FINANCING ENTREPRENEURIAL ACTIVITY IN URUGUAY: TIME TO DEFAULT IN A PUBLIC MICROCREDIT INSTITUTION

María Nela Seijas-Giménez, mnseijas@ccee.edu.uy, University of the Republic, Uruguay

Milagros Vivel-Búa[a], mila.vivel@usc.es, University of Santiago de Compostela, Spain

Rubén Lado-Sestayo, ruben.lado.sestayo@usc.es, University of Santiago de Compostela, Spain

Sara Fernández-López, sara.fernandez.lopez@usc.es, University of Santiago de Compostela, Spain

[a] Corresponding author. Faculty of Economics and Business Studies. University of Santiago de Compostela, Av/ Burgo das Nacións s/n, 15782 Santiago de Compostela (Spain). Email: mila.vivel@usc.es

ABSTRACT

This paper develops a tool to predict the percentage of compliance in the repayment of microloans granted by non-profit microfinance institutions (MFI) of the Uruguayan government. The database consists of 1,357 microloans granted by the Program for the Strengthening of Productive Entrepreneurs (PFEP) of the Uruguayan Ministry of Social Development (MIDES) during the period 2012-2016. The paper uses Cox's (1972) proportional risk model, employing four penalty modes: ENET, LASSO, AENET and ALASSO. The analysis shows that with a reduced set of variables that are easy for the MFI to obtain, it is possible to obtain high predictive power.

KEYWORDS

Microloans, microfinance, default, repayment, Uruguay.

JEL CODES

C14, C41, G21

1. INTRODUCTION

According to the World Bank (2014), financial inclusion is defined as the proportion of individuals and companies using financial services. One of the benefits of financial inclusion is its contribution to reducing the vulnerability of lower-income households and entrepreneurs, since it provides instruments that help limit the variability of consumption and manage adverse shocks more efficiently (De Olloqui et al., 2015; Demirgüç-Kunt et al., 2018). The Global Findex survey on financial inclusion published by the World Bank (2017) shows that 1.7 billion people of working age do not have access to an account in a banking institution, and thus they are considered to be financially excluded (Demirgüç-Kunt et al., 2018). For this reason, since 2015, the World Bank has been striving toward global financial inclusion. The World Bank has acknowledged that the number and quality of initiatives developed in recent years to achieve global financial inclusion has been satisfactory, but more work is still needed, especially after the COVID-19 crisis (Andrews et al., 2021).

One of the instruments of financial inclusion is the microloan, i.e., a loan in a small amount to finance low-income entrepreneurs, who are often excluded from the traditional financial system. On a global level, microloans are considered to be instruments that favor financial inclusion (De Olloqui e al., 2015). Their importance was recognized by the United Nations General Assembly, which declared 2005 to be the International Year of Microcredit, spreading the role of this instrument in the eradication of poverty. Microloans form part of a broader concept, microfinance, which refers to the provision of financial services to the most economically vulnerable ventures. These services are provided by microfinance institutions (MFIs) that also offer training and advisement and, on occasion, social services such as literacy and health care (Robinson, 2001; Ledgerwood, 1999). Thus, the concepts of microfinance and financial inclusion are closely related, since the financial products and services provided by MFIs play a key role in promoting access to full financial inclusion (World Bank, 2017).

Traditionally, MFIs provide microloans using a credit risk assessment methodology, in which the analyst plays a key role. This individual uses information derived from a more direct connection with the individual or company to be financed than that used by banking institutions. The gathering and evaluation of information about the potential borrower is quite informal and subjective, as it depends largely on the discretion of the analyst. Therefore, this credit risk assessment is costly in terms of both time and money, conditioning the efficiency of MFIs, as well as their financial sustainability. The commitment between the social purpose of microcredit and the quest for profitability highlights the importance for MFIs to seek greater efficiency in the process of granting microloans, moving towards credit risk mitigation practices that enable their survival in the market. According to Perossa and Gigler (2015), in recent decades, MFIs have had to pursue their social and financial objectives in an increasingly constrained environment, due to increasing competition, over-indebtedness and economic crises. Therefore, developing efficient risk management tools focused on MFIs is becoming increasingly more relevant in terms of contributing to their sustainability, and this is the focus of this research.

In the specific case of Uruguay, in 2006, the Program for the Strengthening of Productive Entrepreneurs (PFEP) of the Ministry of Social Development (MIDES) was implemented with the aim of providing technical and financial support to national entrepreneurship and contributing to its sustainability. The credit evaluation is carried out by an MIDES officer, who must evaluate personal and financial factors related to the entrepreneur and the venture to be financed. The repayment of these microloans does not involve the payment of interest and is made in cash, with the possibility of MIDES granting subsidies to the borrower. During the 2012-2016 period, the only years for which statistics on the PFEP can be compiled from data published by MIDES, 47.6% of the microloans granted went into default, compromising the viability of the Uruguayan microcredit program. In this context, the objective of this research is to analyze and predict the credit risk of these microloans granted by MIDES, using survival models and information gathered from 1,357 microloans granted during the 2012-2016 period. This research thus contributes to identifying a tool to assess credit risk, reducing default and, consequently, ensuring the viability of a microcredit program over the medium and long term.

In the field of microfinance, the previous literature offers several examples of the application of credit risk analysis through standardized models. In general, logistic regression models have been the most popular. Survival analysis has been used to evaluate, for example, credit risk in consumer loans (Hassan et al., 2018; Cao et al.,2009), although it has not been greatly explored in the field of microfinance. Previous studies have shown its strength as a tool for predicting default in credit relations. For example, Bellotti and Crook (2009), Madorno et al. (2013) and Tong et al.

(2012) find that the probability of default identified through survival analysis is more robust than that found with logistic regression analysis. In fact, Banasik et al. (1999) had previously found survival analysis also to be superior to logistic regression in identifying the probability of default in the first year of the life of a microloan. Survival analysis has two main advantages over other methodologies. First, it models the time to default. Hence, this methodology not only predicts the determining factors of default, it also contributes to predicting the moment in time when it is most likely to occur. Second, it can assess data with incomplete information (censored observations). This is very important, because it is common for MFIs in undeveloped or developing countries not to have complete information about the evolution of their microcredit portfolio.

The main contributions of this research are twofold. First, it is a methodological approach based on the use of survival models with different penalties (LASSO, Adaptive LASSO, Elastic Network and Adaptive Elastic Network) through a predictive approach and geared to need only a small number of variables, which are easily identifiable in the evaluation of the borrower and its environment. To our knowledge, survival models have not been applied in the context of microfinance in Latin America, even though this region has a very relevant position. Janda and Zetek (2013) stress that Latin America is the most experienced region in terms of microfinance, referring to historical government microcredit programs for the agricultural sector. Furthermore, according to the Microfinance Barometer (2012, 2018) the Latin America and Caribbean region ranks first in terms of managed microcredit portfolios.

The second contribution of this research refers to the area of analysis that is focused on Uruguay, which has undergone a major regulatory change in recent years to favor financial inclusion. According to the Uruguayan Ministry of Economy and Finance (MEF), Uruguay was characterized by a low level of financial inclusion, so in 2005 its government began to design structural reforms to enhance it, especially among households with lower income levels and small business. These reforms concluded with a new legislative framework that came into effect in 2014, under the name of the Law on Financial Inclusion and Promotion of Use of Electronic Means of Payment. This new regulation promotes the access of small businesses to bank financing, which redefines the available credit market in which banks and MFIs, as well as other credit institutions, compete. Thus, the Uruguayan market is a relevant focus of research in order to assess an MFI’s financial activity in a new regulatory context aimed at promoting financial inclusion. In particular, as indicated above, this research analyzes a subsidized microcredit program granted by a government MFI for a strictly social purpose (PFEP), which corresponds to the original microfinance paradigm (Robinson, 2001). Finally, it should be noted that the Uruguayan market is also a relevant target for study because statistics show that the growth of microfinance in Latin America was more pronounced in smaller geographical countries, such as Uruguay, than in larger countries, due to more direct government policies to finance entrepreneurs through microloans (Janda & Zetek, 2013). In short, this research aims to generate empirical results that can be extrapolated to other environments or emerging microcredit programs that need standardized tools to assess the credit risk, so that default does not undermine their resources and survival, using information that is easily accessible and cheap to obtain.

The article is organized into four sections. Following this introduction, the theoretical framework of this research is presented by reviewing the studies related to credit risk assessment in microfinance institutions. Next, the third section identifies the empirical study. Finally, the paper presents the main conclusions obtained, identifying their limitations and future lines of work.

2. THEORETICAL FRAMEWORK

Based on the pioneering work of Viganò (1993) and Schreiner (1999a, 1999b), several studies can be identified aimed at evaluating credit risk in MFIs. This research can be characterized based on three parameters: the risk to be predicted, the methodology applied and the variables analyzed as determinants of this risk.

Focusing on the risk to be predicted, the previous literature on microfinance has defined this variable as the delay in the payment of an installment by the borrower, quantified in the number of days (Chakravarty & Jha, 2012; Schreiner, 2000). This temporary delay in microcredit repayment requires extraordinary and costly management at the MFI, with the ultimate goal of getting the borrower to pay the installments, i.e., to no longer be a defaulter. In this sense, there is research in which the risk to be predicted consists of the borrower being 15 or more days late in the payment of an installment (Schreiner, 2004), accumulating a delay of at least 30 days (Dellien and Schreiner, 2005), being in a situation of financial insolvency (Reinke, 1998) or, simply, registering some type of delay without specifying a time interval (Viganò, 1993). For his part, Vogelgesang (2003) also proposes considering the risk to be predicted as a delay that, on average, reaches one day, or defining it as a default situation that, in his opinion, should be measured as the accumulation of an average of ten days of delay in the payment of an installment. Likewise, there are studies that define risk by considering those borrowers who may go into debt a second or more times in an MFI, that is to say, once they have already repaid the microloan. For example, Schreiner (2003) analyzes the probability that a borrower will not apply for a second microloan after paying off the first one (“drop-out scoring”).

From a methodological point of view, it should be noted that credit risk assessment in microfinance has usually used the expert judgment approach, relying on information obtained through the direct contact of the loan officer with the entrepreneur and the business to be financed. This methodology is costly, considering the resources it consumes, in particular, in terms of time, and it also conditions the efficiency of MFIs, which is very relevant in a competitive credit environment like the current one. Thus, the literature has proposed tools and/or methodologies to assess credit risk that contribute to the professionalization of the MFI activity. In particular, the literature has shown that the application of credit scoring techniques in the evaluation of the granting of credit allows the mitigation of credit risk in MFIs (Schreiner, 2000). Credit scoring techniques involve the use of statistical models that estimate a borrower’s risk of default, using qualitative and/or quantitative information. These methodologies may be conceived of as a system that, by means of predictions, qualifies a credit and measures the risk of default. Their use in MFIs contributes to understanding the financial behavior of borrowers and their default, the relation between risk and profitability and the cost of operation. Bumacov et al. (2014) found that the total number of microcredits, the number of microcredits per loan officer are positively and statistically significantly affected by applying credit scoring methodologies. According to Bumacov et al. (2014), credit scoring methodology contributes to faster growth in MFI outreach, defining "outreach" as the number and socio-economic level of microentrepreneurs served by the MFI's microcredit offer. However, Schreiner (2000) and Van Gool et al. (2012), among other authors, indicate that credit scoring is not a substitute for the judgment of credit officers. Hence, they explain that credit scoring has the power to predict credit risk and reduce costs, although this should be complemented with the judgment of the credit officer.

The statistical techniques applied in the early academic work on credit risk in MFIs were limited to discriminant analysis (Viganò, 1993), tobit models (Zeller, 1998; Sharma & Zeller, 1997), probit analysis (Vogelgesang, 2003; Reinke, 1998) and logistic regressions (Schreiner, 2004, 2003; Dellien & Schreiner, 2005; Dellien, 2003). Starting with Bensic et al. (2005), the literature also considered non-parametric techniques, such as neural networks and CART decision trees (Kammoun & Triki, 2016; Ayouche et al., 2017). However, this previous literature includes few and relatively recent examples of the application of survival models in microfinance, although they have been used successfully in the field of banking (Roszbach 2004). In our opinion, the study of the factors that predict the time to default on a microloan is as important as the analysis of those factors that predict the probability of default. As Bourlès and Cozarenco (2018) indicate, an early default has more costs for the bank than a default occurring towards the end of the loan period. In short, survival models applied to microcredit provide an additional and complementary dimension to the empirical study of the phenomenon of default in an MFI, but they have not been widely explored to date. This research helps to fill this gap.

Table 1 identifies the previous literature that applies survival analysis to the study of credit risk in microloans granted by MFIs. As indicated above, there are few empirical studies and they tend to be quite recent, especially as compared to the extensive previous literature in this field that applies other methodologies, in particular to probit and logit models (Seijas-Giménez et al., 2017). In addition, most refer to the African continent, followed by Europe and, finally America, through the study of the USA. Therefore, despite the prominence of Latin America in microfinance, which was highlighted in the Introduction section, there is no empirical evidence in this credit context. This research analyzes a Uruguayan MFI, and thus it also contributes to bridging this gap in the literature.

Another aspect to highlight in this previous literature applying survival models is that the focus of analysis does not exclusively subscribe to the time-to-default analysis, a very important aspect for the medium- and long-term financial sustainability of the MFI, but rather some studies evaluate other aspects of the credit relation. For example, the pioneering work of Pagura (2004) uses a survival model to identify the factors that affect the length of time of the borrowing relationship between the entrepreneur and an MFI in Mali (Africa) during the period 1999-2001. Their results indicate that the probability of the entrepreneur leaving the microcredit program that links them to that MFI generally depends on business profits, the amount of the microloan, their level of education, the number of dependents in their household, their history of default and the use of other sources of credit. In general, the evidence obtained by Pagura (2004) confirms that the entrepreneur has a negative dependency on the duration, that is, the longer they maintain a credit relationship with the MFI, the more likely they are to abandon it.

The research of Bekele and Worku (2008), also focused on Africa, analyzes the long-term survival and viability of small enterprises financed through banks, microfinance, iqqub schemes and personal donations in Ethiopia during the period 1996-2001, paying special attention to women's role as entrepreneurs. Their evidence confirms that being female increases the probability of the company’s bankruptcy. This negative effect is related to their educational level, their technical skills, their business management capacity, their difficulties in accessing credit and the possibilities they have to invest the surpluses generated with their business. Subsequently, Babajide (2011) studied the impact of microfinance on micro and small business survival in Nigeria during the period 2004-2008. This author confirms a positive impact of microfinance on the survival of Nigerian micro and small companies. In fact, the probability of survival of these enterprises increases with their capacity to generate profits on a regular basis, their ability to reinvest the profits in new investments, and having easy access to microcredit. In addition, evidence confirms that financial and non-financial services offered by MFIs contribute positively to the survival of these companies, favoring the exchange of business skills, the generation of innovative ideas, and diminishing problems related to funding shortages.

Focusing on the European continent, Bos et al. (2015) apply a survival model, together with other methodologies, such as linear regression and logistic regression to assess the probability of rejecting applications for microloans presented by micro and small companies from Bosnia-Herzegovina, depending on the existence of a credit registry. These authors also study the temporal evolution of the default probability of microloans. In general, their results show that the credit officer is more conservative when a credit record exists, i.e., the mandatory exchange of information about the borrower in this record increases the probability of rejecting the application, especially when it is the first one.

Recently, referring to France, Bourlès and Corzarenco (2018) used survival models to evaluate whether entrepreneurial motivation, differentiating between entrepreneurs by need and by opportunity, is a predictor of duration of microentrepreneurship. These authors found that there is no significant relationship between entrepreneurship out of necessity and corporate survival. Likewise, they find that entrepreneurs out of necessity experience difficulties with repaying the microloan sooner than opportunity entrepreneurs. The only previous evidence for the Americas refers to the recent work of Hassan et al. (2018) who analyze the factors that contribute to an accelerated failure time in a Southern Louisiana (USA) credit union. The evidence shows that the factors specific to borrowers and local factors play an important role in the duration of a loan.

In conclusion, as can be inferred from the analyzed studies, most authors use the survival model methodology to assess the time in the market of the financed company or the duration of the relationship between the MFI and the borrower, but only Babajide (2011), Bos et al. (2015), Bourlès & Cozarenco (2018) and Hassan et al. (2018) use it to model the risk of default, which is the subject of this research.

INSERT TABLE 1

Regarding the determinants of credit risk in an MFI, the literature has proposed a very broad and heterogeneous set of variables. As can be seen in Table 1, in this research we have proposed to organize them into seven categories, according to whether they are related to the borrower, the MFIs, the credit officer, the business to be financed, the microloan, the microcredit program and the macroeconomic environment. One of the challenges of all the previous empirical work has been, regardless of its methodology, first, to access all this information and, second, to filter the data in order to consider a reduced set of variables that facilitate their management by the credit officer and speed up the credit evaluation process. The criteria for the selection of the variables has usually been based on the elimination of the problems of multicollinearity in the empirical study and/or the researcher’s own judgment. However, this research proposes an approach oriented toward the use of a reduced number of variables that the MFI’s credit officer can easily obtain at a low cost, and whose selection is based on a statistical criterion that goes beyond multicollinearity. Thus, first, only those variables that are available at the time of evaluating the granting of microloan are considered. Next, survival models with four penalty modes (LASSO, Adaptive LASSO, Elastic Network and Adaptive Elastic Network) are then estimated to select the least possible number of variables. These penalties are relatively recent statistical developments that have not yet been explored in the study of credit risk in microfinance. The testing of the results obtained through a cross validation process allows for the selection of a very small number of variables with great predictive power, and which do not offer redundant information. This process provides an efficient model for evaluating microloan applications, optimizing all the available information and reducing analyst biases and those generated by information that becomes known after the loan is granted.

Finally, it should be noted that none of the previous works using survival models simultaneously considered the seven categories of variables proposed, as can be seen in Table 1. In fact, the most complete analytical framework was that applied by Pagura (2004), although this author did not evaluate the time to default in the credit relationship, but rather the duration of that relationship, as indicated above. According to Seijas-Giménez et al. (2017), the variables referring to the characteristics of the borrower, microloan and business are, to a large extent, those most commonly used in studies on credit scoring in microfinance, and those that have shown the best behavior as predictors of default. Recently, authors such as De Cnudde et al. (2015) have proposed complementing sociodemographic and microcredit information with information on client interests, using social networks. Likewise, Baklouti (2014) proposes incorporating representative variables of the psychological characteristics of the borrowers in relation to their entrepreneurial character. However, the difficulties and increased cost of accessing this information may affect its incorporation into MFI analysis and, while useful for assessing causes of default, its use as predictor variables may consume an excessive amount of resources.

3. EMPIRICAL STUDY

This research analyzes the credit risk in a portfolio of microloans granted to Uruguayan entrepreneurs by a non-profit Uruguayan government MFI. The objective is to develop a model to predict their expected degree of repayment, while at the same time, avoiding bias and overadjustment. Specifically, it analyzes the PFEP that operates under the MIDES of the Government of Uruguay, whose objective is to provide economic support, training and assistance to entrepreneurs in a situation of socioeconomic vulnerability or poverty to help improve or complement their family income.

The database consists of 1,357 microloans granted and disbursed by the PFEP between 2012 and 2016 at zero interest rates, and which were amortized or repaid prior to February 2017. According to MIDES, the usual repayment period for microcredits is 12 monthly instalments, which can vary from 1 to 18 instalments depending on the characteristics of the investment and the production process. The majority of microcredits were granted in 12 instalments (54.22%), followed by 18 instalments (30.97%) and 15 instalments (10.54%).

A microloan is accordingly considered to be in default if the borrower has returned less than 100% of its value at the end of all repayments, excluding the effect of subsidies.

Uruguayan entrepreneurs who apply for a microloan from the PFEP have the option of repaying up to 30% of the amount of the microloan in kind (subsidy 1). They can make donations of goods, products or services to the community to repay a portion of their microloan. In addition, PFEP credit officers may approve measures according to which the borrower only has to repay a percentage of less than 100 percent of the amount of the microloan, depending on the characteristics of the enterprise being financed, the viability of the project and/or the characteristics of the entrepreneurs' households. In this case, this second type of subsidy can reach a maximum of 80% of the amount of the microloan (subsidy 2).

3.1. DESCRIPTIVE ANALYSIS

This paper applies a survival analysis in which both the variable that represents survival time, that is, the degree of compliance in the repaying of the microloan, and the variable that indicates non-survival, that is, non-payments, must be defined. In the first case, the variable [%_payment] is defined as the percentage recovered from the microloan, and since these are aliquots, it represents the percentage of installments paid. In relation to default, the variable [default] is dichotomous and has a value of "1" if the microloan was not fully repaid at the end of all repayments and "'0" if it was. Of the 1,357 microloans analyzed, 646 are unpaid or in default, using this criterion. Table 2 shows descriptive statistics corresponding to these variables. The average of [default] indicates that 47.6% of microloans were not fully repaid at the end of all repayments. Likewise, considering the variable [%_payment], on average, almost 70% of the installments are amortized.

INSERT TABLE 2

For descriptive purposes, Figure 1 shows a histogram that considers only unpaid microloans at the end of all repayments, that is, those for which the payment of all installments was not recorded (for which the variable [default] has a value of 1.). In a large number of microloans, no more than 10% of their value is repaid, which is the most frequent interval among all those evaluated, indicating that a significant part of the defaults occur during the first installments. In addition, the percentage of microloans that are in the initial repayment interval [0-10%] practically doubles in the next interval [10%-20%], demonstrating that a significant number of borrowers did not pay any installments or only paid the first installments of the microloan. However, for larger microloan repayment intervals [60%-100%], the corresponding number of microloans is reduced, exhibiting a clear downward slope.

Figure 1 represents the expected non-parametric repayment, that is, without considering the effect of any other variable. Thus, the Kaplan-Meier S^(r) estimator is used, which is defined as (Kaplan & Meier, 1958):

$$\hat{S}(r) = \prod_{r_i<R} \frac{n_i - d_i}{n_i} \text{[Equation 1]}$$

Where:

d_i is the number of unpaid microloans once a percentage r_i has been repaid.

n_i the number of microloans that have no defaults at time r_i

Therefore, Figure 1 shows the average probability of repayment without default at the end of all repayments, considering the repayment percentage. The results indicate that there is a repayment equivalent to 50% of the installments in approximately 70% of the microloans. In addition, about 50% of the microloans comply with the payment of 100% of the installments. Table 3 details the phenomenon indicated in the previous figure, quantifying the number of microloans that reached a certain percentage of repayment. Almost a third of defaulting microloans paid only 10% of the installments, while more than 50% paid off 30% of all the installments. Furthermore, less than 5% of the microloans were in default after 90% of the repayment.

INSERT FIGURE 1 AND TABLE 3

Figure 2 shows the hazard rate once a percentage r of the amount of the microloan has been repaid. This rate represents how much the probability of default increases or decreases from the last payment. The evidence shows that once more than 20% of the microloan is repaid, the hazard rate remains constant up to 80%, after which it drops. There is a large number of microloans that did not register the payment of any installment or only registered the payment of the first installments, as opposed to a clearly smaller number of microloans that registered the payment of all installments. Hence, the follow-up and control of the repayment of the microloan is very relevant in the initial phase, particularly until a repayment of 20% has been ensured.

INSERT FIGURE 2

Once the default behavior has been evaluated, Table 4 presents the explanatory variables that will be used in this empirical study to develop a predictive tool for assessing credit risk. According to the previous literature, these variables refer to the characteristics of the microloan, the entrepreneur and the environment.

INSERT TABLE 4

INSERT TABLE 5

3.2. EMPIRICAL ANALYSIS

The empirical study presents, first of all, the estimation of a Cox model (1972) to identify which variables affect default, considering both the variables that are known a priori (characteristics of the microloan and the borrower) and those that are known after the microloan has been granted (characteristics of the macroeconomic environment). The objective is to determine the existing maximum predictive capacity with the available information on the microloan. In this model, the dependent variable or survival duration variable is the percentage of microloan repayment[%_payment], while the non-survival event is default [default], according to the expression:

$$S(r) = e^{\left(-\int_0^r h(s)ds\right)} = e^{-H(r)} \text{[Equation 2]}$$

where S(r) is the non-default probability (1 – cumulative default probability) at percentage recovery level r.

Despite there is discussion in epidemiologic studies without conclusive results about the optimal time scale considering age or time on study as the main alternatives, in social sciences and particularly in our data the use of an absolute time scale would be problematic. The main reasons are that hazard would depend on calendar time and the use of loan duration in months, weeks or days would not hold the proportionality assumption. Consequently, the alternative of a relative time scale measured in percentage does allow to fulfill assumptions of the Cox model and avoid the above problems.

Based on Equation 2, the probability of default occurring once a percentage of the microloan has been repaid is as follows:

$$h_i(r) = h_0(r)e^{X_i\beta} \quad \text{[Equation 3]}$$

Where h_i (r) is the hazard rate of default after the repayment of a percentage r and h_0 (r) is the baseline hazard, which has an unknown functional form and is common to all microloans. This baseline hazard is only influenced by the cumulative amount the borrower has repaid on the microloan. Likewise, a set of variables $X_{i,}$ their coefficients β, and the baseline hazard determine the total hazard attributed to each variable, through

$e^{\beta X_i}$ (Mills, 2011). Thus, each variable X_i, together with is coefficient β, can increase or reduce the probability of total default h_i (r). A value greater than 1 increases the risk of default, while one less than 1 reduces it. The coefficients β are obtained through a partial maximum likelihood function using the packages survival, rms and hdnom in the R software (Therneau, 2022, 2020; Harrell, 2000; Xiao et al., 2019).

Table 6 below shows the estimation of the parameters in equation [3] for each variable X_i. The results indicate that the variables which increase the probability of default are the year in which the microloan was granted, the number and the value of the installments, the monthly variation in the wage during the repayment term of the microloan, the percentage of type 1 subsidy, the performance of a commercial type of entrepreneurial activity and the location of the enterprise in the Central region. In addition, the number of installments paid on the previous microloan, the age of the entrepreneur when they obtained the microloan and the performance of a productive venture in the Southern region reduce the probability of future default.

INSERT TABLE 6

Next, Equation 3 was estimated, considering all the X_i variables together (Model A). Likewise, since we do not know the future evolution of the macroeconomic environment at the time the microloan is granted, the model is estimated while excluding this information (Model B). According to the results identified in Table 7, the explanatory capacity of Model A reaches high values with a Harrel’s C of 0.624. This value is the maximum capacity that a predictive model would have with the available information. However, the results related to the proportionality test indicate that there may be problems with the variables [amount], [loan_fee] and [wages], and so their impact may be underestimated or overestimated. With regard to Model B, the results of the VIF test indicate that, without the macroeconomic environmental variables, there are problems with multicollinearity, particularly for the variables [number_fee], [amount] and [loan_fee]. It should also be pointed out that Harrel’s C only decreases to 0.612 without the information on the macroenvironment. Therefore, this model will be the starting point for developing the predictive tool in this work.

As an analysis of robustness of the results in the Cox model, as a complement, the modeling of the baseline hazard was carried out through the exponential distributions, Weibull and Gompertz functions (see Appendix 1). These alternative models show values similar to those obtained through the Cox model, which confirms its robustness.

INSERT TABLE 7

In order to evaluate which variable or variables could be omitted from the model because they provide little or no information about the probability of expected default, a variable reduction method has been used, based on including a penalty according to the maximum partial likelihood. The penalty reduces the size of the coefficient estimates shrinking them towards zero in this regularization process and consequently it is useful for parameter selection. When the penalty parameters are exactly zero, we have the original objective function, but higher penalty values would lead coefficients toward zero. This contributes to reducing the number of variables, decreasing problems of multicollinearity and facilitating the application of the model in the MFI. In short, these penalties contribute to selecting the most relevant variables and to improving the predictive capacity by reducing the over-adjustment. They also make it possible to reduce the number of variables required, based on an objective and homogeneous criterion, thus preventing selection biases. The ultimate objective is to facilitate the objective evaluation of microloans in the MFI, reducing the cost and effort needed to collect the information that is relevant for the analysis. In order to implement the penalties, Verweij and Van Houwelingen (1994) proposed a cross-validation criterion for a survival analysis with the aim of determining the appropriate value for the adjustment parameters. In this way, the objective function is estimated is expressed as follows:

$$l_n(\beta) - \sum_{j=1}^{J} \lambda \left[(1-\alpha)\frac{\beta_j^2}{2} + \ \alpha\left|\beta_j\right| \right] \text{[Equation 4]}$$

where λ and α represent the intensity of the penalty. This type of penalty is known as elastic net, and it was developed by Zou and Hastie (2005).

Different alternatives are available, depending on the combinations between λ and α, which can generate different penalties. The ENET elastic net is estimated based on equation 4 (Zou & Hastie, 2005). If in equation 4, it is specified that α is equal to 1, the LASSO penalty is obtained (Tibshirani, 1997). The AENET and ALASSO penalties are versions of the previous two that ensure compliance with the "oracle property." This property implies that the estimator is consistent in the estimation of the parameter, as well as in the variable selection process (Zou, 2006). This thus constitutes a robust measure of penalization applied to empirical studies.

Model B has been re-estimated, taking into account all four penalties proposed, i.e., AENET, ENET, ALASSO and LASSO (see Appendix 2). Figure 3 shows an analysis of the area under the curve (ROC) by means of an internal validation though a bootstrap method where the models are evaluated at different repayment values (Tibshirani, 1996). We can compare the results to Harrel's C, which was previously obtained (0.624 for the model with all the variables – Model A- and 0.612 if the environmental variables are excluded – Model B). The comparison has been made using the hdnom package of the software R (Xiao et al., 2019). AENET shows the best fit, as it has the highest values for ROC, although there are not any very important differences among all the penalties.

INSERT FIGURE 3

Taking into account the AENET penalty, the results of the survival model are identified in Table 8. The selection of variables makes it possible to correct the problems of multicollinearity that were previously identified (Cox – Model B), and Harrel's C is maintained at 0.620. Likewise, it is observed that except for those variables related to subsidy 2 and gender, all variables are significant. A more detailed analysis of each variable indicates that the risk of default is higher if the repayment of the microloan is increased by one installment. Likewise, the results show an increase in the risk of default if the entrepreneur performs an industrial, commercial or service activity as opposed to if they work in the primary sector (primary sector was considered as reference in the econometric model). Furthermore, the age of the entrepreneur and engaging in the activity in the southern and eastern geographical regions are factors reducing the risk of default.

INSERT TABLE 8

The results above could present over-adjustment and therefore are useful for explaining, but not for predicting, the risk of default. For this reason, the model is validated as a predictive tool.

Figure 4 identifies a graphic analysis validating the model estimated in Table 10. The graph shows that the model shows a greater capacity for prediction at the higher microloan repayment levels. On the other hand, the lower capacity for prediction at lower recovered values could be attributed to the existence of an important number of microloans that go into default at the first installment, which complicates the prediction in this segment. Therefore, the ROC area presents a value close to 0.600 at very low microloan repayment values and between 0.625 and 0.650 at higher repayment values.

INSERT FIGURE 4

Below, considering the model presented in Table 10, the expected percentage of repayment has been estimated for each microloan (Table 9). For this estimation, the sample has been divided into three groups of equal size, according to the predicted repayment values: *high risk, medium risk, low risk.* These three categories correspond to three equal sized intervals of probability. The values obtained have been compared to the values actually observed for each microloan. This process has been completed using random sampling that divides the sample into ten groups. This prevents the model from presenting good results as the consequence of potential overadjustment. The results are shown in Table 11 and confirm the appropriate calibration of the model, since the average predicted value of repayment and the value actually observed are very similar to one another.

INSERT TABLE 9

Finally, Figure 5 shows the graphic evolution of defaults, taking into account different microloan repayment percentages in each of the three risk groups. The results show that the model properly classifies the percentage of microloan repayment based on the microloan risk group, as confirmed by the *logrank* test (Cox, 1972).

INSERT FIGURE 5

The previous results confirm that the model developed by cross-validation with AENET penalization, using only the information available prior to the time when the microloan is granted (excluding the characteristics of the macroenvironment) adequately classifies the microloans according to their risk of default. Likewise, due to the specification of the model, it is possible to assess the default risk at different times during the life of the microloan. However, in order for these results to be useful to a credit officer in an MFI, they must be transferred to a tool that is simple to apply and whose output is easily interpretable. For this purpose, the results obtained in this empirical research have been applied to a nomogram (Figure 6). The nomogram is a graphic tool that makes it possible to obtain the value of a complex function based on the value of different input variables. To achieve this, first a score is assigned to the value of each input variable (top line). Second, the sum of the different scores is applied to an indicator of the probability that default will occur. This indicator is evaluated for different percentages of repayment (bottom line). Based on the nomogram, it is also possible to graphically identify the variables with the greatest impact on the future risk of default, given that they are the ones that show the greatest possible variation in the score. Accordingly, the variables with the greatest importance in order to predict the risk of default are, in this order, past-due installments [fee_defaulter], the percent of subsidies [benefit_1, Benefit_2] and the age of the individual [age_entrepreneur], followed by the activity sector [industry] and the number of installments [number_fee]. Likewise, the installment value [loan_fee] and region [region] where the activity takes place are variables of notable importance, albeit less than that of the previously mentioned variables.

INSERT FIGURE 6

4. CONCLUSIONS

Microcredit is a very popular tool for financial inclusion, especially oriented towards financing low-income entrepreneurs who, generally speaking, tend to be excluded from the traditional financial system. In a broader sense, microfinance offers different services of an economic and social nature through the MFIs, which play a fundamental role in achieving global financial inclusion.

Over the last two decades, financial stability and even the very survival of the MFIs has been weakened due to increasing competition, their overindebtedness and the economic crises. Accordingly, in order to ensure their survival and fulfillment of their social and financial purposes, their activity should be accompanied by more efficient operations that optimize the use of their resources, in particular when they must work in less favorable macroeconomic environments. In this context, default has become a very relevant topic for study in the literature, developing models that contribute to explaining and predicting it, so that the MFIs can reduce it and ensure their financial sustainability over the medium and long term.

This work has analyzed the credit risk in a non-profit Uruguayan MFI during the period 2012-2016. In particular, this research has estimated a Cox proportional hazard model using four penalties, namely ENET, LASSO, AENET and ALASSO. To our knowledge, this methodology so far has not been applied in the previous literature focused on Latin America, where microfinance plays a crucial role as a tool for financial inclusion.

The analysis of survival has contributed to modeling the time to default, that is to say, not just identifying the determining factors of default, but also the moment in time when it is most likely to occur. Therefore, the model developed in this paper has presented an important capacity for prediction and makes it possible to properly estimate the expected probability of repayment the microloan, contributing to classifying the loan operation according to the default risk profile (high, medium and low). These results have been demonstrated by the fact that the mean predicted value of repayment and the value actually observed are very similar to one another. The preceding literature was normally focused on estimating the binary models in order to identify the variables that explain and/or predict the risk of default on microloans and at a specific moment in time. In this research, the empirical analysis makes it possible to predict the likelihood of default at different moments during the life of the microloan, thus making it easier to assess the risk of default for a microcredit portfolio. As a result, it helps the MFI adjust its own needs for funds in order to deal with possible defaults, so that they do not weaken its financial position, jeopardizing its survival. Therefore, altogether, these aspects make it possible to reduce costs and increase the availability of funds so that an MFI can efficiently manage its resources.

The methodology applied in this paper has also contributed to being able to evaluate the sample when there is incomplete information, which is fairly common in MFIs in those countries with an immature level of development, such as Uruguay, where data collection by the MFIs is quite informal. As a matter of fact, one of the main challenges encountered in this research was the collection and organization of data on the microloans granted by the Uruguayan MFI analyzed. In fact, the evaluation of these data with a purely descriptive focus has already represented an important contribution to the microloan strategy applied by the Uruguayan government. It made it possible to know in detail the sociodemographic and economic characteristics related to microloans granted in this country, seeing for the first time how the default rates occur that are recorded throughout the life of each microloan.

A third contribution refers to the inclusion of four penalties in the estimation of the Cox model, which has enabled it to select a smaller number of variables of an objective nature, that are easily identifiable by the credit officer prior to granting a microloan. These variables were: the number and value of the monthly installments for the microloan, the number of unpaid installments in a previous microloan, the percentage of subsidy (repayment in kind and cancellation of the debt), the age and gender of the borrower, the activity sector of the business to be financed and the geographical region where it is located. This reduction in the number of variables needed simplifies and streamlines the empirical application of this methodology in the daily operations of the MFI, without compromising the good fit of the predictive model. The results of this research therefore confirm that a small number of variables can allow for an adequate prediction of the likelihood of default on microloans. Furthermore, these variables would be objective in nature, which is to say that they avoid bias on the part if the credit official when collecting and analyzing data, and they could be obtained at a low cost before granting the microloan. Thus, the evaluation of the credit risk in an MFI should not consume an excessive volume of resources by applying this empirical focus, particularly considering the amount of work time that the credit officer must dedicate to evaluating each microloan. This constitutes an important contribution made by this research for application in other MFIs that wish to professionalize the analysis of credit risk, but that do not have an abundance of resources.

Ultimately, and in light of all the results obtained, this research contributes to professionalizing the management of microloans in the MFIs, reducing costs and the subjectivity in the evaluation process prior to granting these loans. Likewise, it reveals the importance of properly selecting the variables that are used in the credit risk analysis, since many of them could have a high cost associated with obtaining them, in spite of the fact that they make a relatively small contribution to the prediction of the likelihood of default. This is especially important in the context of the MFIs, which have a fairly limited volume of financial resources, due to the relationship between the amount of the microloan and the expenses incurred to evaluate both the borrower and the requested loan. In fact, this research has applied the results obtained to a nomogram, thus making them easier to understand and implement by the credit officers in an MFI.

Focusing on the Uruguayan market, this research has contributed to defining a set of variables to be collected by the loan officer for the evaluation of the potential borrower, where the emphasis is not on the number of variables but on which variables. It has also proposed the professionalization of microcredit assessment by implementing credit scoring.

Finally, this research has not been free of certain limitations, in particular those related to the availability of data. It should be noted that, as indicated above, an important effort was made to collect and organize the data on the microloans that were evaluated. In spite of this, some variables could not be considered in the study because they were not available for the entire period and/or for all or most of the borrowers. Some examples of these variables were the marital status or level of formal education of the borrower, their specific intended objective of the microloan or their default history with other MFIs. Another limitation of this research was the impossibility of starting the study period in 2006. The existence of important differences in the variables collected during the period

2006-2012 as compared to the period 2012-2016 meant that the actions of the PFEP could not be included in the first period, which would have been of interest in terms of expanding the sample of microloans.

With regard to future lines of work, it would first be interesting to apply other methodologies that have recently been applied to the study of default in the MFIs, such as neuronal networks, for example. Similarly, another line of future research consists of studying other type of credit risk, such as “drop-out scoring”, which measures the likelihood that the borrower of a current loan will not apply for a new loan in the future. Given the social purpose of microloans, this could help estimate the impact of the PFEP on mitigating the poverty of those entrepreneurs with the greatest economic vulnerability in Uruguay. Along these same lines, “collections scoring” can also contribute a new perspective to the study, as it involves predicting the probability that a borrower with a certain number of late payments on the current microloan could generate a larger number of defaults in the future. Accordingly, their knowledge and prediction would contribute to improving the efficiency of the PFEP in managing collections related to microloans.

5.BIBLIOGRAPHY

Ayouche, S., Aboulaich, R., & Ellaia, R. (2017). Partnership credit scoring classification Problem: A neural network approach. *International Journal of Applied Engineering Research*, 12(5), 693-704.

Babajide, A. A. (2011). Microfinance and micro & small enterprises (MSEs) survival in Nigeria-A survival analysis approach. *Global Journal of Management and Business Research*, 11(11).

Baklouti, I. (2014). A psychological approach to microfinance credit scoring via a classification and regression tree. *Intelligent Systems in Accounting, Finance and Management*, 21(4), 193-208.

Banasik, J., Crook, J. & Thomas, L. (1999). Not if but when will borrowers default. *The Journal of the Operational Research Society*, 50 (12), 1185-1190.

Bekele, E., & Worku, Z. (2008). Women entrepreneurship in micro, small and medium enterprises: The case of Ethiopia. *Journal of International Women's Studies*, 10 (2), 3-19.

Bellotti, T. & Crook, J. (2009). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operational Research Society*, 60 (12), 1699-1707.

Bensic, M., Sarlija, N., & Zekic-Susac, M. (2005). Modelling small-business credit scoring by using logistic regression, neural networks and decision trees. *Intelligent Systems in Accounting, Finance and Management*, 13(3), 133-150.

Bos, J., De Haas, R., & Millone, M. (2015). *Sharing borrower information in a competitive credit market* (No. 180). Working Paper, available at: https://bg.uek.krakow.pl/e-zasoby/siec_lokalna/Ebor/w180.pdf

Bourlès, R., & Cozarenco, A. (2018). Entrepreneurial Motivation and Business Performance: Evidence from a French Microfinance Institution. *Small Business Economics*, 51 (4), 943-963.

Bumacov, V., Ashta, A., & Singh, P. (2014). The Use of Credit Scoring in Microfinance Institutions and Their Outreach. *Strategic Change*, 23(7-8), 401-413.

Cao, R., Vilar, J.M. & Devia, A. (2009). Modeling consumer credit risk via survival analysis. Statistics & Operations Research Transactions, 33 (1), 3-30.

Chakravarty, S., & Jha, A. (2012). Viability of "Credit Scoring in Microfinance" for Developing Countries. *International Review of Social Sciences and Humanities*, 3(1), 104-107.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 34, 187-220.

De Cnudde, S., Moeyersoms, J., Stankova, M., Tobback, E., Javaly, V., & Martens, D. (2015). Who cares about your Facebook friends? *Credit scoring for microfinance*, No. 2015018.

De Olloqui, F., Andrade, G., & Herrera, D. (2015). *Inclusión financiera en América Latina y el Caribe: coyuntura actual y desafíos para los próximos años*. Banco Interamericano de Desarrollo, No. IDB-DP-385.

Dellien, H. (2003). Credit Scoring in Microfinance: Guidelines Based on Experience with WWB Affiliates in Colombia and the Dominican Republic. *Women's World Banking*, 1(2), 1-15.

Dellien, H., & Schreiner, M. (2005). Credit scoring, banks, and microfinance: balancing high-tech with high-touch. *Microenterprise Development Review*, 8(2).

Demirgüç-Kunt, A.; Klapper, L; Singer, D.; Ansar, S. & Hess, J. (2018). *The Global Findex Database 2017: Measuring Financial Inclusion and the Fintech Revolution*. Washington, D.C: Banco Mundial.

DINAPYME - Dirección Nacional De Artesanías, Pequeñas Y Medianas Empresas (2018). Encuesta Nacional de Mipymes comerciales, industriales y de servicios. Edición

2017. Avalaible at: http://www.dinapyme.gub.uy/documents/4694435/0/web%20encuesta%202018.pdf

Harrel JR, Frank E. (2016): rms: Regression modeling strategies. R package version, 2016, vol. 5, no 2.

Hassan, M. K., Brodmann, J., Rayfield, B., & Huda, M. (2018). Modeling credit risk in credit unions using survival analysis. *International Journal of Bank Marketing*, 36(3), 482-495.

Janda, K. & Zetek, P. (2013). Macroeconomic factors influencing interest rates of microfinance institutions in Latin America. MPRA Paper No. 49973. Munich Personal RePEc Archive.

Kammoun, A. & Triki, I. (2016). Credit scoring models for a Tunisian microfinance institution: comparison between artificial neural network and logistic regression. *Review of Economics & Finance*, 6(1), 61-78.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282), 457-481.

Ledgerwood, J. (1999). *Microfinance handbook: an institutional and financial perspective*. World Bank, Washington DC.

Madorno, F., Mecatti, F. and Figini, S. (2013). "Survival models for credit risk estimation", in Carpita, M., Brentari, E. and Qannari, E.M. (Eds), Advances in Latent Variables: Methods, Models and Applications, Springer-Verlag, New York, NY.

Microfinance Barometer (2012). 1st. Edition. Available at: https://www.microfinancegateway.org/sites/default/files/mfg-en-paper-microfinance-barometer-2012-jul-2012.pdf

Microfinance Barometer (2018). 9th. Edition. Avalaible at: http://www.convergences.org/wp-content/uploads/2018/09/BMF_2018_EN_VFINALE.pdf

Mills, M. (2011). *Introducing Survival and Event History Analysis*. Thousand Oaks, CA: Sage.

Pagura, M. E. (2004). *Client Exit in Microfinance: A Conceptual Framework with Empirical Results from Mali*. CSAE Conference, Oxford University.

Perossa, M., & Gigler, S. (2015). Modelos microfinancieros latinoamericanos: una experiencia para la inclusión social y el desarrollo. *Cooperativismo & Desarrollo*, 23(106).

Reinke, J. (1998). How to lend like mad and make a profit: A micro-credit paradigm versus the start-up fund in South Africa. *Journal of Development studies*, 34(3), 44-61.

Robinson, M. (2001). *The microfinance revolution: Sustainable finance for the poor*. World Bank Publications.

Roszbach, K. (2004). Bank lending policy, credit scoring, and the survival of loans. *Review of Economics and Statistics, 86*(4), 946–958.

Schreiner, M. (1999a). *The risk of exit for borrowers from a microlender in Bolivia.* Center for Social Development, Washington University in St. Louis, gwbweb. wustl. edu/users/schreiner.

Schreiner, M. (1999b). *A scoring model of the risk of costly arrears at a microfinance lender in Bolivia.* Microfinance Risk Management and Center for Social Development, Washington University in St. Louis.

Schreiner, M. (2000). Credit scoring for microfinance: Can it work? *Journal of Microfinance/ESR Review*, 2(2), 6.

Schreiner, M. (2003). Scoring Drop-out at a Microlender in Bolivia. *Savings and Development*, 27(2), 101-118.

Schreiner, M. (2004). Scoring arrears at a microlender in Bolivia. *ESR Review*, 6(2), 65.

Seijas-Giménez, M.N.; Vivel-Búa, M.; Lado-Sestayo, R. & Fernández-López, S. (2017). La evaluación del riesgo de crédito en las instituciones de microfinanzas: estado del arte. *COMPENDIUM: Cuadernos de economía y administración*, 4(9), 36-52.

Sharma, M., & Zeller, M. (1997). Repayment performance in group-based credit programs in Bangladesh: An empirical analysis. *World Development*, 25(10), 1731-1742.

Therneau T.M. (2022). A Package for Survival Analysis in R. R. https://CRAN.R-project.org/package=survival.

Therneau, T.M. & Grambsch, P. M. (2000). Modeling Survival Data: Extending the Cox Model. Springer, New York. ISBN 0-387-98784-3.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Statistics in medicine*, 16(4), 385-395.

Tong, E.N., Mues, C. & Thomas, L.C. (2012), "Mixture cure models in credit scoring: if and when borrowers default", *European Journal of Operational Research*, 218 (1), 132-139.

Van Gool, J., Verbeke, W., Sercu, P., & Baesens, B. (2012). Credit scoring for microfinance: is it worth it? *International Journal of Finance & Economics*, 17(2), 103-123.

Verweij, P. J., & Van Houwelingen, H. C. (1994). Penalized likelihood in Cox regression. *Statistics in medicine*, 13(23-24), 2427-2436.

Viganò, L. (1993). A credit scoring model for development banks: An African case study. *Savings and Development*, 17 (4), 441–482.

Vogelgesang, U. (2003). Microfinance in times of crisis: The effects of competition, rising indebtedness, and economic crisis on repayment behavior. *World Development*, 31(12), 2085-2114.

World Bank (2014). *Global Financial Development Report 2014: Financial Inclusion*. Washington, DC.

World Bank (2017). *Revolutionizing microfinance: Insights from the 2017 Global Symposium on Microfinance*. Kuala Lumpur.

Xiao, N. Xu, Q.S. & Li, M.Z. (2019). hdnom: Building Nomograms for Penalized Cox Models with High-Dimensional Survival Data.

Zeller, M. (1998). Determinants of repayment performance in credit groups: The role of program design, intragroup risk pooling, and social cohesion. *Economic development and cultural change*, 46(3), 599-620.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476), 1418-1429.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301-320.

6.APPENDIXES

Appendix 1. Robustness analysis (exponential, Weibull and Gompertz)

Variables	Hazard ratio (standard error)		
	Exponential	Weibull	Gompertz
[year]	1.196*** (-0.061)	1.141*** (-0.047)	1.178*** (-0.057)
[amount]	1.002 (-0.008)	1.003 (-0.007)	1.002 (-0.007)
[loan_fee]	0.997 (-0.008)	0.997 (-0.006)	0.998 (-0.007)
[number_fee]	1.04 (-0.084)	1.014 (-0.067)	1.039 (-0.079)
[benefit_1]	1.025*** (-0.007)	1.022*** (-0.006)	1.025*** (-0.006)
[benefit_2]	1.006** (-0.003)	1.006** (-0.002)	1.006** (-0.003)
[fee_defaulter]	1.128*** (-0.026)	1.090*** (-0.019)	1.116*** (-0.024)
[type_entrepreneur] (2=female)	0.830** (-0.077)	0.854** (-0.064)	0.831** (-0.073)
[age_entrepreneur]	0.981*** (-0.004)	0.985*** (-0.003)	0.982*** (-0.004)
[industry] 2	1.835*** (-0.367)	1.716*** (-0.297)	1.798*** (-0.346)
[industry] 3	2.131*** (-0.415)	1.880*** (-0.316)	2.056*** (-0.385)
[industry] 4	2.236*** (-0.451)	1.986*** (-0.346)	2.158*** (-0.419)
[region] 2	0.852 (-0.112)	0.895 (-0.094)	0.863 (-0.107)
[region] 3	0.903 (-0.147)	0.926 (-0.119)	0.91 (-0.139)
[region] 4	0.787* (-0.105)	0.840* (-0.089)	0.800* (-0.1)
[region] 5	0.772 (-0.125)	0.826 (-0.108)	0.787 (-0.121)
[region] 6	0.934 (-0.162)	0.918 (-0.122)	0.92 (-0.149)
[CPI]	0.809 (-0.199)	0.864 (-0.167)	0.814 (-0.187)
[employment]	0.932 (-0.312)	0.963 (-0.277)	0.934 (-0.299)
[interest]	0.999 (-0.028)	1.003 (-0.023)	0.999 (-0.026)
[water]	1.269 (-0.26)	1.187 (-0.196)	1.26 (-0.243)
[light]	0.891 (-0.177)	0.879 (-0.144)	0.888 (-0.167)
[wages]	1.102 (-0.356)	1.059 (-0.284)	1.065 (-0.319)
Constant	0.000*** (0.000)	0.000*** (0.001)	0.000*** (0.001)
/ln_p	---	-1.717*** (0.071)	---
p	---	0.180	---
1/p	---	5.567	---

Variables	Hazard ratio (standard error)		
	Exponential	Weibull	Gompertz
/gamma	---	---	-1.155*** (0.169)
N (defaults)	1.357	1.357	1.357
Pseudo log. likelihood	-4,704.390	-2,771.784	-4,672.408
AIC	9,456.781	5,593.568	9,394.816
BIC	9,581.893	5,723.894	9,525.142
Wald (χ^2) (p-value)	145.02 (0.000)	130.91 (0.000)	139.28 (0.000)

Note:***,**;* significant at the 1%, 5% and 10% level, respectively. The variables [amount] and [loan_fee] have been included with logarithmic transformation in the model.

Appendix 2. Cox Model with penalties (AENET, ENET, LASSO and ALASSO)

Variable	Hazard ratio			
	AENET	ENET	LASSO	ALASSO
[number_fee]	1.040	1.028	1.038	1.040
[loan_fee]	1.000	0.998	1.000	-
[amount]	-	1.001	-	-
[benefit_1]	1.026	1.026	1.025	1.026
[benefit_2]	1.003	1.004	1.003	1.004
[fee_defaulter]	1.101	1.099	1.097	1.101
[type_entrepreneur] (2=female)	0.884	0.893	0.899	0.888
[age_entrepreneur]	0.982	0.982	0.983	0.982
[industry] 2	1.865	1.626	1.582	1.876
[industry] 3	2.385	2.079	2.036	2.404
[industry] 4	2.268	1.964	1.912	2.293
[region] 2	0.861	0.882	0.926	0.860
[region] 3	0.897	0.919	0.971	0.897
[region] 4	0.776	0.796	0.829	0.774
[region] 5	0.775	0.792	0.834	0.775
[region] 6	0.925	0.945	-	0.928
Best alpha	0.05	0.05	0 is assumed	0 is assumed
Best lambda	0.01765	0.00862	0.00304	0.0028
ROC (C-INDEX) at 90%	0.624	0.603	0.603	0.594
AIC	-92.094	-90.005	-92.643	-93.703
BIC	-14.256	-7.003	-19.973	-21.034

Note:***,**;* significant at the 1%, 5% and 10% level, respectively. The variables [amount] and [loan_fee] have been included with logarithmic transformation in the model.

7. TABLES AND FIGURES

Table 1. Previous literature using survival analysis in microfinance

Authors	Country (Continent)	Period	Sample	Dependent variable	Independent variables Number of input variables		Characteristics related to:						
					Available variables	Included variables	A)	B)	C)	D)	E)	F)	G)
Pagura (2004)	Mali (Africa)	1999-2001	206 microentrepreneurs	Duration of the debt relationship between the entrepreneur and the MFI	32	8	X	X		X	X	X	X
Bekele and Worku (2008)	Ethiopia (Africa)	1996-2001	500 micro, small and medium firms	Duration of the debt relationship between the launch of the business and the end of the study (2001) Duration of the debt relationship between the launch and close of the business	6	6	X			X			X
Babajide (2011)	Nigeria (Africa)	2004-2008	502 firms	Time to default	8	8	X			X			X
Bos *et al*.(2015)	Bosnia-Herzegovina (Europe)	2002-2012	210,044 loan applications (2007-2012) 164,760 approved loans (2002-2012)	Time to default	20	20	X				X		X
Bourlès and Cozarenco (2018)	France (Europe)	2006-2016	574 micro-firms and 294 micro-entrepreneurs	Time to 3rd late payment Time to business closure	18	4	X			X			X
Hassan *et al*.(2018)	United States (North America)	1981-2015	3,976 loans	Time to default	21	11	X				X		X

Notes: characteristics of the microloan evaluated in the research and that are related to: A) Borrower; B) MFIs; C) Credit officer; D) Business; E) Microloan; F) Microcredit program; G) Macroeconomic environment.

Table 2. Descriptive statistics

Variables	Mean	Standard Deviation	Coefficient of variation	Observations
[default]	0.476 (47.6%)	0.500	1.049	1357
[%_payment]	0.691 (69.1%)	0.384	0.556	1357

Figure 1. Evolution of the percentage of microloans that did not register defaults with respect to the percentage of their amount repaid

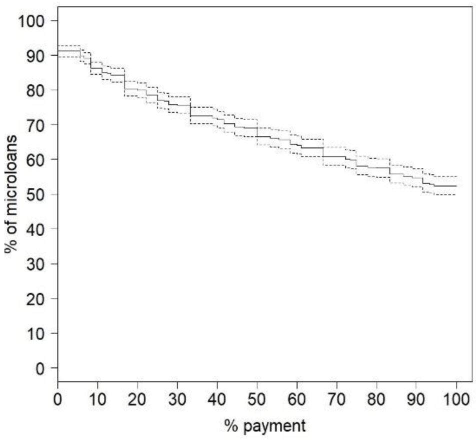


Table 3. Distribution of microloans according to repayment percentage

% of repayment	Number of microloans	% of total	% % cumulative
<10%	186	28.79%	28.79%
10%-20%	87	13.47%	42.26%
20%-30%	57	8.82%	51.08%
30%-40%	57	8.82%	59.91%
40%-50%	66	10.22%	70.12%
50%-60%	34	5.26%	75.39%
60%-70%	44	6.81%	82.20%
70%-80%	46	7.12%	89.32%
80%-90%	38	5.88%	95.20%
>90%	31	4.80%	100.00%
Total	646	100.00%	

Figure 2. Hazard rate according to the percentage repaid of the microloan amount

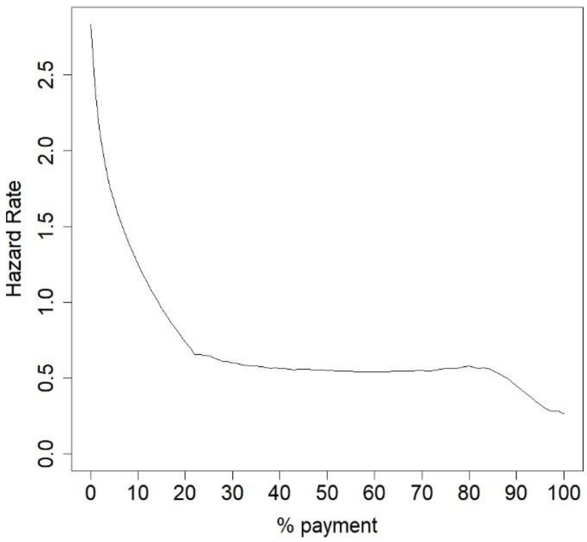


Table 4. Independent variables

Variables	Definition
MICROLOAN CHARACTERISTICS	
[amount]	Amount of the microloan in Uruguayan pesos
[loan_fee]	Value of the monthly microloan installment in Uruguayan pesos
[number_fee]	Number of monthly microloan installments
[benefit_1]	Percentage that subsidy 1 represents out of the total amount of the microloan repayment in kind of up to 30% of the amount of the microloan
[benefit_2]	Percentage that subsidy 2 represents out of the total amount of the microloan cancellation of up to 80% of the amount of microloan, depending on the characteristics of the venture being financed, the viability of the project and/or the characteristics of the entrepreneurs' homes.
BORROWER CHARACTERISTICS	
[fee_defaulter]	Number of total installments of the previous microloan in default
[type_entrepreneur]	Gender of the entrepreneur (1= Male; 2=Female)
[age_entrepreneur]	Age of the borrower at the time the microloan is granted
[region]	Geographic region of the venture (1= North, 2= Northern Coastal, 3= Western Coastal, 4= South, 5= East, 6= Central)
[industry]	Sector of activity to which the financed venture belongs (1=Primary, 2=Industry, 3=Trade, 4=Services)
CHARACTERISTICS OF THE MACROECONOMIC ENVIRONMENT	
[year]	Year of microloan disbursement
[CPI]	Monthly variation in the Consumer Price Index in the microloan repayment period
[wages]	Monthly variation of the Mean Salary Index in the microloan repayment period
[interest]	Monthly variation in the average interest rate for cash loans to microenterprises in the microloan repayment period
[employment]	Monthly variation of the Employment Index in the microloan repayment period
[light]	Monthly variation in electricity rates during the microloan repayment period
[water]	Monthly variation in water rates during the microloan repayment period

Table 5. Descriptive statistics - independent variables

Variable	Mean	Minimum	Maximum	Standard Deviation
[amount]	12,953.56	500	741,000	24,860.28
[loan_fee]	1,738.47	29.37	370,500	14,770.33
[number_fee]	13.96	1	18	3.33
[benefit_1]	0.18	0	0.56	0.08
[benefit_2]	0.32	0	0.8	0.19
[fee_defaulter]	4.42	0	18	5.68
[type_entrepreneur]				
Male	38.10%	0	1	
Female	61.90%	0	1	
[age_entrepreneur]	42.29	18	77	10.81
[region]				
North	19.35%	0	1	
Northern Coastal	21.59%	0	1	
Western Coastal	10.61%	0	1	
South	29.60%	0	1	
East	11.70%	0	1	
Central	7.15%	0	1	
[industry]				
Primary	9.82%	0	1	
Industrial	23.18%	0	1	
Commerce	42.96%	0	1	
Services	24.04%	0	1	
[year]				
2012	25.64%	0	1	
2013	31.76%	0	1	
2014	16.36%	0	1	
2015	24.17%	0	1	
2016	2.06%	0	1	
[CPI]	1.07%	-0.43%	3.94%	0.46%
[wages]	0.25%	-1.25%	3.14%	0.29%
[interest]	1.43%	-18.98%	13.03%	2.60%
[employment]	-0.05%	-2.48%	2.08%	0.27%
[light]	0.57%	0.00%	1.94%	0.32%
[water]	0.71%	0.00%	1.81%	0.28%

Table 6. Univariate survival analysis

Variables	e^{β} (standard error)	Lr test (χ^2) (p-value)
[amount]	1.000 (0.001)	0.80 (0.370)
[loan_fee]	0.999*** (0.000)	8.76*** (0.003)
[number_fee]	1.040*** (0.012)	11.87*** (0.001)
[benefit_1]	1.025*** (0.005)	25.80*** (0.000)
[benefit_2]	1.000 (0.002)	0.03 (0.852)
[fee_defaulter]	1.094*** (0.018)	28.72*** (0.000)
[type_entrepreneur] (2=female)	0.973 (0.780)	0.12 (0.729)
[age_entrepreneur]	0.982*** (0.003)	27.18*** (0.000)
[industry] 2	0.903 (0.094)	1.19 (0.275)
[industry] 3	1.266*** (0.079)	8.89*** (0.003)
[industry] 4	1.113 (0.092)	1.32 (0.250)
[region] 2	1.059 (0.094)	0.37 (0.542)
[region] 3	1.107 (0.126)	0.64 (0.424)
[region] 4	0.795*** (0.090)	6.78*** (0.009)
[region] 5	0.886 (0.126)	0.96 (0.327)
[region] 6	1.252 (0.140)	2.43 (0.119)
[year]	1.187*** (0.037)	29.86*** (0.000)
[CPI]	0.988 (0.081)	0.02 (0.887)
[employment]	1.008 (0.135)	0.00 (0.953)
[interest]	1.009 (0.014)	0.41 (0.522)
[water]	1.059 (0.137)	0.20 (0.659)
[light]	0.820* (0.095)	2.91 (0.088)
[wages]	0.772** (0.090)	4.87** (0.027)

Note: ***,**;* significant at the 1%, 5% and 10% level, respectively. The variables [amount] and [loan_fee] have been included with logarithmic transformation in the model. The variables [industry] and [region] have been treated as dummies for each value.

Table 7.COX Model

	MODEL A (all variables)		MODEL B (excluding macroeconomic variables)	
Variables	e^{β} (standard error)	Proportionality test Rho (p-value)	e^{β} (standard error)	VIF
[amount]	1.002 (0.007)	0.059 (0.034)	1.003 (0.004)	19.87
[loan_fee]	0.997 (0.007)	-0.056 (0.040)	0.996 (0.004)	18.84
[number_fee]	1.037 (0.075)	-0.050 (0.143)	1.01 (0.035)	5.49
[benefit_1]	1.024*** (0.006)	0.005 (0.892)	1.027*** (0.006)	1.23
[benefit_2]	1.006** (0.003)	-0.009 (0.838)	1.004* (0.003)	1.42
[fee_defaulter]	1.101*** (0.015)	0.031 (0.536)	1.101*** (0.018)	1.02
[type_entrepreneur] (2=female)	0.855* (0.071)	0.067 (0.086)	0.884 (0.073)	1.13
[age_entrepreneur]	0.983*** (0.003)	-0.022 (0.612)	0.982*** (0.003)	1.02
[industry] 2	1.768*** (0.332)	0.002 (0.966)	1.878*** (0.348)	2.60
[industry] 3	2.031*** (0.372)	0.007 (0.856)	2.405*** (0.424)	3.18
[industry] 4	2.103*** (0.399)	-0.005 (0.901)	2.287*** (0.427)	2.63
[region] 2	0.867 (0.103)	0.041 (0.295)	0.860 (0.102)	1.74
[region] 3	0.914 (0.134)	0.049 (0.208)	0.894 (0.13)	1.42
[region] 4	0.807* (0.097)	0.024 (0.505)	0.785** (0.092)	1.90
[region] 5	0.784* (0.115)	0.011 (0.781)	0.775** (0.113)	1.47
[region] 6	0.951 (0.146)	0.064 (0.114)	0.921** (0.145)	1.35
[year]	1.170*** (0.054)	0.049 (0.167)		
[CPI]	0.819 (0.182)	0.019 (0.644)		
[employment]	0.938 (0.286)	0.005 (0.870)		
[interest]	0.996 (0.025)	0.029 (0.396)		
[water]	1.232 (0.228)	-0.017 (0.650)		
[light]	0.912 (0.165)	-0.021 (0.556)		
[wages]	1.064 (0.303)	0.065 (0.009)		
N (defaults)	1,357 (646)		1,357 (646)	
LR Chi2 (p-value)	168.09 (0.000)		128.63	
Proportionality test (p-value)	20.95 (0.584)		21.96	
Harrel's C	0.624		0.612	
AIC	8,874.828		8,850.374	
BIC	8,994.728		8,933.783	

Note: ***,**;* significant at the 1%, 5% and 10% level, respectively. The variables [amount] and [loan_fee] have been included with logarithmic transformation in the model.

Figure 3. Comparison of the area under the curve (ROC) among the different penalties taking into account different repayment values

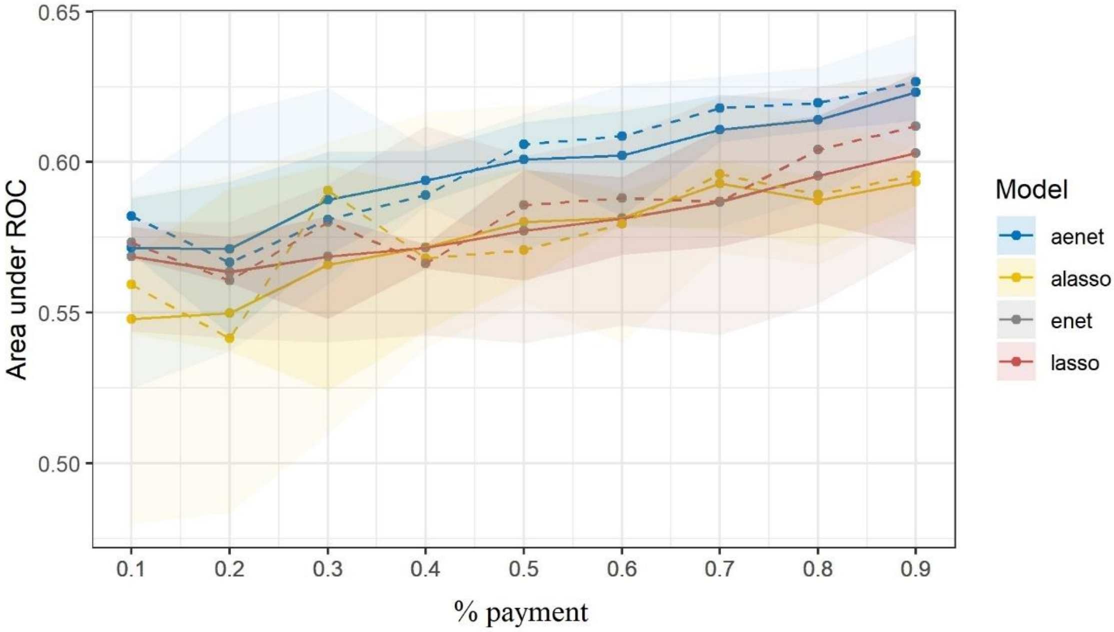


Table 8.COX penalty model (AENET)

Variables	e^{β} (standard error)	VIF
[number_fee]	1.040*** (0.013)	1.11
[loan_fee]	1.000* (0.001)	1.28
[benefit_1]	1.026*** (0.006)	1.21
[benefit_2]	1.003 (0.003)	1.41
[fee_defaulter]	1.101*** (0.017)	1.02
[type_entrepreneur] (2=female)	0.884 (0.072)	1.13
[age_entrepreneur]	0.982*** (0.003)	1.02
[industry] 2	1.865*** (0.346)	2.60
[industry] 3	2.384*** (0.421)	3.16
[industry] 4	2.268*** (0.422)	2.62
[region] 2	0.861 (0.102)	1.73
[region] 3	0.897 (0.131)	1.42
[region] 4	0.776** (0.090)	1.88
[region] 5	0.775* (0.112)	1.46
[region] 6	0.925 (0.144)	1.35
Best alpha	0.05	
Best lambda	0.01765	
N (defaults)	1.357 (646)	
Harrel's C	0.62	
AIC	-92.094	
BIC	-14.256	

Note:(***,**;*) significant at the 1%, 5% and 10% levels, respectively. The variable [loan_fee] has been included with logarithmic transformation in the model.

Figure 4. Area under the curve (ROC), taking into account different repayment values

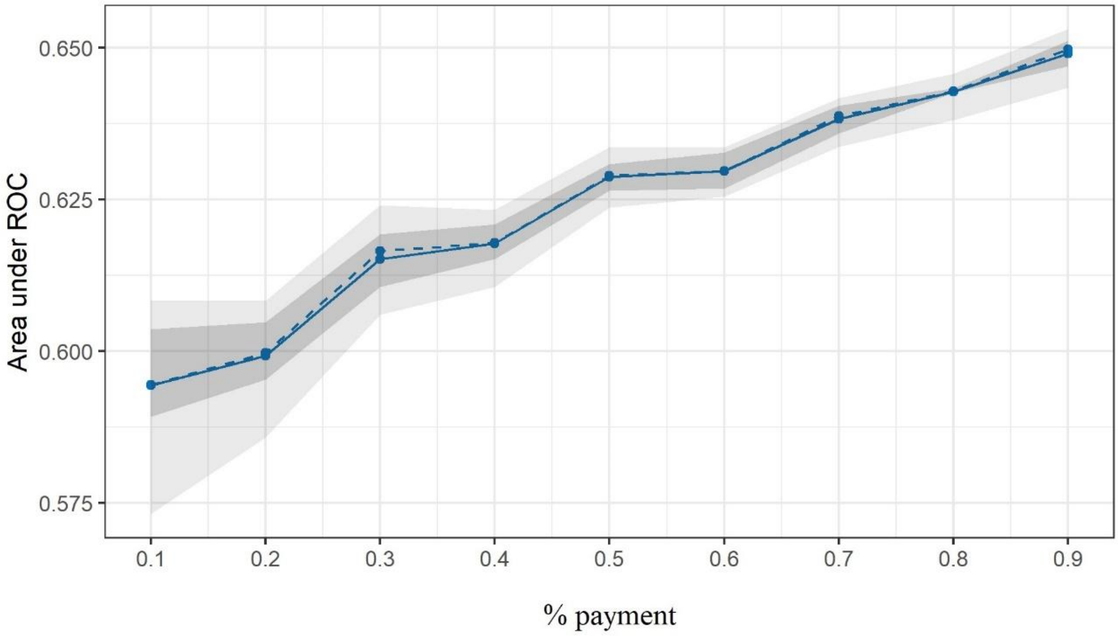
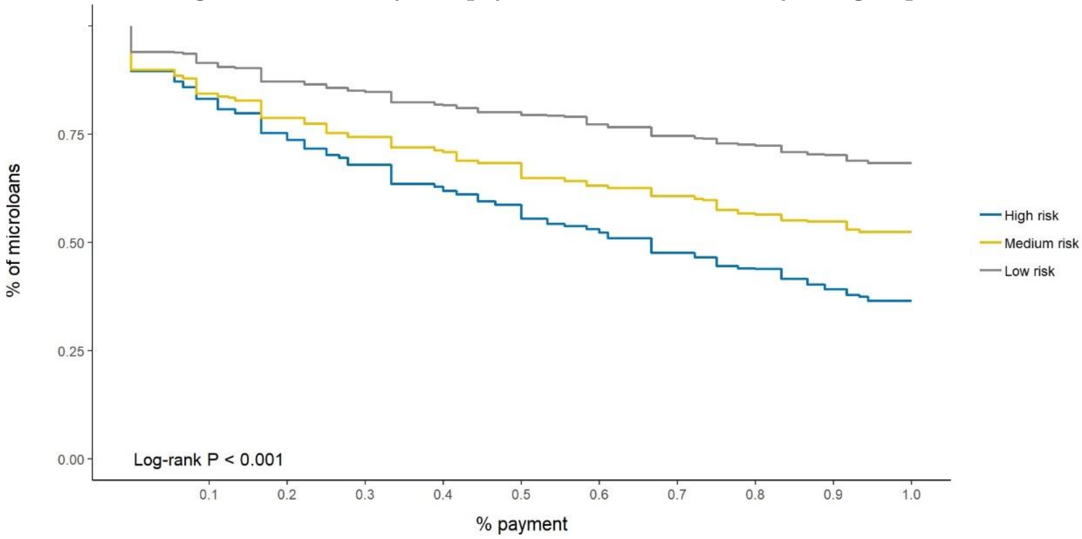


Table 9. Average expected and real repayment percentage by risk group

Risk groups	Average predicted value	Average observed value
High risk	57.06%	55.53%
Medium risk	67.60%	64.82%
Low risk	78.71%	79.42%

Figure 5. Probability of repayment of the microloan by risk group



Number of microloans

Low risk	413	394	384	370	362	349	337	328	317	309
Medium risk	381	356	336	322	309	285	274	256	248	237
High risk	376	340	307	284	265	240	215	199	177	165

Figure 6. Nomogram of the COX model with AENET penalty