Credit scoring – a historic recurrence in microfinance[1]

Vitalie Bumacov, Oxford Brookes Business School, Oxford Brookes University, UK

Arvind Ashta, CEREN, EA 7477, Burgundy School of Business - Université Bourgogne Franche-Comté, France

Pritam Singh, Oxford Brookes Business School, Oxford Brookes University, UK.

By synthesizing a range of insights based on a review of the literature on credit scoring in the developed world, we outline a conceptual framework of credit scoring that enables the use of this technique in micro lending, avoiding the pitfalls of the past.

Key points

Microfinance is a new credit segment that can benefit greatly from the advantages the credit scoring technique can offer.

Since the credit scoring technique did not evolve fast enough to meet the needs of microfinance, its adoption by microfinance institutions is slow and resembles more a historic recurrence rather than a new historical stage.

[1] JEL classification codes: G21, O33, H81.
Acknowledgements: Arvind Ashta thanks the Banque Populaire and Conseil Regional de Bourgogne for partially co-financing this research.

Introduction

Microcredit is the provision of credit facilities to poor and financially excluded people who hitherto were not being provided credit services except by money lenders and loan sharks at high interest rates ranging from 50 to over 1,000 percent per year (Wai, 1957). This appalling state of affairs was owing to asymmetric information about the creditworthiness of the borrowers, high relative transaction costs, borrowers' lack of valuable collateral, and weakness of contract enforcement mechanisms (Armendáriz & Morduch, 2005). The social innovation of microcredit has enabled a reduction of these interest rates to an average of about 28 percent per year (Rosenberg *et al.*, 2009).

The essential social innovation of the microcredit has been that of group lending. In one variant of this group lending model – the Nobel Prize winning Grameen Bank approach – the microfinance institution (MFI) first lends to two women in a group of five. If they repay, another two are given a loan and if all four repay, the fifth also receives a loan (Yunus, 2003). The inside information of the group members enables the MFI to avoid bad borrowers. Group members monitor, advise and help each other in matters dealing with the loans. The repayment of the micro loans is in public, thus further reducing transaction costs and moral hazard. Default rates reached surprisingly low levels in many big MFIs. Based on such low risks and high returns, microcredit has been growing at 20 to 30 percent per year being offered to millions of borrowers in poor countries, most being among the poorest (Maes & Reed, 2012; Reed, 2012; Reed *et al.*, 2014).

The gradual shift from group to individual lending has nurtured problems of loan repayment in many MFIs because the techniques of individual risk assessment lagged behind. Bad credit behavior may have led to aggressive actions on the part of lenders and social manifestations on the part of borrowers, some refusing to repay the loans as did the members of No pago! movement in Nicaragua. The legislator has often clamped down on the sector, as it happened in Andhra Pradesh, India. One technique that MFIs could use for safe individual lending on a large scale is credit scoring.

The use of credit scoring in microfinance has the potential to improve market efficiency, but a large knowledge gap exists between the demand and supply of credit scoring solutions. To avoid the mistakes of the past, we present the history of credit scoring in developed countries, mainly USA, which were confronted with problems in providing credit similar to the ones faced by MFIs nowadays. Although several books present a history of credit scoring (Thomas, 2002; Anderson, 2007), none focused on the lessons microfinance practitioners and academics can extract from it.

This paper divides the evolution of credit scoring into different historical periods, each presented in a different section based on the kind of subjects being studied predominantly in each period. We have characterized the time periods as early works, the commercial era of credit scoring for consumer credit, credit scoring for corporate lending, institutionalization, new dimensions of credit scoring, and microfinance era, with the latter looking more like a historic recurrence rather than a new historical stage.

Early works

The set of studies in consumer “instalment” financing conducted by the US National Bureau of Economic Research (NBER) in the last years of the 1930s and published in 1940 and 1941 represents the base on which credit scoring emerged. The NBER, a not-for-profit organization, was engaged in researching and disseminating knowledge about essential economic facts. Installment financing was one important economic activity in the USA. Loans outstanding doubled from 1934 to 1938 (Chapman *et al.*, 1940).

Personal loans contributed significantly to the growth of consumer credit. A predecessor of the current micro loan, the amount at the end of the 1930s was generally limited to 300 USD – the equivalent of about 5,200 USD in 2017 – in cash and maximum legal interest rates were between 24 and 42 percent per year, depending on the particular state in the USA where the transaction took place. The supply of personal loans was encouraged “to combat the loan-shark evil, which had arisen because the usury statutes prevented the profitable lending of small sums at legitimate rates” (Young *et al.*, 1940, p. 1). This problem is current in numerous developing countries and microfinance was foreseen as the main tool to solve it. At that time the need to reduce transaction costs and enhance the appraisal of risks was obvious in a competitive environment.

Another NBER study in installment financing presented the problems that were experienced by the lenders (Chapman *et al.*, 1940). Credit providers were conscious that some attributes of the applicants were specific to bad borrowers and considered them in their subjective evaluation. Two broad indicators were important when judging a prospective credit risk: willingness and ability to repay the loan. Information on the income of the applicant, her net worth and other financial characteristics were predictors of the ability to pay back, while the willingness to pay depended on the applicant's character.

"On the basis of experience, and to some extent intuition, the loan officer decides which applicants are more likely to default than others or which loans are likely to involve collection costs so great as to render the transaction unprofitable" (Chapman *et al.*, 1940, p. 109). "Lenders need to know the relative importance of as many credit risk factors as can be isolated, and in making a final decision on a loan application the responsible officer must give due weight to each factor" (Chapman *et al.*, 1940, p. 137). The latter citation defines the banking issue that credit scoring is solving using empirical methods. A credit scoring algorithm contains all risk factors that could be identified and their weights – the relative importance of the risk factor in making the subject a bad credit.

Based on samples of good and bad clients coming from different financial institutions, Chapman *et al.* (1940) found the most significant risk factors: possession of a bank account, stability of employment, nature of occupation, permanence of residence, ownership of real estate and industrial affiliation. They made an important qualifying remark about the sampling that poses challenges for researchers and practitioners even today. "Since these borrowers had already passed through a selection process at the hands of credit men, the sample cannot be considered completely representative of the general run of personal loan applicants" (Chapman *et al.*, 1940, p. 111). This implied that the identification of risk factors was based on a previously screened sample of accepted applicants which became consequently good or bad credit risks. The sample could not include applicants who had been refused.

One other NBER study (Plummer & Young, 1940) went beyond identification of credit risk factors and presented functioning credit rating systems. The use of such systems was not uncommon in the USA in the late 1930s. In some cases a rating of good, fair or poor was entered for each factor considered as an indicator of credit risk. The final rating represented the average of favorable and unfavorable indications. “In other cases a specific grade is entered opposite each item and the sum of the grades serves as the index of credit risk” (Plummer & Young, 1940, p. 136). The use of such rating systems is still common today, especially in micro and small business lending, but we note that subjective attribution of weights to identified risk factors makes the technique distinct from credit scoring. Since empirics are not used in deriving the relative weights of risk factors, such techniques are generally referred as credit rating.

The work of Durand (1941), in yet another NBER study, marked the first important step in the development of the technique of credit scoring. The author used a sample of 7,200 consumer loans disbursed by 37 financial institutions which included commercial banks offering personal loans, personal finance companies, automobile finance companies and appliance financers. Available borrowers’ attributes were copied from their loan application forms. These included age, gender, marital status, dependents in the household, stability of employment, permanence of residence and other socio-demographic variables. Also, borrowers’ assets and liabilities, and loan characteristics such as amount and number of instalments were available. No information was collected on past credit behavior or on “matters like physical or mental health, which are certainly germane to risk problem, but which obviously do not lend themselves to analysis in a statistical study of credit risks” (Durand, 1941, p. 20).

The crucial issue that Durand dealt with was the formulation of the concept of the credit scoring algorithm, which he called “credit-rating formula”. The formula combined the most important credit risk factors and respective relative weights. For the first time, identification of risk factors and calculation of their weights was performed empirically. When lenders were scoring the applicants for credit, the weights of risk factors were computed, yielding a sum: the “credit-rating” score. The score was used as the basis for accepting or rejecting loan applications. The advance and diffusion of the technique of statistical discrimination of populations (Fisher, 1936) played a key role in the emergence of credit scoring.

Intuitive-subjective rating formulas had been created before and used in financial institutions. Durand “experimented with deriving purely objective credit formulae by statistical methods” (Durand, 1941, p. 84). He pointed out precisely the advantages of credit scoring – loan officers could assess ordinary loan applications faster and most of the routine evaluation work could be handled by less experienced and relatively low-salaried personnel. Considering that most developing countries suffer from the epidemic of brain drain, the advantages of credit scoring should appeal to many microfinance providers and supervising bodies.

Durand identified several practical aspects to be considered in developing credit scoring algorithms. Two of them are particularly relevant for microfinance: size of the sample and the time frame. About the sampling size, he argued that though large samples were preferable, since collecting thousands of cases could be impossible or too expensive, a sample as small as one hundred good and one hundred bad loans might be adequate for statistical significance. Since many MFIs experience problems with data collection on a large scale due to bad databases or lack of management information systems (MIS), this suggestion of one hundred good and one hundred bad loans in the sample is feasible within the microfinance reality.

About the time frame problem, Durand remarked that risk experience associated with different factors may alter with time. One such example could be the stability of employment. Changing employers too often has been commonly associated with workplace problems and subsequent financial uncertainty. However, in today's work environment changing employers could be associated with exceptional skills required for ad hoc, sophisticated and well-remunerated projects. Durand's solution to the time frame problem was to limit the study to recent homogeneous periods. Good and bad credit risks being included in the sample had to be recent. Further we use the term 'training sample' to identify the set of good and bad credit risk profiles used to develop the credit scoring algorithm.

Durand made the interesting remark that in practice it was difficult to make a precise distinction between confirmed good and bad loans. If the net revenue from a loan did not cover the expenses it generated, it was certainly a bad loan. If the principal and interest were repaid in full and on time, it was a good loan. The problem was with grey cases where borrowers paid late. The financial institution collected penalty fees to compensate delinquency-related costs, such as reminders and specific supervision, but these fees did not always cover all costs. Since it was expensive to determine accurately at what level of delinquent credit behavior a loan ceased to be profitable, Durand proposed several rules to define bad credits: "loan was more than 90 days delinquent; comaker [cosigner] paid all or part of loan after demand by bank; legal action was taken; loan was charged off" (Durand, 1941, p. 38). Currently, many microfinance institutions use the same rules when defining bad borrowers. In some markets however, where micro borrowers make weekly repayments, 30 days delinquency – the equivalent of four missed installments – is the preferred definition of a bad credit risk (Simbaqueba *et al.*, 2011).

Durand reiterated the remarks of Chapman *et al.* (1940) about the limitations of the sample due to initial screening of applicants by the loan officers. To overcome this problem, Durand suggested that financial institutions would initially use the scoring algorithm as a supplementary evaluation tool after the usual screening by the loan officers. Thereafter experimental loans had to be disbursed by lowering temporarily and progressively the approval standards of the loan officers in order to gain new experience about the behavior of applicants that were accepted by credit scoring but would have been refused by the loan officers under normal circumstances. Such a solution could temporarily increase losses but will provide new experience for developing better credit scoring algorithms which should free the loan officer from the task of screening standard applications.

Later on, researchers and practitioners proposed alternative techniques that were easier to implement. For example Hand & Henley (1993) proposed the reject inference method. This involves inferring the credit risk class of rejected applicants – rejected applicants with profiles similar to bad credit risks would be considered bad borrowers and rejected applicants with profiles similar to good credit risks would be considered good borrowers – and including theses profiles in the training sample. Nevertheless, this method also does not completely resolve the problem of biased samples, unless credit is provided to a representative sample of applicants that would have been rejected by the loan officer but approved by the credit scoring algorithm.

It follows from the discussion above that for developing a credit scoring algorithm a representative training sample of good and bad borrowers is needed. One needs to know their detailed profiles: the more characteristics the better. Data has to be structured, so that statistical methods can be used to identify risk factors and their relative weights. These weights are used to generate a credit score, which based on a certain cut-off, indicates the acceptance or the rejection of the scored applicant. Obviously, statistics may contradict some subjective ways of thinking. For example Durand (1941) had discovered that women tended to be good risks, much to the surprise of some lenders who were convinced of the contrary view. This has also been found to be true in microfinance and MFIs in many countries lend predominantly to women.

Durand found another counter-intuitive discovery – for personal loans, there was no direct relationship between the credit risk and the income level of the borrower. This finding highlighting the fact that not all poor persons are bad credit risks stands at the foundation of microfinance (Dowla & Barua, 2006).

The revolutionary approach of Durand took some time to become known to scholars and financial institutions. That explains why there were no other academic publications covering the topic of credit scoring till the 1960s.

The commercial era of credit scoring for consumer credit

The next stage in the evolution of credit scoring was its use on a commercial basis. Consumer lending grew in the USA, affecting large segments of the public, to the point that this growth became unsustainable. Myers & Forgy (1963) stated that increasing demand for credit pushed many financial institutions to expand beyond their capacities to train and maintain experienced loan officers. In such conditions, the role of credit scoring became critically important. In recent years, in many developing countries, for example Morocco, Bosnia and Herzegovina, and Nicaragua, micro lending experienced similar unsustainable high growth, leading to over-indebtedness and consequent degradation of the market (Constantinou & Ashta, 2010). We think that credit scoring could have played a positive role in reducing the default rate of micro borrowers.

Myers and Forgy developed several scorecards using one training sample of good and bad credit risks coming from one financial institution. This was different from Durand who had combined risks from different financial institutions. They used diverse statistical and non-statistical techniques to construct scorecards and then compared their predictive effectiveness.

Myers and Forgy enriched the conceptual framework of credit scoring by introducing the concept of a holdout sample, although they acknowledged that the idea was based on an unpublished thesis by Wolbers H. L. of 1949. The holdout sample was a sub-sample obtained by withholding a part of the available information before developing the credit scoring algorithm. They developed the algorithm using the training sample and tested it on the holdout sample. In the current literature the term test sample has replaced the holdout sample. It is used to validate the robustness of the scorecard.

Being assisted by a computer, Myers and Forgy were able to experiment with different approaches to construct scorecards. In addition to the discriminant analysis used by Durand, they employed stepwise regression, the credit rating method of equal weights for identified risk factors, and double-discriminant analysis. Double-discriminant analysis gave the most satisfactory results. This method first required scoring of good and bad loans using the scorecard developed under the first approach – discriminant analysis – then the scorecard was redeveloped using only good and bad borrowers that received low scores – high credit risk. The purpose was to improve risk discrimination at ranges where good borrowers and bad borrowers have similar low scores.

Myers and Forgy found that a smaller number of risk factors (12) will predict almost as well as all (21) identified relevant factors. This is an important finding when the cost of collecting information is relatively high, as in most microfinance environments.

Myers and Forgy acknowledged the importance of the face validity of the algorithm as perceived by the final users. The financial institution that provided the data for the research chose to implement one of the developed formulas, but not the most predictive one, which was disregarded by the management due to bizarre weights of some risk factors.

Smith (1964) proposed a different approach to construct credit scoring algorithms. He suggested adding together bad account probabilities corresponding to each measured attribute (characteristic) of the applicant and considering the sum as the final credit score. These conditional probabilities measure the odds of a loan being a bad credit given certain attributes. The probabilities, all positive, were calculated using the training sample. If 25 in 100 borrowers who owned family cars were bad credit risks while those without cars were 50 bad risks in 100 and if 20 married borrowers in 100 were bad risks, while single borrowers were 40 bad risks in 100, then a married applicant who was owning a car would receive a tiny score of 25/100 + 20/100 = 45/100, while a single applicant without a car would score the double 50/100 + 40/100 = 90/100.

Smith introduced the practice of multiplying these probabilities by 1,000 for easier computing of the sum and clearer interpretation of the results. This multiplication by 1,000 was consequently adopted by the industry. Applied to the example given above, this practice would give 450 points for the credit riskiness of the married applicant with a car and 900 points for the applicant who is single and without a car. Today’s practice is to have the scores indicating creditworthiness – higher the score, lower the credit risk.

The methodology proposed by Smith is simple and clear because it uses simplifying assumptions and is thus not easily amenable to empirical investigation. Such methodology can still be used in credit granting but in our judgment, due to its weak empirical foundations, it should not be termed credit scoring. An even more serious weakness is found in the third technique proposed by Myers & Forgy (1963) which supposed giving equal weights for identified risk factors, however, the cost of such bias might be significantly lower than the opportunity cost of not using any tool to estimate the credit risk. Many micro and small business lenders do not use credit scoring because statistics look too complicated to them. If simpler techniques can motivate micro lenders in using algorithms for credit granting, then benefits from reduction of transaction and delinquency costs would probably cover the costs associated with bias.

Smith made a valuable contribution in suggesting that the rejected applicants should also be subjected to credit scoring in order to judge the robustness of the credit scoring algorithm and possibly to overcome the bias which resulted in the first instance when differentiating between selected and rejected applicants. If the scoring algorithm rejected most of applicants that were turned down by the loan officers, then such algorithm could immediately replace the loan officer in treating standard loan applications. If, however, the algorithm failed to do so, it had to be used initially as a supplementary tool until the missing information on the credit behavior of refused applicants would be gathered, as proposed by Durand.

Weingartner (1966) attributed the increasing use of credit scoring in the USA to the growing availability of "electronic" computers. He observed that financial institutions were still granting loans to applicants with higher credit risk, provided they were above the rejection cut-off, but these loans were smaller in amount and shorter in maturity. His study did not examine the aspect of risk-based pricing in the form of differential rates of interest charged to such low-scoring applicants, but the appearance of the practice of risk-based adjustments of loan conditions should be mentioned.

Weingartner highlighted the importance of performing pilot tests before credit scoring was used by the institution. He suggested initially testing newly emerging delinquent accounts to observe if they would have received low scores at the time of application.

Weingartner's work, for the first time in the literature, showed the importance of field trials. His procedure involved the use of credit scoring first by only a few loan officers or by only one branch out of the entire network. The procedure was intended for "training as well as for ironing out difficulties that arise" (Weingartner, 1966, p. 52) and, in that sense, though not a constituent of the conceptual framework of the credit scoring, it is a potentially useful optional procedure.

Post-implementation reports certainly are a constituent part of the conceptual framework of credit scoring. Weingartner proposed a continuous "barometer" that would measure the average score that would reflect the quality of new disbursed credits. The trend would indicate the quality of newly engaged portfolio over time. In the same way the overall quality of refused applications could be observed. Significant fluctuations registered by the barometer would indicate changes in the population of applicants – a warning sign.

Research tendencies by the mid-1960s included efforts to build algorithms that would predict the profitability of the loan, rather than the delinquency or default (Greer, 1967). After Hassler *et al.* (1963) provided evidence that current repayment records show strong ability to identify bad risks, Weingartner (1966) popularized the proposal to use the credit scoring technique in loan collection procedures. Collection scoring is useful in credit financing since it identifies current borrowers that are more likely to repay after being delinquent. In this way the financial institution can target them with appropriate collection actions and obtain certain return. Since the lender has already provided the borrower with a loan, collection scoring does not belong to the conceptual framework of credit scoring.

To sum up, it is clear that Durand (1941) laid the foundations of credit scoring but left unexplored areas, some of which were addressed by Meyers & Forgy (1963). Smith (1964) contributed to a practical use of credit scoring and Weingartner (1966) introduced pilot and post-implementation tests. Further improvement on their work and the first description of the implementation and functioning of a credit scoring system in a financial institution in the USA was provided in a path-breaking paper by Boggess (1967). This provides a comprehensive picture to conclude that the main traits of the conceptual framework of modern consumer credit scoring were set by the mid-1960s. Today this framework can be used by MFIs interested in implementing credit scoring.

Boggess explained that the financial institution he studied implemented a credit scoring system in 1964 in which loan applications were automatically scored by a computer if policy limits were respected. Within 24 hours the credit department was accepting or rejecting the loan application based on the credit score. Earlier, this task required up to one week. “The company cut bad debt losses enough to realize a 1.5 million [USD] profit improvement on more than 100 million [USD] in sales in the first full year of the system’s operation” (Boggess, 1967, p. 121). Computerized MIS elements combined with the virtues of credit scoring gave the lender the possibility to operate procedures that adapted the strategy of the company to shifts in the population of loan applicants. The financial institution was developing a new scoring formula every six months and was tracking changes in the weights of risk factors over time.

In Figure 1 we have captured the continuous process of development and use of credit scoring.

[Figure 1 here]: The conceptual framework of credit scoring as it existed during the mid-1960s.

The Population box in Figure 1 indicates the process of selecting the training, test and reject samples. The Algorithm box indicates the process of developing a credit scoring algorithm after required samples have been selected from the population and relevant data have been extracted and structured. This box includes the sub-boxes Empirics, Scoring Formula, Distribution and Cut-off. These sub-boxes are separate for parametric credit scoring algorithms, which are based on scorecards. For non-parametric credit scoring algorithms these sub-boxes are welded in an all-in-one block.

The Empirics sub-box indicates that for developing the algorithm a statistical technique has to be selected. All the statistical techniques that can be used to derive the algorithm belong to the group of multivariate statistical analysis. The algorithm attempts to model post factum risk performance of the borrowers using certain information available at the time of loan appraisal. We assume that a perfect algorithm, which identifies all the bad and all the good risks in the training and test samples, can be developed if the totality of factors that determine the performance of these loans can be observed and measured, and the influence of each factor can be isolated and mathematically represented. Setting the rejection cut-off in a perfect credit scoring algorithm is simple: it is the lowest score observed in good risks. All bad risks obviously score below that level. When using only a few factors – the ones that can be accurately captured by the loan application and structured for computation purposes – the resulting algorithms are less accurate. One can observe that some bad risks have higher scores than the good risks. Nevertheless such algorithms have high utility as the share of bad risks outweighs the share of good risks in low scores and vice-versa.

The sub-box Scoring formula indicates that all identified credit risk factors and their relative weights are grouped in a credit scoring formula – the scorecard. It will be used to score future applicants. The robustness of the credit scoring formula is tested using the test sample and if necessary the reject sample. The Distribution sub-box indicates that scores of past applicants that are already known to be good or bad credit risks (contained in the test and training samples) are compared with their credit risk status (good or bad). One should observe that bad credit risks systematically receive low scores and good credit risks systematically receive high credit scores. The credit scoring distribution indicates the share of bad credit risks to good credit risks for a specific score or score range. The share of bad credit risks decreases as the credit score increases. The credit scoring distribution will serve as the scale by which future applicants will be judged. The Cut-off sub-box concerns the selection of a cut-off score that will correspond to the threshold. Applicants with scores above will be accepted because they are considered good credit risks. Applicants below will be refused because they are considered bad credit risks. Financial institutions may choose to have a grey score area in between the good and bad credit risks. Loan applications which fall in the grey area are scrutinized by the loan officers who decide if they accept or reject the loan application. The reject sample is used to observe if the credit scoring algorithm screens out most of the previously refused applicants.

The Use box indicates that when credit scoring is used, the credit decision has to be consistent with the outcome of the credit scoring algorithm. Overrides should not happen. If the loan officer disregards the output of the credit scoring algorithm, the decision has to be justified. The results of such decisions have to be analyzed, especially when the risk class, good or bad, of accepted through override applicants is revealed.

The Control box is to check if the algorithm is still appropriate with time. If tests indicate deterioration of discrimination power or shifts in the population of applicants, it is necessary to develop a new credit scoring algorithm. Once the accepted applicants pay back their loans or register delinquencies, they reveal their credit risk class. Such new cases come to enrich the population of applicants with known credit risk. These cases are used afterwards to keep the credit scoring algorithm up to date. In fact, credit scoring algorithms built by using artificial intelligence can learn and adjust themselves constantly as new experience of using credit scoring becomes available. In such cases, the Algorithm box in Figure 1 covers also the dashed area – the boxes Use, Control and Population.

Credit scoring for corporate lending

While credit scoring was the exclusive domain of consumer credit, Altman (1968) employed the discriminant analysis approach to predict corporate bankruptcy using financial ratios as credit risk factors. He did not take into account other important attributes such as business-demographic characteristics like the age of the business or professional experience of the management. Other potential risk factors such as the purpose of the loan, maturity and guarantees were not considered either. In spite of its potential for estimating bad credit risk, Altman recommended the use of the algorithm as a supplement procedure in screening potential bad credit risks.

This contribution to the conceptual framework of credit scoring is important as it opens it for use in business loans. Altman enlarged the framework to include firms in addition to individuals. Obviously the attributes that describe firms are different, although some overlap may exist, for example age, address and value of owned assets.

Orgler (1970), inspired in part by the work of Altman (1968), focused on the use of the credit scoring technique in current commercial loans. His scope excluded loan approval, addressing only the periodic review of the quality of already disbursed loans. He made the interesting remark that business borrowers form less homogeneous populations compared with consumer borrowers. This problem had not yet been addressed in the academic literature. Since then, the condition of homogeneity of the targeted population, considered by practitioners and academics to be implicit in consumer credit, became part of the framework of credit scoring. Orgler concluded that commercial borrowers were so diverse and their loan products so varied in terms of maturity, amount and security, that the application of credit scoring for loan approval would be less appropriate. The main advantages of the technique were seen in releasing loan officers' time from routine evaluation of all current credits and allocating it to the small proportion of loans that would be identified by the scoring tool as deteriorating.

As formulated briefly by Altman (1968) and more in detail by Orgler (1970), the use of credit scoring to measure the evolution of the credit risk during the course of a loan was certainly new and useful for the industry, given that corporate loans had longer maturities and meantime changes could affect seriously their credit risk. During the loan reimbursement new information, especially on repayment behavior, came in regularly to enrich the known profile of the borrower. Credit scoring algorithms could integrate such rich data in predicting default even more accurately, but since the credit risk could not be avoided as the money had been already lent, the technique could not be accommodated by the conceptual framework of credit scoring, as in the case of collection scoring.

The distinction between consumer and corporate loans tends to lose its significance in microfinance. Since microcredit addresses generally poor and self-employed people, the borrower is at the same time the person applying for the loan, the household of this person and this person's small informal income-generation activity, all in one. The person – the applicant, and the firm – the business, are in general evaluated together in microfinance. Several scholars, including Orgler, cited the dissertation of Ewert of 1968 that proposed a credit scoring formula to be used by wholesale distributors in granting trade credit to retail stores. As the stores were mostly sole-proprietor small firms needing trade credit in amounts similar to consumer credit, the formula used a combination of risk factors describing the company – legally responsible for the loan, and the owner.

From this experience, we will note that the profile of the owner of the business as well as the profile of the business itself may be important in predicting credit risk, especially in small companies where personal and corporate property confounds. This is certainly the case of micro lending. Informal business contributes to blurring the limits between the natural person and the business.

Institutionalization

The early 1970s saw the credit scoring industry grow. Fair Isaac and Company (FICO), which sold the first commercially built credit scoring algorithm in 1958, started its collaboration with Wells Fargo – a major financial institution in the USA. The credit scoring provider was prospecting the exportation of the technique to Europe (Fair Isaac

Corporation, 2017). An increasing amount of academic literature dealt with different practical and theoretical topics related to the credit scoring technique. Emphasis was put on costs and net present value of loan repayments (Edmister & Schlarbaum, 1974), on best statistical techniques to be used (Muchinsky, 1975a), as well as on better definition of good and bad credit risks (Muchinsky, 1975b).

Some serious works were addressing the use of credit scoring in low-income populations in the USA. Tabor & Bowers (1977) stated that credit scoring systems are not appropriate for evaluating low-income consumers, while Sexton (1975) found that only a few risk factors differentiated between high-income and low-income households, and thus could not conclude that specific algorithms for different income categories were necessary.

In microfinance, the debate on the suitability of credit scoring is far from over. There are academics (Schreiner, 2000; Rayo Canton *et al.*, 2010; Bumacov *et al.*, 2014) and practitioners (Kortenbusch & Hauser, 2010; Simbaqueba *et al.*, 2011) that argue in favor of the adoption of credit scoring by the MFIs while others are against (Balke, 2005).

The debate over the use of credit scoring in low-income populations in the USA was triggered by the report of the US National Commission on Consumer Finance (1972), which had questioned the feasibility of a credit scoring system applicable to low-income consumers and concluded that factors most likely to discriminate the credit risk of low-income consumers were excluded from standard loan application forms.

Muchinsky (1975b) found that two opposite aspects of the borrower's repayment behavior were critical to its classification as good or bad credit risk. One was delinquency. The other was the early repayment of the loan. A premature reimbursement was susceptible of making the account unprofitable because it led to the loss of potential interest. This new element modified the perception of what constituted a bad credit risk.

The Consumer Credit Protection Act of 1968 facilitated the extension of credit to low-income clients and added legitimacy to the concept of credit scoring in the USA. The Equal Credit Opportunity Act of 1974, with the amendments of 1976, had also notable implications. This law prohibited discrimination in the granting of credit on the basis of race, religion, gender, marital status and age. These ethical concerns were formalized in a legal setup and in this way constrained the conceptual framework of credit scoring in the USA.

The in-house knowhow character of credit scoring systems used by the financial institutions represented an increasing problem for scholars, who found it difficult to relate and research how well the US industry incorporated new tendencies and legal requirements into practice. The new regulation required that statistically sound scoring systems be constructed using empirical methodologies, but no precise standards were imposed. The hypothetical obligation to demonstrate the soundness of a scoring system in Court made scholars focus on different technical aspects and assumptions that were ignored before, as long as the algorithm showed evidence of credit risk discrimination.

In the absence of case studies, Eisenbeis (1978) analyzed the credit scoring algorithms developed by academics at that time, hoping that these reflected the systems in use by lenders. Since the majority of the scorecards were developed using discriminant analysis, he pointed out the limitations of the technique and warned the public on the risks of ignoring the inherent statistical assumptions.

With the emergence of credit bureaus selling information on past credit performance of potential borrowers, the cost of extra information was considered in different credit granting schemes (Eisenbeis, 1978). If a scoring formula predicted credit risk accurately using fewer risk factors, as observed by Meyers & Forgy (1963), there was no point in paying for the extra information. On the other hand, if additional information could help discriminate better loan applications near the cut-off limits, then extra costs were clearly justified. Academics were investigating how credit scoring algorithms could accommodate these options.

Credit reporting is either new in many developing countries or non-existent (Doing Business, 2016). More than that, in countries where private or public credit reporting systems exists, these exclude microfinance institutions. This is unfortunate because credit reporting data is widely used in credit scoring algorithms.

Ang *et al.* (1979) were among the first to build a non-parametric credit scoring system. They applied the decision tree technique to a training sample with good and bad credit risks. The result was a tree-like scheme where risk factors did not have weights but acted like nodes and branches indicating at the end if the applicant was a potential good or bad credit risk.

Use of automatic interaction detector analysis confirmed that the relationships between late payments and some borrower attributes were nonlinear (Ang *et al.*, 1979). Take the example of a continuous variable such as age. Most would expect to see credit risk diminishing with age as borrowers become more experienced. In some credit segments, however, older borrowers pay better until a certain age when their credit behavior starts to worsen and resemble the behavior of young borrowers. Linear credit scoring models have a disadvantage in such situations.

Since the technique of decision trees belongs to the class of statistical multivariate analysis, the conceptual framework of credit scoring did not change. We note that besides discriminant analysis and regressions, which became popular at the end of the 1970s, a myriad of parametric and non-parametric statistical techniques was used in identifying good and bad credit risks (Häuβler, 1979).

We conclude that by the end of 1970s, credit scoring was a recognized industry. The concept found its first use in Europe, being implemented by FICO in a bank in 1977 (Fair Isaac Corporation, 2017). The company had by that time delivered approximately five hundred systems to approximately two hundred customers, including half of the fifty largest US banks, according to the testimony of William Fair – head of FICO, during hearings required by a Senate Commission (U.S. Senate, 1979).

New dimensions of credit scoring

From the perspective of the 1980s, looking back at the adoption in 1974 of the Equal Credit Opportunity Act which was originally perceived as a threat to credit scoring, we have a different view. Since the use of discriminatory factors was prohibited, many feared an overall reduction in predictive power of scoring systems in spite of some research findings (Nevin & Churchill, 1979) that showed no impact of the constraints imposed by the law. However, since judgmental systems were strongly criticized as subjective and thus prone to stronger discrimination by loan officers, credit scoring had to gain. Credit scoring was objective by definition and, under the new law, was considered respectful of ethical issues.

Capon (1982), one of the strongest opponents of the use of "brute force empiricism" in credit scoring, remarked that algorithms in use in the USA included factors with no obvious or possible logical relationship to the creditworthiness of the applicant. Other factors directly related to the capacity to pay back a loan, such as the income of the applicant, were often ignored. With computers becoming more popular and powerful, and statistical packages becoming more user-friendly and affordable, many practitioners engaged in a "wild West" conquest of predictive risk factors or in a search for surrogates of prohibited discriminatory factors.

In parallel, the credit bureau industry grew, supported by continuous expansion of the consumer credit market. Databases expanded in volume and complexity of information gathered. In 1981, FICO introduced the first credit bureau score (Fair Isaac Corporation, 2017). For an additional fee, the inquiring financial institution would henceforth receive not only the existing credit record of the applicant but also a credit score. The "bureau score" could be used solely or in combination with the institution's appraisal result in deciding whether to grant the credit or not.

A new concept called "expert system" was supposed to have serious implications in the credit business, particularly in consumer credit. Backed up by developments in the IT industry, expert systems were software designed to imitate the way of thinking of human expert (Holsapple *et al.*, 1988). However, the expert system technique remains judgmental and not empirically founded – a sine qua non condition for a credit scoring system. Nevertheless, the advantage of cyber loan officers consisted in the possibility of providing a credit decision in real time or much faster than a loan officer, leaving to the later the task of treating non-standard applications. Cyber loan officers that can be reached by phone or through Internet have strong potential in cutting transaction costs in microfinance, especially in remote areas where it takes a lot of effort and time to reach to a branch of financial institution.

Credit scoring principles progressively invaded other banking and non-banking sectors such as detection of fraud with credit cards (Rutledge, 1996; Bolton & Hand, 2002), direct marketing (Thrasher, 1992), or credit performance (Crook *et al.*, 1992). We argue, however, that use of the technique outside credit granting remains beyond the scope of the conceptual framework of credit scoring. On the other hand, the application of the scoring technique in approving loans to small and medium businesses (SMEs) and mortgage loans enriched the framework. In 1995, FICO developed a credit scoring tool for granting credit to SMEs. The concept of pooled data was introduced to solve the problem of heterogeneity of profiles of small and medium businesses (Fair Isaac Corporation, 2017).

Microfinance era

In many countries around the world the informal and semi-formal sectors represent important shares of the GDP (Harriss-White, 2003). There are lots of people employed in the informal economy who need financial services. Microfinance is a possible answer to a majority of these needs.

As microcredit evolved from group to individual lending, Viganò (1993) studied the applicability of credit scoring in less developed countries proving the usefulness of the technique in granting credit to small and micro firms in Burkina Faso. After consumer, SME and corporate finance, credit scoring entered microfinance. It was later recognized as one important technique that may affect micro credit (Rhyne & Christen, 1999).

Schreiner (2000) concluded that credit scoring for microfinance can work, but highlighted certain differences compared to its use in mainstream finance. The most important difference he found was in the information available for decision making in financing poor borrowers. In a typical microfinance setting, information is usually qualitative and informal, while credit scoring algorithms need quantitative inputs. The new challenge of credit scoring was then to adapt to this constraint.

A study shows that MFIs which manage to adopt credit scoring have more productive loan officers and by consequence have bigger impact on the financial inclusion of the poor (Bumacov *et al.*, 2014). Unfortunately the credit scoring adoption rate in MFIs remains low and current debates suggest that credit scoring history repeats itself in microfinance instead of immediately flowing with the mainstream. Repetition of past mistakes generates inefficiencies. These mistakes can be avoided if microfinance institutions apply the conceptual framework presented in this paper.

At the same time, microfinance has permitted scoring techniques to go into new areas. It is in the context of microfinance that the concept of poverty scoring emerged (Schreiner, 2010). Similar to credit scoring, poverty scoring estimates a person's chances of being poor without measuring her income and wealth, but observing several easy-measurable factors. It is argued that using credit scoring and poverty scoring together, an MFI can avoid both economic failure and mission drift – tendency to serve richer clients while ignoring the poor (Bumacov, 2012). Since poverty scoring does not estimate the credit risk, it lies outside the conceptual framework of credit scoring.

Recommendations

By synthesizing a range of insights from the literature review, in this paper we construct a conceptual framework of credit scoring that accommodates the use of the techniques in micro lending as well as in any new homogeneous credit market. The aim of this framework is to avoid a historic recurrence and bring the use of credit scoring in microfinance into the mainstream quickly.

The conceptual framework of credit scoring includes the scope, the population, the definition of bad credit risks, the profile of the applicant, statistical multivariate analysis, the algorithm and ethics.

The scope of credit scoring is to help granting credit to safe borrowers. To avoid bad credit risks, credit scoring uses empirics. Credit scoring algorithms estimate the future classification of applicants as good or bad credit risks using revealed applicants' profiles. Applicants belong necessarily to a homogeneous and massive population. The algorithm is derived using a statistical multivariate analysis technique that allows identifying risk factors and respective weights using a training sample that is composed of recent borrowers whose status as good or bad credit risk is known. The profiles of these borrowers are also known. The credit scoring assumption is that loan applicants will have a credit behavior like past borrowers with similar profiles. Due to the fact that past borrowers have already been screened by loan officers during approval, the population with known credit risk status is biased. If the credit scoring algorithm does not screen out a big share of applicants that were refused by loan officers, a sample of such applicants has to receive credit in order to obtain new experience on their credit behavior and incorporate it into a new algorithm. Before that, the current credit scoring algorithm should not be used as a sole screening method, but rather to complement the work of the loan officer. Algorithm's robustness is confirmed using a representative test sample and a sample of rejected applicants. The algorithm's economic significance is also estimated. If satisfactory, the credit scoring system is implemented and used in the process of screening loan applicants. Results of the use of credit scoring are regularly verified through post-implementation reports, unless the algorithm is able to learn constantly from newly available experience.

The population comprises subjects, individuals or businesses, which apply for loan products of the financial institution. The population should be homogeneous for statistical reasons and large for economical reasons. Equilibrium between these dimensions can be reached by adjusting the characteristics of available credit products and application policies.

The definition of bad credit risks represents the rule used to identify credit risks to be avoided. Certainly the profitability of the loan is the best indicator to classify it as good or bad credit risk, but delinquency serves as a good proxy used extensively in identifying bad risks.

The profile of the applicant represents the required information on the subject applying for a credit. This can accommodate five categories: socio-demographic variables (for personal and small business loans), business-demographic variables (for business loans, including small business), financial information, information about the required loan product, and past credit behavior.

Statistical multivariate analysis is at the core of credit scoring. Different techniques are proposed and all are empirical. This represents the big advantage over judgmental-intuitive reasoning of loan offices and credit rating tools.

The algorithm tells the user, human or machine, how to proceed to estimate the classification of the applicant as good or bad credit risk. The accuracy of the estimation is known. The algorithm includes the precise rules for accepting or rejecting the applicant. In non-parametric algorithms the cut-off is predetermined, while in parametric algorithms the cut-off is fixed based on the scoring distribution and economic calculations, such as evaluation of misclassification costs.

The ethics of using credit scoring may raise legal concerns about the information used by the algorithm or moral concerns about the refusal of credit to otherwise creditworthy applicants who are considered bad by credit scoring. In microfinance the opportunity cost of not serving a good client is much higher than in most other credit segments. Similarly, over indebting a micro borrower has high opportunity costs. Credit scoring replaces the subjective judgement of the loan officers with empirics, thus reducing ethical issues.

Conclusions

Our review shows that outlines of the conceptual framework of credit scoring date back to late 1930s in the USA. In the post-Great Depression period, the practice of using judgmental rating systems based on the experience and intuition of loan officers was not uncommon. In this context Durand pioneered the use of statistical methods in identifying credit risk factors in consumer financing. His research set up the pillars of the framework of the credit scoring technique and opened the way for further research.

The resemblance between the debate described by Durand and current post-crisis debates in microfinance reveals important findings for microfinance institutions. This literature review and the enclosed analysis of actions and reactions that led to the current framework of credit scoring provide many lessons for the microfinance sector.

In 1958, FICO registered the sale of the first credit scoring system. Currently 95 percent of the largest financial institutions in the US are FICO clients (Fair Isaac Corporation, 2017). Nevertheless, the commercial success of FICO did not benefit the academic literature until recent publications (Oliver & Wells, 2001; Hand & Kelly, 2002; Zhu *et al.*, 2002) brought to light the results of its best practices and research findings.

Today, credit scoring is not in widespread use in microfinance, just as it was not in extensive use in the USA in the early 1960s, although the traits of the conceptual framework of modern consumer credit scoring were established and publicly available by 1964. With the expansion of the technique in consumer lending, Altman was the first to extend the use of credit scoring to corporate financing.

The Consumer Credit Protection Act of 1968 added legitimacy to credit scoring in the USA, which by the beginning of the 1970s had become a recognized industry. The Equal Credit Opportunity Act of 1974 marked a new stage in the history of the conceptual framework of credit scoring by prohibiting discrimination in credit granting practices on the basis of identities such as ethnicity and gender. Incorporation of such attributes in credit scoring algorithms was forbidden. To the extent that these factors were statistically significant, their exclusion reduced the accuracy of bad credit risk estimation, unless new risk factors could either compensate the lost accuracy or act as proxies for outlawed factors.

Along with the dynamic expansion of credit scoring in consumer lending, small business financing appeared as the next credit segment with potential for the diffusion of the technique. However, fears that commercial lending was not as homogeneous as consumer credit kept academics and practitioners away. It was only in 1995 that FICO in partnership with Robert Morris Associates started to offer pooled data credit scoring for small businesses. This technique allowed a dynamic algorithm construction based on clusters of good and bad companies similar to the SME to be scored.

By the end of the century in the developed world the credit scoring technique was commonly used in granting consumer and small business credit. Simultaneously academics and practitioners started looking for credit risk evaluation techniques for developing countries. Viganò's work proposing the application of credit scoring for evaluating small and micro firms in Burkina Faso was a pioneering development in the literature. Schreiner's work is most significant in extending the use of credit scoring to microfinance.

The consolidated conceptual framework of the credit scoring technique, which we developed on the basis of our critical review of literature, can be applied to personal, retail, mortgage, micro, small, medium, corporate and other new loan markets. This framework can enable both practitioners and academics to address their different approaches to the subject of credit scoring, especially the subject of the application of the credit scoring technique in microfinance.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23,* 589–609.

Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation.* Oxford: Oxford University Press.

Ang, J.S., Chua, J.H., & Bowling, C.H. (1979). The profiles of late-paying consumer loan borrowers: an exploratory study. *Journal of Money, Credit and Banking*, *11*, 222–226.

Armendáriz, B., & Morduch, J. (2005). *The Economics of Microfinance.* Cambridge: The MIT Press.

Balke, D. (2005). The Applicability of Empirical Credit Models in Developing and Transition Countries. *LFS Discussion Papers*. Berlin: LFS Financial Systems GmbH.

Boggess, W.P. (1967). Screen–test your credit risks. *Harvard Business Review*, *45*, 113–122.

Bolton, R.J., & Hand, D.J. (2002). Statistical fraud detection: a review. *Statistical Science*, *17/3*, 235–249.

Bumacov, V. (2012). Mission drift in micro lending: how the joint use of credit and poverty scoring can help MFIs get back on track. *Cost Management*, *26/5*, 29–36.

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

Capon, N. (1982). Credit scoring systems: a critical analysis. *Journal of Marketing*, *46*, 82–91.

Chapman J.M., *et al.* (1940). Commercial Banks and Consumer Instalment Credit. New York: National Bureau of Economic Research.

Constantinou, D., & Ashta, A. (2010). Financial crisis: lessons from microfinance. *Strategic Change: Briefings in Entrepreneurial Finance*, *20/5-6*, 187–203.

Crook, J.N., Hamilton, R., & Thomas, L.C. (1992). A Comparison of a Credit Scoring Model with a Credit Performance Model. *Service Industries Journal*, *12/4*, 558–579.

Doing Business. (2016). Getting credit. Retrieved from http://www.doingbusiness.org/data/exploretopics/getting-credit.

Dowla, A., & Barua, D. (2006). *The poor always pay back: the Grameen II story.* Bloomfield: Kumarian Press.

Durand, D. (1941). *Risk Elements in Consumer Instalment Financing.* New York: National Bureau of Economic Research.

Edmister, R.O., & Schlarbaum, G.G. (1974). Credit policy in lending institutions. *Journal of Financial & Quantitative Analysis*, *9*, 335–356.

Eisenbeis, R.A. (1978). Problems in applying discriminant analysis in credit scoring models. *Journal of Banking and Finance*, *2*, 205–219.

Fair Isaac Corporation. (n.d.). About us. Retrieved from http://www.fico.com/en/about-us.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, *7*, 179–188.

Greer, C.C. (1967). The optimal credit acceptance policy. *Journal of Financial & Quantitative Analysis*, *2/4*, 399–415.

Hand, D.J., & Henley, W.E. (1993). Can reject inference ever work? *IMA Journal of Mathematics Applied in Business & Industry*, *4/1*, 45–55.

Hand, D.J., & Kelly, M.G. (2002). Superscorecards. *IMA Journal of Management Mathematics*, *13/4*, 273–281.

Harriss-White, B. (2003). *India Working: Essays on Society and Economy*. Cambridge: Cambridge University Press.

Hassler, H.W., Myers J.H., & Seldin, M. (1963). Payment history as a predictor of credit risk. *Journal of Applied Psychology*, *47/6*, 383–385.

Häuβler, W. M. (1979). Empirische Ergebnisse zu Diskriminationsverfahren bei Kreditscoringsystemen. *Zeitschrift für Operations Research*, *23/8*, B191–B210.

Holsapple, C.W., Tam K.Y., & Whinston, A.B. (1988). Adapting expert system technology to financial management. *Journal of the Financial Management Association*, *17*, 12–22.

Kortenbusch, M., & Hauser, P. (2010). Credit Scoring in Microfinance – an Unfolding Story. *The Euromoney Environmental Finance Handbook 2010.* London: Euromoney Institutional Investor PLC.

Maes, J.P., & Reed, L.R. (2012). *State of the Microcredit Summit Campaign Report 2012.* Washington, D.C.: Microcredit Summit Campaign.

Muchinsky, P.M. (1975a). A comparison of three analytic techniques in predicting consumer installment credit risk: toward understanding a construct. *Personnel Psychology*, *28*, 511–524.

Muchinsky, P.M. (1975b). Consumer installment credit risk: a need for criterion refinement and validation. *Journal of Applied Psychology*, *60/1*, 87–93.

Meyers, J.H., & Forgy, E.W. (1963). Development of numerical credit evaluation systems. *Journal of American Statistical Association*, *50*, 797–806.

National Commission on Consumer Finance. (1972). *Consumer credit in the United States: the report of the National Commission on Consumer Finance.* Washington, D.C.: Government Printing Office.

Nevin, J.R., & Churchill, G.A. (1979). The Equal Credit Opportunity Act: an evaluation. *Journal of Marketing*, *43/2*, 95–104.

Oliver, R.M., & Wells, M. (2001). Efficient frontier cutoff policies in credit portfolios. *Journal of the Operational Research Society*, *52/9*, 1025–1033.

Orgler, Y.E. (1970). A credit scoring model for commercial loans. *Journal of Money, Credit and Banking*, *2*, 435–445.

Plummer, W.C., & Young, R.A. (1940). *Sales Finance Companies and Their Credit Practices.* New York: National Bureau of Economic Research.

Rayo Canton, S., Lara Rubio, J., & Camino Blasco, D. (2010).Un modelo de credit scoring para instituciones de microfinanzas en el marco de Basilea II. *Journal of Economics, Finance and Administrative Science*, *15/28*, 89–124.

Reed, L.R. (2012). *State of the Microcredit Summit Campaign Report 2013*. Washington, D.C.: Microcredit Summit Campaign.

Reed, L.R., Marsden, J., Rogers, S., Rivera, C., & Ortega, A. (2014). *Resilience: State of the Microcredit Summit Campaign Report 2014.* Washington, D.C.: Microcredit Summit Campaign.

Rhyne, E., & Christen, R.P. (1999). *Microfinance Enters the Marketplace*. Washington, D.C.: USAID.

Rosenberg, R., Gonzalez, A., & Narain, S. (2009). The New Moneylenders: Are the Poor Being Exploited by High Microcredit Interest Rates? In T. A. Watkins & K. Hicks (Eds.), *Moving Beyond Storytelling: Emerging Research in Microfinance*

(Contemporary Studies in Economic and Financial Analysis, Volume 92) (pp. 145–181). Bingley: Emerald Group Publishing Limited.

Rutledge, G. (1996). Taming the fraud monster. *Credit World*, *85/1*, 10–13.

Schreiner, M. (2000). Credit scoring for microfinance: can it work? *Journal of Microfinance, 2*, 105–118.

Schreiner, M. (2010). Seven extremely simple poverty scorecards. *Enterprise Development and Microfinance*, *21/2*, 118–136.

Sexton, D.E. Jr. (1975). Credit riskiness of low–income consumers. *Advances in Consumer Research*, *2*, 197–202.

Simbaqueba, L., Salamanca, G., & Bumacov, V. (2011). The role of credit scoring in micro lending. In A. Ashta (Ed.) *Advanced Technologies for Microfinance*: *Solutions and Challenges*. Hershey: IGI Global.

Smith, P.F. (1964). Measuring risk on consumer instalment credit. *Management Science*, *11*, 327–340.

Tabor, J.S., & Bowers, J.S. (1977). Factors determining the credit worthiness of low income consumers. *Journal of Consumer Affairs*, *11*, 43–52.

Thomas, L.C., Edelman, D.B., & Crook, J.N. (2002). *Credit scoring and its applications*. Philadelphia: Society for Industrial and Applied Mathematics.

Thrasher, R.P. (1992). A note on the use of the K-S statistic as a measure of model strength. *Journal of Direct Marketing*, *6/4*, 40–45.

Wai, U.T. (1957). Interest rates outside the organized money markets of underdeveloped countries. *Staff Papers (International Monetary Fund), 6/1*, 80–142.

U.S. Senate. (1979). *Credit Card Redlining: Hearings before the Subcommittee on Consumer Affairs of the Committee on Banking, Housing, and Urban Affairs, United States Senate.* Washington, D.C.: Government Printing Office.

Viganò, L. (1993). A credit scoring model for development banks: an African case study. *Savings and Development*, *17*, 441–479.

Weingartner, H.M. (1966). Concepts and utilization of credit-scoring techniques. *Banking*, *58*, 51–54.

Young, R.A., *et al.* (1940). *Personal Finance Companies and Their Credit Practices.* New York: National Bureau of Economic Research.

Yunus, M. (2003). *Banker to the Poor: Micro-Lending and the Battle against World Poverty.* New York: PublicAffairs.

Zhu, H., Beling, P.A., & Overstreet, G.A. (2002). A Bayesian framework for the combination of classifier outputs. *Journal of the Operational Research Society*, *53/7*, 719–727.

Vitalie Bumacov is conducting a doctoral research at the Oxford Brookes Business School, Oxford Brookes University, UK, on the use of poverty scoring and credit scoring in microfinance. He is an entrepreneur and consultant in access to small finance.

Arvind Ashta is professor of accounting, finance and law at CEREN, EA 7477, Burgundy School of Business - Université Bourgogne Franche-Comté in Dijon. He has taught Microfinance as visiting faculty in Brussels, Chicago, New Delhi, Pforzheim and Winterthur, and provided student seminars in Barcelona, Hertfordshire, Kathmandu, Mysore and Nancy. He is a member of research associations CERMi, ISEE and RRI. He is part of the European Microfinance Platform – the European association of professionals and researchers in microfinance.

Pritam Singh is a professor in economics at the department of accounting, finance and economics at the Oxford Brookes Business School, Oxford Brookes University, UK. His area of specialization is development and environmental economics. He is Federalism Advisor to the Government of Punjab India and has been Economic Advisor to the British House of Commons' Parliamentary Group on Punjabis in Britain. He has been a visiting professor at Jawaharlal Nehru University in Delhi and Moscow State University.

Correspondence to:
Vitalie Bumacov
Tudor Panfile 38, MD2049, Chisinau, Moldova
Email: bumacov@hotmail.com.