How Do Bank Customers Go Digital? A Random Forest Approach

Carbó-Valverde, Santiago (scarbo@cunef.edu)
CUNEF, Bangor University and Funcas

Cuadros-Solas, Pedro J. (pedro.cuadros@cunef.edu)
CUNEF and Funcas

Rodríguez-Fernández, Francisco (franrod@ugr.es)
University of Granada and Funcas

July 2018

Abstract

This study examines the sequence of decisions that bank customers follow to adopt digital services and diversify the use of those services. The sequential approach relies on a random forest model applied to an in-depth survey on consumer preferences for financial services. The results show that the adoption of digital banking services starts with information-based services (e.g. checking account balance), and it is then followed by transactional services (e.g. online or mobile money transfer). However, the diversification of the use of online channels is mainly explained by the consciousness about the range services available and the perception that they are safe. The findings also reveal that bank customers adopt non-bank payment services only once they are frequent and diversified digital bank customers. This suggests a certain degree of complementary between bank and non-bank digital channels. The random forest model is shown results to outperform the forecasting accuracy of parametric econometric models.

JEL classification: G21, O33
Key words: Technology adoption, banks, machine learning

1. Introduction

Digitalization is changing the shape of many industries and the way companies and clients interact. Banking services is no exception. The banking firm is particularly sensitive to the transformation of information systems, the treatment of personal data, and the emergence of new (fully digital) competitors and delivery channels.

On the supply side, financial institutions have reacted gradually to these changes. Despite incorporating online distribution channels two decades ago, and in spite of the renewed digitalization wave, banks continue to intensify their digital footprints. This effort is driven by both rival precedence (Hernández-Murillo, Llobet, & Fuentes, 2010) and changes in demand (Campbell & Frei, 2010).

A large number of studies on banking organization and technology have addressed the adoption of the most basic electronic banking services developed over the last few decades, including debit and credit cards and, more recently (although partially covered), online banking. These studies have found that perceived security, usefulness, quality, and convenience drive the adoption of those services by consumers (Casaló, Flavián, & Guinalíu, 2007; Hoehle, Scornavacca, & Huff, 2012; Laukkanen, 2016; Maria Correia Loureiro, Rüdiger Kaufmann, & Rabino, 2014; Yoon & Barker Steege, 2013; Yusuf Dauda & Lee, 2015). However, the relevance of each one these factors depends on the stage of the adoption. This is an important lesson for new digital services given the heterogeneous penetration that they have, both geographically and demographically (Montazemi & Qahri-Saremi, 2015). This is particularly relevant considering that socio-demographic characteristics—age, gender, income, or location—(Jaruwachirathanakul & Fink, 2005; Laukkanen, 2016; Tsai, Zhu, & Jang, 2013) as well as customer experience (on other products with varying levels of technological sophistication) are strongly related to the demand of online banking services (Szopiński, 2016).

While most prior studies have focused on the determinants of adopting electronic or mobile banking services, there is little evidence on the decision process that leads bank consumers to go digital. Financially speaking, going digital means predominantly or exclusively using online or mobile banking. This transition is not trivial. The Organization for Economic Co-operation and Development (OECD) identifies some of the core properties and cross-cutting effects of the digital transformation (OECD, 2017) as the most important business challenge currently underway. Furthermore, the OECD recognizes banking as one of the sectors where such transformation is more relevant in economic, organizational, and social terms. This paper aims to examine the process by which consumers adopt digital banking services.

Unlike prior studies, our study use machine learning random forest techniques to predict the sequence of adoption of digital financial services using a wide range of indicators from a comprehensive survey specifically designed for this purpose. We compare the results of the random forest approach with those of conventional econometric methodologies used in other consumer demand studies. These methodologies only identify the determinants of adoption rather than the sequence of adoption. Machine learning, instead of being limited to making strong assumptions about the structure of the data, allows researchers to identify and display complex patterns in a data-driven form (Bishop, 2006). In our case, the use of algorithms that establish a set of decision trees allow us to run random forest regressions that reveal how individuals make their financial digitalization choices. We show that these random forest models outperform standard logit and ordered logit models, not only because they show the sequence of adoption, but also on the forecasting accuracy of the adoption decision.

Our paper offers a twofold contribution to the existing literature on technology adoption. Firstly, we explore the sequence of the adoption of digitalization services by bank customers. Unlike previous studies, ours does not limit its scope to analysis of adoption vs. non-adoption; it explains how consumers make their decisions and how they become frequent and diversified users of digital financial services. Moreover, we show how the adoption process of digital banking services is related to other non-bank digital financial services (e.g. Paypal or Amazon). Secondly, by employing random forest techniques, this paper offers a greater statistical accuracy than earlier studies in describing the main determinants of consumers' choice to adopt digital financial services.

The empirical analysis relies on extensive data collected from a survey on digital banking and payment services of 3,005 consumers between the ages of 18 and 75. The survey included controlled representative quotas from a sociological standpoint based on age, sex, and location. This dataset allowed us to explore the financial digitalization in a developed country with deep internet penetration (84.6% of adults are internet users[1]), a highly banked population (97.2% of adults have a bank account[2]), and a growing use of electronic banking among consumers (62% of sample individuals are e-banking users to some extent, although the degree and scope of the adoption varies substantially across individuals[3]).

By way of preview, the results of our empirical analysis suggest that bank customers need to become familiar with the information content of digital services before they begin making financial transactions. Customers check their bank balances, make inquiries, and explore the possibilities of the digital channels before making payments, transferring money, or engaging in other transactional services. As for the scope of digitalization, the perceived safety of digital bank services by consumers becomes a critical filter for consumers' diversified use of digital bank services. However, there

[1] Survey on Equipment and Use of Information and Communication Technologies in Households (2017) conducted by the Spanish Statistical Office (INE).
[2] G20 Financial Inclusion Indicators. World Bank Data.
[3] The Online Banking Landscape in Europe. GlobalWeb Index (2017Q1).

appear to be noticeable exceptions. In the case of mobile banking, for example, even if perceived safety influences consumers' adoption decision, the speed and easiness of the device appears to be more decisive. The efficiency of this service contrasts with the adoption process of more traditional and more established bank services such as credit and debit cards, which are used on a regular basis only when they are perceived as safe and relatively costless. Finally, our results also indicate that consumers adopt other non-bank digital financial services (e.g. Amazon or Paypal) only after they have already become frequent and diversified digital bank customers.

The remainder of the paper is organized as follows: section II reviews the related literature; section III describes the dataset and the methodology employed; section IV addresses the digitalization dimensions; section V discusses the main empirical results from the random models and classification trees; and section VII concludes.

2. Related Literature

The main relevant studies related to financial technology adoption in the digital age refer to firm management and information systems. A number of theories aim to explain the evolution of these new technologies and the interaction between the consumer and the firm. Among them, the Technology Acceptance Model (TAM) (Davis, Bagozzi, & Warshaw, 1989) and its latter versions (TAM2 and TAM3) have become popular in explaining how people accept and adopt new technology in the context of banking. The TAM model, which is based on the Theory of Reasonable Action (TRA) (Fishbein & Ajzen, 1975) and the Theory of Planned Behaviour (TPB) (Ajzen, 1985, 1991), suggests that a technological adoption depends on customers' perception of the utility and ease of use of the technology. Other theories, such as the Diffusion of Innovations (DIT), the Task-Technology Fit (TTF), the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Technology Resistance Theory (TRT) have complemented the drivers of online adoption. These theories have thereby given prominence to a number of technological components of the service and not only to consumers' perceptions.

From an empirical standpoint, prior studies on customers' perceptions have identified the main factors that explain the adoption and utilization of online banking. These include security (Casaló et al., 2007; Cheng, Lam, & Yeung, 2006; Gerrard, Barton Cunningham, & Devlin, 2006; Hoehle et al., 2012; Vatanasombut, Igbaria, Stylianou, & Rodgers, 2008; Yoon & Barker Steege, 2013), ease of use (Aldás-Manzano, Lassala-Navarré, Ruiz-Mafé, & Sanz-Blas, 2009; Lee, 2009; Maria Correia Loureiro et al., 2014; Yoon & Barker Steege, 2013; Yusuf Dauda & Lee, 2015), convenience (Maria Correia Loureiro et al., 2014; Yoon & Barker Steege, 2013), and cost (Huang, Makoju, Newell, & Galliers, 2003; Laukkanen, 2016). Overall, consumers use e-banking services when they perceive them as safe, useful, convenient, and relatively costless.[4] As for the relative importance of these factors, Hoehle et al. (2012) survey the literature and conclude that security is found to be a major determinant of consumers' use of e-banking services. Additionally, many of these studies highlight that a range of socio-demographic characteristics also influence the adoption of online banking services (Jaruwachirathanakul & Fink, 2005; Laukkanen, 2016; Tsai et al., 2013). Specifically, young people who have a higher income and live in areas of high internet penetration (Laukkanen, 2016; Veríssimo, 2016; Xue, Hitt, & Chen, 2011) are prone to use online services. However, as Montazemi and Qahri-Saremi (2015) underline, the importance of these socio-demographic factors depends on the stage of the adoption of online banking services within each market segment or jurisdiction. It is also worth noting that Hitt and Frei (2002) explore the differences between branch-based and online bank customers.

[4] (Hoehle et al., 2012) and (Dahlberg et al., 2015) provide a detailed coverage of the literature within the last three decades.

They suggest that online banking customers are apparently more profitable, primarily due to unobservable characteristics that existed before the adoption of online banking. Moreover, Szopiński (2016) finds that having other banking products such as mortgages and credit cards also have a significant influence on consumers' use of online banking services.

Closely related to online banking, studies on mobile banking adoption have also recently emerged. The empirical and theoretical approaches in these studies are similar to those of online banking (Alalwan, Dwivedi, & Rana, 2017; Baptista & Oliveira, 2015; Lu, Tzeng, Cheng, & Hsu, 2015; Luo, Li, Zhang, & Shim, 2010; Susanto, Chang, & Ha, 2016; Zhou, Lu, & Wang, 2010). The results of these studies suggest that age is the most decisive factor in mobile banking adoption. However, other determinants such as trust in the device, security, and cost have also been reported to strongly influence the adoption of mobile payments (Dahlberg, Guo, & Ondrus, 2015).

The finance and banking literature has also examined online banking but mainly focuses on its impact on bank competition and performance. In line with the studies shown above, Hernández-Murillo et al. (2010) find that banks' adoption of new technologies, such as online banking services, is also partially triggered by their competitors' adoption of the technology. Xue et al. (2011) find that when consumers go digital, they acquire more products from the bank and make more transactions across different channels. Campbell and Frei (2010) document a positive relationship between the use of online banking and customer retention. DeYoung, Lang, and Nolle (2007), Hernando and Nieto (2007), and He (2015) show that online banking has a positive effect on bank performance, being a complement channel rather than a substitute for bank branches.

3. Data and Methodology

3.1 The Survey

The primary data for this study were collected from a consumer survey that was conducted specifically for this research by IMOP during November and December 2016. The survey participants, from a population of Spanish consumers between the ages of 18 and 75, were asked about their digital preferences and, in particular, about those related to banking and payment services. The main structure of the survey followed the Survey of Consumer Payment Choice (SCPC) conducted by the Federal Reserve Bank of Boston. However, our survey incorporated comprehensive information on consumers' digital preferences and not only on payment services. Controlled quotas for a representative sample of the population were established based on age, sex, and location. The survey was conducted via telephone interviews and resulted in a sample size of 3,005 consumers; participation was voluntary. The sample error is estimated to be ±1.8% for a confidence level of 95.5%.[5]

Spain seems to be a good laboratory for this study, as it has overcome the initial implementation phase of electronic banking[6] and ranks third in the world for annual growth in mobile banking adoption.[7]

As Table 1 illustrates, the gender breakdown was 49.7% men and 50.3% women. The largest percentage of participants fell into the age bracket of 35–44 years old (22.8%), followed by 45–54 years old (21%). In terms of the employment, roughly 60% of participants were employed. The median number of the household members was three.

Consistent with official statistics, 92% of participants were frequent internet users, connecting mainly from home; 75% of them reported having a laptop, 97% reported having a mobile phone (85.3% a smartphone), and 47.2% reported having a tablet.

---

[5] All the variables extracted from the survey questionnaire are listed in the appendix.
[6] There are over 15 million of e-banking services users in Spain according to Arellano & García, (2017)
[7] Ditrendia Mobile report in Spain and the world (2016)

Table 2 provides insight on the degree of digitalization by gender, age, and employment situation. Importantly, there seems to be a gap (common to most advanced countries) between the availability of the online services and their (partial or exclusive) use by consumers. In any event, the figures suggest that Spanish consumers have attained a medium-high degree of digitalization and a medium degree of financial digitalization. In general, it seems that adults under the age of 45 (working or studying) are the most digitalized.

3.2 Descriptive Statistics

3.2.1 Degree of Banking Digitalization

Figure 1 plots statistics regarding the number of accounts and the number of financial entities per customer. On average, each banking client had 2 bank accounts and operates with 1.5 entities. It is worth noting that while 79.6% of respondents had an online bank account, only 13% were exclusively online account users. Regarding the type of financial activities conducted online, internet users reported accessing online banking services to check account their balance/transactions (68.7% of respondents), to receive online communications from their bank (51.4%), and to make payments or transfer money (50.9%). In the case of mobile banking, the activities lean even more toward checking and communication rather than transactional services.

Figure 2 illustrates the degree to which consumers use various financial services. Debit cards (78.1% of respondents reported using) seem to dominate over credit cards (50.8%). As for the most common uses, 56% of internet users check the balance of their accounts weekly, either by mobile, tablet, or computer, while only 32.4% check their credit card balance weekly. Table 2 also illustrates the degree of financial digitalization by gender, age, and employment situation. Young and employed people exhibit the largest degree of financial digitalization. Furthermore, accounting for all the socioeconomic features, the typical profile of a digital banking consumer would be an employed woman, under 39 years old, with children, living in a large residential area of more than 200,000 inhabitants, and with a monthly household income between €3,000 and €5,000.

3.2.2 Consumer Perceptions

According to the results of the survey, 88.8% of respondents considered cash to be safe or very safe, while such a statement was only made by 58.8% of respondents regarding online banking and only 44.2% of respondents regarding mobile banking. As for perceived cost, 63.2% of respondents considered online banking to be a low-cost or costless service; 58.8% of respondents said the same of mobile banking. While more than 90% considered it to be easy or very easy to withdraw cash at ATMs or pay by debit card, this was only the case for 67.8% and 64.4% of online and mobile banking users, respectively. However, online banking and mobile banking were perceived as high-quality services by 86.2% and 84% of users, respectively.

3.2.3 Non-Banking Services and Social Networks

Importantly, 38% of respondents indicated that they used at least one non-banking method of payment (Amazon Pay, Google Wallet, PayPal, Apple Pay, etc.). Consumers that reported using non-bank services had on average more than one non-bank account (1.47 account per person). Moreover, 20.6% of respondents also reported installing a mobile app in order to make payments. Although 70% of respondents had a Facebook account, and 28% had a Twitter account, users preferred email as the main channel to communicate with (30.5%) or make complaints to their bank (17.7%).

3.3 A Random Forest Approach

Most previous studies have employed discrete choice models to examine consumer preferences on payment and other financial services (Dick, 2008; Hernández-Murillo et al., 2010; Honka, Horta, & Vitorino, 2017; Yusuf Dauda & Lee, 2015). These models, derived from utility theory, are based on maximizing consumers' utility. Other studies have used structural equations. These structural equations are useful to impute relationships between latent variables that affect e-banking adoption (Aldás-Manzano et al., 2009; Maria Correia Loureiro et al., 2014; Montazemi & Qahri-Saremi, 2015).

In addition to these traditional approaches, machine learning offers an alternative to statistical approaches for modeling consumers' financial choices. The development of computational engineering and big data analysis has allowed the growth of a scientific discipline where algorithmic systems learn automatically. These algorithms are able to identify complex patterns among millions of data points in order to make inferences and predictions. Among these techniques, the random forest approach has proved particularly accurate (Varian, 2014). It exhibits several advantages for our purposes. First, no pre-established or strict assumptions are required regarding the structure of the data. Second, by generating hundreds of random decision trees, it allows to reveal the most common decision sequences. Therefore, the final outcome improves our understanding of what factors are the most commonly considered in a decision-making process. Additionally, by identifying these characteristics, we are able to build classification trees that illustrate the sequence of consumers' decision-making actions.

Statistically, random forests are an ensemble of tree predictors in which each tree depends on the values of a random vector sampled independently and with the same distribution for all trees within the forest (Breiman, 2001). The algorithm follows these steps:

1. A forest of many trees is grown—1,000 trees in our research; Each tree is grown from an independent bootstrap sample derived from the data.

2. For each node of the tree, m variables are independently selected at random out of all M possible variables. Then, on the selected m variables it finds the best split.

3. The algorithm grows each tree to largest extent possible.

4. These steps are iterated over all trees in the ensemble, and the average vote of all the trees is reported as the random forest prediction.

The use of random forest regressions in economics is gaining ground. In the case of banking, Miguéis, Camanho, and Borges (2017) use a random forest model to predict responses to direct banking marketing. They find that the forecasting power of random forest models outperforms many other methods. Other finance-related studies, such as De Moor, Luitel, Sercu, and Vanpée (2018) also suggest a greater accuracy of random forests models (compared with other standard approaches) when examining the determinants of sovereign credit ratings. Similarly, Long, Song, and Cui (2017) analyze the influence of capital operations on the performance of listed companies and conclude that random forest algorithms have the highest classification accuracy and are more stable under different threshold definitions.

Some macroeconomic studies have used a random forest approach to predict the likelihood of default of some European countries (Behr & Weinblat, 2017), the euro area gross domestic product (GDP) forecasting (Biau & D'Elia, 2011) and predict the probability of occurrence of a banking and currency crisis in developed countries (Joy, Rusnák, Šmídková, & Vašíček, 2017).

In microeconomics and consumer theory, Bajari, Nekipelov, Ryan, and Yang (2015) survey and apply several techniques for demand estimation. They conclude that a random forest approach is both adequate and effective in estimating changes in demand.

Such an approach has also been used to estimate consumer preferences for technology products (Chen, Honda, & Yang, 2013) and travel choices (Hagenauer & Helbich, 2017).

4. Dimensions of the Digitalization Process

Going digital is a much broader concept than is commonly understood. While literature on the global digitalization of societies utilizes a multidimensional approach to explore the digitalization (Cruz-Jesus, Oliveira, & Bacao, 2012; Vehovar, Sicherl, Hüsing, & Dolnicar, 2006), previous studies on the financial digitalization of consumers have mainly focused on the adoption of online channels. However, our study assumes a broad definition of adoption that considers not only the first use of a certain service, but also its scope and frequency. As the OECD suggests, it is convenient to apply a multidimensional approach to explore the digital transformation of bank customers. Figure 3 plots the main dimensions that we identified from earlier studies: adoption of digital banking, diversification of use, and adoption of bank and non-bank payment instruments (Campbell & Frei, 2010; Montazemi & Qahri-Saremi, 2015; Szopiński, 2016; Xue et al., 2011; Yusuf Dauda & Lee, 2015).

4.1 Adoption of Digital Banking

What drives becoming a digital customer of banking services on a regular basis? Making use of the comprehensive set of variables in our survey on general digitalization and financial digitalization, we classified individuals into three categories: non-users (F), occasional users (N), and frequent users (S). Non-users are defined as those who have not adopted any financial digitalization, including those who are not even digitalized consumers (i.e., do not use the internet). Respondents who checked their account balance online and carried out at least one other online financial activity over the course of the year were classified as occasional users. Finally, frequent users were those who checked their account balance and carried out other (transactional) online activities at least once a month. Figure 4 shows that 1,772 out of the 3,005 respondents (58.9%) were frequent users of online financial services, a figure that is consistent with the growth of online banking in Spain officially reported in the so-called European Digital Agenda.[8]

4.2 Diversity of Digital Use

While the initial phase of the digital transformation of consumers involves regular online access, going digital is also related to consumers' use of diverse digital services. Then, going digital means conducting a number of financial activities online and not a single online activity—as it is usually the case of checking the account balance. Within this dimension, we acknowledge that there is a transition between beginning to go digital and becoming an "omni-digital" bank customer.

The factors that drive consumers' digital diversification might be different depending on the capabilities of the electronic device used in access the service. Therefore, we differentiate between the diversification of online banking users and mobile banking users. In doing so, survey respondents were classified according to their variety of uses they carry out (check account balances, pay bills, make transfers, or receive communications) for each type of terminal used to conduct these activities (computer or mobile). Based on these factors, respondents were then sorted into four categories: no digital users, non-users of digital financial services, incipient users of digital financial services, and diversified users of digital financial services.

Individuals who are outside of the digitalization process (i.e. who had no access to the internet) were classified as no digital users. Individuals who are frequent internet users but do not conduct any financial activity online were classified as non-users of digital financial services. Incipient users are those who perform some but not all online financial activities at least once a month. Finally, those users that carry out all financial

[8] http://www.agendadigital.gob.es/digital-agenda/Paginas/digital-agenda-spain.aspx

activities online at least once a month are classified as diversified users of digital financial services. Figure 4 reveals that most of the respondents are incipient users—a finding that reflects the worth of exploring this dimension. Bank customers appear to also be customers of digital financial services, but they are still far from being considered as "omni-digital" users.

4.3 Use of Banks' Payment Instruments

Another dimension that determines the financial digitization process relates to a consumer's method of payment. Although debit and credit cards cannot be considered fully new electronic payment instruments, we also consider them as there has been a technological and safety evolution. There were new, varied, and easy ways (such as contactless technology) of using them.

The sample was then divided into two groups: non-debit (non-credit) card users and debit (credit) card users. As Figure 4 shows, there was a larger use of debit cards in comparison to credit cards.

4.4 Use of Non-Bank Payment Instruments

While banks have traditionally offered non-cash payment instruments, some technology companies, particularly BigTech and FinTech, have begun offering non-banking alternatives to pay bills or transfer money (Amazon Pay, PayPal, Google Wallet, Apple Pay, etc.). The adoption of these new means of payment whose provider is not a financial entity has gained ground. Since most of the technological transformation is being led by the irruption of high-tech companies, it is interesting to analyze how consumers adopt these alternative means of payments. Therefore, this paper considers what factors drive consumers to use non-bank payment instruments.

In our research, customers were classified as non-digital users, non-users of non-bank payment instruments, and users of non-bank payment instruments. Consumers without regular internet use were classified as non-digital users. Consumers of online financial services who did not use non-bank means of payment were classified as non-users of non-bank payment instruments. Finally, users of non-bank payment instruments include consumers that utilized payment methods of non-bank providers. As illustrated in Figure 4, most respondents were non-users of non-bank payment instruments, despite being digitalized.

5. Results

In this section we present the random forest regression results of each of the abovementioned dimensions as well as the classification trees that outline the sequential digitalization of bank consumers.

Firstly, we used 1,000 decision trees, randomly constructed, for each dimension, using the set of variables provided in the survey in order to obtain the random forest output. Then, we reported the plots showing the relative statistical importance of each factor in the classification of individuals by their digital profile. The determinants and characteristics are plotted on the *y*-axis ranked by their absolute level of importance while their relative importance is charted on the *x*-axis. The mean decrease in accuracy reflects the mean loss in accuracy when each specific variable is excluded from the regression algorithm. Therefore, the determinants and characteristics with the greater mean decrease in accuracy are the most relevant for the classification of bank customers. Additionally, the mean decrease in Gini is a measure of how each feature contributes to the homogeneity between the decision trees that were used in the resulting random forest.

Secondly, we used the characteristics and determinants with the largest discriminant power for each of the digital dimensions to build a decision tree. A conditional inference tree was estimated. This technique estimated a regression relationship by binary recursive partitioning in a conditional inference framework. The algorithm tested 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. Then, the algorithm implemented a binary split in the selected input variable and recursively repeated this process for the each of the remaining variables. The classification tree inferred the sequencing of customers' decision-making process, which helped in explaining how bank customers go digital. This is particularly relevant since those trees do not require any linearity assumptions, which is important because many of the digitalization determinants could be nonlinearly related.

5.1 Random Forest Regression Results

5.1.1 Adoption of Digital Banking

The matching learning algorithm revealed that the following bank customers' features stand out as first-order factors that differentiate between non-users (F), occasional users (N), and frequent users (S):

- *Online check balance:* indicates whether account balances are checked online. As it is easier, faster, and less costly than physically going to the bank branch, it fosters going digital.
- *Number of online bank accounts*: indicates the scope of digital banking. Offering online access to bank customers when they open a bank account increases the probability of the customer going digital.
- *Online transfers:* indicates whether the customer has made an online bank transfer over the last three months. Online bank-to-bank transactions are a driver of transactional financial digitalization.
- *Consciousness*: it is the ratio of the number of bank accounts that the customer believes have online access to the total number of accounts with online access. It indicates the degree to which each customer is aware of the existence of online

financial services at his or her disposal as in practical terms all the accounts offer the possibility of online access. Honka, Horta, and Vitorino (2017) argue that customer awareness is a relevant factor in the use of banking services.

Bank customers' perceptions of security, cost, or ease of use of banking services were found to be secondary factors in going digital. The decision to adopt a digital profile did not seem to be primarily motivated by customers' perceptions. Our results suggest that the relevant factors in going digital are those related to customers becoming accustomed to the online channels by checking their bank account balances or transferring money and related to being aware that these activities can be conducted online.

As in other industries, consumers tend to go through several stages of adoption: awareness, consideration, and choice. Our results confirm the significance of awareness in the multistage process of going digital.

5.1.2 Diversity of Digital Use: Online Banking and Mobile Banking

Figures 6 and 7 show the baseline random forest results in terms of the diversification of online and mobile banking services, respectively. We found that the following features have the largest influence on increasing customers' adoption of online banking services:

- *Number of online bank accounts:* in addition to adopting a financial digital profile, bank customers' degree of online diversification depends on how many of their accounts offer digital access.
- *Consciousness:* which is to say, being aware of the possibility of having access to online services, is essential for the customers to diversify their financial activities.
- *Safety of online banking:* indicates how customers perceive the level of security of online banking.

- *Online banking communication:* indicates whether customers have used online services or e-mail as their communication method with their bank.

Considering both the adoption and diversification of digital use, we argue that the digitalization process originates from the customers' need to check their bank account balances and transfer money. However, being aware of the possibility of accessing financial services through online banking and the perceived safety of operating online were the main factors in determining whether customers diversified their use of online banking services. Furthermore, the digitalization of the communication channel between customers and banks also fostered the diversification of customers' online activities.

Regarding the diversification of the use of mobile banking, the following factors had the greatest predictive power:

- *Number of online bank accounts*: as is the case with online banking, the degree of mobile banking diversification depends on how many online accounts are available to the customer.
- *Safety mobile banking:* regards bank customers' perception of the level of security of mobile banking, which is also relevant to them deciding to go broadly digital with mobile banking.
- *Consciousness:* being aware of the possibility of having access to financial services is again relevant, influencing the diversification of mobile-related services by bank customers.
- *Transferring money via mobile:* rather than information checking (as it was the case with online banking), mobile banking diversification seems to be driven by transactional services:

Overall, the algorithm reveals that online and mobile diversification were driven by common features: consciousness of the possibilities offered by digital banking, perceived level of security of the channel used, and the number of digital bank accounts available. However, it is worth noting that transferring money was a distinct factor in determining diversification of mobile banking. One of the firsts steps to become "omni-digital" in mobile banking seems to begin by transferring money. It seems that money transferring via mobile may turn into the gateway for using other digital financial activities. This finding partially explains the importance of the irruption of FinTech companies in the payment sector compared to other financial services (Eickhoff, Muntermann, & Weinrich, 2017; Jagtiani & Lemieux, 2017).

5.1.3 Use of Banks' Payment Instruments: Debit and Credit Cards

Consistent with prior estimations, employing 1,000 randomly computed forest trees, we determined the main factors that influence the use of debit and credit cards, respectively (see Figures 8 and 9):

- *Cost:* Customers' perceived cost affects the usage of both types of cards, although it has a greater impact on the use of debit cards.
- *Safety:* The perceived safety of the transactions conducted with debit and credit cards is relevant in determining the their use by customers, although to a slightly greater extent with credit cards.
- *Acceptance:* Merchants' acceptance of debit and credit cards as payment instruments determines their utility, which could explain why bank customers are concerned with ensuring their acceptance before adopting them as regular payment instruments.
- *Convenience:* Customers' perception regarding the convenience of using these banking payments (easiness, time saving, etc.) also influences customers' use of them.

Unlike the adoption and penetration of online and mobile banking, the use of debit and credit cards seems to be dominated by bank customers' perceptions of cards' cost, safety, and acceptance.

5.1.4 Use of Non-Bank Payment Instruments

Figure 10 illustrates the most relevant factors in explaining customers' adoption of non-bank payment services.

- *Mobile payment app*: Customers' use of mobile apps to make payments has a large predictive power in determining whether customers will use non-bank payment instruments.
- *Frequency and degree of online banking:* The scope and frequency of customers' use of online banking services was also found to explain the use of non-bank payment instruments, suggesting a complementarity between the bank and the non-bank payment alternatives.
- *Online banking complaint:* Customers' use of online channels to lodge a complaint with the bank also appears to drive their use of non-bank services. In other words, unsatisfied bank customers making online complaints are more prone to adopt non-bank means of payment.
- *Twitter and Facebook user:* Being a user of social media also appears to be related to the use of non-bank payments.

5.2 Accuracy: Random Forest vs. Logit Models

The previous section described the main factors that influence customers' online adoption and diversification. Then, we used random classification trees to determine the sequence in which these factors operate. However, before estimating these trees, it is important to determine the prediction accuracy of the random forest regressions. Consistent with prior studies (e.g. De Moor et al., 2018), we randomly selected 70% of the data as training data (2,104 observations) and designated the remaining data (901 observations) as test data. Through this process, we aimed to show the out-of-sample fit precision of the random forest technique. The machine learning algorithm was able to accurately predict 88.41% of bank customers' online banking adoption profile, 70.11% of diversity of digital use of online banking, 70.01% of diversity of digital use of mobile banking, 85% (74.89%) of debit (credit) card adoption, and 76.14% of non-bank payment instruments adoption.

We also compared the baseline results obtained using a random forest technique with the standard discrete choice models used in most of the previous studies to analyze consumer preferences of financial services. We used ordered logit regressions for the adoption decision and the diversification of digitalization usages because they rank consumers according to certain classifications (as shown above). However, the decision between bank or non-bank payment instruments is binary, so it is estimated using a simple conditional logit. The general form of the logit model is as follows:

$$\mathrm{E}(\mathrm{Y}\,|\mathrm{X}=\mathrm{x}) = \Pr(Digital\ dimension = y \mid \mathrm{X}) = \Lambda\,(\beta_0 + \beta_1 X_{personal\ features} + \beta_2 X_{socio-demographics\ feautures} + e_i) \quad (1)$$

For each digital dimension tested, we included as regressors bank customers' features ($X_{personal\ features}$), which according to prior theoretical and empirical studies are the most relevant factors for going digital. These variables were classified into four different subsets: degree of digitalization, financial profile, perceptions, and social profile. The vector of variables $X_{socio-demographics\ feautures}$ accounts for gender, age, people living at home, and geographical location. The models were estimated considering the sampling weights. Furthermore, the digitalization process may be influenced by customers' choice of bank, as some banks are more digitalized than others. In order to address this issue, errors are clustered on the main bank of each bank customer. The estimation results are reported in Appendix B.

The ordered logit and simple logit models were able to accurately predict 79.27% of bank customers' online banking adoption, 55.01% of diversity of digital use of online banking, 59.57% of diversity of digital use of mobile banking, 84.23% (70.62%) of debit (credit) card adoption, and 73.46% of non-bank payment methods adoption. Table 3 compares the forecasting accuracy obtained using the random forest regressions and the logit models. Random forests models (both in out-of-sample and whole-sample tests) outperformed logit models. The greater predictive power was particularly relevant for the adoption of online banking whose fitting ability is close to 90% for both the whole-sample and out-of-sample predictions in the random forest while it was 80% for the logit model. Similarly, the accuracy of the prediction of the diversified use of online (and mobile) banking was 70% with the random forest model approach compared to a 55% (and 59%) for the logit models. Consistent with prior studies (Cui, Moreno, & Zhang, 2017; De Moor et al., 2018; Krauss, Do, & Huck, 2017; Long et al., 2017; Miguéis et al., 2017), the random forest algorithm presented a higher classification accuracy compared to alternative econometric models.

5.3 Classification Trees

In order to obtain a sequence of customers' financial digitalization decision, we used those variables identified by the random forest as having larger predictive power to build a decision tree for each of the dimensions analyzed. We initially tested whether the decision trees maintained the prediction accuracy of the baseline random forest models. The trees were able to accurately predict around 70–85% of individual choices.

5.3.1 Tree: Adoption of Digital Banking

Figure 11 plots the decision tree of customers' adoption of digital banking. Although the range of services available online is wide, the adoption of online banking seems to emerge from customers checking their account balances. It was only after customers checked their account balances that they moved into transferring money online. Bank customers who did not perform either of these activities were classified as occasional or low-frequency users (node 5). Comparing those individuals who only check their account balances (node 10) with those who only transfer money (node 7 and 8), checking account balances appears to be the more decisive step in becoming a frequent user of online banking services. Furthermore, when customers begin to make transactions and are largely aware of the online possibilities, they become frequent users (nodes 14 and 15).

In conclusion, an overview of the random models and the classification trees suggests that the main channel by which bank customers become frequent users of online banking services is by their need to check their account balances and, subsequently, transfer money. Consciousness about the availability of online possibilities available is also important for the customer to become a frequent digital bank user. Furthermore, the perceived safety of online banking services was not a primary determinant in becoming a frequent user. This finding is relevant since most of the literature has concluded that adoption is mainly driven by consumers' perceptions, including their perception of safety. As we show in the next sub-section, safety only becomes influential when customers consider to conduct a wide range of services.

5.3.2 Tree: Diversity of Digital Banking Use

Figure 12 illustrates the classification tree for the diversity of digital use in online banking with four main variables. This tree reveals the relevance of the perceived security of online banking in influencing customers' use of online financial services (branch 2). Customers who considered online banking to not be safe were not likely to become diversified users of online services (node 14-21). Together with safety, customers' use of digital channels for information purposes and their awareness of the range of online services were key determinants of the diversification of digital services demanded (node 11). However, consciousness did not compensate for the perceived lack of safety. At most, being conscious made customers switch from non-users to incipient users (node 17-21).

Overall, the results suggest that while being a regular online banking user is driven by customers' needs (e.g. checking account balances and transferring money) as well as by having a certain level of consciousness about the online possibilities, becoming a diversified digital user depended largely on the perceived level of safety.

Figure 13 plots the classification tree for the diversity of digital use of mobile banking. The results suggest that the diversity of online and mobile banking use is driven by similar factors. The perceived level of safety of mobile banking is also relevant (node 7). It is also shown that is unlikely to find diversified users who have not transferred money with their phones even if they perceive mobile banking as not safe (node 5).

5.3.3 Tree: Adoption of Bank Payment Instruments

Figures 14 and 15 plot the classification trees for the debit and credit card adoption, respectively. Both trees show that safety and cost are the main drivers of adoption.

Regarding the adoption of debit cards, customers' perception that debit cards are a convenient payment instrument was a primary determinant of their use. Debit card users could be classified into users who consider debit cards safe, accepted but not very convenient regardless of their cost (node 11), and users who consider the method convenient, costless, and safe (node 24 and 26). It could be then argued that a costless perception could compensate a lack of perceived convenience.

In the case of credit cards, the most influential factor was the perceived safety. Customers who perceived credit cards as unsafe regardless of their cost were less likely to use them (nodes 14-19). Similar to debit cards, users who perceive credit cards as safe and relatively costless made up the majority of credit card users (node 12). The probability of adoption dropped to 12% if the credit cards were considered costly.

5.3.4 Tree: Use of Non-Bank Payment Instruments

The classification tree for the adoption of non-bank payment instruments is shown in Figure 16. This tree reveals that the adoption of non-bank payment methods occurred when customers were frequent and diversified digital banking users. For occasional and incipient online users, the likelihood of using non-bank payment instruments was quite small. However, as the frequency and diversity of use increased, being active on social media and making mobile payments increased the likelihood that customers would use non-bank payment channels. Furthermore, being active on social media and using apps for mobile payments were also relevant factors. However, it is worth noting that frequent online user do not use non-bank payment methods if they are just incipient users (node 23); it is necessary for customers to undertake several digital financial uses to jump into non-bank payments. Similarly, digital banking users who do not have frequent online access are not regular adopters of non-bank payment methods (node 7, 16, 17, and 28).

6. Conclusion

Modern societies are undergoing a rapid digital transformation. A sizeable part of this change is related to the demand of financial services. The use of electronic devices such as smartphones, laptops, and tablets to conduct many financial activities has risen sharply. While the banking industry is aware of this transformation, adjusting the supply side depends on related changes in demand.

In this paper, we aim to offer a multi-dimensional comprehensive picture of the process by which bank customers become digitalized. While most previous studies discuss the determinants of certain adoption decisions, we outline the sequence of steps that customers follow to adopt and become a diversified user of digital financial services. We consider various dimensions: the adoption of online banking, the diversification of the use of online services, and the choice of bank vs. non-bank payment instruments. Our approach benefits from the use of machine learning techniques applied to an in-depth consumer survey specifically designed for the purpose of this study. Specifically, we run random forest models and regression classification trees.

The empirical results suggest that the digitalization process seems to be originated from customer need to get information on basic aspects of its banking accounts (e.g. checking their account balances), and this facilitates a transition to transactional services (e.g. transferring money). We also find that once the initial adoption has taken place the diversification of online and mobile services adopted by the customers became larger when they became conscious about the range of possibilities provided by the bank and when they perceive those options as safe. Furthermore, we show that the adoption of non-bank payment instruments (e.g. PayPal and Amazon) happens when consumers are already diversified bank digital customers. This suggests that a certain degree of complementary between bank and non-bank digital services.

Overall, these results confirm the need to conduct research that covers the entire digitalization process rather than focusing on a single dimension. In addition, our research confirms that the application of matching learning techniques on consumer research provides accurate results that improve the understanding of complex topics.

References

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. *Action Control*, 11–39.

Ajzen, I. (1991). The theory of planned behavior. *Orgnizational Behavior and Human Decision Processes*, *50*, 179–211.

Alalwan, A. A., Dwivedi, Y. K., & Rana, N. P. (2017). Factors influencing adoption of mobile banking by Jordanian bank customers: Extending UTAUT2 with trust. *International Journal of Information Management*, *37*(3), 99–110.

Aldás-Manzano, J., Lassala-Navarré, C., Ruiz-Mafé, C., & Sanz-Blas, S. (2009). Key drivers of internet banking services use. *Online Information Review*, *33*(4), 672–695.

Arellano, A., & García, J. R. (2017). Determinantes de la adopción de la banca digital en España. *BBVA Resarch*, 1–26.

Bajari, P., Nekipelov, D., Ryan, S. P. S. P., & Yang, M. (2015). Demand estimation with machine learning and model combination. *Working Paper*, 1–45.

Baptista, G., & Oliveira, T. (2015). Understanding mobile banking: The unified theory of acceptance and use of technology combined with cultural moderators. *Computers in Human Behavior*, *50*, 418–430.

Biau, O., & D'Elia, A. (2011). Euro Area GDP Forecasting Using Large Survey Datasets: A Random Forest Approach. *Euroindicators Working Papers*.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. *Information Science and Statistics* (Vol. 53).

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

Campbell, D., & Frei, F. (2010). Cost Structure, Customer Profitability, and Retention Implications of Self-Service Distribution Channels: Evidence from Customer Behavior in an Online Banking Channel. *Management Science*, *56*(1), 4–24.

Casaló, L. V., Flavián, C., & Guinalíu, M. (2007). The role of security, privacy, usability and reputation in the development of online banking. *Online Information Review*, *31*(5), 583–603.

Chen, H. Q., Honda, T., & 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), 61007.

Cheng, T. C. E., Lam, D. Y. C., & Yeung, A. C. L. (2006). Adoption of Internet Banking: An Empirical Study in Hong Kong. *Decision Support Systems*, *42*(3), 1558–1572.

Cruz-Jesus, F., Oliveira, T., & Bacao, F. (2012). Digital divide across the European Union. *Information and Management*, *49*(6), 278–291.

Cui, R., Moreno, A., & Zhang, D. (2017). The Operational Value of Social Media Information. *Production and Operations Management*, 1–33.

Dahlberg, T., Guo, J., & Ondrus, J. (2015). A critical review of mobile payment research. *Electronic Commerce Research and Applications*, *14*(5), 265–284.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003.

De Moor, L., Luitel, P., Sercu, P., & Vanpée, R. (2018). Subjectivity in sovereign credit ratings. *Journal of Banking & Finance*, *88*(July 2010), 366–392.

DeYoung, R., Lang, W. W., & Nolle, D. L. (2007). How the Internet affects output and performance at community banks. *Journal of Banking and Finance*, *31*(4), 1033–1060.

Dick, A. A. (2008). Demand estimation and consumer welfare in the banking industry. *Journal of Banking and Finance*, *32*(8), 1661–1676.

Eickhoff, M., Muntermann, J., & Weinrich, T. (2017). What do FinTechs actually do? A Taxonomy of FinTech Business Models. *Working Paper*, 1–19.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley.*

Gerrard, P., Barton Cunningham, J., & Devlin, J. F. (2006). Why consumers are not using internet banking: a qualitative study. *Journal of Services Marketing*, *20*(3), 160–168.

Hagenauer, J., & Helbich, M. (2017). A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*, *78*, 273–282.

He, Z. (2015). Rivalry , Market Structure and Innovation : The Case of Mobile Banking. *Review of Industrial Organization*, *47*(2), 219–242.

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

Hernando, I., & Nieto, M. J. (2007). Is the Internet delivery channel changing banks' performance? The case of Spanish banks. *Journal of Banking and Finance*, *31*(4), 1083–1099.

Hoehle, H., Scornavacca, E., & Huff, S. (2012). Three decades of research on consumer adoption and utilization of electronic banking channels: A literature analysis. *Decision Support Systems*, *54*(1), 122–132.

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

Huang, J., Makoju, E., Newell, S., & Galliers, R. D. (2003). Opportunities to learn from "failure" with electronic commerce: A case study of electronic banking. *Journal of Information Technology*, *18*(1), 17–26.

Jagtiani, J., & Lemieux, C. (2017). *Fintech Lending: Financial Inclusion, Risk Pricing, and Alternative Information*. *Working Paper, Federal Reserve Bank of Philadelphia*.

Jaruwachirathanakul, B., & Fink, D. (2005). Internet banking adoption strategies for a developing country: the case of

Thailand. *Internet Research*, *15*(3), 295–311.
Joy, M., Rusnák, M., Šmídková, K., & Vašíček, B. (2017). Banking and Currency Crises: Differential Diagnostics for Developed Countries. *International Journal of Finance and Economics*, *22*(1), 44–67.
Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702.
Laukkanen, T. (2016). Consumer adoption versus rejection decisions in seemingly similar service innovations: The case of the Internet and mobile banking. *Journal of Business Research*, *69*(7), 2432–2439.
Lee, M. C. (2009). Factors influencing the adoption of internet banking: An integration of TAM and TPB with perceived risk and perceived benefit. *Electronic Commerce Research and Applications*, *8*(3), 130–141.
Long, W., Song, L., & Cui, L. (2017). Relationship between Capital Operation and Market Value Management of Listed Companies Based on Random Forest Algorithm. *Procedia Computer Science*, *108*(June), 1271–1280.
Lu, M. T., Tzeng, G. H., Cheng, H., & Hsu, C. C. (2015). Exploring mobile banking services for user behavior in intention adoption: using new hybrid MADM model. *Service Business*, *9*(3), 541–565.
Luo, X., Li, H., Zhang, J., & Shim, J. P. (2010). Examining multi-dimensional trust and multi-faceted risk in initial acceptance of emerging technologies: An empirical study of mobile banking services. *Decision Support Systems*, *49*(2), 222–234.
Maria Correia Loureiro, S., Rüdiger Kaufmann, H., & Rabino, S. (2014). Intentions to use and recommend to others. *Online Information Review*, *38*(2), 186–208.
Miguéis, V. L., Camanho, A. S., & Borges, J. (2017). Predicting direct marketing response in banking: comparison of class imbalance methods. *Service Business*, *11*(4), 831–849.
Montazemi, A. R., & Qahri-Saremi, H. (2015). Factors affecting adoption of online banking: A meta-analytic structural equation modeling study. *Information and Management*, *52*(2), 210–226.
OECD. (2017). *OECD Digital Economy Outlook 2017*.
Susanto, A., Chang, Y., & Ha, Y. (2016). Determinants of continuance intention to use the smartphone banking services. *Industrial Management & Data Systems*, *116*(3), 508–525.
Szopiński, T. S. (2016). Factors affecting the adoption of online banking in Poland. *Journal of Business Research*, *69*(11), 4763–4768.
Tsai, C. H., Zhu, D. S., & Jang, Y. M. (2013). A study on the consumer adoption behaviors of internet bank. *Journal of Networks*, *8*(12), 2956–2963.
Varian, H. R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28.
Vatanasombut, B., Igbaria, M., Stylianou, A. C., & Rodgers, W. (2008). Information systems continuance intention of web-based applications customers: The case of online banking. *Information and Management*, *45*(7), 419–428.
Vehovar, V., Sicherl, P., Hüsing, T., & Dolnicar, V. (2006). Methodological Challenges of Digital Divide Measurements. *The Information Society*, *22*(5), 279–290.
Veríssimo, J. M. C. (2016). Enablers and restrictors of mobile banking app use: A fuzzy set qualitative comparative analysis (fsQCA). *Journal of Business Research*, *69*(11), 5456–5460.
Xue, M., Hitt, L. M., & Chen, P. (2011). Determinants and Outcomes of Internet Banking Adoption. *Management Science*, *57*(2), 291–307.
Yoon, H. S., & Barker Steege, L. M. (2013). Development of a quantitative model of the impact of customers' personality and perceptions on Internet banking use. *Computers in Human Behavior*, *29*(3), 1133–1141.
Yusuf Dauda, S., & Lee, J. (2015). Technology adoption: A conjoint analysis of consumers' preference on future online banking services. *Information Systems*, *53*, 1–15.
Zhou, T., Lu, Y., & Wang, B. (2010). Integrating TTF and UTAUT to explain mobile banking user adoption. *Computers in Human Behavior*, *26*(4), 760–767.

Table 1. Sample demographics

Gender	n	%
Male	1,493	49.68
Female	1,512	50.32
Age		
18 - 24 years	282	9.38
25 - 34 years	498	16.57
35 - 44 years	686	22.83
45 - 54 years	631	21.00
55 - 64 years	500	16.64
> 65 years	408	13.58
Habitat		
0 – 10,000 inhabitants	637	21.20
10,001 – 50,000 inhabitants	806	26.82
50,001 – 200,000 inhabitants	696	23.16
> 200,000 inhabitants	866	28.82
Nº People at home		
1 person	644	21.4
Two people	850	28.3
Three people	757	25.2
More than three people	754	25.1
Employment situation		
Working	1,815	60.4
Pensioner/retired	500	16.6
Unemployed	338	11.2
Student	193	6.4
Unpaid domestic work	159	5.3
Sample size	3,005	100

Figure 1. Degree of Financial Digitalization: nº bank accounts, nº of financial entities by customer and number of online bank accounts (% of sample users)

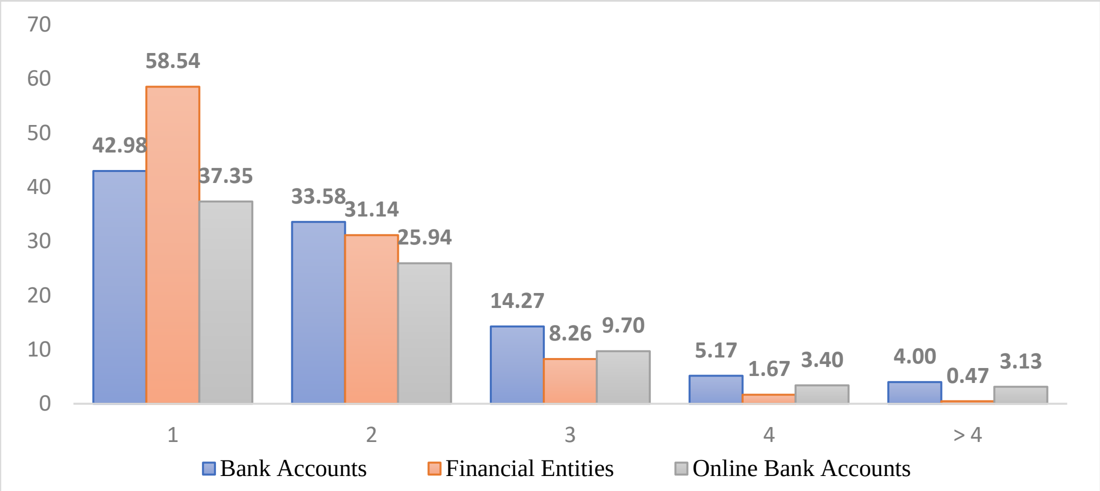


Figure 2. Degree of Financial Digitalization: Financial online activities (% of internet users)

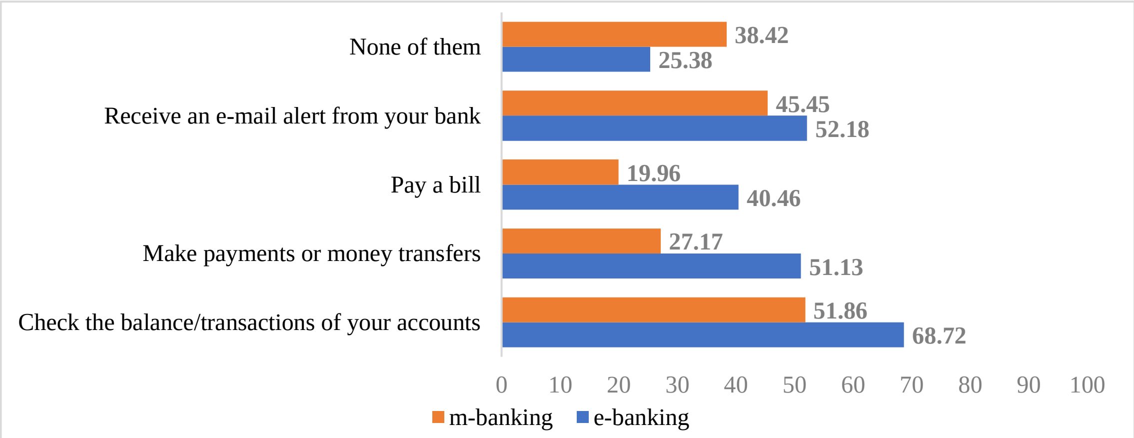


Figure 3. Dimensions of the financial digitalization

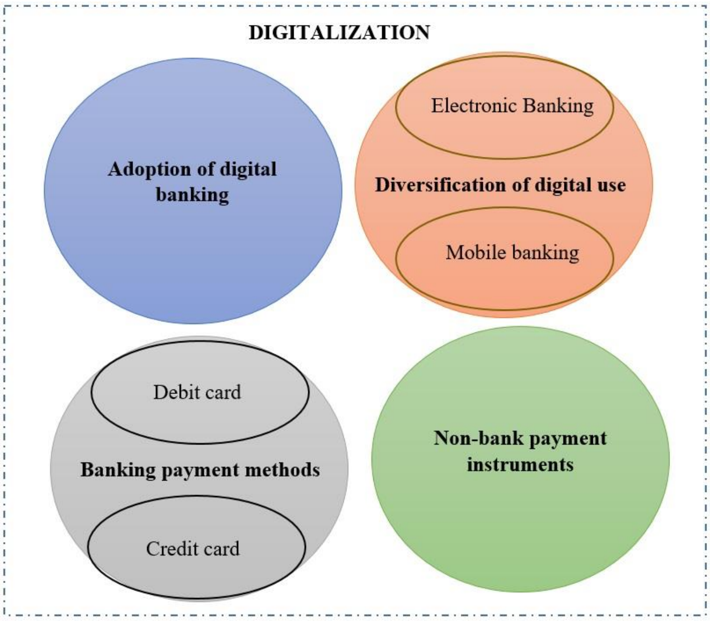


Table 2. Sample Matrix (Heatmap) by Dimensions and Socio-Demographics features

	Degree of Digitalization				Degree of Financial Digitalization				Consumers' perceptions						Non-banking services and social networks			
	% Frequent Internet Users	% Mobile Phone	% Laptop	% Tablet	% Online Bank Account	% Exclusive Online Users	% Debit Card	% Credit Card	Low or Very Low Cost: Online Banking	Low or Very Low Cost: Mobile Banking	Safe or Very Safe: Online Banking	Safe or Very Safe: Mobile Banking	Easy or Very Easy: Online Banking	Easy or Very Easy: Mobile Banking	Non-Bank account	Social Network User	Facebook or Twitter Communication	Facebook or Twitter Complain
Male	93.03	98.00	76.76	46.48	80.44	14.43	78.77	55.12	65.51	61.49	60.35	46.62	67.52	63.43	39.45	64.37	3.33	18.94
Female	89.74	97.00	72.95	47.16	78.31	11.18	78.37	47.09	60.05	56.15	57.61	42.00	66.20	63.76	30.29	66.93	3.75	16.60
18 - 24 years	100	100	91.84	43.62	85.11	5.67	75.18	30.14	76.95	75.53	58.51	45.39	80.85	83.33	50.71	91.13	2.33	21.40
25 - 34 years	100	99.59	83.73	51.41	88.35	21.89	84.14	38.35	79.72	77.91	68.67	57.43	79.12	78.11	50.60	89.36	5.17	19.10
35 - 44 years	98.39	99.71	76.82	56.85	87.32	18.22	82.94	56.71	72.45	70.26	67.64	53.79	74.34	73.62	41.98	76.38	2.86	18.51
45 - 54 years	96.35	97.29	78.76	50.08	82.41	13.15	79.87	49.60	60.22	54.83	61.49	45.01	69.10	63.87	32.96	63.55	2.49	14.71
55- 64 years	86.60	97.87	68.20	39.60	74.00	11.20	75.00	49.08	51.60	43.60	54.80	34.60	57.20	48.40	21.40	46.40	4.74	15.95
> 65 years	61.27	91.88	50.98	30.39	52.94	6.37	69.12	50.00	33.58	29.41	34.07	22.30	37.99	33.58	12.01	27.94	4.39	14.91
Working	97.41	98.90	80.66	52.56	87.05	16.75	83.58	56.25	71.13	67.05	67.82	51.90	74.77	71.40	40.55	72.67	3.94	18.20
Pensioner/retired	70.00	94.11	56.00	35.00	59.60	7.40	72.80	55.00	39.20	33.20	40.60	27.00	42.80	37.20	16.20	34.20	4.68	16.37
Unemployed	92.30	97.39	67.46	38.76	71.89	7.69	67.16	35.50	54.14	50.89	46.75	35.80	60.65	58.58	27.81	68.05	1.74	15.22
Student	100	100	93.78	39.38	87.05	6.74	76.17	25.91	80.83	78.76	60.62	49.74	83.42	85.49	53.37	88.60	2.34	18.71
Unpaid domestic work	77.36	92.98	60.38	44.65	60.38	3.14	66.67	43.40	37.74	37.74	39.62	23.27	45.28	41.51	20.75	51.57	2.44	18.29
Mean	91.38	98.50	74.84	46.82	79.57	12.81	78.57	51.08	62.76	58.80	58.97	44.29	66.86	63.59	34.84	65.66	3.55	17.74

Note: If the box is colored green means the value is above the sample mean

If the box is colored red means the value is below the sample mean

Figure 4. Consumers classification by dimensions (number of surveyed individuals)

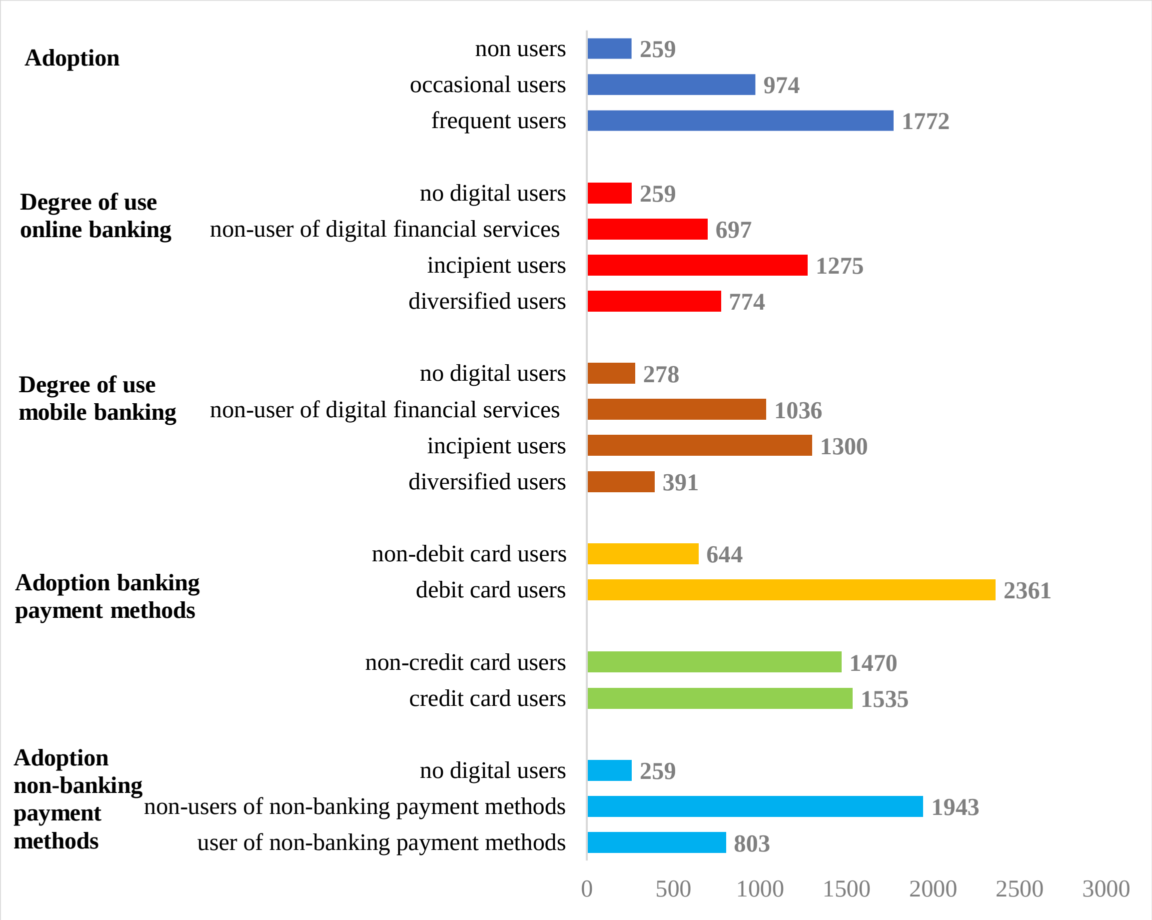


Table 3. Predictive validity of the random forest and logit models

	Out-of-sample fit's precision (70/30% split)		Whole sample fit's precision		Logit fit's precision	Random forest vs Logit
	Accuracy	**95% Conf. Interval**	**Accuracy**	**95% Conf. Interval**	**Accuracy**	**Most Accurated**
Adoption of online banking	88.41%	(86.11% - 90.45%)	87.35%	(85.62% - 89.08%)	79.27%	Random forest
Diversity of digital use: online banking	70.11%	(66.97% - 73.12%)	69.15%	(67.42% - 70.88%)	55.01%	Random forest
Diversity of digital use: mobile banking	70.01%	(65.92% - 72.13%)	69.35%	(67.62% - 71.08%)	59.57%	Random forest
Debit card	85.00%	(82.47% - 87.30%)	84.89%	(82.77% - 87.01%)	84.23%	Equally
Credit card	74.89%	(71.88% - 77.72%)	75.17%	(73.29% - 77.05%)	70.62%	Random forest
Adoption of Non-bank payment methods	76.14%	(73.18% - 78.92%)	76.27%	(74.36% - 78.18%)	73.46%	Random forest

Figure 5. Variable importance for the random forest model on online banking adoption

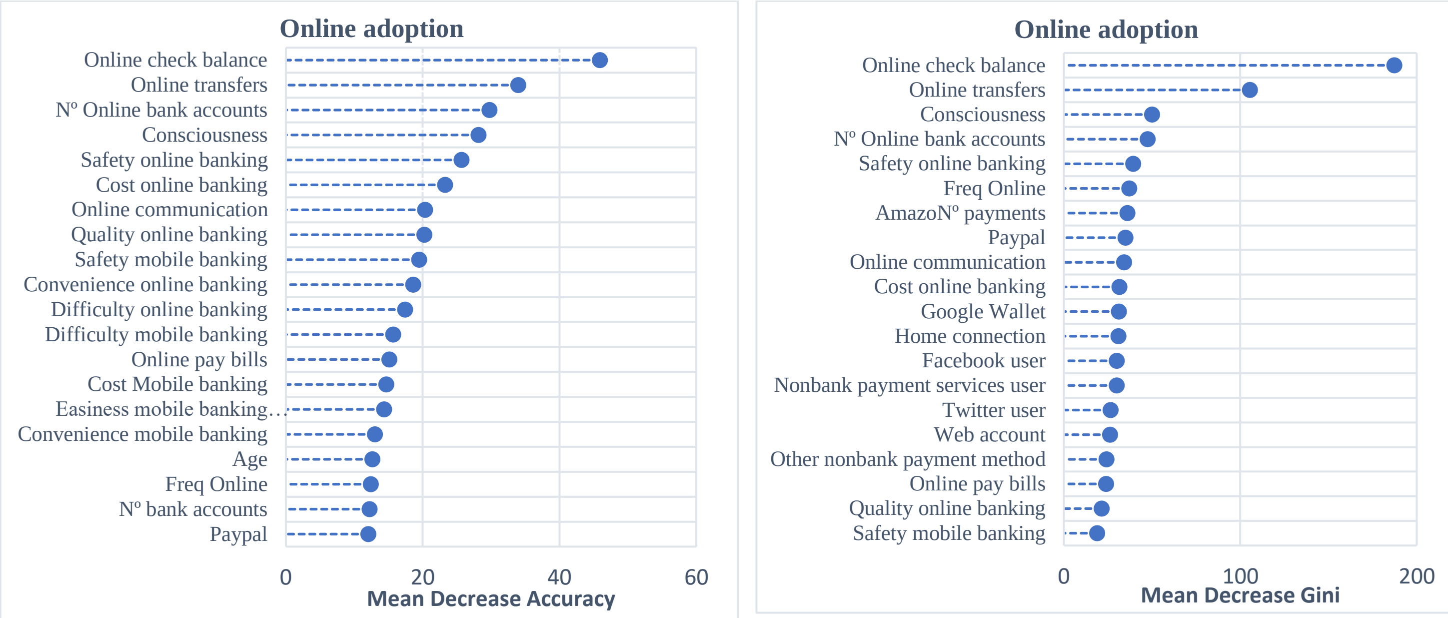


Figure 6. Variable importance for the random forest model on diversification of online banking uses

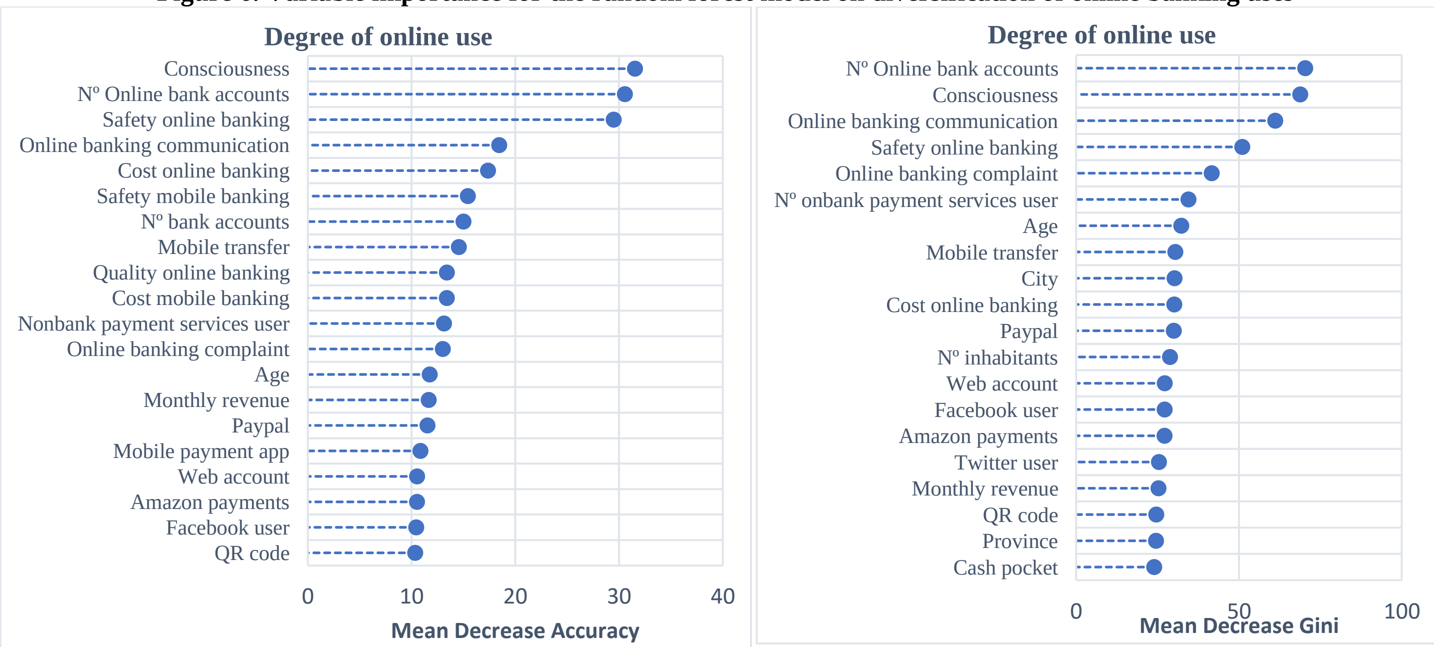


Figure 7. Variable importance for the random forest model on diversification of mobile banking uses

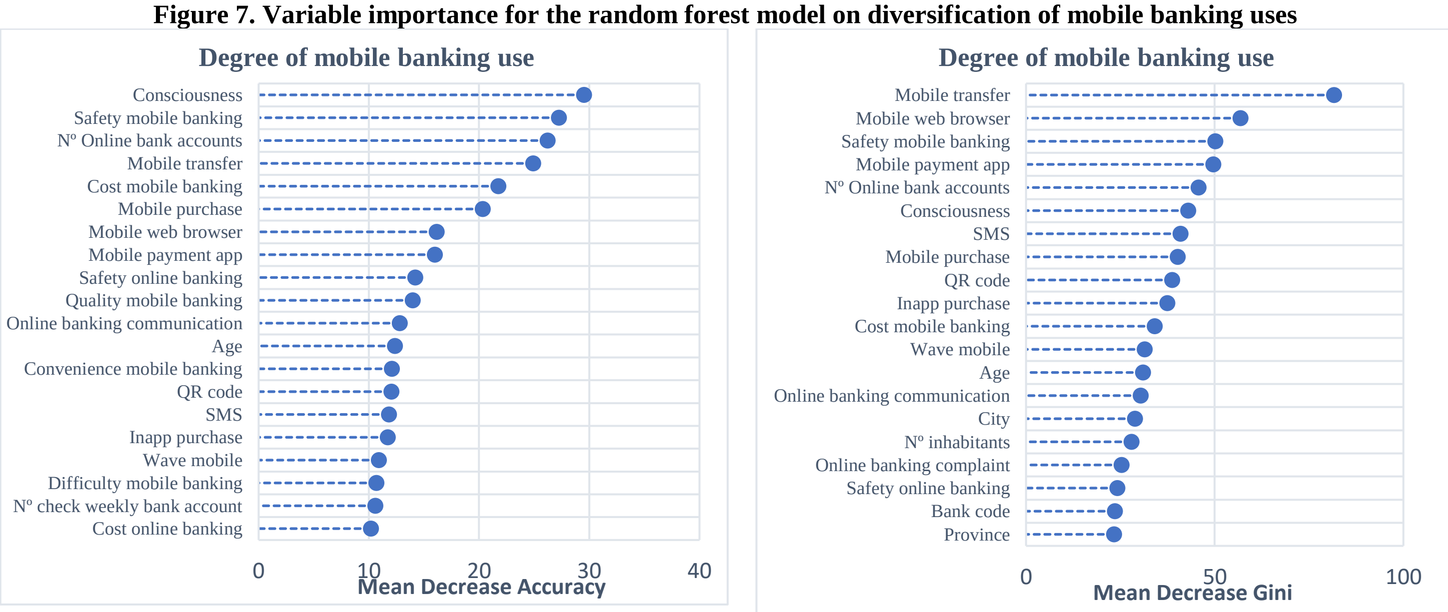


Figure 8. Variable importance for the random forest model on debit card adoption

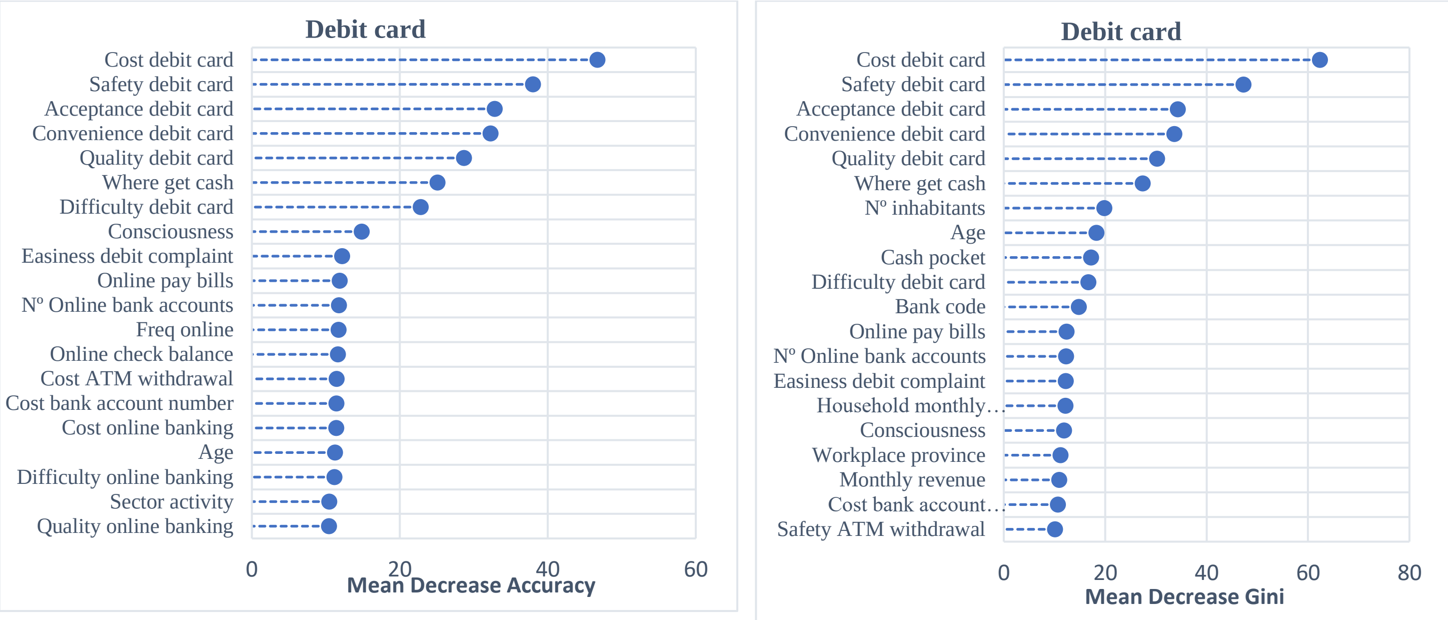


Figure 9. Variable importance for the random forest model on credit card adoption

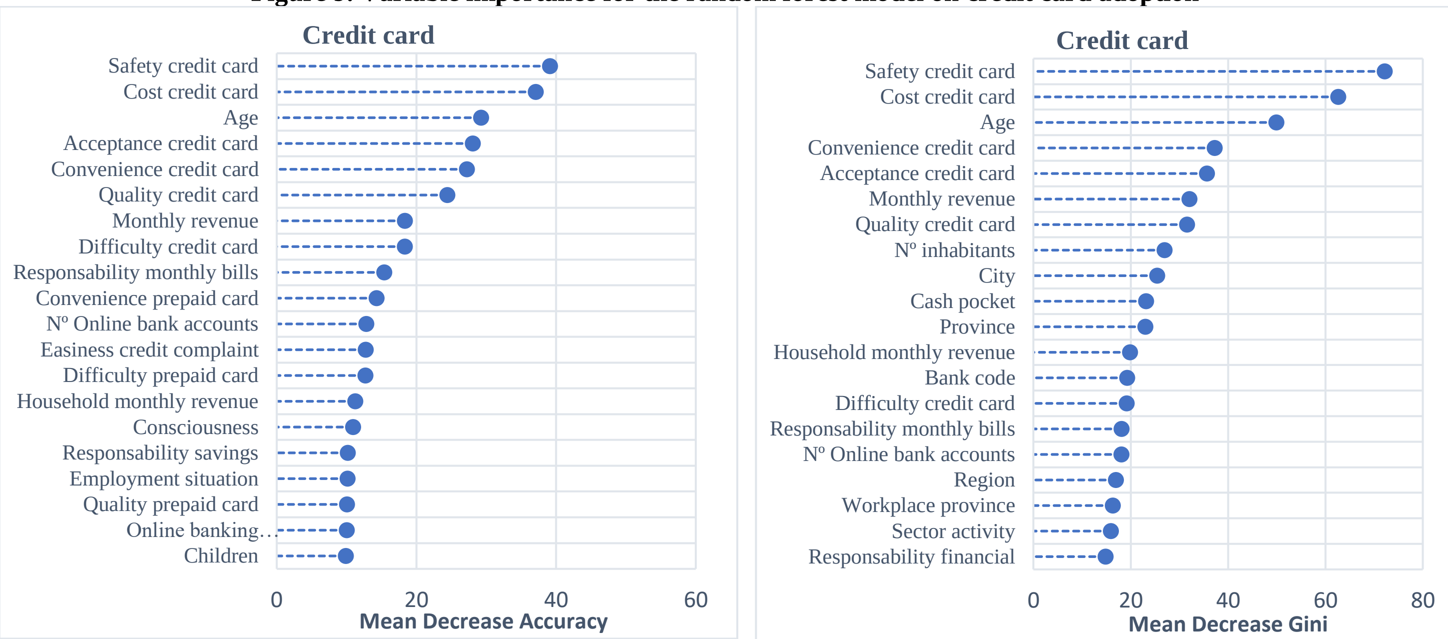


Figure 10. Variable importance for the random forest model on adoption of non-bank payment methods

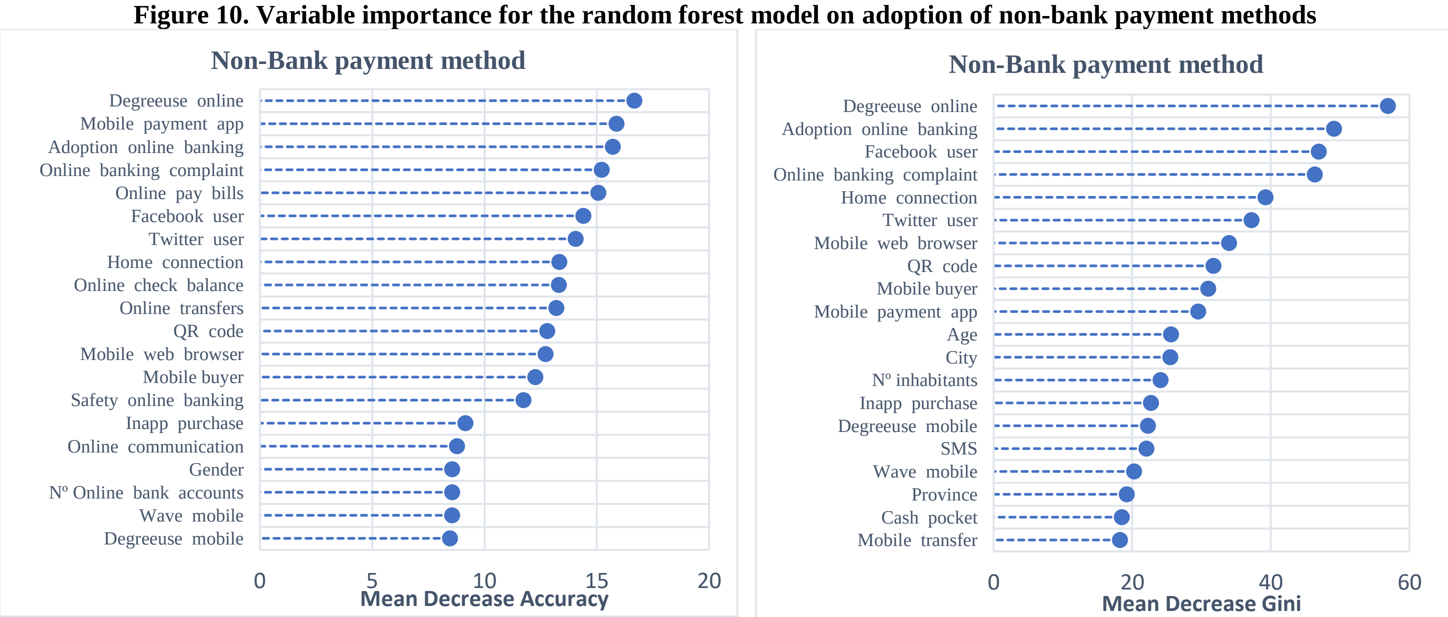


Figure 11. Tree: Adoption of digital banking

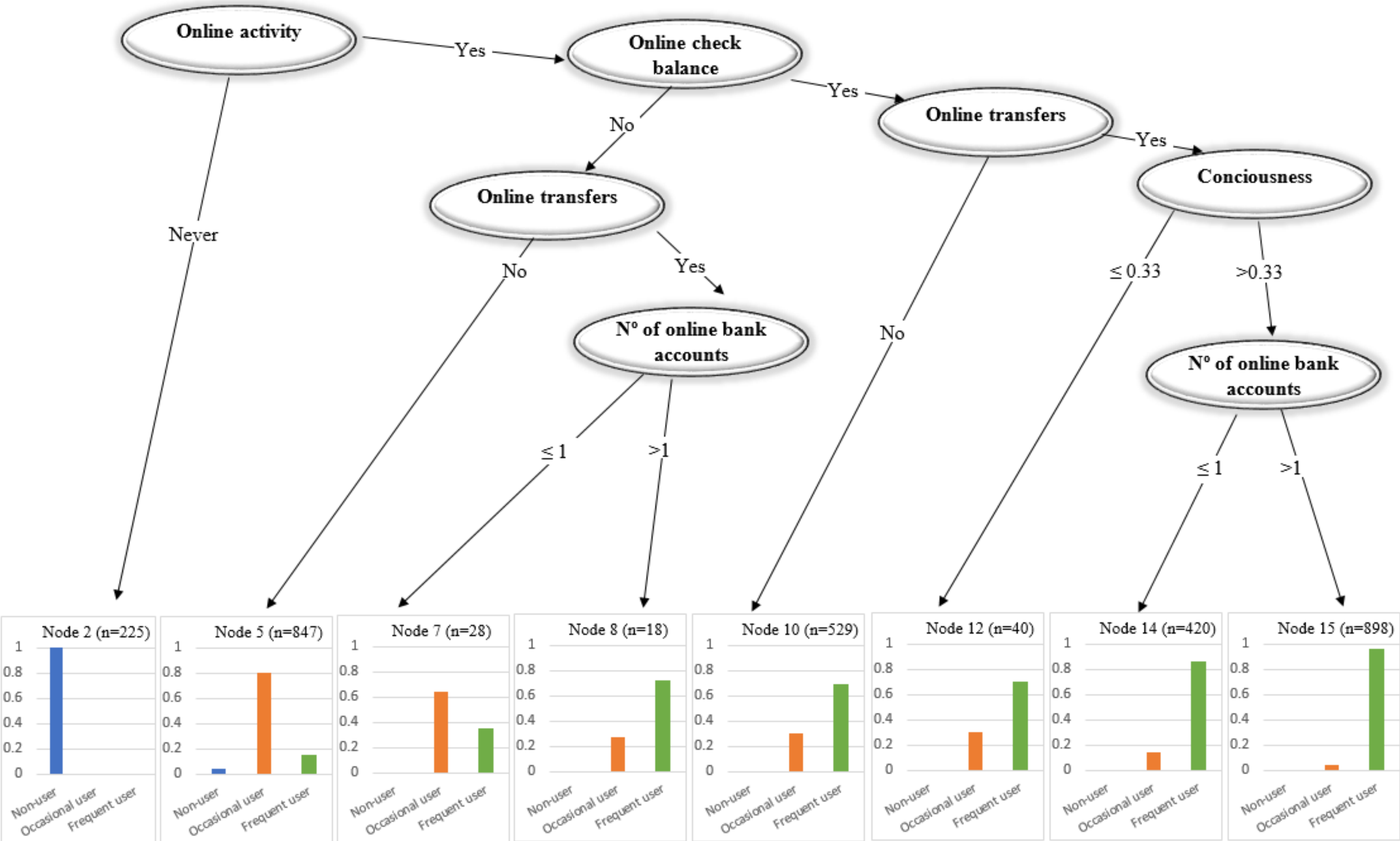


Figure 12. Tree: Diversity of digital use - online banking

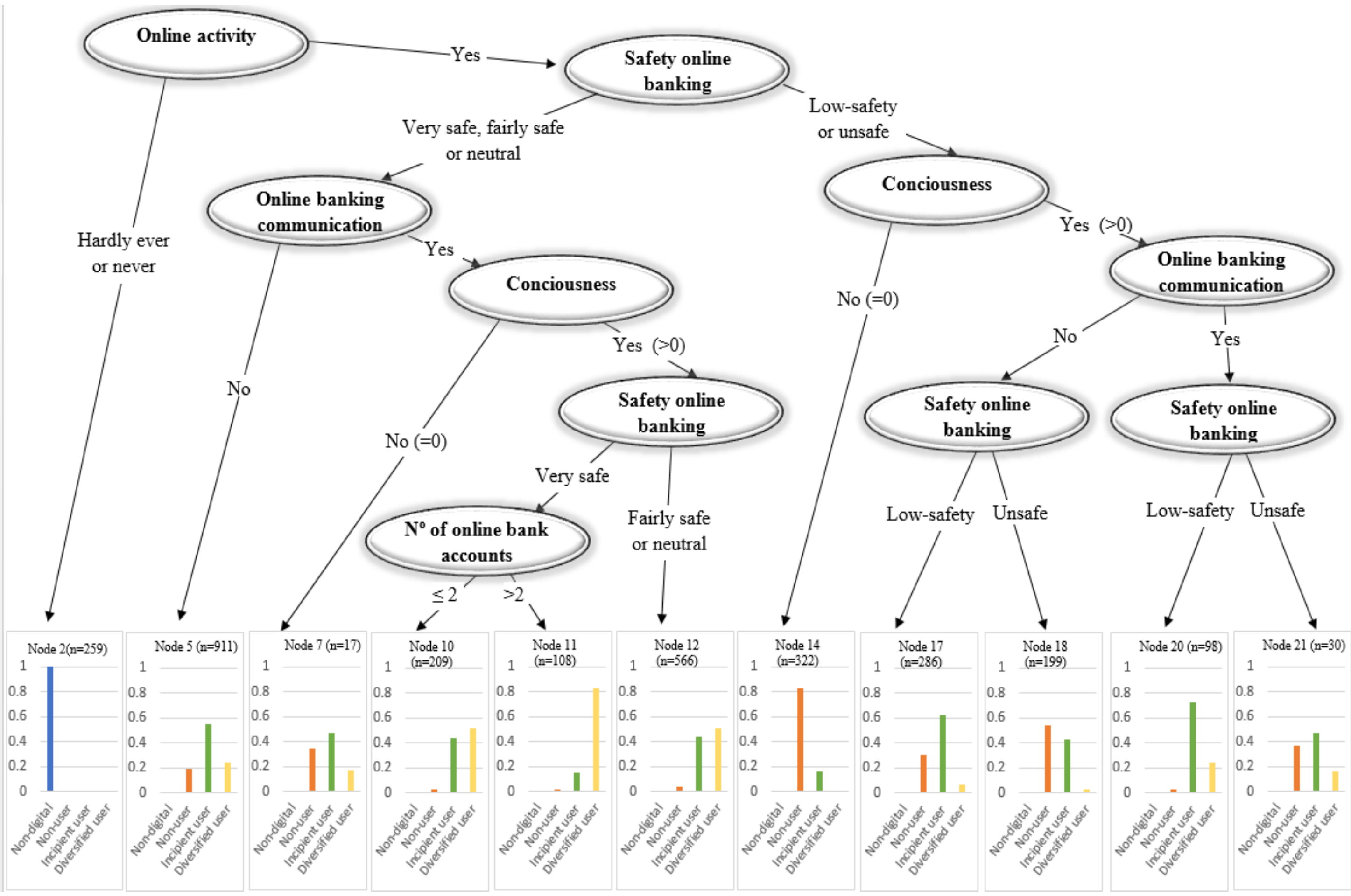


Figure 13. Tree: Diversity of digital use - mobile banking

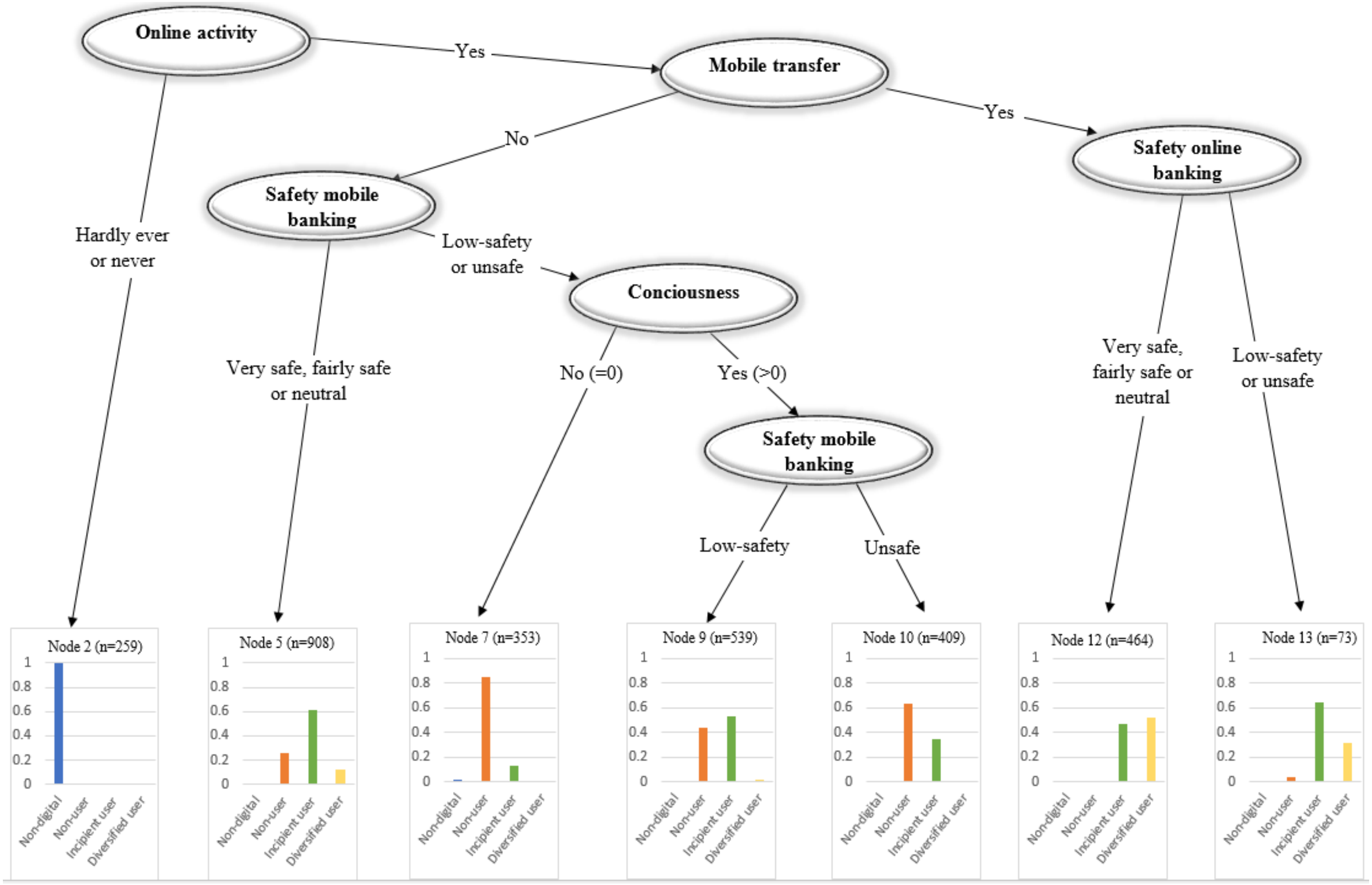


Figure 14. Tree: Debit card use

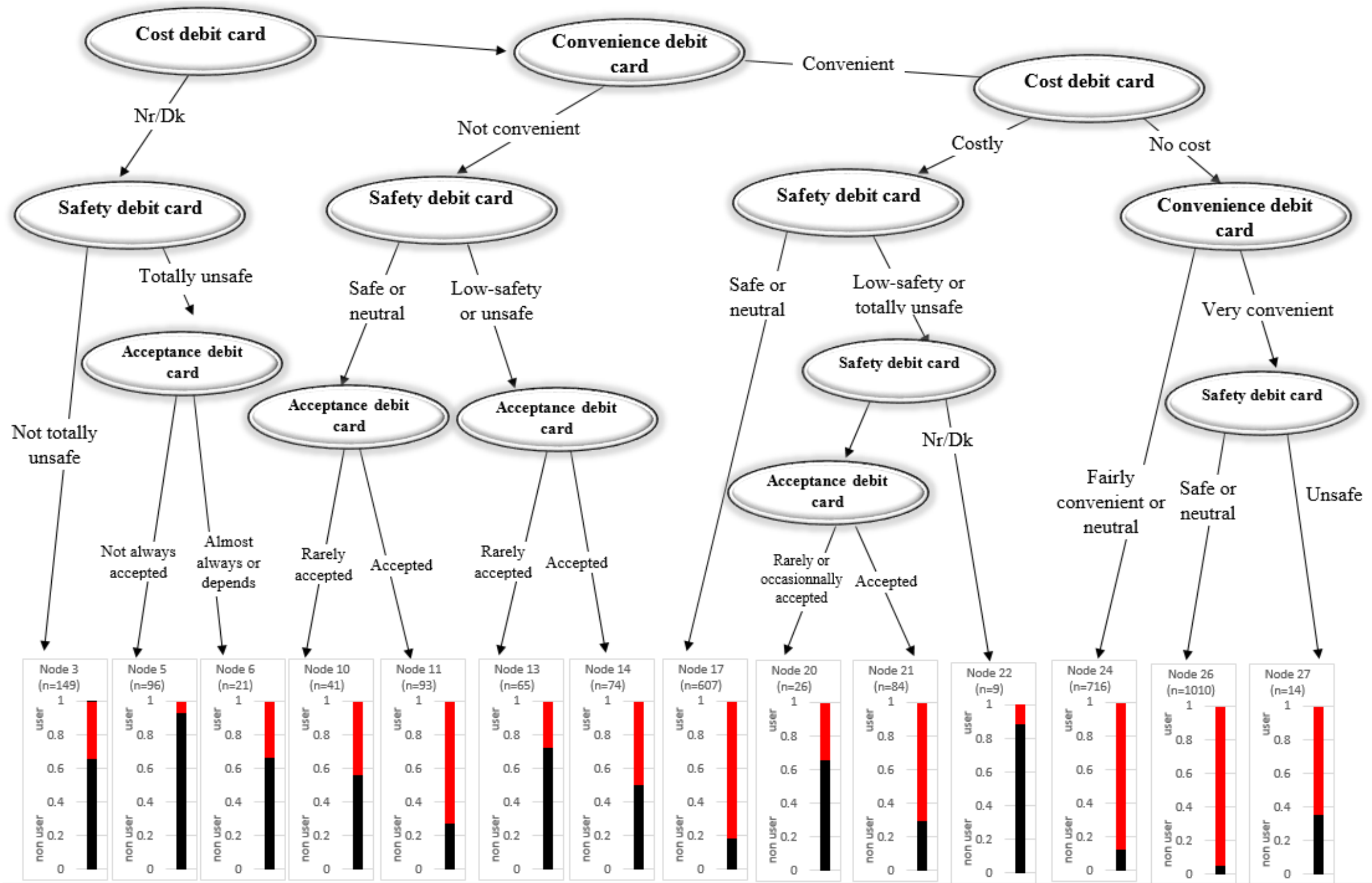


Figure 15. Tree: Credit card use

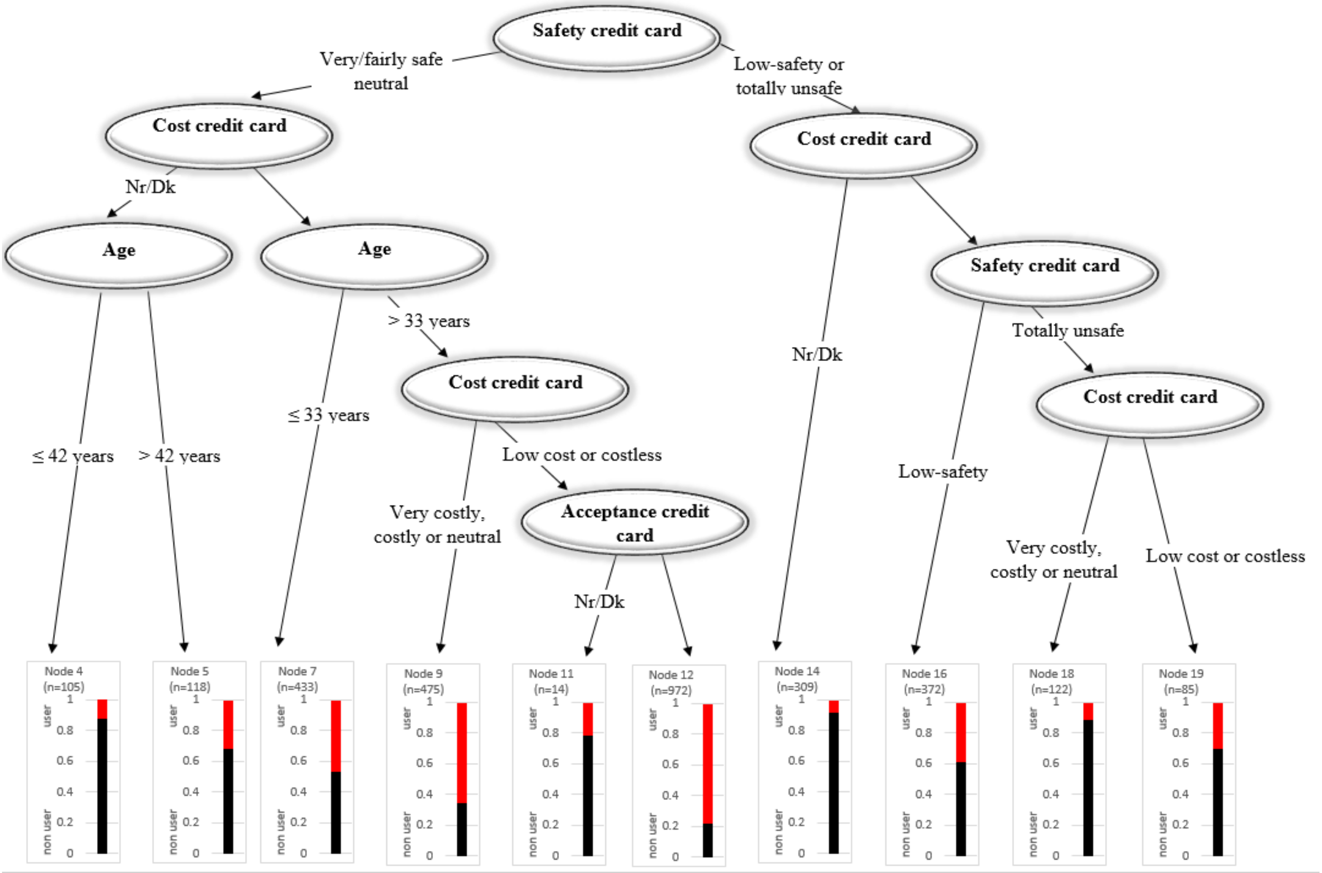
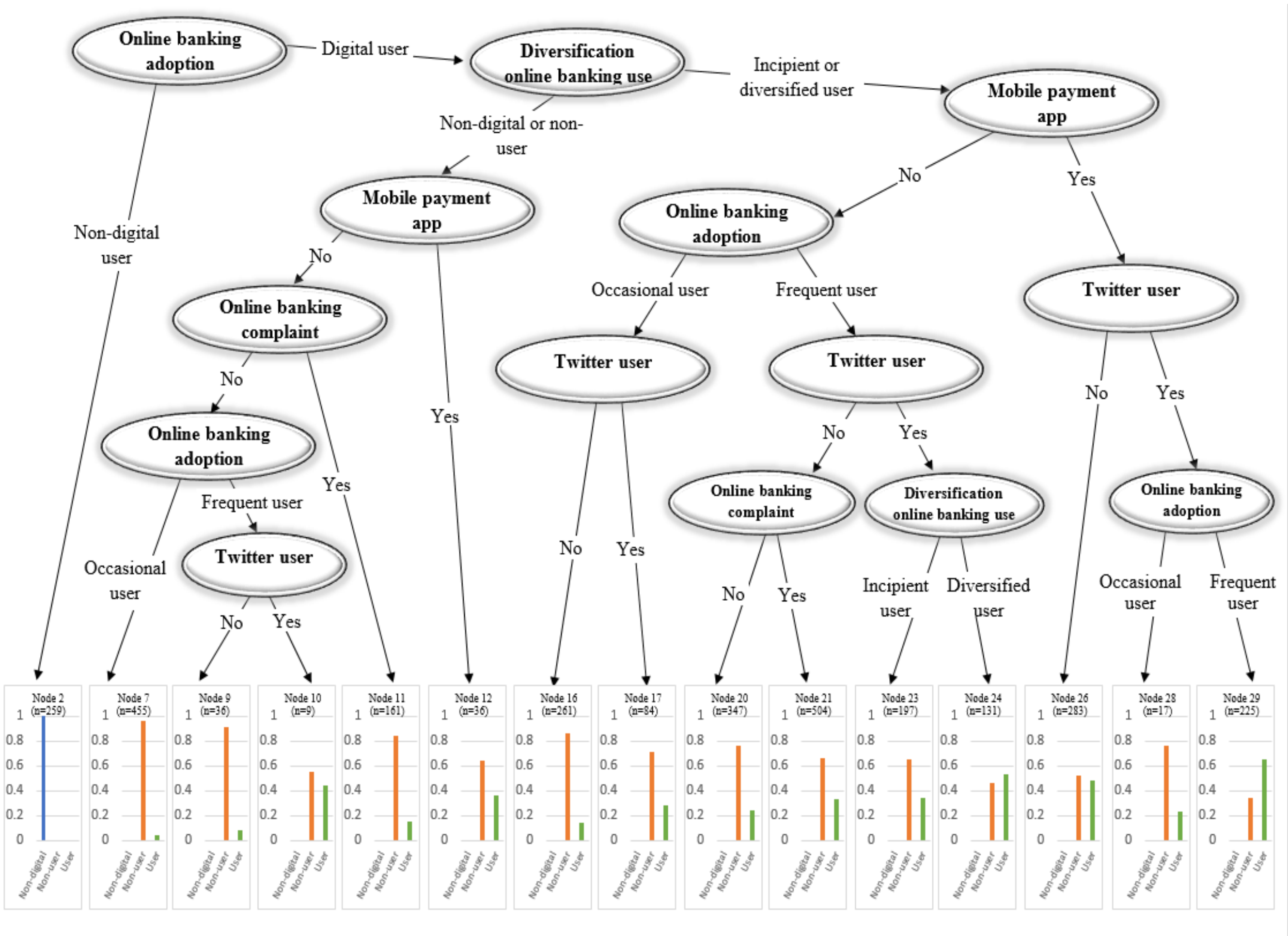


Figure 16. Tree: Use of non-bank payment instruments



Appendix A. List of Survey Questionnaire Variables

Socio-demographic characteristics	Age
	Gender
	Province
	City
	Nº inhabitants
	Household size
	Children
	Employed worker
	Employment situation
	Sector activity
	Unemployment period
	Full-time job
	Permanent job contract
	Monthly revenue
	Household employees
	Household monthly revenue
	Workplace province

Digital media	Landline or mobile phone
	Tablet
	Computer or laptop
	Nº computers
	Nº household laptops
	Exclusive computer
	Home connection
	Workplace connection
Financial status	Nº bank accounts
	Nº banks
	Savings bank account
	Current bank account
	Nº Online bank accounts
	Nº Online only bank accounts
	Consciousness
	Bank code
	Credit card holder
	Debit card holder

Financial activities	Online check balance
	Online communication
	Online pay bills
	Online transfers
	Nº online uses
	Mobile check balance
	Mobile communication
	Mobile pay bills
	Mobile transfers
	Nº mobile uses
	Mobile transfer
	Mobile payment app
	Mobile web browser
	Mobile purchase
	In-app purchase
	QR code
	SMS
	Wave mobile
	Online banking communication
	Online banking complaint
	Phone complaint
Social profile	Facebook user
	Twitter user
	Twitter FB bank comm
	Twitter FB bank complaint
Non-bank alternatives	Nonbank payment services user
	Google Wallet
	Amazon payments
	Paypal
	Web account
	Other non-bank payment method
Are you responsible for?	Responsibility financial
	Responsibility monthly bills
	Responsibility savings
	Responsibility shopping

Have you been?	Victim fraud
Do you know?	Know interest rate
	Know prepaid card
Cash	Cash home
	Cash pocket
	Annual cash payments
Where do you get cash?	ATM cash
	Bank branch cash
	Family cash
	Cashback cash
	Never cash
	Nº where cash
	Times withdrawal
How often do you . . . ?:	Freq bank branch check
	Freq check bank account
	Freq check credit
	Freq check prepaid
	Freq online
	Freq online check
	Freq phone complaint
	Freq use online
	Freq withdrawal
How many times do you . . . ?:	Nº bank branch check
	Nº check credit weekly
	Nº check monthly bank account
	Nº check monthly credit
	Nº check prepaid
	Nº check prepaid weekly
	Nº check weekly bank account
	Nº check weekly credit
	Bank branch check
	Check monthly prepaid
	Check weekly prepaid

Bank customers' perceptions	Acceptance	Bank account number, cash, credit card, debit card, prepaid card
	Convenience	Bank account number, cash, credit card, debit card, mobile banking, online banking, prepaid card
	Cost	ATM withdrawal, bank account number, credit card, debit card, mobile banking, online banking, prepaid card
	Difficulty	Bank account number, credit card, debit card, mobile banking, online banking, prepaid card
	Easiness	Bank account number, credit card, debit card, mobile banking, online banking, prepaid card
	Quality	ATM withdrawal, bank account number, cash, credit card, debit card, mobile banking, online banking, prepaid card
	Safety	ATM withdrawal, bank account number, cash, credit card, debit card, mobile banking, online banking, prepaid card
	Risk pay	App, email, online, personally, SMS
	Value	Confidentiality, easiness, protection losses, speed deduction, speed payment, speed registration

Appendix B.

Table B.I. Ordered logit estimation results for the determinants of online banking adoption

Dependent variable: Adoption of online banking		
Online check balance	2.059***	(0.183)
Online pay bills	0.188	(0.175)
Online communication	0.699***	(0.162)
Online transfers	1.113***	(0.154)
Consciousness	-0.0406	(0.0387)
Nº Online bank accounts	0.0152	(0.0153)
Safety online banking	0.212***	(0.0459)
Cost online banking	0.0351	(0.0787)
Quality online banking	0.176***	(0.0325)
Difficulty online banking	-0.104*	(0.0614)
Convenience online banking	0.0587	(0.0663)
Nº check weekly bank account	0.256***	(0.0406)
Social network user	1.120***	(0.0854)
Gender: Woman	-0.355***	(0.0924)
Nº inhabitants	1.08e-07	(7.45e-08)
Household monthly revenue	0.0621**	(0.0317)
Age	-0.243***	(0.0331)
Household size	0.0133	(0.623)
Constant cut1	-0.611**	(0.268)
Constant cut2	3.395***	(0.329)
Observations	3,005	
Pseudo R^2	0.4598	
Errors	Clustered at the bank-level	
Log Likelihood	-1537.1433	

Note: *** p<0.01, ** p<0.05, * p<0.1

Table B.II. Ordered logit estimation results for the determinants of the diversification of online banking uses

Dependent variable: Diversification of online banking uses		
Consciousness	-0.0109	(0.0442)
Nº Online bank accounts	0.00848	(0.0104)
Online communication	1.420***	(0.0858)
Nº check weekly bank account	0.177***	(0.0334)
Safety online banking	0.454***	(0.0471)
Cost online banking	0.0386	(0.0364)
Quality online banking	0.163***	(0.0252)
Difficulty online banking	-0.111***	(0.0397)
Convenience online banking	-0.0168	(0.0238)
Non-bank payment user	0.631***	(0.106)
Social network user	0.982***	(0.143)
Gender: Woman	-0.408***	(0.0656)
Nº inhabitants	3.06e-08	(8.02e-08)
Household monthly revenue	0.130***	(0.0195)
Age	-0.0795***	(0.0220)
Household size	0.0699	(0.0471)
Constant cut1	0.262	(0.263)
Constant cut2	2.667***	(0.252)
Constant cut3	5.605***	(0.306)
Observations	3,005	
Pseudo R^2	0.2738	
Errors	Clustered at the bank-level	
Log Likelihood	-2838.2248	

Note: *** p<0.01, ** p<0.05, * p<0.1

Table B.III. Ordered logit estimation results for the determinants of the diversification of mobile banking uses

Dependent variable: Diversification of mobile banking usages		
Consciousness	-0.0241	(0.0470)
Nº Online bank accounts	0.0204*	(0.0104)
Online transfers	1.852***	(0.109)
Mobile purchase	1.093***	(0.108)
Safety mobile banking	0.343***	(0.0401)
Cost mobile banking	0.0189	(0.0452)
Quality mobile banking	0.131***	(0.0252)
Difficulty mobile banking	-0.0709*	(0.0412)
Convenience mobile banking	-0.0280	(0.0458)
Nº check weekly bank account	0.176***	(0.0360)
Non-bank payment user	0.409***	(0.123)
Social network user	1.038***	(0.0945)
Gender: Woman	-0.178*	(0.108)
Nº inhabitants	1.85e-07***	(5.75e-08)
Household monthly revenue	0.0383	(0.0248)
Age	-0.161***	(0.0354)
Household size	0.0595	(0.0469)
Constant cut1	-0.248	(0.302)
Constant cut2	2.297***	(0.318)
Constant cut3	6.033***	(0.383)
Observations	3,005	
Pseudo R	0.294	
Errors	Clustered at the bank-level	
Log Likelihood	-2704.581	

Note: *** p<0.01, ** p<0.05, * p<0.1

Table B.III. Logit estimation results for the determinants of the debit and credit card adoption

Dependent variable:	Debit card adoption		Credit card adoption	
Safety debit card	0.389***	(0.0466)	0.453***	(0.0335)
Cost debit card	-0.0431	(0.0375)	0.0718	(0.0480)
Quality debit card	0.265***	(0.0592)	0.144***	(0.0489)
Difficulty debit card	-0.115	(0.0784)	-0.117**	(0.0512)
Convenience debit card	-0.163**	(0.0667)	-0.106***	(0.0378)
Acceptance debit card	0.310***	(0.0484)	0.187***	(0.0350)
Easiness debit card complaint	0.0189	(0.0594)	0.0739*	(0.0410)
Bank branch cash	-1.202***	(0.195)	-0.855***	(0.118)
Cashback cash	-2.791**	(1.185)	-0.571	(1.059)
Salary cash	-0.841	(0.891)	0.336	(0.623)
Family cash	-1.163***	(0.422)	0.309	(0.465)
Never cash	1.525	(1.843)	-	-
Credit card	-0.0818	(0.159)	-0.0391	(0.172)
Mobile phone	-0.112	(0.386)	0.761*	(0.404)
Non-bank payment user	0.440**	(0.176)	0.214	(0.137)
Social network user	0.281**	(0.139)	0.284***	(0.0920)
Gender: Woman	0.0542	(0.0912)	-0.272***	(0.0816)
Nº inhabitants	2.30e-07***	(7.17e-08)	-2.62e-08	(8.18e-08)
Household monthly revenue	0.0445*	(0.0245)	0.0895***	(0.0144)
Age	0.0684	(0.0485)	0.356***	(0.0444)
Household size	0.00971	(0.0699)	-0.0368	(0.0406)
Constant	-2.160***	(0.662)	-4.987***	(0.467)
Observations	3,005		3,005	
Pseudo R	0.2834		0.203	
Errors	Clustered at the bank-level		Clustered at the bank-level	
Log Likelihood	-1169.3937		-1650.7946	

Note: *** p<0.01, ** p<0.05, * p<0.1

Table B.IV. Logit estimation results for the determinants of non-bank payment adoption

Dependent variable: Adoption of Non-bank payment instruments		
Adoption online banking	2.333***	(0.197)
Diversification online banking uses	0.948***	(0.128)
Mobile payment app	0.391***	(0.0864)
Online check balance	-0.786***	(0.137)
Twitter user	0.271**	(0.125)
Facebook user	0.782***	(0.108)
Credit Card	0.0786	(0.159)
Debit Card	0.167	(0.120)
QR code	0.346**	(0.142)
SMS	0.0344	(0.0828)
Wave mobile	-0.204	(0.152)
Online communication	-0.0329	(0.0927)
Mobile Phone	0.804**	(0.333)
Gender: Woman	-0.261***	(0.0785)
Nº inhabitants	-7.92e-08*	(4.20e-08)
Household monthly revenue	0.0573**	(0.0244)
Age	-0.213***	(0.0338)
Household size	0.0886**	(0.0432)
Constant cut1	4.659***	(0.456)
Constant cut2	11.39***	(0.517)
Observations	3,005	
Pseudo R	0.4291	
Errors	Clustered at the bank-level	
Log Likelihood	-1513.3098	

Note: *** p<0.01, ** p<0.05, * p<0.1