The effects of customer segmentation, borrower behaviors and analytical methods on the performance of credit scoring models in the agribusiness sector

Citation for published version:
Lazo, D, Calabrese, R & Bravo, C 2021, 'The effects of customer segmentation, borrower behaviors and analytical methods on the performance of credit scoring models in the agribusiness sector', *The Journal of Credit Risk*, vol. 16, no. 4, pp. 119-156. https://doi.org/10.21314/JCR.2020.272

Digital Object Identifier (DOI):
10.21314/JCR.2020.272

Published In:
The Journal of Credit Risk

Journal of Credit Risk **16**(4), 119–156
DOI: 10.21314/JCR.2020.272

Research Paper

The effects of customer segmentation, borrower behaviors and analytical methods on the performance of credit scoring models in the agribusiness sector

Daniela Lazo,[1] Raffaella Calabrese[2] and Cristián Bravo[3]

[1]Programa de Magíster en Gestión de Operaciones, Facultad de Ingeniería, Universidad de Talca, Camino a Los Niches Km. 1, Curicó, Chile; email: dlazo09@alumnos.utalca.cl

[2]Business School, University of Edinburgh, 29 Buccleuch Place, Room 2.11, Edinburgh EH8 9JS, UK; email: raffaella.calabrese@ed.ac.uk

[3]Department of Statistical and Actuarial Sciences, Western University, Western Science Centre, Room 262, 1151 Richmond Street, London, Ontario N6A 5B7, Canada; email: cbravoro@uwo.ca

(Received May 17, 2019; revised January 23, 2020; accepted August 18, 2020)

ABSTRACT

The main aim of this study is to analyze the joint effects of customer segmentation, borrower characteristics and modeling techniques on the classification accuracy of a scoring model for agribusinesses. To this end, we used data provided by a Chilean company on 161 163 loans from January 2007 to December 2013. We considered random forest, neural network and logistic regression models as analytical methods. Regarding borrowers' profiles, we examined the effects of sociodemographic, repayment behavior, agribusiness-specific and credit-related variables. We also segmented the customers as individuals, small and medium-sized enterprises, and large holdings. As the segments show different risk behaviors, we obtained a better performance when we estimated a scoring model for each segment instead of using a

Corresponding author: D. Lazo

Print ISSN 1744-6619 | Online ISSN 1755-9723

segmentation variable. In terms of the value of each set of variables, behavioral variables increased the predictive capability of the model by double the amount achieved by including agribusiness-related variables. The random forest is the model with the best classification accuracy.

Keywords: agribusiness finance; credit scoring; repayment behavior; random forests; logistic regression.

1 INTRODUCTION

The main focus of this paper is credit risk assessment in the agricultural sector. We use the definition of agriculture provided by the International Standard Industrial Classification of All Economic Activities (ISIC); the definition includes crops and livestock production, forestry, and hunting and fishing (Department of Economic and Social Affairs Statistics Division 2016). Agricultural production is an inherently risky business, the risks of which also affect the lenders who provide financial leverage to the sector. The variability of farm outputs is mainly explained by production risk (Tiedemann and Latacz-Lohmann 2013), which comprises not only financial risks but also damage by pests, diseases and weather effects that can result in borrowers' inability to repay their loans (Hazell 1992). Moreover, agriculture has long production cycles, during which the market prices of agricultural products may diverge rapidly from initial projections (Becerra 2004). Finally, agricultural lending is subject to a relatively higher moral hazard risk, both because farmers have more knowledge about their production risks than their lenders do and because information about borrowers with low incomes, which is common in small farming, is difficult to obtain (Becerra 2004).

Agribusinesses have a pressing need for funding in order to sustain their operations. Given that the production cycle can be a year or more, the necessary working capital and the funding to acquire the supplies to operate within the cycle are usually achieved through loans. The providers of these loans must carefully control their credit risk, yet detailed studies about how to deal with this risk are not common.

Most of the studies on credit risk for primary producers have used financial ratios such as liquidity, profitability and leverage (Jouault and Featherstone 2011; Miller and LaDue 1988; Novak and LaDue 1999; Rambaldi *et al* 1992). Gallagher (2001) reported that the inclusion of nonfinancial agribusiness-related characteristics brought significant improvements to the model. Hou *et al* (2005) included demographic statistics and loan information such as loan size and lending year, providing a higher number of significant variables. Limsombunchai *et al* (2005) defined the lending decision as a function of borrower characteristics, relationship indicators and dummy variables about the agricultural sector and loan information. Aruppillai and Phillip (2014) showed that considering socioeconomic characteristics, such as number of family members, amount of loan disbursement and secondary education, improves the efficiency of the lending decision.

In addition to choosing the variable sets to include, another important aspect of credit risk assessment in the agribusiness sector is segmentation, that is, dividing the clients into groups according to a specific variable or set of variables. In some cases, using several scorecards on different customer segments provides better risk differentiation than using just one scorecard on everyone (Siddiqi 2007). In credit scoring for the agribusiness sector, there are various possible segmentations, eg, current and noncurrent loans (Ziari *et al* 1995), loan size (Miller and LaDue 1988), type of activity or produce (Bandyopadhyay 2008), or loan type (Bandyopadhyay 2008).

We performed a credit risk analysis on a data set provided by a Chilean company that grants credit to farmers. This company is a major distributor of agricultural supplies, machinery and services to support farmers (businesses and people) in Chile. The company's business has an important seasonal component. In effect, income and costs are more concentrated in the second half of each year.

One of the most important services offered by this company is the granting of credit to pay for supplies. The company gives financial alternatives that fit farmers' needs, considering, for example, the seasonality of crops. To manage the credit risk, the company has credit and collection policies that are controlled regularly, but it does not have any automatic model to support the credit risk process.

The company offers installment loans in payment structures equivalent to agricultural cycles, and most of them have terms of less than one year (99%). Around 90% of the loans are insured; however, depending on the credit policies, additional guarantees, such as mortgages or personal guarantees, could be required.

There are a few studies on credit risk research in Chile. Romani *et al* (2002) used different techniques to predict bankruptcy in Chilean companies and found that neural networks performed better than logistic regression and discriminant analysis. Fica *et al* (2018) concluded that a credit scoring model allowed greater flexibility and objectivity in the credit management process in a company dedicated to the production, marketing and distribution of asphalt products in the southern zone of Chile. Madeira (2019) indicated that the default rate of the total consumer loan portfolio of all Chilean banks has a high covariance risk and recommended that banks reduce the default rate of their loan portfolio by choosing customers that suffer fewer shocks during economic downturns.

To the best of our knowledge, our study is the first to analyze the joint effects of modeling techniques, segmentation and borrower characteristics on the performance of scoring models in the agribusiness sector. Specifically, the borrower information available constitutes sociodemographic and repayment behavior data in addition to agribusiness-specific and credit-related variables. We use the data of a company that grants credit and distributes supplies. Funding sources that also serve as input suppliers (with multiple offices close to their customers) have the advantage of being geographically close to customers and having a knowledge of different agricultural specialities (ODEPA 2013).

Because the customers had different sizes of agricultural crops and varied incomes, we could segment them and compare the different types of clients. We also analyzed the effects of the available information on farmers, measuring the contribution of these variables to default prediction. Finally, we examined the main classification techniques used in the industry and in the literature (Thomas *et al* 2017) to understand the value of using a complex, nonlinear technique such as a neural network or a random forest, instead of a simpler technique such as a logistic regression.

Each of these factors was a determinant in building an efficient model. In order of importance, better information was, unsurprisingly, the top factor that can be used to improve predictive models, followed by the choice of model (analytical technique) and the segmentation (size of the company). One caveat relating to the last factor is that holdings require their own model because they are structurally different from both large (nonholding) companies and small farmers.

This paper is organized as follows. First, we review the literature on agricultural credit scoring and explain the main financing sources available for farmers. Second, we describe the data, and we present the main credit scoring methodologies. Third, we show the empirical results. Finally, we draw conclusions, state the limitations of the research and suggest directions for future work.

2 MEASURING CREDIT RISK IN THE AGRIBUSINESS SECTOR

Credit risk is the primary source of risk for retail-oriented financial institutions. Information about past financial performance is the most critical signal that agricultural borrowers can send to distinguish their level of credit risk (Miller *et al* 1993). However, data limitations are a major impediment in assessing farm financial performance (Zhang and Ellinger 2006). Regarding small farmers, their business scale, geographic remoteness, informal accounting practices, and business and financial risks indicate high information needs in order to allow lenders to adequately manage credit risks (Barry and Robison 2001).

Several studies have examined credit risk in agribusiness. A number of these studies used portfolio credit risk management models, seeking to estimate capital requirements for agricultural lenders. Katchova and Barry (2005) developed credit value-at-risk methods to calculate probability of default, loss given default, and expected and unexpected losses. Featherstone *et al* (2006) used credit scoring techniques to rate a portfolio of loans. Sherrick *et al* (2000) and Dressler and Tauer (2016) developed credit risk valuation models for measuring credit risk to estimate expected and unexpected losses. Other studies have assessed the credit risks of individual loans through credit scoring models (Hou *et al* 2005; Miller and LaDue 1988; Novak and LaDue 1999; Turvey 1991). However, the literature on credit scoring is very limited compared with that on portfolio analysis (Thomas *et al* 2017).

With regard to credit models that have been used for assessing the agricultural sector, those included are logistic regression (Durguner and Katchova 2007; Hou *et al* 2005; Limsombunchai *et al* 2005; Miller and LaDue 1988; Novak and LaDue 1999; Rambaldi *et al* 1992; Römer *et al* 2017; Savitha *et al* 2016), discriminant analysis (Bonazzi and Iotti 2014; Rambaldi *et al* 1992; Ziari *et al* 1995) and machine learning techniques such as decision trees and neural networks (Limsombunchai *et al* 2005; Novak and LaDue 1999).

Logistic regression is a classic, and the most widely used, technique due to its simplicity and explanatory power (Siddiqi 2017). Ziari *et al* (1995) found that both mathematical programming techniques and statistical models performed equally well and that mixed integer-programming models perform better than parametric models. An advantage of nonparametric models is that they can fit several distribution functions. Further, when the data sample is small or if it is too dirty, nonparametric models such as neural networks may generate better results (Gustafson *et al* 2005).

Logistic regression is the technique most frequently applied in agricultural credit scoring (see Table 1), with isolated studies showing a comparison of similar general linear classification techniques. Turvey (1991) used data from Canada's Farm Credit Corporation to compare the performance of four credit scoring models (linear probability model, discriminate analysis, logit and probit) and found similar classification accuracies (between 71.5% and 67.1%) for these models. Nonlinear techniques have also been benchmarked: Odeh *et al* (2006) compared logistic regression, artificial neural networks and the adaptive neuro-fuzzy inference (ANFI) system to predict default using data from the Farm Credit System in the United States, identifying slight differences in prediction accuracies. ANFI gave better results than the other methods, particularly in terms of sensitivity and specificity measures.

The types of variables used in the literature on credit scoring for farmers mainly describe financial ratios such as liquidity, profitability and leverage (Durguner and Katchova 2007; Jouault and Featherstone 2011; Ziari *et al* 1995); farmer characteristics (educational level, age, goods, etc; Limsombunchai *et al* (2005)); farm characteristics, such as types of crops and farm size (Limsombunchai *et al* 2005; Miller and LaDue 1988; Novak and LaDue 1999; Onyenucheya and Ukoha 2007); and credit features, including credit history (Aruppillai and Phillip 2014; Eyo and Ofem 2014; Hou *et al* 2005; Jouault and Featherstone 2011). Other studies have used weather data (Pelka *et al* 2015; Römer *et al* 2017) and variables related to the sustainability of crops (Henning and Jordaan 2016). No studies have measured the relative impact of these sets of variables; each study apparently used what was available to them.

Turvey (1991) stressed the importance of including qualitative and quantitative attributes in credit scoring models. Gallagher (2001) indicated that a prediction model without nonfinancial variables could have model misspecification issues. Zech and Pederson (2003) identified the debt-to-asset ratio as a major predictor of repayment ability. Zech and Pederson (2003) also argued that both the total asset turnover ratio and family living expenses are strong predictors of the financial performance of a farm. Further, it is a well-known fact that better sources of information are more useful than better models when it comes to making predictions. This is discussed at length in Baesens *et al* (2016) and shown empirically for newer so-called alternative data in a peer-to-peer and retail credit risk environment by Calabrese *et al* (2019) and Óskarsdóttir *et al* (2019).

In relation to the definition of default, the literature on credit scoring models for agribusiness takes different approaches. Jouault and Featherstone (2011) used the definition of ninety days past due, in concordance with the Basel Accords (Basel Committee on Banking Supervision 2004). Miller and LaDue (1988) defined default as whenever a loan was refinanced. An alternative to traditional credit scoring is to use the coverage ratio directly as a measure of creditworthiness (Novak and LaDue 1994).

Regarding the purpose of the models, there are two categories: application scoring and behavioral scoring. The former relates to the decision on whether to grant the loan, and the latter is about the decision on the credit limit or new product offers (when the credit is already granted). Most of the literature on credit scoring for agribusiness is related to application scoring. Miller and LaDue (1988) evaluate existing borrowers using only financial ratios; their analysis did not use behavioral variables.

Table 1 presents a summary of previous work on lending in agribusiness, in terms of model types, variables used and the country in which the study was conducted. There are only a few analyses of all factors affecting the failure of farmers to repay, as most studies focus on the analysis of different types of models or variables, without considering the impact of the factors simultaneously. Limsombunchai *et al* (2005), Eyo and Ofem (2014) and Savitha *et al* (2016) analyze two different models and types of variables but do not take the size of the company and behavioral variables into account. This paper presents a simultaneous analysis of the impact of the creation of specialized variables (agribusiness and repayment behavior), the type of classification techniques and company size. We perform this analysis to determine the most important factors when predicting the default of farmer debt and to make recommendations to agricultural lenders in relation to credit risk.

TABLE 1 Credit scoring models for farmers.

Author (year)	Models	Variables	Country
Miller and LaDue (1988)	LR	Farm size, liquidity, solvency, profitability, capital efficiency, operating efficiency	USA
Rambaldi *et al* (1992)	DA, LR	Liquidity, debt utilization, profitability, assets, operational efficiency	USA
Ziari *et al* (1994)	DA, FLDA, LDA	Financial ratios	Canada
Novak *et al* (1999)	RPA, LR	Debt-to-asset ratio, current ratio	USA
Hou *et al*(2005)	LR	Demographic statistics, business and loan information	USA
Limsombunchai *et al* (2005)	LR, ANN	Borrower characteristics, credit risk proxies, relationship indicators	Thailand
Durguner and Katchova (2007)	LR	Financial ratios	USA
Onyenucheya and Ukoha (2007)	RM, DA	Farmer characteristics, credit features, ratios, distance (home loan source)	Nigeria
Jouault and Featherstone (2011)	LR	Ratios, credit information	France
Eyo and Ofem (2014)	DA, RM	Borrower features, loan information, financial ratios, farm size	Nigeria
Aruppillai and Phillip (2014)	RM	Borrower features, loan information	Sri Lanka
Bonazzi and Iotti (2014)	MDLA	Financial ratios	Italy
Savitha *et al* (2016)	LR, MLR	Borrower characteristics (both financial and nonfinancial) and relationship indicators	India
Römer (2017)	LR	Socioeconomic characteristics of clients, financial ratios, credit related, working experience	Madagascar

The models applied were the following: logistic regression (LR), multinominal logistic regression (MLR), discriminant analysis (DA), variations of discriminant analysis (MDLA, LDA and FLDA), a recursive partitioning algorithm (RPA) equivalent to decision trees, and regression models (RM).

3 FINANCING FARMERS IN DEVELOPING COUNTRIES

According to Klein *et al* (2001), the types of rural lenders found in developing countries are the following.

- Formal lenders: banks, agricultural development agencies, rural branches of commercial banks, cooperative banks and rural banks/community banks.
- Semiformal lenders: credit unions, other cooperatives, semiformal local or community banks and nongovernmental organizations (NGOs).
- Informal lenders: relatives and friends, independent moneylenders, rotating savings and credit associations.
- Credit interconnected systems: suppliers of agricultural inputs/crop buyers, and agroindustries.

The sources of formal financing, such as commercial banks, have a strong aversion to lending to small farmers because of the characteristics of this sector, ie, relatively higher and complex risk profiles (ODEPA 2009). Other sources of funding, particularly interconnected systems (suppliers of agricultural inputs/crop buyers), "have an advantage in relation to customer closeness and knowledge of different fields, attributes that are valued beyond the rate interest charge" (ODEPA 2013).

In the source country of our data, 17.9% of farmers use some form of credit to finance their business (EME 2014). Table 2 shows the sources of financing used by these farmers (ODEPA 2013). Most of the farmers chose bank credits (84.4%), with the second most important source of financing corresponding to suppliers of agricultural inputs (11.6%).

Using data from farmers seeking loans in credit interconnected systems can permit the determination of relevant factors in this segment, with reference to their repayment behavior. This is due to the knowledge of the agricultural area and the proximity of these institutions to their customers.

4 DATA

This section describes the data set used in this analysis. In particular, we provide details on the data preparation, and we present the variables used in the scorecard. Moreover, we explain the transformations applied to the data.

4.1 Data preparation

We used data provided by a Chilean company that grants loans to farmers for the supply of inputs and provides support services. The data was anonymized to protect

TABLE 2 Number of farmer loans in Chile (original data from ODEPA (2013)).

Source	Amount (US$ million)	Share (%)
INDAP (Agricultural Development Institute)	69.81	1.1
Input suppliers	711.16	11.6
Agriculture contract	68.90	1.1
Commodity exchange	53.87	0.9
Foreign investment	39.51	0.6
Credit unions	11.35	0.2
Factoring	4.17	0.1
Subtotal	958.77	15.6
Banks	5192.60	84.4
Total	6151.37	100.0

customer confidentiality and identity. The data relates to 6658 customers who were approved between January 2007 and December 2013. The data includes a subset of customers' application characteristics and full subsequent repayment behavior up to December 2014. We considered a sample of 161 613 credit sales, splitting the data set into three segments. The person (independent farmers) segment has 48 875 cases; the company segment has 58 443 cases; and the holding segment has 54 295 cases. The default percentage across the sample is 2.55%, and the rates by segment are 2.56%, 2.48% and 2.64%, for people, companies and holdings, respectively.

The data time period reflects an entire economic cycle, including the end of an economic expansion, a recession and a recovery; thus, given that our objective is to study the effects of our factors on modeling agricultural loans, we consider the data both sufficient to cover the application of these technologies under most conditions and robust to changes in the economic conditions. We can possibly extrapolate this to multiple countries, as Chile is an upper-middle-income country with very large holding corporations (represented in the holdings data set) that are more competitive than many companies from high-income countries (Schwab 2017). We also study small farmers whose reality is much closer to that of a low-to-middle-income country; in particular, we look at farmers within the supply chains of large companies (Reardon *et al* 2009). The studies of small and medium-sized (SME) agribusinesses lie somewhere in the middle: they are much more representative of the Chilean upper-middle-income reality, given that they are much more dependent on the local economy than large holding companies. Thus, we believe our data set and segmentation create an interesting profile of the use of models for risk management in agribusiness and represent different conditions and realities worldwide.

Best practice, according to the literature (Siddiqi 2007), is to consider default as occurring when one payment is more than ninety days in arrears during the first twelve months after granting the loan. We use the same definition for this study. The ninety-days definition of the target variable corresponds to the definition of a good/bad borrower within the Basel Accord (Basel Committee on Banking Supervision 2004): this considers an obligor "bad" if the bank determines that the obligor is unlikely to pay its credit obligations, or if any material credit obligation is past due by more than ninety days. The definition of default can be applied at the level of a loan (a particular facility) for retail exposures, that is, a default by a borrower on one loan does not imply that all other loans are in default (Basel Committee on Banking Supervision 2004). In this sense, the definition is applied at loan level because most of the company's loans can be classified as retail exposures, especially the loans granted to people and small companies.

Given that some borrowers have a history with the company, we also need to study past behavior during a set period of time. This requires setting up a performance window during which each loan is studied: a period that, again, is usually considered to be from six to twelve months. Considering the periodicity of crops, a twelve-month performance window, by capturing an entire period, gives the best chance of capturing the borrowers' behavior.

We do not add macroeconomic variables to this study because the idea of the model was first to consider the standard approach of estimating scores with no macroeconomic variables in an unconditional model, and then to calibrate this model over macro variables to meet the provisioning and capital requirements in the International Financial Reporting Standard 9's expected credit loss framework. In this framework, the probability of default can be obtained by using internal historical data adjusted by forward-looking information and according to different possible macroeconomic scenarios.

In terms of data preprocessing, we removed variables that had low variability (if more than 95% of the observations showed the same category) and more than 30% of their values missing.

The variables selected for this study fall into the following categories.

- Sociodemographic variables: the region of the borrower's residence; the economic sector in which the farmer operates, according to the company's internal classification (agricultural and others); the level of purchases made during the last year; and the type of client (person, company or holding company).

- Agribusiness variables: the reported income of the borrower, the cost of operation, the types of crops (cherry, plum, corn, apple, walnut, meadows, wheat, wine grapes and others) and information about the customer's properties (related to location, plantation area and number of properties).

- Credit variables: the attributes of the loan and the history of the customer in the company (eg, the client's length of tenure, the branch office region of the credit application, the installment and loan amount, and the payment type or term type according to payment frequency of the loan).

- Behavioral variables related to payment behavior, which can be divided into three time windows: the last three months, the period of the last three to six months and the period of the last six to nine months. As the values of behavioral variables change over the performance window, we computed the maximum, the minimum, the average and the number of increments and decrements in the standing balance and various ratios, such as amounts of arrears and days in arrears.

In total, the data set is composed of five sociodemographic variables, seventeen agribusiness-related variables, nineteen credit variables and forty-two behavioral variables.

4.2 Variable selection and transformation

The variable selection process was developed in two stages. To test the independence of the explanatory variables with the target variable, we used the χ^2 test for categorical variables and the Kolmogorov–Smirnov test for continuous variables. We removed the variables that did not show a relationship with the target variable at a 95% confidence level. Afterwards, we created clusters of the independent variables in order to reduce the dimensionality of the data set using the `ClustOfVar` algorithm (Brida *et al* 2014). This algorithm applies K-means clustering to categorical and continuous variables using a synthetic variable calculated by principal component analysis as a center (Kiers 1991).

We also performed a multicollinearity analysis by removing variables with a variance inflation factor higher than five (Mansfield and Helms 1982). Finally, we used a stepwise selection procedure, and we removed the variables that had a significance level higher than 0.05 in each iteration. We finally obtained thirty variables for the whole sample, thirty-three variables for the data on individuals, thirty-two variables for the companies and twenty-nine variables for holding companies.

To normalize the data set and center it using a common scale, we applied the weight of evidence (WOE) transformation on the variables, computed as follows:

$$\mathrm{WOE}_{c_v,v} = \ln\left(\frac{\mathrm{DistrGood}_{c_v,v}}{\mathrm{DistrBad}_{c_v,v}}\right), \tag{4.1}$$

where v is the index of the variables that are available, and c_v is the index of each variable's categories. $\mathrm{DistrGood}_{c_v,v}$ and $\mathrm{DistrBad}_{c_v,v}$ are the proportions of cases of the attribute that belong to the good and bad classes, respectively, across the total cases of the class. We used this transformation because it is a common procedure in credit scoring models (Siddiqi 2007). To apply this transformation to the continuous variables, we discretized them using classification trees. For the categorical variables, we aggregated categories in order to have at least 5% of the total cases in each category.

The resulting data set is free of outliers, centered and discretized to better capture behavior. We now proceed with the experimental design to test our hypotheses.

5 EXPERIMENTAL DESIGN

The experimental design of this study consists of a factorial experimental setup to assess the effects of three different factors on the performance of default prediction for farmers. The first factor represents the type of explanatory variables and consists of four possible levels, given by the credit variables, the behavioral variables, the sociodemographic variables and the agribusiness variables. This factor both reflects the amount of information that a company must store and supports the complexity analysis, since more complex patterns require more data; it also allows us to study the diversity of these patterns. If more data sources are needed, it suggests that a mix of different risks affects the ability of borrowers to satisfy their obligations.

Formally, let $\boldsymbol{x}$ be the set of all the independent variables, x_{ag} the subset of agribusiness variables, x_{sd} the subset of sociodemographic variables, x_{ap} the subset of credit variables and x_{bh} the subset of behavioral variables. We estimate the probability of default $P(y = 1 \mid \boldsymbol{x})$ as a function of the four subsets of variables:

$$P(y = 1 \mid \boldsymbol{x}) = f(x_{\mathrm{ag}}, x_{\mathrm{ap}}, x_{\mathrm{sd}}, x_{\mathrm{bh}}). \tag{5.1}$$

The second factor of the experimental design concerns the classification techniques. It has three possible levels, given by logistic regression, random forest and neural network analysis. The main question to be answered by this factor is the relevance of complex, nonlinear patterns in the behavior of the borrowers. If a more complex model results in a much higher discrimination capacity, then we can conclude that there is a much more complex structure among the borrowers' behavior, which impacts the ability of small lenders to model risk effectively.

The first model selected is given by logistic regression, a widely used approach in credit risk analysis (Baesens *et al* 2003b). This model is the basic generalized linear model and can correctly represent the relationship between linear combinations of variables in the sample and the logit or the logarithm of the odds that a borrower presents the event being studied (Hosmer and Lemeshow 2000). Formally, the logistic regression models the probability of the event as

$$f(x) = \frac{1}{1 + \exp(-\beta^t \cdot x)}, \tag{5.2}$$

where $x = (x_{\text{ag}}, x_{\text{ap}}, x_{\text{sd}}, x_{\text{bh}})$, the vector of the variables in the model, and β is the vector of weights that each variable has in the model.

We also use a neural network, a powerful but difficult to interpret nonlinear model (Hassoun 1995) that can capture a more complex structure in a single expression. We use a shallow model representation, which is effective when looking at general nonlinear patterns in the data. We use one hidden layer as well as sigmoid transfer and output functions in the architecture. The number of neurons in the hidden layer is obtained by maximizing the area under the receiver operating characteristic (ROC) curve (AUC) in a validation set.

The last approach is the random forest method, which is a robust alternative for predicting default due to its ability to detect complex patterns following a deep analysis of all the subsets of the input space (Breiman 2001). The random forest approach combines decision trees so that they all use a separate sample of cases and variables simultaneously. This produces diverse trees that create, when evaluated jointly, a very detailed analysis of the input space (deep search). A bootstrapped sample of the data, usually of size 64.2%, for each tree as well as a sample of the variables, usually $1/3$, for each split within that tree is selected to train it. Assuming that each tree produces a binary output given by o_i, we can generate a valid output by simply averaging each individual tree. As shown in Probst and Boulesteix (2018), the best strategy for selecting the number of trees is to simply train as many as possible. We chose a number of trees such that no improvement occurred in the out-of-bag sample when adding a new one.

In previous studies, these three methods have been identified as the most accurate for building credit scorecards for each level of complexity (Lessmann *et al* 2013). The chosen models are from very different areas of the interpretability/complexity spectrum. A logistic regression will only account for linear relationships between variables; however, it will provide a very clear picture of the way the variables have an effect on the target, in terms of both the magnitude of the impact and its direction, that is, if a larger value of a given variable implies an increase or decrease in borrower risk. A random forest is exactly at the other extreme, because the only information that can be extracted is the contribution of each variable. This is done by comparing metrics (usually AUC or accuracy) between trees that include a certain variable and those that do not. Neural networks lie somewhere in between because they are a model represented by a unique function – as opposed to random forests, which are ensembles of decision trees – and it is possible to extract rules from their output (Baesens *et al* 2003a); otherwise, they are black box models. These three models give a very broad picture of the technological abilities that are currently available to extract patterns for structured data, and they allow us to profile the usefulness of complex models versus simpler solutions.

The third factor represents the type of clients over three possible levels, given by company, holding company or person. Here, a person is a customer who applies for credit individually and is not associated with or does not belong to any company. The remaining categories – enterprises and holding companies – are clients who represent a company or a business organization that controls a number of companies. This factor illuminates not only the differences that arise from multiple organizational structures but also how their composition, from single farmers operating on their own to large holdings, affects the lender's ability to capture credit risk through a statistical procedure. If each segment is completely different, then more scorecards, and therefore more independent systems, need to be kept in parallel to serve the customers. This, again, has managerial implications because the lender has a more complex risk area, thereby increasing the cost of sustaining proper operations.

For each possible combination, we estimated a scoring model and computed the AUC, a common index reported in the literature for analyzing predictive accuracy (Lobo *et al* 2008) and for comparing the predictive capabilities of the model. In the next section, we consider all the possible combinations given by a total of 135 models ($15 \times 3 \times 3$).

6 EMPIRICAL RESULTS

This section presents the main results of the study in relation to the analysis of the impact of various factors – explanatory variables, modeling techniques and segmentation – in default prediction in the agribusiness sector.

6.1 Explanatory variables

After applying the WOE transformation, we analyzed the ability of the explanatory variables to predict the good and bad cases. We determined whether the relationship of the independent variable coheres with expectations. For each variable, the information value (IV) was computed, a measure that comes from information theory (Kullback 1997) and that reveals the predictive power of the attributes. According to Siddiqi (2007), a variable is highly predictive if its information value is greater than 0.3. The results for all the customers are presented in Table 3. We denote by "sd", "ap", "ag" and "bh" the sociodemographic variables, the credit variables, the agribusiness variables and the behavioral variables, respectively. We also report the strength of the relationship between each explanatory variable and the dependent variable in Table 3 following Siddiqi (2007). The variables ArrearsLast3M, Arrears3to6Months, TimelyPayLast3M, CropTypeG2 and TimelyInstLast3M, belonging to the behavioral and agribusiness groups, show the higher information values. These results show that the behavioral variables represent the strongest predictors of capacity to repay, as is the case in consumer lending. The signal given by the most recent payment behavior (previous three to six months) is of greater relevance within this subset. A more important variable for this segment, highly ranked in the sample and with very strong explanatory power, is the type of crop. This indicates that the seasonality of crops will be a very strong indicator of future performance, but at the same time the inclusion of this variable brings the risk that the model can be affected by an external impact on the crops (eg, a particular climate event); thus, the predictive capability of the model might be affected. Usually, a recalibration using more recent data is all that is needed to recover from this circumstance, so this risk should not discourage a potential user from including the variable.

6.2 Predictive accuracy

Because the data sets were imbalanced with respect to the classes of the target variable, we applied the synthetic minority oversampling technique (SMOTE), a method that combines oversampling and undersampling to generate balanced data sets (Chawla *et al* 2002). To avoid overfitting, we estimated the models and validated them both on an out-of-sample set, generated by randomly drawing 30% of the customers, and on an out-of-time sample, given by the credit sales after January 2014.

Both neural networks and random forest analysis require tuning certain parameters in order to find the choices that better represent the patterns in the data. Neural networks require the number of neurons in the hidden layer and the number of training epochs, while random forests require the maximum depth per tree and the number of variables per tree. These parameters are adjusted by grid search, finding the optimal parameter for each of the 135 models using 20% of the training sample.

To measure the predictive accuracy, we used the AUC. This curve corresponds to the plotted values of the probability of true positives (correctly predicted defaults) and the probability of false positives (incorrectly predicted good loans), illustrating a trade-off between the captured response fraction and the false positive fraction. Each point on the ROC chart corresponds to a specific fraction of cases, ranked by their

TABLE 3 Information values (IVs) and the strength of the relationship for each explanatory variable on all customer segments. [Table continues on next page.]

Variable	Description	Group*	IV	Strength
ArrearsLast3M	Avg. days in arrears in last 3 months	bh	0.56	Strong
Arrears3to6Months	Avg. days in arrears in last 3–6 months	bh	0.44	Strong
TimelyPayLast3M	Avg. amount paid on time and total paid in last 3 months	bh	0.30	Strong
CropTypeG2	Crop type	ag	0.29	Strong
TimelyInstLast3M	Avg. ratio between payment and installment in last 3 months	bh	0.29	Strong
TotalBalance	Total amount owed	ap	0.24	Strong
TimelyPay3Mto6M	Avg. of ratio of amount paid on time and total paid in last 3–6 months	bh	0.21	Strong
TimelyPay6Mto9M	Avg. of ratio of amount paid on time and total paid in last 6–9 months	bh	0.21	Strong
TimelyInstLast3M	Avg. ratio between payment and installment in last 3 months	bh	0.20	Strong
NrTimelyLast3M	No. of installments paid on time in last 6–9 months	bh	0.20	Medium
RegionG1	Geographic region	sd	0.18	Medium
Cost	Agricultural investment	ag	0.17	Medium
LevelPurchases	Purchases level	sd	0.13	Medium
IncomeHectare	Income per hectare	ag	0.13	Medium
CostProperty	Ratio between agricultural investment and no. of properties of the customer	ag	0.13	Medium

TABLE 3 Continued.

Variable	Description	Group*	IV	Strength
AmountArrears3Mto6M	Avg. arrears amount in last 6–9 months	bh	0.12	Medium
CropsNumber	No. of different crop types	ag	0.12	Medium
Income	Agricultural activity income	ag	0.11	Medium
AmountArrears6Mto9M	Avg. arrears amount in last 6–9 months	bh	0.11	Medium
ArrearsIncreaseLast3M	No. of increases of the arrears amount in last 3 months	bh	0.10	Medium
CostHectare	Agricultural activity cost per hectare	ag	0.10	Medium
NrPastDueLast3M	No. of past due installments in last 3–6 months	bh	0.10	Medium
ArrearsIncrease3Mto6M	No. of increases of the arrears amount in last 3-6 months	bh	0.09	Weak
Tenure	If the credit applicant is a client	ap	0.09	Weak
PropertyLocationN	No. of different property locations	ag	0.08	Weak
PropertyDistance	Avg. distance between each property and its nearest branch office	ag	0.07	Weak
PreviousPurchasesN	No. of previous purchases	ap	0.06	Weak
NrPastDue3Mto6M	No. of past due installments in last 6–9 months	bh	0.04	Weak
ArrearsDecrease3Mto6M	No. of decreases in arrears amount in last 3–6 months	bh	0.02	Weak
ArrearsDecrease6Mto9M	No. of decreases in arrears amount in last 6–9 months	bh	0.02	Weak
TimeLastMaturity	Months since the most recent maturity	ap	0.01	Unpredictive

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

predicted value. The AUC is the probability that a randomly chosen positive case is correctly rated; it comes with greater suspicion than a randomly chosen negative case (Hanley and McNeil 1982).

The AUCs calculated for all combinations are reported in Tables 4, 5, 6 and 7. If we constrain each model to include only one type of variable, then behavioral variables, followed by agribusiness-related characteristics, give the best performance. In most of the combinations, the highest accuracy is achieved using information from different groups of variables. For example, Table 4 shows that the higher AUC for all customers for each modeling technique is achieved by using all of the explanatory variables on the out-of-sample test set.

To measure the contribution of each group of variables in terms of performance, we computed the normalized AUC, dividing the AUC by the maximum out-of-time AUC for each segment. We show the results in Tables 8, 9, 10 and 11. In general, behavioral variables increase the AUCs from 5% to 20%, whereas agribusiness variables contribute from 5% to 10% in extra predictive capability. In particular, behavioral variables show the highest impact on the AUC for all the customers in a logistic regression model. We obtained similar results for all the other segments of customers.

Applying segmentation to the customers can increase the AUC by up to 2.7% on the out-of-sample data. Conversely, the accuracy decreases by 2.5% if the segmentation is implemented on the out-of-time sample, indicating that using a one-size-fits-all model can deliver a more stable result. Tables 8, 9, 10 and 11 show that the best model with all the available variable types is the random forest approach, followed by the logistic regression approach. The neural network model shows the worst performance on the out-of-time sample for all the customers, as is also displayed in Figure 1.

In order to check the prediction stability for each of the applied techniques, using the models that consider all the types of variables, we plotted the predicted default rate of the models versus the real one. To check how the crop periodicity influences the outcome, we used an out-of-time sample of one year. Based on a profitability criterion, namely the expected maximum profit measure (Verbraken *et al* 2014), we obtained the optimal cutoff point for each model.

The results can be seen in Figure 2. In general, the three applied techniques were able to capture the default rate periodicity. The random forest technique gave the best performance during all periods, with a default rate that was closest to the real default rate.

To sum up, logistic regression performs well in predictive accuracy compared with machine learning techniques (random forests and neural networks). While neural networks demonstrate good performance out-of-sample, they produce unstable results in out-of-time samples. Random forests perform significantly better for out-of-time

TABLE 4 AUC indexes for all borrowers.

Variables*	Logistic regression		Neural networks		Random forests	
	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time
sd	0.650	0.595	0.648	0.609	0.630	0.589
ag	0.675	0.663	0.767	0.668	0.717	0.633
ap	0.692	0.686	0.707	0.689	0.695	0.689
bh	0.736	0.806	0.819	0.761	0.762	0.797
sd + ag	0.706	0.675	0.820	0.693	0.818	0.720
sd + ap	0.714	0.681	0.746	0.711	0.726	0.702
sd + bh	0.756	0.802	0.828	0.783	0.842	0.830
ag + ap	0.733	0.720	0.823	0.727	0.820	0.743
ag + bh	0.779	0.816	0.854	0.769	0.882	0.846
ap + bh	0.779	0.821	0.848	0.773	0.840	0.830
sd + ag + ap	0.745	0.716	0.847	0.720	0.870	0.774
sd + ag + bh	0.786	0.813	0.869	0.774	0.902	0.865
sd + ap + bh	0.785	0.816	0.851	**0.784**	0.873	0.844
ag + ap + bh	0.800	**0.828**	0.861	0.762	0.899	0.871
sd + ag + ap + bh	**0.803**	0.824	**0.871**	0.783	**0.917**	**0.879**

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 5 AUC indexes for the subset of people.

Variables*	Logistic regression		Neural networks		Random forests	
	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time
sd	0.556	0.640	0.662	0.559	0.643	0.635
ag	0.740	0.731	0.806	0.673	0.812	0.694
ap	0.792	0.738	0.783	0.683	0.800	0.671
bh	0.736	0.774	0.821	0.776	0.759	0.796
sd + ag	0.741	0.737	0.861	0.735	0.886	0.803
sd + ap	0.798	0.760	0.820	0.694	0.834	0.748
sd + bh	0.734	0.784	0.855	0.771	0.836	0.735
ag + ap	0.819	0.770	0.832	0.666	0.909	0.796
ag + bh	0.807	0.820	0.863	0.779	0.896	0.785
ap + bh	0.820	0.795	0.880	**0.837**	0.898	0.869
sd + ag + ap	0.820	0.773	0.836	0.705	0.924	0.788
sd + ag + bh	0.806	0.819	0.905	0.770	0.919	0.774
sd + ap + bh	0.820	0.799	**0.905**	0.799	0.907	**0.854**
ag + ap + bh	0.845	0.822	0.897	0.775	0.933	0.844
sd + ag + ap + bh	**0.845**	**0.826**	0.898	0.813	**0.938**	0.833

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 6 AUC indexes for the subset of companies.

Variables*	Logistic regression		Neural networks		Random forests	
	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time
sd	0.681	0.571	0.697	0.569	0.680	0.572
ag	0.699	0.646	0.819	0.746	0.791	0.734
ap	0.737	0.636	0.785	0.638	0.762	0.622
bh	0.733	0.798	0.816	0.745	0.788	0.793
sd + ag	0.726	0.653	0.881	0.805	0.884	0.796
sd + ap	0.755	0.637	0.834	0.646	0.827	0.675
sd + bh	0.771	0.782	0.876	0.758	0.875	0.812
ag + ap	0.769	0.713	0.891	0.708	0.891	0.769
ag + bh	0.789	0.805	0.902	0.776	0.907	0.825
ap + bh	0.810	0.799	0.872	0.732	0.883	0.821
sd + ag + ap	0.777	0.708	0.886	0.682	0.917	0.840
sd + ag + bh	0.798	0.807	0.896	0.772	0.933	0.848
sd + ap + bh	0.813	0.795	0.867	0.722	0.907	0.830
ag + ap + bh	0.825	**0.822**	0.913	0.756	0.931	0.853
sd + ag + ap + bh	**0.826**	0.821	**0.925**	**0.830**	**0.939**	**0.870**

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 7 AUC indexes for the subset of holding companies.

Variables*	Logistic regression		Neural networks		Random forests	
	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time	Out-of-sample	Out-of-time
sd	0.654	0.574	0.656	0.575	0.654	0.574
ag	0.749	0.616	0.835	0.711	0.809	0.687
ap	0.705	0.599	0.724	0.581	0.760	0.598
bh	0.778	0.807	0.847	0.771	0.801	0.813
sd + ag	0.770	0.628	0.900	0.767	0.874	0.748
sd + ap	0.738	0.625	0.834	0.712	0.853	0.708
sd + bh	0.789	0.802	0.883	0.762	0.847	0.802
ag + ap	0.778	0.642	0.907	0.722	0.920	0.789
ag + bh	0.826	0.784	**0.911**	**0.788**	0.907	0.818
ap + bh	0.810	**0.816**	0.883	0.714	0.889	0.816
sd + ag + ap	0.792	0.655	0.880	0.677	0.928	0.799
sd + ag + bh	0.832	0.782	0.875	0.717	0.929	0.825
sd + ap + bh	0.817	0.810	0.905	0.736	0.922	0.828
ag + ap + bh	0.842	0.802	0.852	0.689	0.937	0.852
sd + ag + ap + bh	**0.847**	0.798	0.858	0.634	**0.944**	**0.863**

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

FIGURE 1 The ROC curves of all customers on an out-of-time sample for different classification techniques.

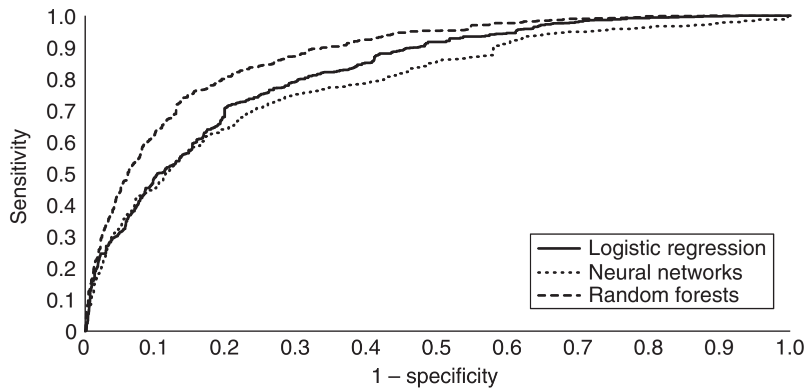


TABLE 8 Normalized AUC indexes, all customers, out-of-time sample.

Variables*	Logistic regression	Neural networks	Random forests
sd	0.677	0.693	0.670
ag	0.754	0.760	0.720
ap	0.781	0.784	0.783
bh	0.917	0.866	0.906
sd + ag	0.767	0.788	0.819
sd + ap	0.775	0.809	0.799
sd + bh	0.912	0.890	0.944
ag + ap	0.819	0.827	0.846
ag + bh	0.928	0.875	0.963
ap + bh	0.935	0.879	0.945
sd + ag + ap	0.815	0.819	0.881
sd + ag + bh	0.925	0.881	0.984
sd + ap + bh	0.928	0.892	0.961
ag + ap + bh	0.942	0.867	0.991
sd + ag + ap + bh	0.937	0.891	1.000

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

samples; this can be explained by the fact that random forests use multiple decision trees and different samples and variables to generate robust results and avoid overfitting.

TABLE 9 Normalized AUC index, people, out-of-time sample.

Variables*	Logistic regression	Neural networks	Random forests
sd	0.736	0.643	0.730
ag	0.841	0.774	0.798
ap	0.849	0.786	0.772
bh	0.890	0.893	0.916
sd + ag	0.848	0.846	0.923
sd + ap	0.875	0.799	0.860
sd + bh	0.902	0.887	0.845
ag + ap	0.885	0.766	0.916
ag + bh	0.943	0.896	0.903
ap + bh	0.914	0.963	1.000
sd + ag + ap	0.889	0.811	0.907
sd + ag + bh	0.942	0.886	0.891
sd + ap + bh	0.919	0.919	0.982
ag + ap + bh	0.945	0.892	0.971
sd + ag + ap + bh	0.951	0.935	0.958

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 10 Normalized AUC indexes, companies, out-of-time sample.

Variables*	Logistic regression	Neural networks	Random forests
sd	0.656	0.653	0.658
ag	0.742	0.857	0.843
ap	0.731	0.733	0.715
bh	0.917	0.857	0.911
sd + ag	0.750	0.925	0.915
sd + ap	0.732	0.742	0.776
sd + bh	0.898	0.871	0.934
ag + ap	0.819	0.813	0.884
ag + bh	0.925	0.892	0.948
ap + bh	0.919	0.841	0.944
sd + ag + ap	0.813	0.784	0.965
sd + ag + bh	0.928	0.887	0.974
sd + ap + bh	0.913	0.829	0.954
ag + ap + bh	0.944	0.869	0.980
sd + ag + ap + bh	0.943	0.954	1.000

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 11 Normalized AUC indexes, holding companies, out-of-time sample.

Variables*	Logistic regression	Neural networks	Random forests
sd	0.665	0.666	0.665
ag	0.714	0.824	0.796
ap	0.694	0.674	0.693
bh	0.935	0.893	0.942
sd + ag	0.728	0.889	0.867
sd + ap	0.724	0.824	0.821
sd + bh	0.930	0.883	0.929
ag + ap	0.744	0.836	0.914
ag + bh	0.908	0.913	0.948
ap + bh	0.946	0.828	0.945
sd + ag + ap	0.759	0.784	0.926
sd + ag + bh	0.906	0.830	0.956
sd + ap + bh	0.939	0.853	0.960
ag + ap + bh	0.929	0.798	0.987
sd + ag + ap + bh	0.924	0.734	1.000

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

FIGURE 2 Default rate by month.

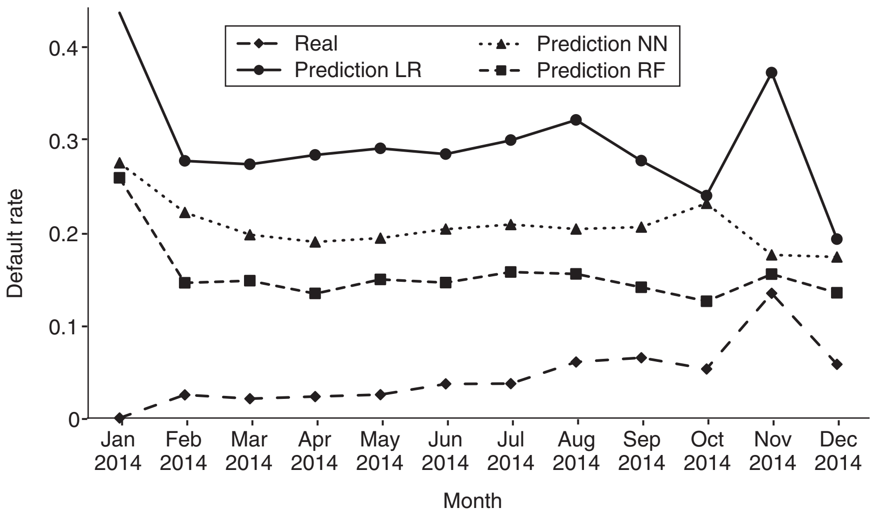


TABLE 12 The most significant variables for the logistic regression model by segmentation of customers.

Variable	Variable group*	All	People	Comp.	Hold.	Total
Arrears3to6Months	bh	1	1	1	1	4
TimelyPay6Mto9M	bh	1	1	1	1	4
ArrearsLast3M	bh	1	0	1	1	3
TimelyPayLast3M	bh	1	0	1	1	3
TimelyPay3Mto6M	bh	1	0	1	1	3
RegionG1	sd	1	0	1	1	3
LevelPurchases	sd	1	1	0	1	3
ArrearsIncreaseLast3M	bh	1	1	1	0	3
Tenure	ap	1	0	1	1	3
PropertyDistance	ag	0	1	1	1	3
CropTypeG2	ag	1	1	0	0	2
Cost	ag	1	1	0	0	2
CropsNumber	ag	1	0	1	0	2
TimelyInstLast3M	bh	0	1	0	0	1
TotalBalance	ap	1	0	0	0	1
IncomeHectare	ag	1	0	0	0	1
Income	ag	0	1	0	0	1

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

6.3 Joint analysis of modeling techniques, explanatory variables and segmentation

In this section, we use a different approach to measure the impact of each group of variables on the performance of the model. Since the logistic regression is the most widely used approach in the literature, we focus our attention on this model for various segments of customers. We use the random forest approach to measure the importance of each explanatory variable because this model performs an implicit feature selection, using a subset of strong variables for the classification (Breiman 2004). In particular, the Gini criterion (equivalent to the AUC) is used for measuring how well a split separates the samples in the two classes. The Gini criterion uses the Gini index, which is often used as a measure of income inequality. This index can be calculated as one minus twice the area between the Lorenz curve and the diagonal line representing perfect equality (values in the interval [0, 1]). In this way, a higher Gini index indicates greater discrimination between two classes.

The random forest model provides two measures of variable importance: the mean decrease Gini (MDG) and the mean decrease accuracy (MDA; Calle and Urrea (2011). The MDG is the sum of all the decreases in the Gini impurity due to a given variable, normalized by the number of trees. The MDA is the average accuracy of the predictor minus the decrease in the accuracy after the permutation of that predictor. We prefer using the MDG because its rankings are more robust than those generated using the MDA (Calle and Urrea 2011).

Figure 3 displays the ranking of the fifteen most important variables from the lowest to the highest MDG. For all the segments of customers, the crop type and the term type (payment frequency of the loan) are the most relevant variables, followed by various ones belonging to the credit and behavioral groups. The main differences between the segments of clients, in relation to the importance of the variables, are shown by the term type and purchase level. The term type is important for companies and people, but not for holdings companies. In contrast, the purchase level is relevant for companies and holding companies. Holding companies in particular show more significant variables in the agribusiness group, which means that the economic conditions related to crops are more prevalent in this segment than in the others. This makes sense since the segment is oriented mostly to the SME, which tends to have higher variability and perceived risk (Maurer 2014).

Table 12 shows that in logistic regression models the variables selected in most cases belong to the behavioral set, followed by the sociodemographic characteristics. Even if the agribusiness variables have been chosen in each segmentation, different segments are related to different agribusiness variables. For example, the crop type and the cost appear in the "people" segment, and the property distance shows up in all three segments. This hints at a diversity among the different segments that needs to be captured by different models.

Regarding neural networks, we used the variable importance method proposed by Garson (1991). This method is based on connection weights to measure the relative importance of the explanatory variables in relation to the response variable. Table 13 presents the most important variables for neural network models that consider all the different types of variables. In this case, various sociodemographic and agribusiness variables related to incomes are the most important in all segments.

Another conclusion that can be drawn is that the risk brought by the liability amount is higher and relevant only for companies and holdings, since people tend to have liabilities concentrated in a narrower and thus less significant range. However, the term is far more relevant for people, which can be explained by the income uncertainty brought about by the extended time between sowing and harvesting/selling crops, which for small borrowers has far more of an impact on their solvency. This also affects their liquidity in the face of unexpected events influencing profitability, which is not the case when they are compared with companies.

TABLE 13 The most important variables for the neural network model by segmentation of customers.

Variable	Variable group	All	People	Comp.	Hold.
LevelPurchases	sd	0.045	0.047	0.046	0.040
IncomeHectare	ag	0.048	0.037	0.040	0.038
RegionG1	sd	0.041	0.036	0.038	0.037
CropTypeG2	ag	—	0.040	0.043	0.039
CompanyTime	ap	—	0.046	0.043	—
PropertyDistance	ag	—	0.046	0.040	—
ArrearsLast3M	bh	0.043	—	—	0.039
RecentAccounts	ap	—	—	0.034	0.042
RatioArrearsAmountLast3M	bh	0.053	—	—	—
OfficeClientDist	ap	0.047	—	—	—
Income	ag	—	0.042	—	—
ArrearsAmount6Mto9M	bh	0.042	—	—	—
OfficeRegion	ap	—	0.042	—	—
Tenure	ap	0.042	—	—	—
ProductGroupNumber	ap	—	0.041	—	—
TermTypeG	ap	—	—	0.041	—
Cost	ag	—	0.041	—	—
TimelyPay3to6Months	bh	—	—	—	0.040
ArrearsAmount3Mto6M	bh	0.039	—	—	—
PayAmount6Mto9M	bh	0.039	—	—	—
ArrearsIncreaseAmount3Mto6M	bh	—	—	—	0.039
CurrencyG1	ap	—	—	—	0.037
MinArrearsAmountLast3M	bh	—	—	—	0.037
PropertyLocationN	ag	—	—	0.037	—
ArrearsIncreaseAmountLast3M	bh	—	—	0.036	—

*The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

6.4 Cost–benefit analysis

This section presents an analysis of the costs and benefits of using the model. These costs and benefits have been measured with a base scenario developed using Verbraken *et al* (2013) as a reference. The base scenario is the situation in which there is no classification model. In the case of credit scoring, this scenario occurs when all loans are granted; this comparison ensures consistency when evaluating different credit scoring models (Verbraken *et al* 2014).

We calculated the profit of using a model with the average classification profit per borrower (Verbraken *et al* 2014):

$$P(t; b_1; c_0; c^*) = (b_1 - c^*)\pi_1 F_1(t) - (c_0 + c^*)\pi_0 F_0(t), \tag{6.1}$$

where functions $F_1(t)$ and $F_0(t)$ are the cumulative density functions of the scores of the cases and noncases, respectively. The prior probabilities of classes 1 and 0 are π_1 and π_0, respectively. In relation to benefits and costs, b_1 is the benefit of correctly identifying a defaulter, $c_0 \geqslant 0$ is the cost of incorrectly classifying a good applicant as a defaulter, and c^* is the cost of the action. We used the methodology developed by Bravo *et al* (2013) to calculate each of these parameters.

b_1 is calculated as the fraction of the loan amount that is lost after default:

$$b_1 = \frac{\text{LGD} \cdot \text{EAD}}{A}, \tag{6.2}$$

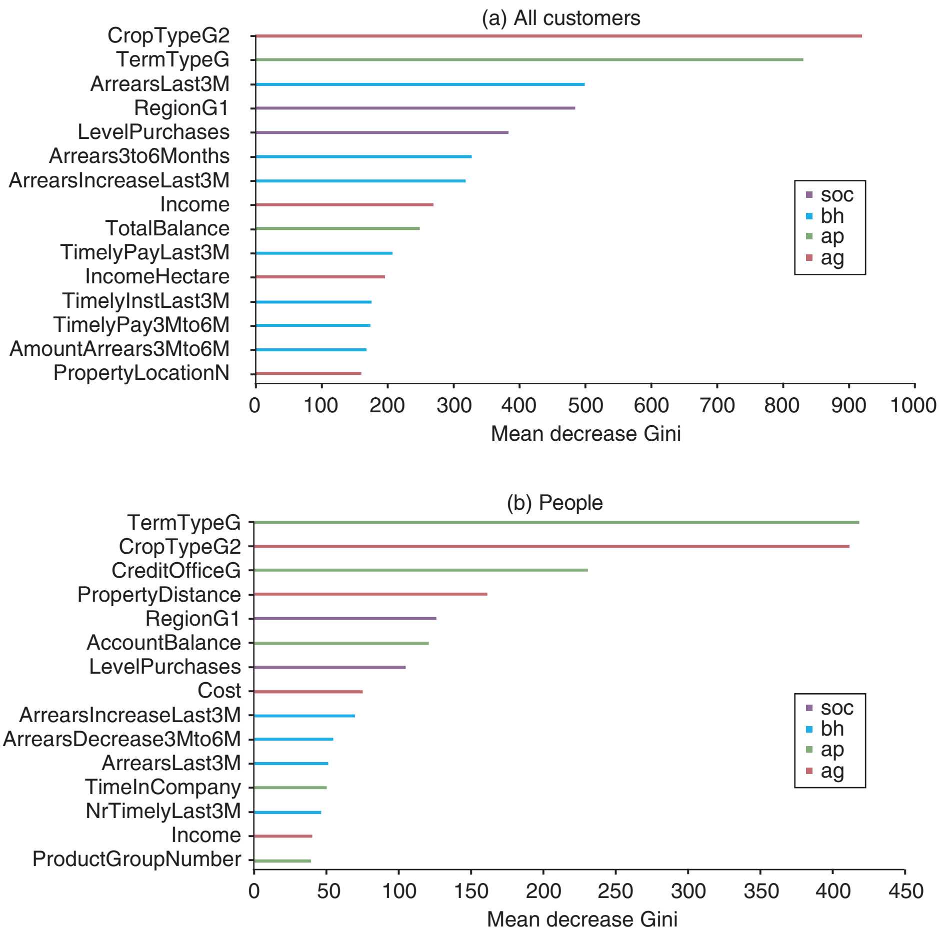
where A is the principal, LGD is the loss given default and EAD is the exposure at default; c_0 is equal to the return on investment (ROI) of the loan and is calculated by the cost of the funds and all operational costs; and c^* is assumed to equal 0 because rejecting a customer does not generate costs.

The ROI of the company (c_0) is 0.05. Because the company does not have any sort of advanced internal ratings-based approach (IRB), that is, its own internal estimates of risk components (Basel Committee on Banking Supervision 2004), we set LGD equal to 1, defaulting to foundational IRB parameters (which mandate LGD $= 1$ for unsecured retail loans).

The results can be seen in Table 14. They demonstrate that making use of a model leads to a utility greater than zero, that is, using a credit scoring tool is beneficial in economic terms. Specifically, the technique that has the biggest total profit is the random forest (RF), and the best profit per loan (granted loans) is achieved by logistic regression (LR). In this sense, the technique can be selected according to the business objective, costs and efforts of model development and implementation.

According to the results, using a credit score model is a good option in economic terms, regardless of the technique chosen. Scoring models can be used at different levels as a support tool in the lending decision, acting as everything from a guide to classifying clients to the main method of evaluation by automatically accepting or rejecting clients according to their credit. First, an easy-to-interpret model could be better than a "black box" as a support tool in the loan decision. Logistic regression is the most interpretable technique of the three analyzed, and this technique exhibits a competitive performance compared with machine learning techniques. However, the random forest approach is the most profitable option, despite the loss of interpretability. Therefore, the decision concerning which model to use depends on the purpose and the level of use of the credit score model.

FIGURE 3 Importance of variables. [Figure continues on next page.]

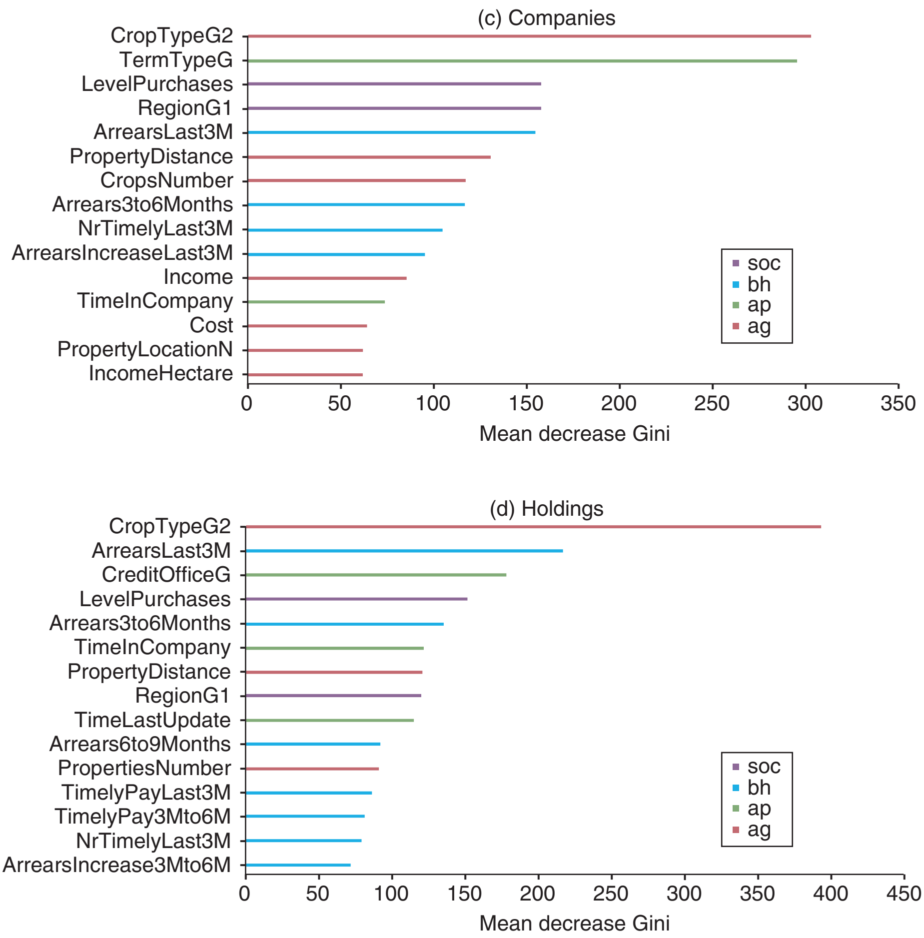


The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

Another important cost to consider is the cost of implementation. This cost could be divided into different aspects: computer infrastructure for the training model process and for future evaluations, expert knowledge for building the model, and training for the organization in order to properly use the credit scoring. This cost increases as the complexity of the model increases.

A general recommendation is to start with an easy-to-interpret technique such as logistic regression and then migrate to a machine learning technique. The random

FIGURE 3 Continued.



The abbreviations sd, ap, ag and bh stand for sociodemographic variables, credit variables, agribusiness variables and behavioral variables, respectively.

TABLE 14 Profit by model.

Model	Profit (US$)	Granted loans	Profit per loan (US$)
LR	230 067	1995	115.32
NN	104 000	2120	49.06
RF	234 930	2554	91.99

forest model is a good option because it offers a good performance and the possibility of computing the importance of variables (mean decrease Gini and mean decrease accuracy), and its training process is less complex than other techniques such as neural networks.

7 CONCLUSIONS AND FUTURE WORK

The credit risk assessment for the agricultural sector shows specific characteristics created by the uncertainty of successful crops and the lack of reliable information. Using data provided by a Chilean company, this study shows that the repayment behavior characteristics and agribusiness variables are some of the most important aspects in causing farmer repayment defaults. Among these groups, the most relevant variables are the days in arrears and the type of crop.

With regard to the chosen modeling techniques, the random forest approach shows the best performance, followed by the widely used logistic regression model. There is a 6% gain between the best logistic regression and the best random forest, which suggests that the gain realized by exploiting more complex patterns is minor when compared with the gain of using better variable segments.

Concerning the segmentation of customers, the model estimated on the out-of-time sample of all the customers shows more stable results than those estimated on the segments of borrowers (people, companies and holding companies). The main differences between these segments are in the importance of the level of purchases and the agribusiness variables. The results clearly show how the patterns are structurally different among these segments, with some variables having distinct relevance. However, the predictive accuracy of a combined model is in line with a differentiated one, so a lender who does not desire to obtain the relevant information that comes from having various models for each segment may choose to use only one model while their sophistication increases.

The previous result also leads to an interesting conclusion: given that we can draw a parallel between the size of the company and the reality of different countries (ie, the purpose of the loans, the type of borrowers and access to bank loans, among other aspects), we can see that, in general, while the models need to be different for each reality (ie, they require different variables), the statistical performance measures are similar. This is surprising, because one would expect that the greater data availability of larger, more sophisticated companies would lead to better capabilities to detect default, but the results seem to indicate that a dedicated lender who collects correct data will be able to detect this correctly across many segments (ie, realities).

The main conclusion that can be drawn from this study is that a lender for agribusinesses does not face an extremely different scenario from that of a traditional lender. As long as the variables regarding the particular business are collected and care is taken regarding which segments the lender serves, it is possible to use existing credit-scoring technologies without further complexity. Doing so should provide an equivalent risk coverage to that of a lender serving a wider segment of the population and not facing any additional risks. Coverage in this segment should, then, be equivalent to that of other groups of the population.

Regarding the limitations of the study, two can be pinpointed. First, the database probably underrepresents very-low-income and low-income countries. Even though there are low-income agribusinesses present in the data, they are sophisticated enough to gain access to formal suppliers, which occurs in low-to-middle-income countries and above (Reardon *et al* 2009). Second, we are using traditional, structured databases without any unstructured data (eg, text, images, psychometric) that would require more sophisticated machine learning approaches. If this data were publicly available, perhaps the potential gains shown here could be more significant.

Future work could include additional factors in the analysis, such as the impact of macroeconomic variables on the stability of the scoring models for the agribusiness sector. Another future development could be to improve the estimates of the agricultural incomes and costs to obtain estimates closer to actual values and to measure the impact of these estimates on the performance of the model.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

The first and last authors acknowledge the support of Conicyt Fondecyt Initiation into Research No. 11140264. The third author acknowledges that this research was undertaken, in part, thanks to funding from the Canada Research Chairs program.

REFERENCES

Aruppillai, T., and Phillip, P. M. G. (2014). Farmers characteristics and its influencing on loans resettlement decision in Sri Lanka. *International Journal of Economics and Finance* **6**(4), 110–117 (https://doi.org/10.5539/ijef.v6n4p110).

Baesens, B., Setiono, R., Mues, C., and Vanthienen, J. (2003a). Using neural network rule extraction and decision tables for credit-risk evaluation. *Management Science* **49**(3), 312–329 (https://doi.org/10.1287/mnsc.49.3.312.12739).

Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., and Vanthienen, J. (2003b). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society* **54**(6), 627–635.

Baesens, B., Roesch, D., and Scheule, H. (2016). *Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS*. Wiley & SAS Business Series. Wiley (https://doi.org/10.1002/9781119449560).

Bandyopadhyay, A. (2008). Credit risk models for managing bank's agricultural loan portfolio. *ICFAI Journal of Financial Risk Management* **5**(4), 86–102.

Barry, P. J., and Robison, L. J. (2001). Agricultural finance: credit, credit constraints, and consequences. *Handbook of Agricultural Economics* **1**, 513–571 (https://doi.org/10.1016/S1574-0072(01)10013-7).

Basel Committee on Banking Supervision (2004). Basel II: international convergence of capital measurement and capital standards: a revised framework. Standard, June 10, Bank for International Settlements.

Becerra, N., Fiebeig, M., and Wisniwski, S. (2004). *Agricultural Production Lending: A Toolkit for Loan Officers and Loan Portfolio Managers*. Pact Publications, Eschborn, Germany.

Bonazzi, G., and Iotti, M. (2014). Agricultural cooperative firms: budgetary adjustments and analysis of credit access applying scoring systems. *American Journal of Applied Sciences* **11**(7), 1181–1192 (https://doi.org/10.3844/ajassp.2014.1181.1192).

Bravo, C., Maldonado, S., and Weber, R. (2013). Granting and managing loans for micro-entrepreneurs: new developments and practical experiences. *European Journal of Operational Research* **227**(2), 358–366 (https://doi.org/10.1016/j.ejor.2012.10.040).

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

Breiman, L. (2004). Consistency for a simple model of random forests. Technical Report, University of California at Berkeley.

Brida, J. G., Fasone, V., Scuderi, R., and Zapata-Aguirre, S. (2014). ClustOfVar and the segmentation of cruise passengers from mixed data: some managerial implications. *Knowledge-Based Systems* **70**, 128–136 (https://doi.org/10.1016/j.knosys.2014.06.016).

Calabrese, R., Osmetti, S. A., and Zanin, L. (2019). A joint scoring model for peer-to-peer and traditional lending: a bivariate model with copula dependence. *Journal of the Royal Statistical Society* A **182**(4), 1163–1188 (https://doi.org/10.1111/rssa.12523).

Calle, M. L., and Urrea, V. (2011). Letter to the editor: stability of random forest importance measures. *Briefings in Bioinformatics* **12**(1), 86–89 (https://doi.org/10.1093/bib/bbq011).

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* **16**, 321–357 (https://doi.org/10.1613/jair.953).

Department of Economic and Social Affairs Statistics Division (2016). International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4. United Nations.

Dressler, J. B., and Tauer, L. W. (2016). Estimating expected and unexpected losses for agricultural mortgage portfolios. *American Journal of Agricultural Economics* **98**(5), 1470–1485, (https://doi.org/10.1093/ajae/aaw049).

Durguner, S., and Katchova, A. L. (2007). Credit scoring models in Illinois by farm type: hog, dairy, beef and grain. American Agricultural Economics Association Meeting, July 29–August 1, Portland, OR.

EME (2014). Acceso a financiamiento en los emprendimientos [Access to financing in enterprises]. Report, January, División de Estudios, Ministerio de Economía.

Eyo, E., and Ofem, U. (2014). Analysis of creditworthiness and loan repayment among bank of agriculture loan beneficiaries (poultry farmers) in Cross River State, Nigeria. *International Journal of Livestock Production* **5**(9), 155–164 (https://doi.org/10.5897/IJLP2014.0218).

Featherstone, A. M., Roessler, L. M., and Barry, P. J. (2006). Determining the probability of default and risk-rating class for loans in the seventh farm credit district portfolio. *Applied Economic Perspectives and Policy* **28**(1), 4–23 (https://doi.org/10.1111/j.1467-9353.2006.00270.x).

Fica, A. L. L., Casanova, M. A. A., and Mardones, J. G. (2018). Análisis de riesgo crediticio, propuesta del modelo credit scoring [Credit risk analysis, credit scoring proposal]. *Revista de la Facultad de Ciencias Económicas: Investigación y Reflexión* **26**(1), 181–207 (https://doi.org/10.18359/rfce.2666).

Gallagher, R. L. (2001). Characteristics of unsuccessful versus successful agribusiness loans. *Agricultural Finance Review* **61**(1), 20–35 (https://doi.org/10.1108/00214730180001114).

Garson, D. (1991). Interpreting neural network connection weights. *AI Expert* **6**(4), 47–51.

Gustafson, C. R., Pederson, G. D., and Gloy, B. A. (2005). Credit risk assessment. *Agricultural Finance Review* **65**(2), 201–217 (https://doi.org/10.1108/00214660580001173).

Hanley, J. A., and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**(1), 29–36 (https://doi.org/10.1148/radiology.143.1.7063747).

Hassoun, M. H. (1995). *Fundamentals of Artificial Neural Networks*. MIT Press.

Hazell, P. B. (1992). The appropriate role of agricultural insurance in developing countries. *Journal of International Development* **4**(6), 567–581 (https://doi.org/10.1002/jid.3380040602).

Henning, J. I., and Jordaan, H. (2016). Determinants of financial sustainability for farm credit applications: a DELPHI study. *Sustainability* **8**(1), 77 (https://doi.org/10.3390/su8010077).

Hosmer, D., and Lemeshow, H. (2000). *Applied Logistic Regression*. Wiley (https://doi.org/10.1002/0471722146).

Hou, J., Skees, J., and Wang, W. (2005). Potential of credit scoring in microfinance institutions in US (Community Venture Corp. of Kentucky taken as case study). Conference Paper, 2005 Annual Meeting, February 5–9, 2005, Little Rock, Arkansas. Southern Agricultural Economics Association.

Jouault, A., and Featherstone, A. M. (2011). Determining the probability of default of agricultural loans in a French bank. *Journal of Applied Finance and Banking* **1**(1), 1–30.

Katchova, A. L., and Barry, P. J. (2005). Credit risk models and agricultural lending. *American Journal of Agricultural Economics* **87**(1), 194–205 (https://doi.org/10.1111/j.0002-9092.2005.00711.x).

Kiers, H. A. (1991). Simple structure in component analysis techniques for mixtures of qualitative and quantitative variables. *Psychometrika* **56**(2), 197–212 (https://doi.org/10.1007/BF02294458).

Klein, B., Meyer, R., Hannig, A., Burnett, J., and Fiebig, M. (2001). *Mejores Prácticas Del Financiamiento Agrícola* [Better Practices in Agricultural Finance]. FAO.

Kullback, S. (1997). *Information Theory and Statistics*. A Wiley Publication in Mathematical Statistics. Dover Publications.

Lessmann, S., Seow, H.-V., Baesens, B., and Thomas, L. C. (2013). Benchmarking state-of-the-art classification algorithms for credit scoring: a ten-year update. *European Journal of Operational Research* **247**(1), 124–136 (https://doi.org/10.1016/j.ejor.2015.05.030).

Limsombunchai, V., Gan, C., and Lee, M. (2005). An analysis of credit scoring for agricultural loans in Thailand. *American Journal of Applied Sciences* **2**(8), 1198–1205 (https://doi.org/10.3844/ajassp.2005.1198.1205).

Lobo, J. M., Jiménez-Valverde, A., and Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**(2), 145–151 (https://doi.org/10.1111/j.1466-8238.2007.00358.x).

Madeira, C. (2019). Measuring the covariance risk of consumer debt portfolios. *Journal of Economic Dynamics and Control* **104**, 21–38 (https://doi.org/10.1016/j.jedc.2019.05.005).

Mansfield, E. R., and Helms, B. P. (1982). Detecting multicollinearity. *American Statistician* **36**(3a), 158–160 (https://doi.org/10.1080/00031305.1982.10482818).

Maurer, K. (2014). Where is the risk? Is agricultural banking really more difficult than other sectors? In *Finance for Food*, Köhne, D. (ed), pp. 139–165. Springer (https://doi.org/10.1007/978-3-642-54034-9_7).

Miller, L. H., and LaDue, E. L. (1988). Credit assessment models for farm borrowers: a logit analysis. Technical Report, Department of Applied Economics and Management, Cornell University.

Miller, L. H., Ellinger, P. N., Barry P. J., and Lajili, K. (1993). Price and non-price management of agricultural credit risk. *Agricultural Finance Review* **53**, 28–41.

Novak, M. P., and LaDue, E. (1994). An analysis of multiperiod agricultural credit evaluation models for New York dairy farms. *Agricultural Finance Review (USA)* **54**, 55–65.

Novak, M. P., and LaDue, E. (1999). Application of recursive partitioning to agricultural credit scoring. *Journal of Agricultural and Applied Economics* **31**, 109–122 (https://doi.org/10.1017/S0081305200028818).

Odeh, O. O., Featherstone, A. M., and Sanjoy, D. (2006). Predicting credit default in an agricultural bank: methods and issues. Conference Paper, 2005 Annual Meeting, February 5–8, 2006, Orlando, Florida. Southern Agricultural Economics Association.

ODEPA (2009). *Estudio de Financiamiento Agrícola: Instrumentos disponibles, coberturas, innovaciones y desafíos pendientes* [Available instruments, coverage, innovations and pending challenges]. Report, July, Clarke & Asociados. URL: https://bit.ly/3qGRjMs.

ODEPA (2013). *Financiamiento Agrícola: ¿Por Qué Surco Debe Seguir?* [Agricultural financing: which groove should it follow?]. Report, May, ODEPA. URL: https://bit.ly/3oub4VC.

Onyenucheya, F., and Ukoha, O. (2007). Loan repayment and credit worthiness of farmers under the Nigerian Agricultural Cooperative and Rural Development Bank (NACRDB). *Agricultural Journal* **2**(2), 265–270.

Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J., and Baesens, B. (2019). The value of big data for credit scoring: enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing* **74**, 26–39 (https://doi.org/10.1016/j.asoc.2018.10.004).

Pelka, N., Musshoff, O., and Weber, R. (2015). Does weather matter? How rainfall affects credit risk in agricultural microfinance. *Agricultural Finance Review* **75**(2), 194–212 (https://doi.org/10.1108/AFR-10-2014-0030).

Probst, P., and Boulesteix, A.-L. (2018). To tune or not to tune the number of trees in random forest. *Journal of Machine Learning Research* **18**(181), 1–18.

Rambaldi, A. N., Zapata, H. O., and Christy, R. D. (1992). Selecting the "best" prediction model: an application to agricultural cooperatives. *Southern Journal of Agricultural Economics* **24**, 163–163 (https://doi.org/10.1017/S008130520002608X).

Reardon, T., Barrett, C. B., Berdegué, J. A., and Swinnen, J. F. (2009). Agrifood industry transformation and small farmers in developing countries. *World Development* **37**(11), 1717–1727 (https://doi.org/10.1016/j.worlddev.2008.08.023).

Romani, G. A., Bravo, A., Aroca, P., Aguirre, N. A., Vega, P. L., and Carrazana, J. M. (2002). Modelos de clasificación y predicción de quiebra de empresas: una aplicación a empresas chilenas [Corporate bankruptcy prediction and classification models: an application to Chilean companies]. *Forum Empresarial* **7**(1), 2–20 (https://doi.org/10.33801/fe.v7i1.3371).

Römer, U., Römer, U., Musshoff, O., and Musshoff, O. (2017). Can agricultural credit scoring for microfinance institutions be implemented and improved by weather data? *Agricultural Finance Review* **78**(1), 83–97 (https://doi.org/10.1108/AFR-11-2016-0082).

Savitha, B., Savitha, B., Kumar, K. N., and Kumar, K. N. (2016). Non-performance of financial contracts in agricultural lending: a case study from Karnataka, India. *Agricultural Finance Review* **76**(3), 362–377 (https://doi.org/10.1108/AFR-01-2016-0001).

Schwab, K. (2017). The global competitiveness report 2017–2018. Insight Report, World Economic Forum.

Sherrick, B. J., Barry, P. J., and Ellinger, P. N. (2000). Valuation of credit risk in agricultural mortgages. *American Journal of Agricultural Economics* **82**(1), 71–81 (https://doi.org/10.1111/0002-9092.00007).

Siddiqi, N. (2007). *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*, Volume 3. Wiley.

Siddiqi, N. (2017). *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards*. Wiley (https://doi.org/10.1002/9781119282396).

Thomas, L., Crook, J., and Edelman, D. (2017). *Credit Scoring and Its Applications*, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (https://doi.org/10.1137/1.9781611974560).

Tiedemann, T., and Latacz-Lohmann, U. (2013). Production risk and technical efficiency in organic and conventional agriculture: the case of arable farms in Germany. *Journal of Agricultural Economics* **64**(1), 73–96 (https://doi.org/10.1111/j.1477-9552.2012.00364.x).

Turvey, C. G. (1991). Credit scoring for agricultural loans: a review with applications. *Agricultural Finance Review (USA)* **51**, 43–54.

Verbraken, T., Verbeke, W., and Baesens, B. (2013). A novel profit maximizing metric for measuring classification performance of customer churn prediction models. *IEEE Transactions on Knowledge and Data Engineering* **25**(5), 961–973 (https://doi.org/10.1109/TKDE.2012.50).

Verbraken, T., Bravo, C., Weber, R., and Baesens, B. (2014). Development and application of consumer credit scoring models using profit-based classification measures. *European Journal of Operational Research* **238**(2), 505–513 (https://doi.org/10.1016/j.ejor.2014.04.001).

Zech, L., and Pederson, G. (2003). Predictors of farm performance and repayment ability as factors for use in risk-rating models. *Agricultural Finance Review* **63**(1), 41–54 (https://doi.org/10.1108/00214990380001140).

Zhang, T., and Ellinger, P. N. (2006). Credit risk and financial performance assessment of Illinois farmers: a comparison of approaches with farm accounting data. American Agricultural Economics Association Annual Meeting, 23–26 July, Long Beach, CA.

Ziari, H. A., Leatham, D. J., and Turvey, C. G. (1995). Application of mathematical programming techniques in credit scoring of agricultural loans. *Agricultural Finance Review* **55**, 74–88.