash (Bolt and Chakravorti, 2008 and 2012).

There have been some payment adoption and usage studies that analyze how incentives affect both consumer and merchant usage of digital payments. Some of these incentives include merchants to steering consumers with differentiated prices based on payment instrument used or

by payment service providers offering usage rewards such as cash back. For example, Schuh *et al.* (2012) reviews some evidence that supports consumers responding more to merchant surcharges than to discounts depending on the payment instrument used. Shy and Stavins (2015) also examine how U.S. merchants used their ability to offer price discounts and other incentives to steer customers to pay with methods that are less costly to merchants.[8] Various studies have also looked at the impact of card rewards issued by issuers to increase card payments especially for credit card purchases (Agarwal, Chakravorti and Lunn, 2010 and Ching and Hayashi, 2008).

Other consumer steering incentives have been related to the adoption of mobile payment technology. For example, the case of rewards and/or cash back offered by card companies as well as other NFC payment providers (Apple Pay, Samsung Pay, and Android Pay) for adoption and usage of mobile payments. Zhao *et al.* (2017) find that these incentives have a positive effect on the decision to adopt NFC mobile payments. While these and other similar incentives may provide some interesting insights, more evidence would be needed on wider (public policy or private-public partnerships) attempts to increase merchant acceptance and consumer usage. Our study also attempts to complement this wider view on incentives especially merchant incentives.

We also contribute to the empirical literature on technology diffusion of new payment technologies. Technology has lowered costs, increased access, and expanded features. Mobile phone technology can be leveraged to make payments. Mobile phone technology along with quick response (QR) codes significantly reduce the cost of payment acceptance from traditional payment card technologies. In some cases, these technologies allow consumers and merchants to adopt mobile payments instead of payment cards. Han and Wang (2021) argue that mobile payments can be card-complementing or card-substituting. In countries with high adoption of payment cards and

---

[8] While not as common as before, some payment networks do not allow merchants to surcharge their branded payment instruments.

card acceptance infrastructure, mobile payments complement cards whereas in countries with low adoption of cards, mobile payments substitute for card transactions. Han and Wang (2021) construct a model where payment technologies arrive sequentially. Under certain cost assumptions, using simple empirical analysis, they explain why advanced economies have not embraced mobile payments unlike developing countries.

Additional incentives may be necessary to reduce the reluctance to use digital payments that are not due to the direct cost of acceptance. The reluctance to use digital payments may be due to non-payments related reasons such as tax evasion. Using a merchant survey of Indian small merchants in Jaipur, Ligon *et al.* (2019) considers other factors that prevent adoption of digital payments by merchants, including demand-side factors and taxes. Some researchers have also investigated a more direct intervention such as restricting cash use to reduce the size of the shadow economy or mandating electronic payments for certain types of transactions (Rogoff, 2016 and Schneider, 2019).

Finally, recent academic studies suggest that structural, as well as policy-related factors, such as channeling government benefit payments through transaction accounts, play an important role in improving financial inclusion (Barajas *et al.*, 2020). Account ownership is not only critical to access formal payments (Allen *et al.*, 2016), it also reduces corruption (Duryea and Schargrodsky, 2007), and increases consumption and productive investment of entrepreneurs (Dupas and Robinson, 2009). Additionally, shifting payments from cash to those using transaction accounts allows for more transparent and efficient payments especially for payments between individuals and governments (Demirguc-Kunt *et al.*, 2017). Government interventions such as demonetization intended to reduce the shadow economy may have had positive implications for adoption for digital payments (Agarwal *et al.*, 2020).

# 3    Empirical Methodology and Data

In this section, we describe our empirical approach and our dataset. We use machine learning algorithms to identify predictors, provide insights into sequencing and interactions of various incentives, and quantifying treatment effects of individual incentives. We also describe how we combine different sources to create our dataset.

## 3.1    Empirical Aim and Approach

Most existing studies in the area of digital financial services use discrete choice models to examine consumers' preferences for various types of payments and other financial services (Hernández-Murillo *et al.*, 2010; Honka *et al.*, 2017; Yusuf, Dauda and Lee, 2015). However, recent studies have shown that capturing changes in behavior from traditional choices to digital options involves several factors, including socio-economic, behavioral and institutional characteristics (Pousttchi and Dehnert, 2018). These multifaceted patterns suggest that a multidisciplinary approach combining economic, behavioral and data analytics is required to address such changes (Verhoef *et al.*, 2019). Machine learning methods are powerful data analysis tools that hold promise for generating new insights into payment behavior (Cui *et al.*, 2006; Lecun *et al.*, 2015; Witten *et al.*, 2019). Bajari *et al.* (2015) surveys a number of methods used in studies of changes in demand behavior and concludes that machine learning techniques are effective and, often, more powerful for this type of analysis.

The advantages of a machine learning approach where there are many potential influences at the macroeconomic and microeconomic levels, such as digital payments, has increased its use for such analysis. For example, Miguéis *et al.* (2017) uses a random forest model to discover

10

hidden patterns that may be valuable for decision-making in bank marketing. Machine learning approaches are also employed to estimate consumer preferences for technology products (Chen *et al.,* 2013), to examine travel choices (Hagenauer and Helbich, 2017), and, more generally, to model consumer response (Cui *et al.*, 2006).

Building on this growing literature, our empirical analysis employs machine learning techniques to analyze a combination of payment ecosystem (macro) and payment system participant specific (micro) factors. Payment ecosystem factors include improvements to payment and nonpayment infrastructure, national ID programs, technological advancements in communication and computing technologies, and policies to reduce the shadow economy. Payment system participant factors include government policies to increase transaction account ownership, payment service provider (PSP) incentives, fiscal incentives to consumers and merchants to adopt digital payments, and technological advances that allow PSPs to provide payment services to more consumers and businesses often at lower costs that impact electronic payment acceptance (EPA) and usage.[9]

The machine learning approaches that we employ carry three primary advantages over traditional parametric statistical methods. First, machine learning is a data-driven, bottom-up analytical approach or commonly referred to as "let the data speak" approach. It requires no preestablished or strict assumptions regarding the structure of the data or the functional relationships. Second, because we rely on tree-based machine learning mechanisms, our analysis may offer some insights on combinations of incentives and decision sequencing. While standard parametric models in econometrics (e.g. multilevel logit and propensity score matching) can identify behavioral patterns among clusters of individuals of firms, they are unable to offer

---

[9] These incentives are defined in line with previous World Bank research. See World Bank Group (2020) for a comprehensive survey.

predictions based on a sequence of actions. In other words, for some of the results obtained, the order of implementation (or a specific combination) of factors matters. In our context, previous empirical analyses and the policy experience in several countries seem to suggest that some actions or conditions can only be effective in promoting merchant EPA when other actions and conditions are already in place. For example, technology such as mobile phones or QR codes or payment infrastructure such as fast payment or card networks as a pre-condition for some payments to be accepted. However, when the number of actions and conditions is large (as in our setting) machine learning can be used to try to find sequences that increase EPA. Taken together, these first two advantages yield an analytical approach that uncovers patterns in real world interactions that induces greater adoption of merchant digital payment acceptance infrastructure and digital payment usage than what could be captured with a top-down statistical model that imposes a global structure on the data. Such an approach is also important to extract policy lessons for further merchant acceptance of electronic payments in countries with current low levels of acceptance as the global treatment of the data somehow elucidates a number of benchmarks or paths for these countries to follow.[10] Lastly, from a predictive standpoint, our machine learning approaches are more accurate than traditional econometric models used in the banking and payments literature.

However, the primary disadvantage of machine learning is that it typically emphasizes prediction over inference. Economists have traditionally been interested in studying what factors cause a change in a dependent variable as opposed to what factors predict the movement in a dependent variable. In our analysis, we are not only interested in what factors or preconditions predict digital payment acceptance and usage but also what incentives would change the behavior

---

[10] Given the treatment of data in our machine learning approach, it would not make much statistical sense to conduct specific sub-sample analyses for a smaller set of countries (e.g., low-income countries) as variability within a single group would not provide much information on benchmarking and potential routes for merchant acceptance development.

of merchants and consumers to increase their usage and adoption of digital payments. We partially overcome this limitation primarily because of data limitations by utilizing three complementary machine learning models, the respective features of which allow us to draw inferences about the drivers of digital payment acceptance.[11]

Specifically, our analysis follows a three-step process. In the first step, we use the random forest algorithm (Breiman, 2001) to analyze and narrow down the list of variables that are most important for predicting our dependent variables. For digital payment acceptance, we use POS terminal adoption because merchants generally must have POS terminals to accept payment cards.[12] Unfortunately, we do not have reliable data for non-card merchant acceptance infrastructure. To account for the growth of non-card-based payments, we also use MSMRs' share of P2B digital payments. While not a direct approach, the share of digital payments also captures the adoption of payment infrastructure because acceptance infrastructure must be in place before payments can be made. In addition, share of digital payments allows us to capture the usage of digital payments which is the eventual goal of policymakers and digital PSPs.

In the second step, we use the most important variables identified by the random forests to build conditional inference trees (Hothorn *et al*., 2006), which more specifically identify the sequential paths of adoption. In other words, we identify the likelihood of greater POS terminal adoption and usage based on a combination or a sequence of factors. These conditional inference trees isolate important sets of factors for prediction.

In the last step, we leverage the recently developed causal forest (Athey, Tibshirani, and Wager 2019 and Athey and Wager 2019)—an adaptation of random forests for more inferential

---

[11] Ideally, we would like to use a panel dataset where we could also study adoption over time within a country. Also as mentioned above, we would prefer greater merchant heterogeneity across the incentives that we study.
[12] Cards can be accepted in other means such as quick response codes but card acceptance at POS generally still requires a terminal.

purposes—to estimate treatment effects of relevant incentives on payment acceptance. Causal forests are frequently implemented as randomized experiments in a controlled sample. In these cases, causal forests may assist to sort the information and calibrate the average treatment effects when the number of covariates is large. For example, Jacob (2021) explores the impact of a randomized implementation of a microcredit program on the money borrowed and pension plan eligibility. While randomized experiments may be applied in the assessment of public policies, they cannot always be implemented. This occurs when the experimentation protocol is too expensive or complex. However, causal forests can also be used (with the necessary interpretation caveats) in the context of quasi-natural experiments. In particular, if some policies are exogenously implemented and identified as a potential shock or driver of change of certain behavior. We study the effects of groups that are treated or receive a type of incentive to increase EPA versus groups that do not receive that incentive. For the country-level sample, we study the impact of incentives on a treated country and on an untreated country. For the merchant-level sample, we study the impact on merchants within a treated country and merchants in an untreated country.[13]

In exploring the mechanisms underlying digital payment acceptance and usage, we take a general-to-specific approach. That is, we first analyze a "general mechanism," that consists of identifying and rank the predicting power of country-level indicators to identify incentive mechanisms broadly. In particular, we strive to explain the absolute or conditional weight of some factors on the success of digital payment acceptance and usage incentives. This exercise provides a general framework with respect to successful implementation steps for different incentives. From a data perspective, the country level data allows us to study 106 countries rather than only the seven countries that we have merchant survey data for.

---

[13] See Tiffin (2019) on the use of causal forest to evaluate policies at the country-level.

Second, in analyzing a merchant-level sample, we explore a more "detailed mechanism," which describes the individual path followed by merchants in accepting various forms of payments and identifies the role of some incentives specific to merchants). As with the country-level analysis, the merchant level analysis investigates the effects of incentives on EPA. The following subsections expand on the data, we use and the three sequential analytical approaches that we employ—the random forest, the conditional inference tree, and the causal forest.

## 3.2    Data

In this subsection, we describe how we construct our dataset. We combine a cross-section of country level explanatory variables with merchant survey data to identify and quantify the impact of incentives on electronic payment acceptance (EPA) infrastructure and usage.

### 3.2.1   Country-Level Data

The country-level data are cross-sectional and mostly from 2014 with some exceptions because data was not availabile for 2014. However, we do not believe these differences will affect the identification of merchant EPA. While our primary response variables for both the country-level and merchant-level data are POS terminal adoption and MSMRs share of digital payments, they are measured differently. At the country-level, we use POS terminals per 100,000 adults at year-end 2014, which we obtain from the World Bank's Global Payment System Survey (GPSS).[14] For MSMRs' share of P2B digital payments at the country-level, we use a country's estimated MSMRs' share of digital payments. World Bank (2016) (hereafter, the "Sizing Study") estimated the MSMRs' share of digital payments for mid-to-late 2015.

---

[14] As mentioned before, a natural limitation of modeling POS terminals is they focus primarily on card-based payments and do not capture the emerging non-card-based methods of electronic payments. P2B electronic payments, however, do capture all forms of electronic payments.

As discussed further below, the Sizing Study is based on primary research in seven representative countries. The research included in-person trade interviews and pulse surveys to complement existing research and data collection by Euromonitor International.[15] The trade interviews, which were conducted with government agencies, retail associations,

financial institutions, and non-bank financial service providers, provide a top-down perspective of the retail payments market. The pulse surveys, which were conducted with individual retailers and suppliers, provide a bottom-up perspective of the retail payments market. Information from both sources were consolidated to produce country-level estimates for the value and volume of P2B, business-to-business (B2B),[16] and business-to-person (B2P) payments by MSMRs such as paying workers. The estimates for the seven representative countries were paired by Euromonitor with selected assumptions and other macroeconomic and financial data to estimate these variables for 161 other countries using simulation methods.[17]

As already indicated, the country-level share of electronic P2B payments from the Sizing Study serves as one of our two primary country-level response variables. We use the B2B and B2P electronic payment estimations as predictor variables. We consider the whole MSMR payment chain because it has important implications for incentive design and effectiveness. In particular, incentives may encourage consumers to use bank accounts rather than cash alternatives. Digital payment acceptance by MSMRs is a critical factor to allow consumers to use digital payments. At the same time, MSMRs may also benefit from incorporating digital payment infrastructure to pay

---

[15] Detailed information on this survey is provided in World Bank Group (2016).

[16] In our dataset, B2B payments include immediate supplier payments only, and not the entire supply chain.

[17] Simulations of the Sizing Study conducted by Euromonitor are explained in Annex A4.3 of World Bank Group (2016). Several models were used for the simulation of the main variables. The 30 best performing models based on R2 were used to predict an outcome variable. Then, these 30 models are evaluated for the out-of-sample fit. The predictions from these 30 models for one outcome variable are then averaged to construct the final prediction for that outcome variable, thereby relying on out-of-sample predictions of the best-fitting 30 models for each of the 61 variables to be predicted.

salaries and suppliers digitally. Hence, policies such as mandated digital payments for MSMR B2B and B2P payments can be strongly connected to P2B digital payments.

The other country-level predictor variables are drawn from various sources. The aim is to obtain as much information as possible on potential determinants of digital payment acceptance and usage to feed the random forest algorithm and ultimately narrow down a core group of predictors. With some exceptions, most of our observations are from 2014, corresponding with the Sizing Study timeframe. A first reference for these variables is the GPSS, which surveys national and regional central banks and monetary authorities on the status of payment system development (e.g., e-money accounts per 1,000 adults and agents of payment service providers per 100,000 adults).

A number of variables are also drawn from the World Bank's Global Findex database. The database collects information on how adults save, borrow, make payments, and manage risk through nationally representative surveys of more than 150,000 adults in over 140 economies. The data employed are primarily from 2014. We also use the 2011 Findex database to compute the change from 2011 to 2014 for some variables (in particular, those reflecting payment usage).

A third database that we use in our analysis is the Global Financial Inclusion and Consumer Protection (FICP) Survey from the World Bank, which tracks the prevalence of key policy, legal, regulatory, and supervisory approaches for advancing financial inclusion and consumer protection, including national financial inclusion strategies, the issuance of e-money by nonbanks, agent-based delivery models, simplified customer due diligence, institutional arrangements for financial consumer protection, disclosure, dispute resolution, and financial capability. Financial sector authorities in 124 jurisdictions - representing 141 economies and more than 90 percent of the world's unbanked adult population - responded to the survey. The data were not available for 2014

and the closest date that they were available was 2017. We believe these variables are quite structural and the time difference does not affect our results.

We also use data from the Financial Access Survey (FAS) compiled by the International Monetary Fund (IMF). The FAS provides data on access to and use of financial services aimed at supporting policymakers to measure and monitor financial inclusion and benchmark progress against peers. FAS covers 189 countries spanning more than 10 years and contains 121 time-series on financial access and use. We use FAS 2014 data in our analysis.

We also use two other databases for very specific information on information technology development (the ICT Development Index, IDI, provided by the United Nations' International Telecommunications Union, ITU) and crime level information (from the United Nations Office for Drug and Crime, UNDOC).

In the end, 81 variables are selected for 106 countries, resulting in 8,586 cross-section datapoints. Appendix A1 provides the list of countries and variable definitions.

### 3.2.2 Merchant-Level Data

Merchant-level data is also cross-sectional and allows us to study merchant heterogeneity within MSMRs while allowing for some country heterogeneity. The merchant-level data is based on the Sizing Study primary research pulse surveys, which are introduced above. Primary research was conducted in seven countries in 2015—Colombia, France, Kenya, Lithuania, Morocco, Pakistan, and Turkey. Similar to our constructed country-level dataset, our two response variables capture POS adoption and P2B electronic payments. However, because these are merchant-level indicators, they take on different forms from the country-level data. The POS variable is a binary (0/1) indicator capturing whether an individual merchant has a POS terminal or not. The P2B

electronic payments indicator captures the estimated share of P2B payments made electronically at the individual retail establishment.

The Sizing Study database contains rich retailer-level data, which we use as predictor variables. These include indicators such as retailer size, customer profiles, merchant and consumer preferences, and whether retailers are part of a larger network, among other characteristics. As with the country-level data, we also use retailer-level data on electronic B2B and B2P payments as predictors. Further, we combine the merchant-level information with a number of country-level indicators similar to those employed in the country-level sample. In total, the merchant-level database consists of 576 merchants and 111 variables, resulting in 63,936 cross-section and time series datapoints. All the variables are defined in Appendix A1. While the merchant-level data corresponds to 2015 and the country-level data to 2014 we believe a one-year difference is not relevant for our empirical analysis.

We use the merchant-level data for two primary empirical purposes. First, we focus on the detailed merchant-level information for the seven countries to check if the machine learning results from the microeconomic structure are similar to those obtained at the country level. Second, we add information on policy incentives for the seven countries to analyze their impact on digital payment acceptance and usage.

In matching the different sources of information for the country-level and merchant-level data, we were particularly careful in using homogenous measures. In order to make the economic interpretations of the results more tractable and, at the same time, ensuring sufficient heterogeneity within the country- and merchant-level data, all quantitative variables in the database were transformed into four-level variables with values 1, 2, 3 and 4 corresponding to percentiles 0-25$^{th}$ (low), 25th-50$^{th}$ (lower mid), 50th-75$^{th}$ (upper mid) and 75th-100 (high). Additionally, a number

of variables contained some missing or not available (NA) values for some observations. However, the different algorithms employed in the machine learning techniques deal well with it treating them as missing values.

### 3.3    Identifying Variable Importance Using Random Forest

The random forest (Breiman, 2001) is a tree-based, recursive partitioning machine learning approach. Within the forest, each tree depends on the values of a random vector sampled independently and with the same distribution for all other trees. The algorithm splits trees with the goal of reducing impurity between clusters of observations. Impurity is generally measured by information content metrics, such as the Gini score and residual sum of squares. Ultimately, the random forest gathers hundreds or thousands of trees to make predictions.

In predicting POS terminal adoption and share of P2B electronic payments at both the country- and merchant-level, we feed the random forest models with all the country- and merchant-level predictor variables, respectively (see Appendix A1). This approach is common when data comes from surveys specifically designed to examine changes in banking and payment instruments (see Asadi *et al.*, 2017 and Kesharwani, 2019).

With the random forest estimation, our primary goal is to identify the variables that are most predictive of digital payment acceptance and usage. As explained further in the results section, the algorithm generates variable importance metrics for the model's predictor variables. These model diagnostics provide inferential value, but they lack directionality and economic interpretation, as they are not in the units of response variables. Thus, perhaps the most important function of the random forest is that it helps us narrow down a core group of variables that are most predictive of digital payment acceptance and usage. We are able to use these predictors to

more thoroughly explore the drivers of digital payment acceptance and usage by using the conditional inference tree and the causal forest.

**3.4     Decision Sequencing with Conditional Inference Trees**

After identifying predictors using random forest algorithms, we use the characteristics and determinants with the largest discriminant power to build a decision tree for each dimension by estimating a conditional inference tree (Hothorn *et al.*, 2006). This technique estimates a regression relationship through binary recursive partitioning in a conditional inference framework. In particular, the algorithm tests the global null hypothesis of independence between each of the input variables and the response and selects the input variable with the strongest association to the response. The algorithm then implements a binary split in the selected input variable and recursively repeats this process for each of the remaining variables. Importantly for our purposes, the conditional inference tree shows the sequential combination of factors that explain the EPA and usage decision process. This approach does not require any linearity assumptions, which is important because many of the digital payment acceptance and usage determinants could be related nonlinearly.

The conditional inference tree is similar in many ways to typical regression and classification trees, including those underlying random forests, in that it is a tree-based, recursive partitioning machine learning approach. A key difference, though, is that it is more statistical in nature, since it uses chi-square tests of independence to determine tree splits. As its name implies, the conditional inference tree is more geared toward "inference" than other tree-based methods. This technique provides insights into examining the tree structure in addition to focusing on predictive accuracy.

### 3.5 Estimating Impact of Incentives with Causal Forest Estimation

Since machine learning models have not, historically, been designed to estimate causal effects, a new field of study has emerged over the last few years that combines the advances from machine learning with the theory of causal inference (Athey, 2017). The aim of these techniques is to complement, rather than to serve as a substitute for, traditional machine learning methods, by helping researchers leverage the data-driven nature of machine learning to estimate causal effects (Athey and Imbens, 2016 and Wager and Athey, 2018). The main advantage of causal machine learning is that it can be used after the modeling phase in order to confirm some of the relations between predictors and the response variables.

In our context, by employing a causal learning method, we aim to examine the impact of incentives or conditions with the largest predictive power on the digitalization process. In a broad sense, it consists of comparing an outcome in a treated group (e.g., countries/merchants exposed to an incentive) with an untreated group (e.g., countries/merchants not exposed to an incentive). This departs to some extent from the sampling and treatment analysis in field experiments where sample selection and treatment implementation is under the control of the researchers, using randomization. In our setting, the treatment is associated with a quasi-natural experiment as we can combine a sample of merchants in different countries where some EPA incentives have been implemented as opposed to other countries or sub-samples of merchants with no access to such incentives. While the extension of our results to other jurisdictions not considered in our analysis should be done cautiously, we believe the causal forest analysis may help understand the impact

of the incentives and conditions that impact EPA and usage controlling for a large number of covariates at both the country and the merchant level.

Among the recent methods developed in the causal machine learning literature, causal forest has gained particular relevance (Athey *et al.*, 2019; Athey and Imbens, 2016; Wager and Athey, 2018). Knaus *et al.* (2020) shows that causal forests perform consistently well across different data generating processes and aggregation levels. The algorithm allows for a tractable asymptotic theory and valid statistical inference by extending the random forest algorithm. Methodologically, causal forests maintain the main structure of random forests, including recursive partitioning, subsampling, and random split selection. However, instead of averaging over the trees, causal forests allow for the estimation of heterogeneous treatment effects (Athey and Wager, 2019) by identifying how different treatments (e.g., incentive vs no incentive) affect the outcome (e.g., digital payment acceptance and usage). One important requirement for a proper identification is the so-called 'honesty' condition. This is the basic idea is that you cannot use the same outcome data to both partition the tree and estimate the average impact. This is particularly important when, rather than the standard randomization of samples, we use data to explore an exogenous change (i.e. policy) on a number of subsamples (Wager and Athey, 2018). The 'honesty' condition is satisfied in our sample.

Compared to a normal decision tree, the causal tree uses a splitting rule that explicitly balances two objectives: (1) finding the splits where treatment effects differ the most; and (2) estimating the treatment effects most accurately. In order to obtain consistent estimates of the treatment effects (in our case, the features that may have an impact on digital payment acceptance and usage), the algorithm splits the training data into two subsamples: a splitting subsample and an estimating subsample (Athey *et al.*, 2019; Wager and Athey, 2018). The splitting subsample is

used to perform the splits and grow the tree, while the estimating subsample is used to make predictions. All observations in the estimating subsample are dropped down the previously grown tree until they fall into a terminal node. Ultimately, the prediction of the treatment effects is given by the difference in the average outcomes between the treated and the untreated observations of the estimating subsample in the terminal nodes. Athey and Wager (2019) provide a full mathematical explanation on how causal forests are built for causal inference.

Using this empirical methodology, we are able to examine the impact of those features with the largest predictive power on digital payment acceptance and usage.[18] In running the algorithm, in the case of the country-level sample, we take a conservative approach by assuming that the level of digital payment acceptance and usage can be arbitrarily correlated within a country. Hence, the errors are clustered at the country-level.

## 4        Results

In this section, we discuss the results of the three empirical parts of our analysis—random forest, conditional inference trees, and causal forest.

### 4.1        Random Forest Results

Employing the random forest algorithm, we identify the best predictors for POS terminal adoption and MSMRs' share of P2B digital payments. From over the 190 factors, we identify 14 variables with the largest predictive power of MSMRs adoption of POS terminals and consumer usage of digital payments at MSMRs. The random forest algorithm generates a variable importance ranking. The relative statistical importance of each factor in predicting the impact on

---

[18] All analyses are carried out using the R package grf (Tibshirani *et al*., 2018).

the dependent variables is estimated. We measure the importance of each predictor by mean decrease in accuracy and mean decrease in Gini. The mean decrease in accuracy reflects the mean loss in accuracy when each specific variable is excluded from the algorithm. Therefore, the determinants and characteristics with the greater mean decrease in accuracy are the most relevant. Additionally, the mean decrease in Gini is a measure of how each feature contributes to the homogeneity between the decision trees used in the resulting random forest.

This analysis provides eight variable importance plots from multiplying two accuracy methods (mean decrease in accuracy and mean decrease in Gini) by two dependent variables (POS terminal adoption and share of P2B electronic payments) covering two samples (country-level and merchant-level). For simplicity, Table 1 offers the factors with the largest prediction power for the two dependent variables that are consistently shown at the country- and merchant-level sample.[19] In addition, Table 2 shows the predictors with mean decrease in accuracy larger than 10 percent and mean decrease Gini larger than 2 percent.[20]

<<<<<<<<<<<<INSERT TABLE 1 HERE>>>>>>>>>>>>>>

For POS terminal adoption, the main predictors correspond to three variable groups: merchant payment chain, ICT infrastructure and account ownership, and institutional and policy actions. Merchant payment chain includes the MSMRs' share of P2B digital payments, merchant perceptions on consumers payment instrument preferences, and the percentage of wages paid digitally at the merchant level. For actual usage of P2B digital payments at MSMRs, the merchant's perception of consumer willingness to use payment cards, and merchant usage of

---

[19] The detailed variable importance plots are available upon request.
[20] Furthermore, our selection is consistent with the procedure proposed by Han *et al.*, (2016). It consists of 1) running the random forest algorithm and returns the mean decrease in accuracy and the mean decrease in Gini of each variable, 2) ranking every variable using the mean decrease in accuracy and the mean decrease in Gini, respectively, 3) scoring each variable, 4) computing the total score of each variable, 5) reordering them by the total score.

digital payments to pay their workers are the main predictors of greater POS terminal adoption and greater share of digital payments. In addition, we might expect that over time, as the total MSMR's share of P2B payments increases, more MSMRs would adopt POS terminals given the popularity of payment cards. However, in some countries, digital payments not requiring POS terminals have leapfrogged payment cards.

For infrastructure, our results suggest that ICT infrastructure, transaction account ownership, and national ID programs are strong predictors of POS terminal adoption. ICT infrastructure increases the ability to open transaction accounts especially remotely and access accounts to make digital payments. Because debit cards are linked to transaction accounts, account ownership is necessary for the adoption of debit cards by both consumers and MSMRs. Not surprisingly, national IDs are a strong predictor of POS terminal adoption because IDs enable widespread ownership of transaction accounts which are necessary for consumer card adoption which is a critical factor for merchants when deciding to install POS terminals.

We also find institutional and policy actions taken by policymakers and the size of the shadow economy are important predictors of POS terminal adoption.[21] Public authorities have implemented financial inclusion programs to increase transaction account ownership often with access to debit cards. Fiscal incentives for merchants are also strong predictors for greater POS terminal adoption. Furthermore, payment of wages into transaction accounts is also a strong predictor of greater adoption of POS terminals by MSMRs. Finally, the size of the shadow

---

[21] Our empirical results regarding the impact of the shadow economy are robust to the use of some alternative measures of economic informality. In particular, our results remain very similar when we use alternative indicators of informality as the informality measures based on dynamic general equilibrium models and on the combination of multiple indicators, as provided by Ohnsorge and Yu (2021). Correlation across the 106 countries in our sample between our shadow economy metric and these economic informality indicators ranges from 87 percent to 89 percent.

economy is also a strong predictor MSMR POS terminal adoption. The larger the size of the shadow economy the lower the likelihood of POS terminal adoption.

We also identify strong predictors of the MSMRs' share of P2B electronic payments of which four of them are the same for POS terminal adoption. We categorize these predictors into four groups: merchant payment chain, payment instrument developments, ICT infrastructure and account ownership, policy variables, and access points. In the MSMR payment chain category, three variables are strong predictors of the MSMRs' share of P2B electronic payments: MSMRs' beliefs about consumer payment preferences, the percentage of the total value of electronic wage payments, and the proportion of the electronic payments made to suppliers. As discussed before, as MSMRs become more digitally liquid, they will tend to adopt digital payments for all incoming and outgoing payments. Given that cards are still a popular digital payment, POS terminal adoption continues to be a strong predictor of MSMRs' share of P2B digital payments.

In the payment instruments group, previous (debit and credit) card penetration (measured as the penetration in 2011) are strong predictors of greater the share of MSMRs' P2B electronic payments. This result suggests that consumers may require time to change their payment habits and merchants will install POS terminals as consumer demand increases over time. There may also be spillover effects between some merchant sectors into others. For example, in many countries, high-end merchants and merchants located in tourist locations are likely to be early adopters of payment cards. As consumer and merchant awareness of digital payments increases, the adoption of POS terminals by other types of MSMRs may also increase.

As expected, the level of development of ICT infrastructure is a strong predictor for the MSMRs' share of P2B digital payments. This result is more general than the result with POS terminal adoption because usage of noncard digital payment instruments is included. In the future,

we would expect that payment cards will continue to face greater competitive pressure from alternative payments such as those based on fast payment networks and closed-loop digital payment networks.

In the policy and access points categories, wages being paid in a transaction account, presence of a killer app, POS terminal adoption, and presence of agents for payment service providers are strong predictors for MSMR P2B digital payments. This result suggests that there are adoption and usage synergies between digital payments across the MSMRs' payment chain even if the payment instruments themselves may differ. In addition, workers that are receiving payments digitally may prefer to pay digitally at MSMRs if given the opportunity.

One key finding is that each of the dependent variables is a predictor for the other dependent variable. In the case of MSMRs' share of electronic P2B payments, POS terminal adoption provides a means to accept payment cards, the most popular digital payment option at MSMRs during our sample period. For predicting POS terminal adoption, the MSMRs' share of P2B electronic payments is a strong predictor. The result can be interpreted as a confirmation of the feedback mechanism, whereas POS adoption increases, more MSMRs adopt POS terminals. In other words, as more MSMRs install POS terminals, other MSMRs also adopt because if they do not, they may lose business to MSMRs that accept digital payments. However, as new digital payment instruments that do not rely on POS terminals have greater market penetration, we would expect this effect to lessen.

## 4.2 Conditional Inference Tree Results

While the random forest estimation techniques identified predictors in isolation, conditional inference trees allow us to study how a combination of factors and their sequence

increase the likelihood of POS terminal adoption and MSMRs' digital share of P2B payments. Based on the predictors from the variance importance analysis, we estimate conditional inference trees that identify the interactions among predictors and their sequences.

An example of a conditional inference tree is shown in Figure 1. In this example, we observe that if MSMRs believe that consumers prefer electronic payments and percent of wages paid electronically is greater than 70 percent, the likelihood of MSMR POS terminal adoption is 70 percent. Alternatively, if MSMRs believe that consumers do not prefer digital payments and less than 40 percent of wages are paid electronically by MSMRs, the likelihood of MSMR POS terminal adoption is only 10 percent. If consumers do not prefer electronic payments, the percentage of wages paid digitally is greater than 40 percent and if long-term fiscal incentives are implemented, the likelihood of MSMR POS adoption is 60 percent suggesting the importance of fiscal incentives especially in the absence of merchant beliefs that consumers do not prefer digital payments.

<<<<<<<<<INSERT FIGURE 1 HERE>>>>>>>

For simplicity, we offer a summary of the most impactful prediction relationships in Tables 3 and 4. These estimates are derived from numerous conditional classification tree analysis similar to Figure 1 that are not shown but available upon request. In Table 2, we summarize the conditional paths that have the most impact on MSMR POS terminal adoption. Interestingly, POS terminal adoption is 200 percent more likely when the MSMRs' share of electronic P2B payments and bank account ownership are above the median country (23.2 percent and 62 percent, respectively).[22] The median percentage of electronic P2B transactions at MSMRs is a relatively low threshold.

---

[22] These thresholds may be challenging for a number of countries within the sample. The 25th percentile for share of electronic P2B payments and bank account ownership are 9.2 percent and 44.2 percent respectively.

This result reinforces the dominant role of payment cards among the digital payment choices generally available.

We also find that if wages are paid electronically above the median (35 percent) and merchants believe that consumers prefer electronic payments is also above the median (52 percent), POS terminal adoption is 100 percent more likely. In addition, there may be a feedback loop whereby as more consumers prefer digital payments, more merchants install POS terminals leading to more consumers adopting and so forth.[23] Consumer preferences toward using digital payments have likely improved because of their access to transaction accounts and payment cards.

Also, if at least 35 percent of wages are paid by a given MSMR into a transaction account at a bank (the median in our sample) and ICT infrastructure is higher than the median country (above 5.3 on a 10-point scale)[24], POS terminal adoption is twice as likely. This result suggests that if merchants pay a sufficient number of workers digitally into a bank account (35% of more), they are more likely to accept card payments from their customers if there is a sufficient level of ICT infrastructure. Our results suggest that public policy should not only focus on P2B but also consider incentives to increase B2P payments such as wages.

We also find that with ICT infrastructure above 5.3 on a 10-point scale or national ID implementation being above median (94 percent) and the proportion of the shadow economy to the whole economy being no greater than 15 percent result in a 50 percent greater likelihood of MSMR POS terminal adoption. As we discussed above, ICT infrastructure is important for POS

---

[23] As consumers become more accustomed to using payment cards, the demand for merchants that did not previously adopt payment infrastructure increases. In addition, payment card acceptance may be a strategic tool to steal customers from merchants that do not accept cards. Unfortunately, our data does not allow us to study business stealing. For more on business stealing in a theoretical context, see Rochet and Tirole (2002).

[24] The 25th percentile values for wages paid in a financial institution account and the ICT index are 10 percent and 3.5, respectively.

terminal adoption, but other factors may be necessary. In this case, a national ID system enables greater ownership of transaction accounts debit card.

Furthermore, implementing a national financial inclusion strategy or merchant fiscal incentive initiative (at a national level) and having a proportion of the shadow economy to the whole economy below 20 percent results in the likelihood of POS terminal adoption increasing by 20 percent. The implementation of a financial inclusion strategy is another variable that captures a necessary condition for card ownership which in turn increases the likelihood of MSMR POS terminal adoption. The merchant fiscal incentive reduces the benefits of tax evasion.

<<<<<<<<<<<<INSERT TABLE 2 HERE>>>>>>>>>>>>>>

In Table 3, we report our results on the factors that impact the MSMRs' share of electronic P2B payments. When wages are paid digitally and card penetration in the previous five years are above the median (35 percent and 11 percent, respectively), the likelihood of P2B electronic payments increases by 100 percent. This result suggests that greater awareness by consumers in the form of greater access to payment cards or by greater payment of their wages digitally, increases the likelihood of digital payments increases substantially.

Given the popularity of cards as a digital alternative to cash, if POS terminal adoption is over the median (10,005 terminals per 100,000 adults) the predicted share of P2B digital payments increases by 60 percent. Alternatively, it is less likely that merchants will adopt POS terminals if they believe that consumers do not have access to them or will not use them.

If the level of ICT is over the sample median (5.4 out of 10) is combined with the development of killer apps, or with a significant use of agents of payment service providers (1.2 per 1,000 inhabitants), P2B electronic payments' likelihood increases by 50-60 percent. Killer apps allow payments to be made by mobile phones and may serve as an alternative to payment

cards. However, sufficient ICT infrastructure is likely necessary along with innovative mobile phone-based solutions to increase MSMR's P2B digital share of payments.[25] Also, the importance of agents suggests that digital payments do not immediately lead to digital liquidity, but cash agents are generally required for consumers and businesses to convert digital funds to cash and vice versa. We would expect a greater reliance on cash agents when significant parts of the population are unbanked or do not use banks, or a lack of merchant acceptance of digital payments.

When payments to direct suppliers of MSMRs are above the median sample value (45.9 percent) and the wages are paid electronically (over the median value), the MSMRs' share of P2B electronic payments is predicted to increase by 30 percent. This result suggests that digital payments in other parts of the merchant's payment chain likely increases usage of digital P2B payments.

<<<<<<<<<<<<INSERT TABLE 3 HERE>>>>>>>>>>>>>

## 4.3    Causal Forest Results

While the random forest and conditional inference trees offer insights into strong predictors of digital payment acceptance and usage along with their sequencing and interactions with each other, we also investigate causal effects. Specifically, we estimate the impact of a change in each explanatory variable on one of our two dependent variables. The treatment effect is given by the difference in the average outcomes between the treated and the untreated observations. The standard procedure is to compute the "average treatment effects" (ATE) when the treatment is defined based on a binary outcome as is the case for some incentives in our database. For example, if we compute the impact of financial inclusion policies on POS terminal adoption, we can estimate

---

[25] In the treatment effects section, our results suggest a negative relationship between leapfrogging and POS terminal adoption. However, card payments could also benefit from mobile phone technology and QR codes but this likely occurs when a robust card ecosystem already exists.

how POS adoption is impacted on average in countries with these policies compared with countries where no such policies exist, controlling for the rest of explanatory factors. Our analysis is not based on a randomized experiment but, instead, our analysis is a natural experiment based on exogeneous EPA incentive policies in various countries for different types of merchants.

We need to consider that our data also have a number of continuous variables. In this case, instead of the ATE, we compute the "average partial effect" (APE), which shows the percent change in the dependent variable due to a unit change in the treatment variables. Given the four-level (1 = low; 2 = mid-low; 3 = mid-high; and 4 = high) transformation of the continuous explanatory variables in our setting, the APE will show the average change in the dependent variable of going from one level of the treatment to the next one. For example, if we select the "ICT development index" as a treatment for POS adoption, we will be showing the average partial effect on POS adoption of a country moving from one level to a higher one (e.g., from mid-low to high ICT group).

As an initial test, we consider the top thirty explanatory factors in the variable importance plot as a reference. We find significant relationships for 23 variables. We group these variables into six categories: access points, economic formality, ICT infrastructure and account ownership, policy variables, instruments, and merchant payment chain. We run a causal forest for each one of the 23 variables and replicate the process for both dependent variables in the country-level and the merchant-level samples. All the results are shown as point estimates with 95 percent level confidence intervals.

In Figure 2, we report the results for the country-level sample where POS terminal adoption is the dependent variable. In the case of access points, the use of "mobile phone or the internet to access a financial institution account" is found to have a significant impact (16.2 percent) on

average across the four levels of POS terminal adoption and treated and untreated countries, suggesting that consumers' financial digitalization significantly influences merchants' adoption of POS terminals. Among more traditional channels, bank branches per capita in 2011 seem to have a more limited positive effect.

MSMRs may have a greater incentive to accept cash as a means to evade taxes. On the other hand, greater reliance on cash transactions may result in MSMRs and consumers. In the case of economic formality, the treatment effect of the "shadow economy over GDP" on POS adoption is, as expected, negative with a relatively tight confidence interval illustrating that a large shadow economy reduces the incentive for MSMRs to adopt POS terminals on average. The effect of "crime" (measured as robbery and assaults per inhabitant) is positive with very tight confidence interval suggesting that safety concerns for consumers and MSMRs leads to greater adoption of digital payment acceptance infrastructure.

Account ownership indicators are among those with the larger marginal impact on POS terminal adoption, suggesting that financial inclusion policies are impactful and should be encouraged. We capture long-term transaction account ownership at banks by considering 2011 data.[26] For short-term bank transaction account ownership, we use 2014. Interestingly, we find that the impact for long-term account ownership is higher than short-term account ownership suggesting that consumers need time to change from their preference for cash transactions at POS. As for technology-driven changes, while there is a positive APE on POS terminal adoption of mobile money accounts per capita possibly suggesting that similar infrastructures are used for noncard and card payments.

---

[26] This is a proxy for a time dimension. Changing consumers' and MSMRs' payment behavior requires time and these variables are able to provide some insights.

Two policy actions are economically significant and positive. These include mandates to pay wages into transaction accounts and the presence of a national financial inclusion strategy. The marginal impacts of both these public policy initiatives are significant.

From a price regulation perspective, we tried to capture the impact of regulating interchange fees, the fees that the merchant's bank pays the cardholder's bank and comprises the bulk of the merchant cost to accept payment cards. The two variables that we consider are: "authorities have taken actions or are considering taking action to address interchange fees" and "authorities consider interchange fees to be high." We would expect both of these variables to have negative impact on POS terminal adoption because high fees would deter merchant adoption of POS terminals. However, the treatment effects of these variables do not have the expected sign and are insignificant. This result suggests that interchange fees may not be the main deterrent for the lack of POS terminal adoption. However, we caution that these results are averages across countries in different stages of financial development. In some cases, interchange fee regulation occurs in countries where merchant adoption is near complete.[27] In other cases, interchange fee regulation occurs in countries where the adoption and usage of cards by consumers is very low.[28] This result is also consistent with the premise that there are other factors besides payment card fees that determine whether consumers will use payment cards even if merchants accept them.

Our next category is payment instruments. While debit and credit card ownership have a significant impact on POS terminal adoption, the latter seems to have a larger effect and to increase

---

[27] In these cases, interchange fee regulation is not likely to increase adoption by merchants or usage by consumers. In some cases, interchange fee regulation was implemented to reduce certain types of payment card transactions.

[28] In these cases, decreasing interchange fees may not provide incentives to consumers to adopt and use payment cards. Alternatively, the lack of a payment card infrastructure may result in new payment technologies being adopted such as noncard based mobile payments.

over time.[29] In many countries, credit cards are accepted at tourist locations and certain high-end stores well before adoption of debit cards by the masses. Consumer awareness and card infrastructure grows from credit card acceptance in certain sectors which in turn may allow for broader acceptance of debit cards.

As we have seen before, digital payment usage in other parts of the MSMR's payment chain positively impacts POS terminal adoption. The impact of the share of P2B, the share of B2B and the share of B2P electronic payments over total transactions on POS terminal adoption is positive and significant. The P2B results suggest a reinforcing feedback loop.

<<<<<<<<<<<<INSERT FIGURE 2 HERE>>>>>>>>>>>>>>

The impact on the MSMRs' share of P2B electronic payments at MSMRs for the country-sample are shown in Figure 3. For access points, not surprisingly, using the internet or mobile phone to access a transaction account, POS terminal per capita, and ATMs per capita in 2011 are significant and positively impact MSMR share of digital P2B payments. However, branches per capita is found to have a negative impact, suggesting the persistence of cash usage has a significant effect in countries with large bank physical networks also the lower shoe leather costs may help maintain the demand for cash transactions.

<<<<<<<<<<<<INSERT FIGURE 3 HERE>>>>>>>>>>>>>>

For the infrastructure variables, there is one main difference between the POS terminal adoption and MSMR digital share of P2B payments. The effect of ICT infrastructure for the latter is large and positive suggesting that all digital payments require sufficient ICT infrastructure.

---

[29] The link between payment cards and POS terminals should not be surprising since payment card acceptance for the most part requires a terminal. Alternative acceptance infrastructure have recently been introduced but the card form factor still remains the dominant one.

For government policy actions, the national financial inclusion strategy variable does not remain significant suggesting that other factors such as technological advancements can reduce the importance of a national financial inclusion strategy.[30] In addition, the percentage of population with a bank account is insignificant suggesting that bank accounts may not be a precursor in all countries for greater digital payment usage.

In the case of payment instruments, debit card ownership in 2011 does not remain significant while the other variables remain positive and significant. Given new data that captures noncard based payments, we would expect the importance for card adoption to decrease even more.[31]

The magnitudes for the payment chain variables increase suggesting that merchants may adopt other digital payments besides cards. As for the merchant payment chain, P2B payments seem also to be substantially driven by B2B payments and, in particular, by B2P payments.

<<<<<<<<<<<<INSERT FIGURE 4 HERE>>>>>>>>>>>>>>

The merchant-level sample allows us to compare treated and untreated merchants for each incentive. Because our incentives occur at the national level, our comparison will be between merchants receiving the incentive in one or more countries (treated group) against merchants not receiving the incentive in other countries (untreated group).

For expositional simplicity, we only report the impact of those incentives not previously included in the country-level sample (Figure 4). Implementing merchant fiscal incentives will increase the likelihood of POS terminal adoption. Consumer fiscal incentives also have a positive and significant effect. The impact is lower for mandated acceptance and for cash limits. We also

---

[30] While we discuss the impact of a considerable number of incentives, others (such as terminal subsidies) or lotteries could not be empirically analyzed due to data availability.
[31] Card networks are also processing other types of payments such as fast payments using their infrastructure.

find that in those cases where a leapfrogging strategy develops, e.g. a killer mobile app for payments, the effect is negative because POS terminal adoption may no longer be necessary. The adoption of dongles on mobile phones and quick response codes may eventually eliminate the need for terminal adoption.

<<<<<<<<<<<<INSERT FIGURE 5 HERE>>>>>>>>>>>>>>

Figure 5 reports the merchant-level effects of the different incentives on the share of P2B electronic payments. In this case, the effect of the killer app is positive and the largest among the incentives considered showing the effects of mobile-related technology adoption on payment usage. Merchant and consumer fiscal incentives are also found to have significant effects and the effect is also considerable in the case of mandated use of electronic payments. The impact of short-term and long-term cash limits are also positive but not as large.

Overall, the empirical analysis provides some generalized conclusions to increase merchant acceptance and usage of digital payments. First, merchant and consumer decisions are interlinked when it comes to adoption and usage of digital payments. Second, digital payment acceptance and usage are positively impacted when any part of the MSMR's payment chain uses digital payments. Third, in cases where well-functioning payment infrastructure does not exist or is not widely used, certain technologies help MSMRs bypass traditional acceptance infrastructure and adopt digital payments via alternative routes (e.g., leapfrogging via mobile payments). Fourth, while several strategies and incentives can be effective to increase digital payment acceptance and usage, their impact depends critically on the formality of the economy.

All the estimated treatment effects at the country-level and merchant-level are shown in Tables 4 and 5, respectively. We have also conducted a number of robustness tests to check the

accuracy and stability of our results compared to other methods and model specifications. These are shown in Appendix A2 for exposition simplicity.

## 5    Policy Implications

In this section, we discuss different government policies that would encourage greater acceptance and usage of digital payments. Merchant adoption of acceptance infrastructure does not necessarily translate into usage of payment cards. Interestingly, we find that newer technologies using the mobile phone channel increases adoption and usage of digital payments especially in countries where card penetration is low. In addition, we are able to identify which public sector incentives along with private sector enhancements such as killer apps result in greater usage of digital payments. Furthermore, our results suggest that a set of actions may be necessary by the public and private sectors to encourage the acceptance and usage of digital payments. We find the following government strategies may be successful to increase digital payment acceptance and usage:

1. Our analysis suggests that transaction account ownership whether at a bank or not generally increases merchant adoption of POS terminals and increases the share of MSMRs' digital payments received from their customers. Countries should implement policies to promote greater financial inclusion.

2. Our analysis suggests that improvements in ICT infrastructures are likely to increase acceptance and usage in countries with high transaction account ownership. Countries should continue to invest in ICT infrastructure to enable greater digital payment acceptance and usage.

3. Our analysis suggests that when other parts of their payment chain are using digital payments, MSMRs are likely to increase their adoption of POS terminals resulting in

more digital payments. Governments should encourage merchants to pay their workers electronically along with mandating digital payments for government wages.

4. Our analysis suggests that leveraging mobile phones to deliver payment services via killer apps may enable leapfrogging of card acceptance infrastructure using less expensive QR technology. Governments should encourage greater adoption of such technologies while maintaining adequate consumer protections and fraud prevention protocols.

5. Our analysis suggests that consumer and merchant fiscal incentives can be effective in increasing merchant acceptance and usage of digital payments. Reducing taxes paid by merchants and consumers are some of the most effective incentives to increase digital payment acceptance and usage. Governments should consider using fiscal incentives to encourage greater digital payment usage.

6. Our analysis suggests that cash limits and mandated digital payment acceptance are likely to increase digital payment acceptance and usage. Our analysis suggests that government policies targeted at reducing the shadow economy such as mandated use of electronic payments or transaction limits for cash transactions can be effective tools to increase EPA and usage although enforcement of these policies may be challenging.

## 6    Conclusion

In this paper, we provide an alternative approach to study how to identify key predictors to increase adoption and usage of digital payments by consumers and merchants. We consider hundreds of predictors and identify certain factors that increase the likelihood of adoption and usage. Finally, we are able to quantify the impact of certain incentives. Our results suggest that government initiatives such as increasing access to transaction accounts at banks and nonbanks,

encouraging digital wage payments, and improvements to ICT technologies increase the acceptance and usage of P2B digital payments at MSMRs. Furthermore, the presence of killer apps that reside on mobile phones along with greater awareness of the benefits of digital payments increases P2B digital payments. Finally, our results suggest that policies targeted at reducing the size of the shadow economy, such as consumer and merchant fiscal incentives, cash thresholds on transactions, and mandated electronic payment acceptance, can be effective policies but often the least studied.

To better inform policymakers on effective policies to enable greater digital payment acceptance and usage, longitudinal data collection and analysis are critical to identify the effectiveness of various public and private sector initiatives. Unfortunately, we lack a time-series component to study how such incentives impact acceptance and usage over time within a country. Furthermore, more standardized cross-country merchant surveys should be encouraged to better understand why certain incentives work better in some countries than others. Such data would allow better estimation of impact of an initiative within a country which is likely to be more important to base policy upon.

Finally, we would encourage policymakers to consider the role of nonbanks in the provision of payment services to increase access to digital payment services while adopting proper safeguards to ensure a safe and efficient payment system with adequate end-user protections especially when a large proportion of the population is unbanked. Technological advances often brought to the marketplace by nonbanks are likely to increase digital connectivity among consumers and businesses and increase the acceptance and usage of digital payments.

# References

Agarwal, S., Chakravorti, S., and Lunn, A. (2010). Why Do Banks Reward Their Customers to Use Their Credit Cards? Federal Reserve Bank of Chicago Working Paper No. 2010-19.

Agarwal, S., Ghosh, P., Li, J., and Ruan, T. (2020). Digital Payments and Consumption: Evidence from the 2016 Demonetization in India, SSRN working paper.

Agarwal, S., Qian, W., Yeung, B. Y., and Zou, X. (2019). Mobile Wallet and Entrepreneurial Growth, AEA Papers and Proceedings, *American Economic Review*, 109, 48-53.

Allen, F., Demirguc-Kunt, A., Klapper, L. and Martinez Peria, M.S. (2016). The Foundations of Financial Inclusion: Understanding Ownership and Use of Formal Accounts, *Journal of Financial Intermediation*, 27, 1-30.

Amromin, G. and S. Chakravorti (2009), "Whither Loose Change?: The Diminishing Demand for Small Denomination Currency," *Journal of Money, Credit, and Banking*, 41 (2-3), 315-335.

Arango, C., Huynh K., P. and Sabetti, L. (2015). Consumer Payment Choice: Merchant Card Acceptance versus Pricing Incentives. *Journal of Banking and Finance*, 55, 130–141.

Asadi, S., Nilashi, M., Husin, A. R. C., and Yadegaridehkordi, E. (2017). Customers perspectives on adoption of cloud computing in banking sector. *Information Technology and Management*, 18 (4), 305–330.

Athey, S. (2017). Beyond Prediction: Using Big Data for Policy Problems. *Science*, 485 (February), 483–485.

Athey, S., and Imbens, G. (2016). Recursive Partitioning for Heterogeneous Causal Effects. Proceedings of the National Academy of Sciences, 113 (27), 7353–7360.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized Random Forests. *Annals of Statistics*, 47 (2), 1148–1178.

Athey, S., and Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. Working Paper No. 3786, 1–15.

Bajari, P., Nekipelov, D., Ryan, S. P., and Yang, M. (2015). Machine Learning Methods for Demand Estimation. *American Economic Review*, 105 (5), 481–485.

Barajas, A., Beck, T., Belhaj, M. and Ben Naceur, S. (2020). Financial Inclusion: What Have We Learned So Far? What Do We Have to Learn?, WP 20/157. International Monetary Fund.

Bolt, W. and Chakravorti, S. (2012). Digitization of Retail Payment Systems, in Martin Peitz and Joel Waldfogel (eds), *Oxford Handbook of the Digital Economy*, Oxford: Oxford University Press, 2012, 108-135.

Bolt, W. and Chakravorti, S. (2008). Consumer Choice and Merchant Acceptance of Payment Media. Federal Reserve Bank of Chicago Working Paper, WP-2008-11.

Bounie, D. François, A. (2016). Cash, Check or Bank Card? The Effects of Transaction Characteristics on the Use of Payment Instruments. Telecom Paris Working Paper Series.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.

Carbo-Valverde, S., Chakravorti, S., Rodriguez-Fernandez, F. (2016). The Role of Interchange Fees in Two-sided Markets: An Empirical Investigation on Payment Cards. *The Review of Economics and Statistics*, 98, 367-381.

Chakravorti, S. and Roson, R. (2006). Platform Competition in Two-Sided Markets: The Case of Payment Networks. *Review of Network Economics*, 5 (1), 1-25.

Chen, T., and Guestrin, C. (2016). XGBoost : A Scalable Tree Boosting System. ArXiv:1603.02754v3, 1–13.

Chen, H. Q., Honda, T., and Yang, M. C. (2013). Approaches for Identifying Consumer Preferences for the Design of Technology Products: A Case Study of Residential Solar Panels. *Journal of Mechanical Design*, 135 (6), 061007.

Ching, A. T., and Hayashi, F. (2008). Payment Card Rewards Programs and Consumer Payment Choice. Available at SSRN: https://ssrn.com/abstract=1114247.

Cœuré, B. (2019). *Payments for the People*, Introductory remarks by the Chair of the CPMI and Member of the Executive Board of the ECB, at the High-level meeting on financial inclusion, Basel, Switzerland, 27 May.

Cohen, M. A. and Rysman, M. (2013). Payment Choice with Consumer Panel Data. FRB of Boston Working Paper No. 13-6.

Cui, G., Wong, M. L., and Lui, H. K. (2006). Machine Learning for Direct Marketing Response Models: Bayesian Networks with Evolutionary Programming. *Management Science*, 52(4), 597–612.

Delen, D., and Zolbanin, H. M. (2018). The Analytics Paradigm in Business Research. *Journal of Business Research*, 90 (May), 186–195.

Demirguc-Kunt, A., Klapper, L. and Singer, D. (2017). Financial Inclusion and Inclusive Growth A Review of Recent Empirical Evidence. Policy Research Working Paper 8040. The World Bank.

Dupas, P., and J. Robinson. 2009. Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. National Bureau of Economic Research Working Paper 14693.

Duryea, S., and Schargrodsky, E. (2007), *Financial Services for the Poor: Savings, Consumption, and Welfare*, IABD, Washington.

European Commission (2015). Survey on Merchants' Costs of Processing Cash and Card Payments. Technical report, European Commission, Directorate-General for Competition.

Hagenauer, J., and Helbich, M. (2017). A Comparative Study of Machine Learning Classifiers for Modeling Travel Mode Choice. *Expert Systems with Applications*, 78, 273–282.

Han, H., Guo, X., and Yu, H. (2016). Variable Selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS, 219–224.

Han, P., and Wang, Z. (2021), Technology Adoption and Leapfrogging: Racing for Mobile Payments. SSRN Working Paper.

Hernández-Murillo, R., Llobet, G., and Fuentes, R. (2010). Strategic Online Banking Adoption. *Journal of Banking and Finance*, 34 (7), 1650–1663.

Honka, E., Horta, A., and Vitorino, M. A. (2017). Advertising, Consumer Awareness and Choice : Evidence from the U.S. Banking Industry. *RAND Journal of Economics*, 48 (3), 611–646.

Hothorn, T., Hornik, K., Van DeWiel, M. A., and Zeileis, A. (2006). A Lego System for Conditional Inference. *The American Statistician*, 60 (3), 257–263.

Jack, W., and Suri, T. (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution. *American Economic Review*, 104 (1), 183-223.

Jacob, D. (2021). CATE meets ML. Conditional average treatment effect and machine learning, *Digital Finance*, 3, 99–148.

Kesharwani, A. (2019). Do (how) Digital Natives Adopt a New Technology Differently than Digital Immigrants? A Longitudinal Study. *Information and Management*, August 2016, 103170.

Klee, E. (2008). How People Pay: Evidence from Grocery Store Data. *Journal of Monetary Economics*, 55 (3), 526-41.

Knaus, M. C., Lechner, M., and Strittmatter, A. (2020). Machine Learning Estimation of Heterogeneous Causal Effects: Empirical Monte Carlo Evidence. *The Econometrics Journal*, Forthcoming.

Koulayev, S., Rysman, M., Schuh, S., and Stavins, J. (2016), Explaining Adoption and Use of Payment Instruments by U.S. Consumers. *RAND Journal of Economics*, 47, 293-325.

Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521 (7553), 436–444.

Ligon, E., Malick, B., Sheth, K., and Trachtman, C. (2019). What Explains low Adoption of Digital Payment Technologies? Evidence from Small-Scale Merchants in Jaipur, India. PLoS One, 14 (7).

Miguéis, V. L., Camanho, A. S., and Borges, J. (2017). Predicting Direct Marketing Response in Banking: Comparison of Class Imbalance Methods. *Service Business*, 11 (4), 831–849.

Ohnsorge, F. and S. Yu (2021). *The Long Shadow of Informality: Challenges and Policies*. The World Bank.

Pousttchi, K., and Dehnert, M. (2018). Exploring the Digitalization Impact on Consumer Decision-Making in Retail Banking. *Electronic Markets*, 28 (3), 265–286.

Rochet, J-C and Tirole, J. (2006). Two-Sided Markets: A Progress Report. *RAND Journal of Economics*, 37 (3), 645-667.

Rochet, Jean-Charles and Jean Tirole (2002). Cooperation among Competitors: Some Economics of Payment Card Associations. *Rand Journal of Economics*, 33 (4), 1-22.

Rogoff, Kenneth (2016). *The Curse of Cash*, Princeton, NJ: Princeton University Press.

Rysman, Marc (2009). The Economics of Two-Sided Markets, *Journal of Economic Perspectives*, 23 (3), 125-143.

Rysman, M. and Schuh, S. (2017). New Innovations in Payments. In Greenstein, S., Lerner, J., and Stern, S., editors, *Innovation Policy and the Economy*, 17, 27–48. University of Chicago Press.

Shy, O. and Stavins, J. (2015). Merchant Steering of Consumer Payment Choice: Evidence from a 2012 Diary Survey. *Journal of Behavioral & Experimental Economics*, 55, 1-9.

Schneider, Friedrich (2019). Restricting or Abolishing Cash: An Effective Instrument for Eliminating the Shadow Economy, Corruption and Terrorism?" *SUERF Policy Note*, 90.

Schuh, S., Shy, O., Stavins, J. and R. Triest (2012). An Economic Analysis of the 2011 Settlement Between the Department of Justice and Credit Card Networks. *Journal of Competition Law and Economics*, 8 (1), 107–144.

Tibshirani, J., Athey, S., Friedberg, R., Hadad, V., Miner, L., Wager, S., and Wright, M. (2018). grf: Generalized Random Forests (Beta). URL: https://Github.Com/Grf-Labs/Grf. R Package Version 0.10.2.

Tiffin, A. (2019). Machine Learning and Causality: The Impact of Financial Crises on Growth. IMF Working Papers WP19/228.

Valiant, L. (2013). *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. Basic Books, Inc.

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Qi Dong, J., Fabian, N., and Haenlein, M. (2019). Digital Transformation: A Multidisciplinary Reflection and Research Agenda. *Journal of Business Research*, July.

Wager, S., and Athey, S. (2018). Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, 113 (523), 1228–1242.

Witten, I. H., Frank, E., and Hall, M. A. (2019). Data Mining: Practical Machine Learning Tools and Techniques (Third Edition). In Morgan Kaufmann (The Morgan, 53 (9)). Morgan Kaufmann.

World Bank Group (2021). *Electronic Payments Acceptance Toolkit: Examples of Incentive Implementations*. Washington DC: International Bank for Reconstruction and Development/The World Bank.

World Bank Group (2020). *Electronic Payments Acceptance Incentives Literature Review and Country Examples*, Washington DC: International Bank for Reconstruction and Development/The World Bank.

World Bank Group (2016). *Cash vs. Electronic Payments in Small Retailing: Estimating the Global Size*, Washington DC: International Bank for Reconstruction and Development/The World Bank.

Yusuf Dauda, S., and Lee, J. (2015). Technology Adoption: A Conjoint Analysis of Consumers' Preference on Future Online Banking Services. *Information Systems*, 53, 1–15.

Zaidi, N. A., Carman, M. J., and Webb, G. I. (2013). Alleviating Naive Bayes Attribute Independence Assumption by Attribute Weighting. *Journal of Machine Learning Research*, 14, 1947–1988.

Zhao, H., Anong, S.T. and Zhang, L. (2019). Understanding the Impact of Financial Incentives on NFC Mobile Payment Adoption: An Experimental Analysis, *International Journal of Bank Marketing*, 37 (5), 1296-1312.

**Table 1: Main Predictors of POS Adoption and Share of P2B Electronic Payments**
**(combined results from the country-level and merchant-level samples)**

| Response variable | Category | Predictor | Variable importance confirmed in: | |
|---|---|---|---|---|
| | | | Country-level sample | Merchant-level sample |
| POS Terminal Adoption | Merchant Payment Chain | Share of P2B electronic payments | ✓ | ✓ |
| | | Merchants' beliefs about consumer payment preferences | n.a. | ✓ |
| | | Percentage of wages paid electronically at the merchant level | n.a. | ✓ |
| | Infrastructure | Information and Communication Technologies | ✓ | ✓ |
| | | Account ownership | ✓ | ✓ |
| | | National ID | ✓ | ✓ |
| | Institutional and Policy | Merchant fiscal incentives | | ✓ |
| | | National financial inclusion strategy | ✓ | ✓ |
| | | Wages paid into a transaction account | ✓ | ✓ |
| | | Shadow Economy | ✓ | ✓ |
| Share of P2B Electronic Payments | Merchant Payment Chain | Share of P2B electronic payments | ✓ | ✓ |
| | | Merchants' beliefs about consumer payment preferences | n.a. | ✓ |
| | | Percentage of wages paid electronically at the merchant level | n.a. | ✓ |
| | Instruments | Previous card penetration | ✓ | ✓ |
| | Infrastructure | Information and Communication Technologies | ✓ | ✓ |
| | Institutional and Policy | Wages paid into a transaction account | ✓ | ✓ |
| | | Killer app | | ✓ |
| | Access Points | POS adoption | ✓ | ✓ |
| | | Agents of payment services providers | ✓ | n.a. |

"n.a." notes that the variable was not available for that sample

**Note:** Predictors selected from the variable importance plots obtained from the random forest algorithm. The grouping of the variables follows XXXXX (cite EPA incentives report here).

**Figure 1: Example of a conditional classification tree**
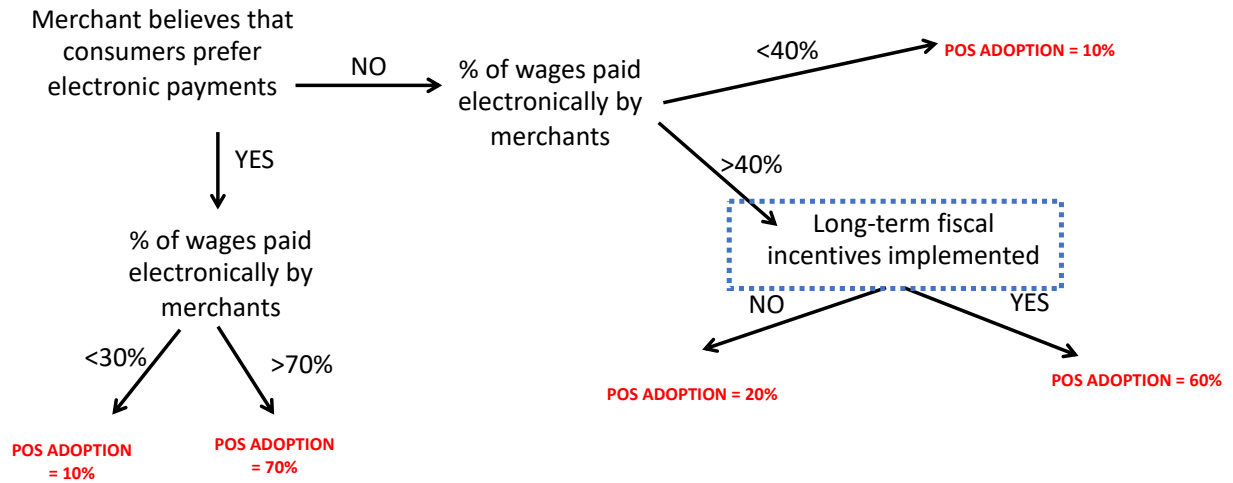


**Note:** This conditional inference tree is obtained from a merchant-sample estimation using the parameters provided by the variable importance plot from the random forest.

**Table 2: Factors Influencing POS Terminal Adoption**

| Factors for POS Terminal Adoption | Increase in Likelihood of POS Adoption |
|---|---|
| Share of P2B electronic payments and bank account ownership is above the median country (country level) | 200 percent |
| Merchants believe consumers prefer electronic payments and wages are paid electronically (mid-high or high preference) (merchant level) | 100 percent |
| Wages are paid electronically and combined with account ownership or with ICT infrastructures above median (country level) | 100 percent |
| ICT or national ID implementation are over the median value and shadow economy below 15% (country level) | 50 percent |
| Shadow economy below 25% is combined with implementing a national financial inclusion strategy or with merchant fiscal incentives (country level) | 20 percent |

**Table 3: Factors Influencing Share of Electronic P2B Payments**

| Factors for P2B Electronic Transactions | Increase in Likelihood of P2B Electronic Share |
|---|---|
| Wages paid electronically and card penetration in previous 5 years are over median value (country level) | 100 percent |
| POS terminal adoption is mid high or high (country level) | 60 percent |
| ICT above median is combined with killer apps or with a significant use of agents of payment services providers (country level) | 50-60 percent |
| Merchants believe consumers prefer electronic payments and electronic payments to suppliers above median value (merchant level) | 30 percent |

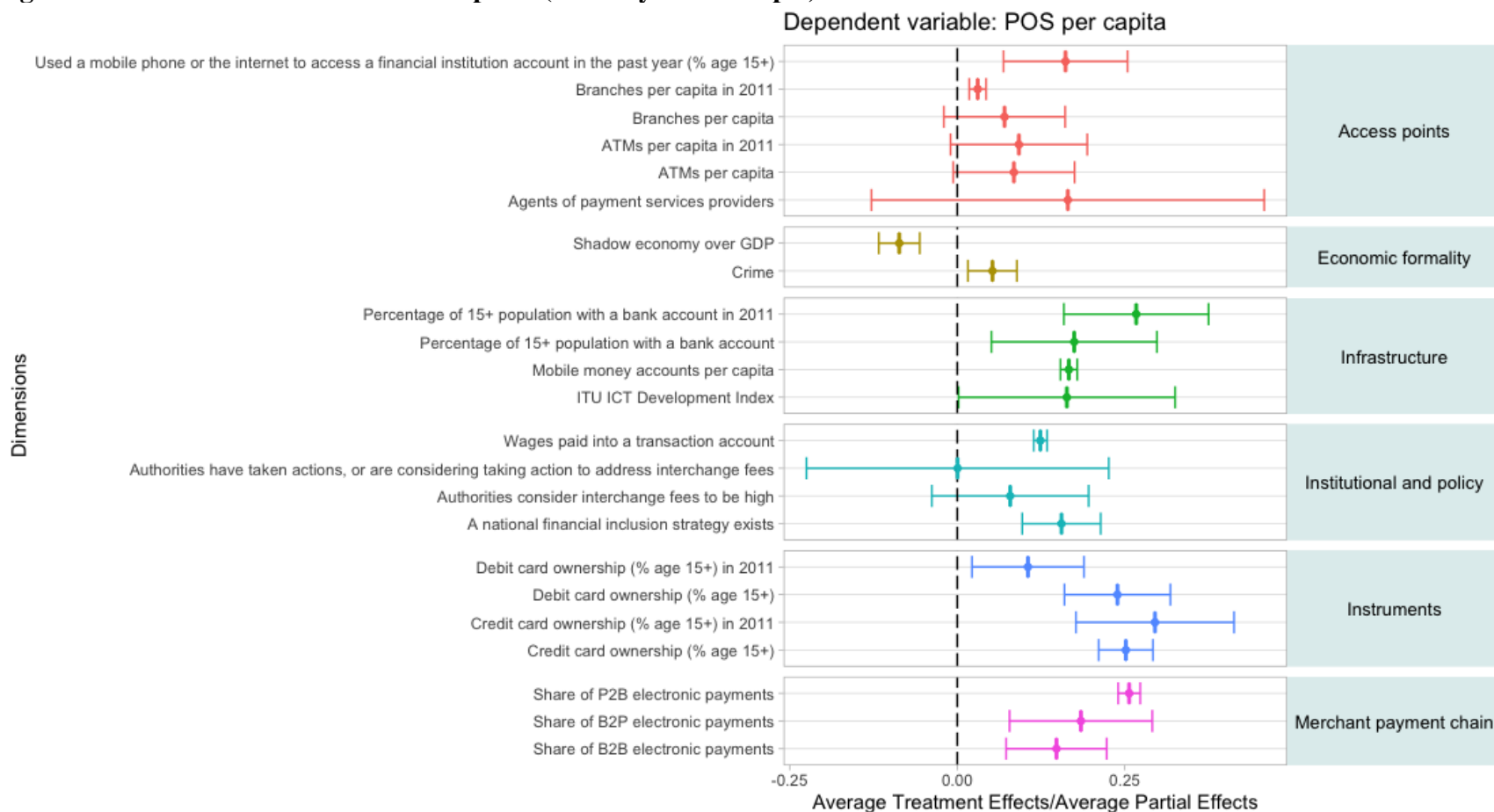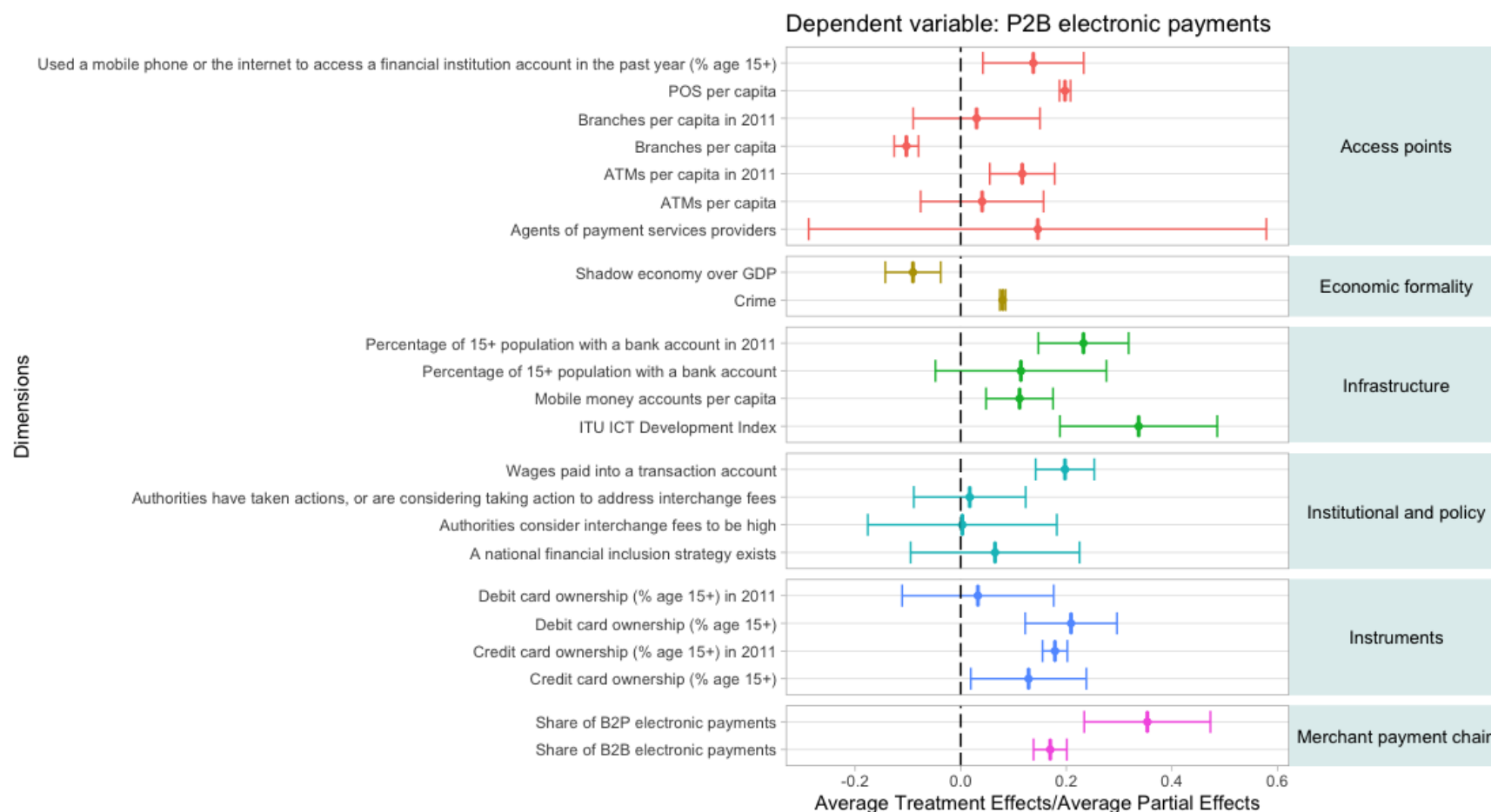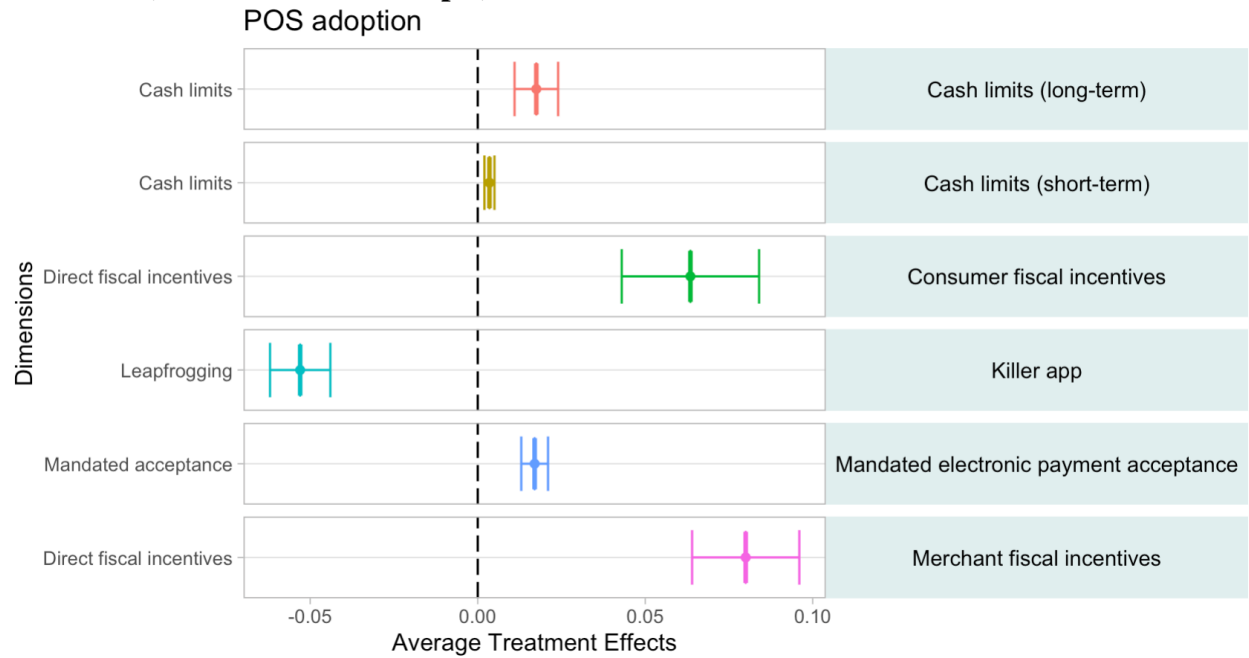**Figure 2: Treatment effects on POS adoption (Country-level sample)**

**Figure 3: Treatment Effects on the Share of P2B Electronic Payments (Country-level sample)**

**Figure 4: Impact of Incentives: Treatment Effects on POS Terminal Adoption (Merchant Level Sample)**



**Figure 5: Impact of Incentives: Treatment Effects on the Share of Merchant P2B Electronic Payment Transactions (merchant-level sample)**
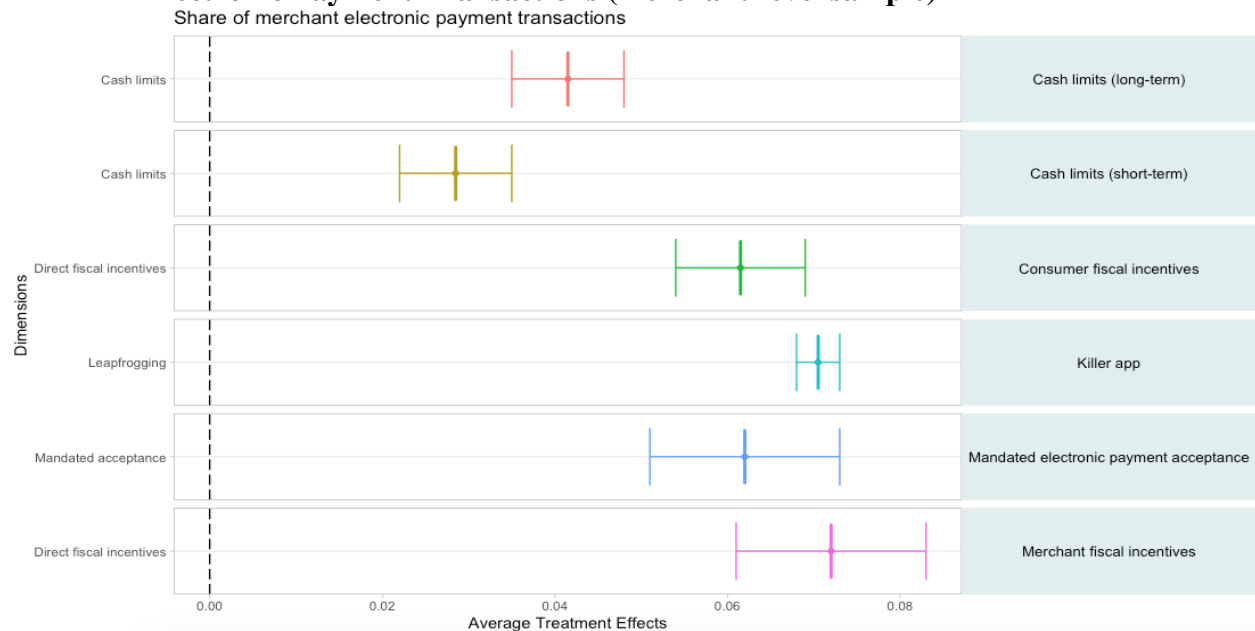
**Table 4: Summary of treatment effects from the country-level sample**

| POS adoption | | | | | | Share of electronic P2B payments | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| variable | level | ATE/APE | Lower bound | Upper bound | | variable | level | ATE/APE | Lower bound | Upper bound |
| Merchant payment chain | Share of P2B electronic payments | 0.257 | 0.24 | 0.273 | | Merchant payment chain | Share of B2B electronic payments | 0.170 | 0.138 | 0.201 |
| Merchant payment chain | Share of B2B electronic payments | 0.148 | 0.073 | 0.223 | | Merchant payment chain | Share of B2P electronic payments | 0.354 | 0.234 | 0.473 |
| Merchant payment chain | Share of B2P electronic payments | 0.185 | 0.078 | 0.291 | | Institutional and policy | Wages paid into a transaction account | 0.198 | 0.142 | 0.253 |
| Institutional and policy | Wages paid into a transaction account | 0.124 | 0.114 | 0.134 | | Institutional and policy | A national financial inclusion strategy exists | 0.065 | -0.095 | 0.225 |
| Institutional and policy | A national financial inclusion strategy exists | 0.156 | 0.097 | 0.214 | | Institutional and policy | Authorities consider interchange fees to be high | 0.003 | -0.176 | 0.182 |
| Institutional and policy | Authorities consider interchange fees to be high | 0.079 | -0.038 | 0.196 | | Institutional and policy | Authorities have taken actions, or are considering taking action to address interchange fees | 0.017 | -0.089 | 0.123 |
| Institutional and policy | Authorities have taken actions, or are considering taking action to address interchange fees | 0.001 | -0.225 | 0.226 | | Access points | Used a mobile phone or the internet to access a financial institution account in the past year (% age 15+) | 0.138 | 0.042 | 0.233 |
| Access points | Agents of payment services providers | 0.165 | -0.128 | 0.458 | | Access points | Agents of payment services providers | 0.146 | -0.288 | 0.579 |
| Access points | Branches per capita | 0.071 | -0.02 | 0.161 | | Access points | POS per capita | 0.198 | 0.187 | 0.208 |
| Access points | ATMs per capita in 2011 | 0.092 | -0.01 | 0.194 | | Access points | ATMs per capita | 0.041 | -0.076 | 0.157 |
| Access points | ATMs per capita | 0.085 | -0.006 | 0.175 | | Access points | Branches per capita | -0.103 | -0.126 | -0.08 |
| Access points | Branches per capita in 2011 | 0.031 | 0.043 | 0.018 | | Access points | ATMs per capita in 2011 | 0.117 | 0.055 | 0.178 |
| Access points | Used a mobile phone or the internet to access a financial institution account in the past year (% age 15+) | 0.162 | 0.069 | 0.254 | | Access points | Branches per capita in 2011 | 0.030 | -0.09 | 0.15 |
| Infrastructure | Mobile money accounts per capita | 0.167 | 0.154 | 0.179 | | Infrastructure | Mobile money accounts per capita | 0.112 | 0.048 | 0.175 |
| Infrastructure | ITU ICT | 0.164 | 0.002 | 0.325 | | Infrastructure | ITU ICT | 0.337 | 0.188 | 0.486 |
| Infrastructure | Percentage of 15+ population with a bank account | 0.175 | 0.051 | 0.298 | | Infrastructure | Percentage of 15+ population with a bank account | 0.114 | -0.048 | 0.276 |
| Infrastructure | Percentage of 15+ population with a bank account in 2011 | 0.267 | 0.159 | 0.375 | | Infrastructure | Percentage of 15+ population with a bank account in 2011 | 0.233 | 0.147 | 0.318 |
| Instruments | Credit card ownership (% age 15+) | 0.252 | 0.211 | 0.292 | | Instruments | Credit card ownership (% age 15+) | 0.129 | 0.019 | 0.238 |
| Instruments | Debit card ownership (% age 15+) | 0.239 | 0.16 | 0.318 | | Instruments | Debit card ownership (% age 15+) | 0.209 | 0.122 | 0.296 |
| Instruments | Credit card ownership (% age 15+) in 2011 | 0.295 | 0.177 | 0.413 | | Instruments | Credit card ownership (% age 15+) in 2011 | 0.179 | 0.155 | 0.202 |
| Instruments | Debit card ownership (% age 15+) in 2011 | 0.106 | 0.022 | 0.189 | | Instruments | Debit card ownership (% age 15+) in 2011 | 0.033 | -0.111 | 0.176 |
| Economic formality | Shadow economy over GDP | -0.087 | -0.117 | -0.056 | | Economic formality | Shadow economy over GDP | -0.091 | -0.143 | -0.038 |
| Economic formality | Crime | 0.053 | 0.016 | 0.089 | | Economic formality | Crime | 0.080 | 0.074 | 0.085 |

Note: the table shows the average treatment effect (ATE) for binary variables and the average partial effect (APE) for four-level (1, 2 ,3, 4) variables, along with the lower and upper bound of the confidence intervals.

**Table 5: Summary of treatment effects from the merchant-level sample**

| POS adoption | | | | | | Share of P2B electronic payment value | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| variable | level | ATE | Lower bound | Upper bound | | variable | level | ATE | Lower bound | Upper bound |
| Merchant fiscal incentives | Direct fiscal incentives | 0.080 | 0.064 | 0.096 | | Merchant fiscal incentives | Direct fiscal incentives | 0.050 | 0.032 | 0.068 |
| Consumer fiscal incentives | Direct fiscal incentives | 0.064 | 0.043 | 0.084 | | Consumer fiscal incentives | Direct fiscal incentives | 0.042 | 0.029 | 0.054 |
| Mandated electronic payment acceptance | Mandated acceptance | 0.017 | 0.013 | 0.021 | | Mandated electronic payment acceptance | Mandated acceptance | 0.047 | 0.028 | 0.066 |
| Cash limits (long-term) | Cash limits | 0.018 | 0.011 | 0.024 | | Cash limits (long-term) | Cash limits | 0.062 | 0.041 | 0.082 |
| Cash limits (short-term) | Cash limits | 0.004 | 0.002 | 0.005 | | Cash limits (short-term) | Cash limits | 0.046 | 0.022 | 0.07 |
| Killer app | Leapfrogging | -0.053 | -0.062 | -0.044 | | Killer app | Leapfrogging | 0.061 | 0.050 | 0.071 |

| Share of P2B electronic payment transactions | | | | |
|---|---|---|---|---|
| variable | level | ATE | Lower bound | Upper bound |
| Merchant fiscal incentives | Direct fiscal incentives | 0.072 | 0.061 | 0.083 |
| Consumer fiscal incentives | Direct fiscal incentives | 0.062 | 0.054 | 0.069 |
| Mandated electronic payment acceptance | Mandated acceptance | 0.062 | 0.051 | 0.073 |
| Cash limits (long-term) | Cash limits | 0.042 | 0.035 | 0.048 |
| Cash limits (short-term) | Cash limits | 0.029 | 0.022 | 0.035 |
| Killer app | Leapfrogging | 0.071 | 0.068 | 0.073 |

Note: the table shows the average treatment effect (ATE), along with the lower and upper bound of the confidence intervals.

# Appendix A1: The Data

**List of countries in the country-level sample:** Afghanistan, Albania, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belgium, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burundi, Cambodia, Canada, Central African Republic, Chile, China, Colombia, Costa Rica, Croatia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, France, Georgia, Germany, Greece, Guatemala, Haiti, Honduras, Hong Kong SAR, China, Hungary, India, Indonesia, Iran, Iraq, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kuwait, Kyrgyz Republic, Latvia, Lebanon, Lesotho, Liberia, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Malta, Mauritania, Mauritius, Mexico, Moldova, Montenegro, Morocco, Myanmar, Namibia, the Netherlands, New Zealand, Nigeria, Norway, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Poland, Portugal, Russia, Rwanda, Saudi Arabia, Serbia, Slovak Republic, South Africa, Korea, Spain, Sri Lanka, Sudan, Sweden, Tajikistan, Thailand, Tunisia, Turkey, Uganda, Ukraine, the United Kingdom, Uruguay, Vietnam, Zambia, and Zimbabwe.

**List of countries in the merchant-level sample:** Colombia, France, Kenya, Lithuania, Morocco, Pakistan, and Turkey

## Variables from the country-level sample

| Variable | Definition | Source |
|---|---|---|
| account | Account (% age 15+) | WBG Findex |
| agents_psp_pc | Agents of payment service providers per 100,000 adults | WBG GPSS |
| aml_cft_risks | Are AML/CFT risks assessed during authorization process for modified or new financial products? | WBG FICP |
| atms_pc | Number of ATMs per 100,000 adults | IMF FAS |
| branches_pc | Number of commercial bank branches per 100,000 adults | IMF FAS |
| consumer_risks | Are consumer risks assessed during authorization process for modified or new financial products? | WBG FICP |
| credit_card | Credit card ownership (% age 15+) | WBG Findex |
| credit_cards_pc | Number of credit cards per 1,000 adults | IMF FAS |
| debit_card | Debit card ownership (% age 15+) | WBG Findex |
| debit_cards_pc | Number of debit cards per 1,000 adults | IMF FAS |
| deposit | Deposit in the past year (% with a financial institution account, age 15+) | WBG Findex |
| deposit_accounts_pc | Number of deposit accounts with commercial banks per 1,000 adults | IMF FAS |
| emoney_accounts_pc | E-money accounts per 1,000 adults | WBG GPSS |
| fi_account | Financial institution account (% age 15+) | WBG Findex |
| if_action | Authorities have taken actions, or are considering taking action to address interchange fees | WBG GPSS |
| if_high | Authorities consider interchange fees prevailing in the card industry to be high | WBG GPSS |
| internet_bill_pay | Used the internet to pay bills in the past year (% age 15+) | WBG Findex |
| itu_idi | ITU ICT Development Index | ITU |
| made_received_dig_pay | Made or received digital payments in the past year (% age 15+) | WBG Findex |
| main_withdrawal_atm | Main mode of withdrawal: ATM (% with a financial institution account, age 15+) | WBG Findex |
| main_withdrawal_teller | Main mode of withdrawal: bank teller (% with a financial institution account, age 15+) | WBG Findex |
| max_cost_account | Is the maximum cost to open a savings/current account regulated in this country? | WBG FICP |
| mm_account | Mobile money account (% age 15+) | WBG Findex |
| mm_accounts_pc | Number of registered mobile money accounts per 1,000 adults | IMF FAS |
| mm_agents_pc | Number of registered mobile money agent outlets per 100,000 adults | IMF FAS |
| mobile_internet_access_fi | Used a mobile phone or the internet to access a financial institution account in the past year (% age 15+) | WBG Findex |
| national_id | Has a national identity card (% age 15+) | WBG Findex |
| nfcs | Has a national financial capability/literacy/education strategy (NFCS/NFLS/NFES) already been launched? | WBG FICP |

| Variable | Definition | Source |
|---|---|---|
| nfis | Has a national financial inclusion strategy (NFIS) already been launched? | WBG FICP |
| nms | Has a national microfinance strategy (NMS) already been launched? | WBG FICP |
| no_account_distance | No account because financial institutions are too far away (% age 15+) | WBG Findex |
| no_account_documentation | No account because of lack of necessary documentation (% age 15+) | WBG Findex |
| no_account_expensive | No account because financial services are too expensive (% age 15+) | WBG Findex |
| no_account_family | No account because someone in the family has an account (% age 15+) | WBG Findex |
| no_account_funds | No account because of insufficient funds (% age 15+) | WBG Findex |
| no_account_religion | No account because of religious reasons (% age 15+) | WBG Findex |
| no_account_trust | No account because of lack of trust in financial institutions (% age 15+) | WBG Findex |
| operational_risks | Are operational risks assessed during authorization process for modified or new financial products? | WBG FICP |
| pension_fi | Received a public sector pension: into a financial institution account (% age 15+) | WBG Findex |
| pos_pc | POS terminals per 100,000 adults | WBG GPSS |
| school_fees | Paid school fees in the past year (% age 15+) | WBG Findex |
| sent_received_dom_remit | Sent or received domestic remittances in the past year (% age 15+) | WBG Findex |
| tax_incentives_savings | Are tax incentive savings schemes in place to promote financial inclusion? | WBG FICP |
| wages | Received wages in the past year (% age 15+) | WBG Findex |
| wages_fi | Received wages: into a financial institution account (% age 15+) | WBG Findex |
| withdrawal | Withdrawal in the past year (% with a financial institution account, age 15+) | WBG Findex |
| m_fiscal_incent_long | Merchant fiscal incentive implemented long time ago | Authors' own elaboration |
| m_fiscal_incent_recent | Merchant fiscal incentive implemented recently | Authors' own elaboration |
| c_fiscal_incent_long | Consumer fiscal incentives (VAT reductions, income tax reductions) long time ago | Authors' own elaboration |
| c_fiscal_incent_recent | Consumer fiscal incentives (VAT reductions, income tax reductions) recently | Authors' own elaboration |
| lott_long | Lotteries implemented long-time ago | Authors' own elaboration |
| lott_recent | Lotteries implemented recently | Authors' own elaboration |
| mandated_epa_long | Mandated acceptance of electronic payments implemented long time ago | Authors' own elaboration |
| mandated_epa_recent | Mandated acceptance of electronic payments implemented recently | Authors' own elaboration |
| subsid_POS_long | Subsidized POS terminals implemented long-time ago | Authors' own elaboration |
| subsid_POS_recent | Subsidized POS terminals implemented recently | Authors' own elaboration |
| cash_limits_long | Cash transaction limits (or significant cash disincentives) implemented long time ago | Authors' own elaboration |
| cash_limits_recent | Cash transaction limits (or significant cash disincentives) implemented recently | Authors' own elaboration |
| killer_app | A "killer app" exists for mobile payment adoption | Authors' own elaboration |
| m_fees_reduced | Lowering of merchant fees | Authors' own elaboration |
| crime | Serious assaults per 100,000 population | United Nations Office for Drug and Crime (UNDOC) |
| shadow_econ | Shadow economy over GDP | Authors' own estimations |

# Variables from the merchant-level sample

| Variable | Definition | Units |
|---|---|---|
| repeat_customers | Repeat customers | Proportion of customers |
| transactions_pd | Transactions per day | Gross volume |
| transactions_pa | Transactions per annum | Gross volume |
| cash_val | Value of P2B cash payments | Proportion of all P2B payments |
| cheque_val | Value of P2B cheque payments | Proportion of all P2B payments |
| transfer_val | Value of P2B bank transfers | Proportion of all P2B payments |
| mobile_val | Value of P2B mobile transfers | Proportion of all P2B payments |
| cards_val | Value of P2B card payments | Proportion of all P2B payments |
| cash_vol | Volume of P2B cash payments | Proportion of all P2B payments |
| cheque_vol | Volume of P2B cheque payments | Proportion of all P2B payments |
| transfer_vol | Volume of P2B bank transfers | Proportion of all P2B payments |
| mobile_vol | Volume of P2B mobile transfers | Proportion of all P2B payments |
| cards_vol | Volume of P2B card payments | Proportion of all P2B payments |
| epay_val | Value of P2B electronic payments | Proportion of all P2B payments |
| epay_vol | Volume of P2B electronic payments | Proportion of all P2B payments |
| c_paper_pref | Customers prefer to pay with paper-based methods? | Binary (0 - No / 1 - Yes) |
| m_paper_pref | Merchant prefers paper-based payment methods? | Binary (0 - No / 1 - Yes) |
| pos | Merchant has a POS terminal? | Binary (0 - No / 1 - Yes) |
| bank_account | Merchant has a bank account? | Binary (0 - No / 1 - Yes) |
| epay | Electronic payments at retailer? (Yes/No) | Binary (0 - No / 1 - Yes) |
| region | Region | |
| group | Income Group | |
| grocery | Grocery? | Binary (0 - No / 1 - Yes) |
| type | Type of retailer | |
| years_service | Years in business | |
| formal | Formal? | Binary (0 - No / 1 - Yes) |
| urban | Urban? | Binary (0 - No / 1 - Yes) |
| location | City/region, etc. | |
| branches | Number of Retailer Branches | |
| num_employees | Number of Employees | |
| female_owned | Female owned? | Binary (0 - No / 1 - Yes) |
| days_open_month | Days open per month | |
| consumer_income | Income demographic of customers | |
| epay_val_country | Value of P2B MSME e-payments in the country | Proportion of all P2B payments at MSMEs |
| epay_vol_country | Volume of P2B MSME e-payments in the country | Proportion of all P2B payments at MSMEs |
| avg_num_suppliers_pw | Average number of suppliers per week | |
| avg_num_supplier_pmts_pm | Average number of supplier payments per month | |
| avg_num_supplier_pmts_pa | Average number of supplier payments per annum | |
| cash_val_b2b | Value of B2B cash payments | Proportion of all B2B payments |
| cheque_val_b2b | Value of B2B cheque payments | Proportion of all B2B payments |

| Variable | Definition | Units |
|---|---|---|
| transfer_val_b2b | Value of B2B bank transfers | Proportion of all B2B payments |
| mobile_val_b2b | Value of B2B mobile transfers | Proportion of all B2B payments |
| cards_val_b2b | Value of B2B card payments | Proportion of all B2B payments |
| cash_vol_b2b | Volume of B2B cash payments | Proportion of all B2B payments |
| cheque_vol_b2b | Volume of B2B cheque payments | Proportion of all B2B payments |
| transfer_vol_b2b | Volume of B2B bank transfers | Proportion of all B2B payments |
| mobile_vol_b2b | Volume of B2B mobile transfers | Proportion of all B2B payments |
| cards_vol_b2b | Volume of B2B card payments | Proportion of all B2B payments |
| epay_val_b2b | Value of B2B electronic payments | Proportion of all B2B payments |
| epay_vol_b2b | Volume of B2B electronic payments | Proportion of all B2B payments |
| employee_pmts_pm | Number of employee payments per month | |
| employee_pmnts_pa | Number of employee payments per annum | |
| cash_val_b2p | Value of B2P cash payments | Proportion of all B2P payments |
| cheque_val_b2p | Value of B2P cheque payments | Proportion of all B2P payments |
| transfer_val_b2p | Value of B2P bank transfers | Proportion of all B2P payments |
| mobile_val_b2p | Value of B2P mobile transfers | Proportion of all B2P payments |
| cards_val_b2p | Value of B2P card payments | Proportion of all B2P payments |
| epay_val_b2p | Value of B2P electronic payments | Proportion of all B2P payments |

Additionally, in each one of the exercises (country-level and merchant-level) we use variables from the other one. In the country-level exercise, for example, we use the share of digital payments made to MSMRs by individuals as a dependent variable but also as an explanatory variable when POS is the dependent variable. Similarly, we use the share of B2B and BTP digital payments and we also use a breakdown of these variables for grocery and non-grocery establishments (as estimated by Euromonitor from merchant data and then extrapolated to a country-level database).

**Appendix A2: Robustness Checks**

As for alternative models, comparisons can be made with parametric standard approaches or with other machine learning algorithms. As for parametric model, a natural alternative would be running an ordered logit model, as the dependent variables take the values 1, 2, 3, 4 depending on the level of POS adoption or the share of electronic P2B payments. This also involves considering three constant cuts to evaluate the estimation at the four levels. As our benchmark model is the country-level sample, we run the logit model with the same set of variables used in the machine learning methods are employed. The results are shown in Table A2.1. Only the most relevant variables are shown although the 23 indicators reported in the earlier section are included. The findings are in line with those of the random and causal forests. Alternative specifications included different combinations of some of the most relevant (incentive-related) variables excluding others to study their effect both with and without other influences controlled for. The results were in line to those shown in the table.

| Table A2.1  Ordered logit estimation results | | | | |
|---|---|---|---|---|
| **Dependent variable** | **POS adoption** | | **Share of P2B electronic payments** | |
| | Coeff. | Std. error | Coeff. | Std. error |
| Wages paid into a transaction account | 0.0283*** | 0.009 | 0.0418** | 0.018 |
| Share of P2B electronic payments | 0.0357** | 0.016 | - | - |
| Information and Communication Technologies (ICT) | 0.1205** | 0.049 | 0.1933*** | 0.044 |
| Account ownership | 0.0742* | 0.041 | - | - |
| National ID | 0.0205** | 0.010 | - | - |
| National financial strategy exists | 0.0928*** | 0.033 | - | - |
| Shadow Economy | -0.0089** | -0.004 | - | - |
| Share of B2B electronic payments | - | - | 0.0419*** | 0.015 |
| Previous card penetration (2011) | - | - | 0.0083* | 0.048 |
| POS adoption | - | - | 0.0902*** | 0.029 |
| Agents of payment services providers | - | - | 0.0351*** | 0.010 |
| Constant cut1 | 0.0362*** | 0,008 | 0.0403*** | 0.009 |
| Constant cut2 | 1.1929*** | 0,382 | 1.0023*** | 0.363 |
| Constant cut3 | 2.9402*** | 0.928 | 3.1987*** | 0.103 |
| Other controls | Yes | Yes | Yes | Yes |
| Errors clustered at the country level | Yes | Yes | Yes | Yes |
| Observations | 8,586 | | 8,586 | |
| Pseudo $R^2$ | 0.4830 | | 0.5775 | |
| Log Likelihood | -1512.19 | | -1728.30 | |
| Note: *** p<0.01, ** p<0.05, * p<0.1 | | | | |

We also benchmark our results to those of two alternative machine learning algorithms: extreme gradient boosting and Bayesian networks. Gradient boosting is based on the idea of whether a weak learner can be modified to become better. As Valiant, (2013) argues, the weak learning method is used several times to get a succession of hypotheses, each iteration is refocused on the examples that the previous ones found difficult and misclassified. Then, using a training sample (y, x) the goal of the algorithm is to obtain an estimate of the function F(x) that minimizes the expected value of a loss function over the joint distribution of all the observed values. Among the gradient boosting methods used in practice, Extreme Gradient Boosting, is widely used due to

its efficiency. Compared to other gradient boosting methods, extreme gradient boosting uses a more regularized model formalization to control over-fitting (Chen and Guestrin, 2016).

The second alternative, the Bayesian network, is a direct acyclic graph encoding assumptions of conditional independence. The nodes are stochastic variables, and the arcs show the dependency between the nodes. A Bayesian network is defined by a finite set N = {A,B,...} of nodes (vertices), a set L of arcs (edges) and a joint probability density function. In this sense, a Bayesian network classifier is simply a Bayesian network applied to classification, that is, the prediction of the probability of some discrete (class) variable Y given some features X (Zaidi et al, 2013). When we apply both extreme gradient boosting or Bayesian networks, we identify that the most relevant factors that explain POS adoption and the share of P2B electronic payments coincide with those identified by random forests. However, as shown in Table A2-2, when we use a subsample of data (30%) to check the out-of-sample accuracy of the predictions, the random forest outperforms the alternative machine learning algorithms and the logit model.

**Table A2.2: Alternative Model Performance in terms of Predictive Accuracy**

|  | Out-of-sample accuracy (70/30% split) | |
| --- | --- | --- |
|  | POS adoption | Share of P2B electronic payments |
| **Random forest** | 89.91% | 92.14% |
| **Extreme Gradient Boosting** | 80.17% | 82.29% |
| **Bayesian Networks (Naive Bayes)** | 59.94% | 42.27% |
| **Logit** | 76.12% | 58.15% |

**Table A2.3: Random Forest Hyperparameters and Cross-Validation of the Algorithm**

|  | Panel A. Random Forest Hyperparameters | |
| --- | --- | --- |
|  | POS adoption | Share of P2B electronic payments |
| **Number of Trees** | 1,000 | 1,000 |
| **Number of Features for each Tree** | 9 | 12 |
| **Maximum Depth of the Tree** | 20 | 20 |
|  | Panel B. Cross-validation accuracy | |
| **K-fold cross validation** | 88.31% | 71.88% |
| **Repeated K-fold cross-validation** | 89.04% | 73.35% |

Additionally, as the random forest require selecting some hyperparameters (i.e., number of features for each tree), which are tuned to obtain the optimal parameter values for higher accuracy. The performance of all machine learning methods is computed after optimizing the hyper-parameters for each method.[32] Finally, in order to check the stability of the accuracy of the results, we employ two cross validation methods: the k-fold cross-validation and the repeated K-fold cross-validation. In doing so, the dataset is split into 10 groups (k=10), since this value has been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance. In case of the repeated K-fold cross-validation, the data is split into 10-folds,

---

[32] We employed the following R packages: tune, caret, tuneRF and xgboost.

repeating the process five times. The results reported in Table A2.3 confirm the validity of the random forest model.