A COGNITION-DRIVEN RISK EVALUATION FRAMEWORK FOR CONSUMER LOANS

*Fernando A. F. Ferreira**
ISCTE Business School, BRU-IUL, University Institute of Lisbon,
Avenida das Forças Armadas, 1649-026 Lisbon, Portugal
&
Fogelman College of Business and Economics, University of Memphis,
Memphis, TN 38152-3120, USA
fernando.alberto.ferreira@iscte.pt or fernando.ferreira@memphis.edu

Marjan S. Jalali
ISCTE Business School, BRU-IUL, University Institute of Lisbon,
Avenida das Forças Armadas, 1649-026 Lisbon, Portugal
marjan.jalali@iscte.pt

Ieva Meidutė-Kavaliauskienė
Faculty of Business Management, Vilnius Gediminas Technical University,
Saulėtekio al. 11, LT-10223 Vilnius, Lithuania
&
BRU-IUL, University Institute of Lisbon,
Avenida das Forças Armadas, 1649-026 Lisbon, Portugal
ieva.meidutekavalauskiene@vgtu.lt

Edmundas K. Zavadskas
Faculty of Civil Engineering, Vilnius Gediminas Technical University,
Saulėtekio al. 11, LT-10223 Vilnius, Lithuania
edmundas.zavadskas@vgtu.lt

Sandra M. J. Catarino
School of Management and Technology, Polytechnic Institute of Santarém,
Complexo Andaluz, Apartado 295, 2001-904 Santarém, Portugal
smjcatarino@gmail.com

Unpublished Manuscript

Please do not quote without permission from the authors

* Corresponding author.

ABSTRACT

Credit to personal consumption is an important activity of the financial system and crucial to the socio-economic development of a country. It is important, therefore, that the methods and techniques used to evaluate consumer credit risk be as efficient and informative as possible, in order to strengthen decisions to approve or reject credit and promote sustainable economic growth. This study aims to create a multiple criteria expert system which integrates cognitive maps and the measuring attractiveness by a categorical based evaluation technique (MACBETH) to create a complementary framework for consumer credit risk assessment. The results show that this integrated approach allows the evaluation process of consumer credit risk to be more informed and transparent, providing value for the evaluation processes of this type of credit application as a result of the privileged contact established with a panel of credit analysts. Limitations and managerial implications are also discussed.

Keyword: Credit to Consumption; Risk Analysis; Cognitive Maps; MCDA.

INTRODUCTION

In wake of the recent global financial crisis, organizations from almost all sectors of economic activity have had to adjust to a new reality, in a bid to withstand and overcome the challenges created therefrom. Financial institutions, whose main function is to arbitrate financial resources, are no exception. By hosting savings in exchange for relatively low returns, and granting credit to households and businesses in exchange for higher returns, these institutions assume a leading role in a country's investment process. Indeed, it is generally accepted that the financial market is of great importance to country development (Mari and Renò, 2005; Ferreira *et al.*, 2014b).

According to Alcarva (2011), the banking business is based on three key pillars: security, profitability and liquidity; and its main objective it to align these three pillars. In the aftermath of the crisis, however, financial markets' approach to this alignment has changed. Financial institutions have had to reform their investment and lending policies, increasing restrictions on access to finance in a comprehensive and wide reaching manner. This has included access to personal consumption credit, despite recognition that such credit constitutes a driver to economic growth. Carvalho (2009: 32) notes that *"the contribution of loans to the smooth functioning of economies is such that whenever credit institutions, as a whole, adopt measures to constrain lending, economic growth is stunted"*.

In addition, more restrictive lending measures affect not only financial institutions and customers, but all intervening stakeholders, governmental authorities included. As such, the development of new and more rigorous risk evaluation methodologies is of interest not only to financial institutions providing credit, but to society at large.

Indeed, economic growth and increased competition within the financial sector have already led to organizational improvements, and to the enhancement and streamlining of financial institutions' credit evaluation and loan approval processes. Better scoring techniques for credit risk analysis have also been developed. Despite this progress, however, there is still much room for the development of more complete frameworks, able to support decision making and promote sustainable economic growth, by improving and/or innovating credit scoring techniques to minimize the risk of default.

In this context, the use of structuring techniques and multiple criteria evaluation methods seems particularly relevant, insofar as these approaches have been shown to be able to clarify complex problems (Belton and Stewart, 2002; Zavadskas and Turskis, 2011; Ferreira *et al.*, 2011). As such, this paper proposes and develops a valuation model based on weights that assess the risk of loan default, through the integrated use of cognitive maps and the measuring attractiveness by a categorical based evaluation technique (MACBETH). Our stance is complementary – aiming to add to, not replace – existing frameworks; and puts forward a multiple criteria system for the analysis of consumer credit risk aimed at minimizing some of the gaps in current approaches. The hope is that such an expert system can contribute to making the lending process more informed, transparent and robust.

The remainder of this article is structured as follows. The following section discusses the importance of lending, and presents some of the main limitations of existing valuation models. Section two presents our proposed methodology, which brings together cognitive maps with the MACBETH technique for the elaboration of a credit evaluation framework. Section three describes the procedures followed to develop and test our credit risk assessment model; and the final section presents our conclusions and suggestions for further research.

LITERATURE REVIEW

Personal consumption patterns have been changing in recent times, marked by a decrease in savings and an increase in the demand for credit (*cf.* Carvalho, 2009; Ferreira *et al.*, 2014b). According to Costa (2004), the liberalization of the banking system has contributed significantly to the changes that have been observed in recent years. Ideally, however, balanced solutions should be sought, bearing in mind the common interests of the client and of the bank – solutions that will solve citizens' problems, without jeopardizing the balance sheets and banks' ability to finance the economy.

Indeed, granting credit is a consequential decision, and should encompass a detailed examination of applications, based on predefined assumptions to determine the risk of default and the feasibility of the loan. Yu *et al.* (2007: 942), for instance, note that *"in credit risk evaluation, credit scoring is one of the key analytical techniques.* […] *credit scoring is a technique that helps some organizations, such as commercial banks and credit card companies, determine whether or not to grant credit to consumers, on the basis of a set of predefined criteria"*. Crook *et al.* (2007: 1447) define credit scoring as *"the assessment of the risk associated with lending to an organization or an individual"*; and in effect, a wide range of credit risk assessment systems have been developed and are used to support decision making in this area (for discussion, see Altman and Saunders, 1998; Doumpos *et al.*, 2002; Grunert *et al.*, 2005).

Although the origins of credit date back to ancient times, the assessment of credit only begins in the mid-twentieth century. Since then, there has been a continued effort to improve the tools and techniques used for this purpose, in order to make them increasingly impartial and effective in the classification and analysis of risk (Crook *et al.*, 2007). According to Wang *et al.* (2010: 223), *"the accuracy of credit scoring is critical to financial institutions' profitability. Even 1% of improvement on the accuracy of credit scoring of applicants with bad credit will decrease a great loss for financial institutions* […]*"*.

Credit scoring models use scoring systems that assign weights to different variables, to then arrive at a weighted sum for each credit application. As pointed out by Scarpel and Milioni (2002: 62), *"discriminant analysis has historically been the most commonly used quantitative method in determining the weights of the indices in credit scoring models.* […] *It is a statistical technique that allows for the study of differences between two or more groups, based on a set of information variables available for all elements of the groups"*. In parallel with this development, qualitative indicators were also advanced; and their combined use with quantitative indicators has been proposed to help to alleviate the risk associated with credit assessments (Avery *et al.*, 2004; Costa, 2004). Indeed, as Lopez and Saidenberg (2000: 152) note, *"over the past decade, banks have devoted many resources to developing internal risk models for the purpose of better quantifying the financial risks they face and assigning* *the necessary economic capital. These efforts have been recognized and encouraged by bank regulators".*

The scenario is thus of a continuous search for improvements and a constant demand for technically more accurate evaluation mechanisms; but with an acute awareness of the difficulty of obtaining these improved evaluation mechanisms, in particular in what pertains to the use of qualitative data. This has supported the call for new approaches (*cf.* Doumpos and Zopounidis, 2001; Zopounidis *et al.*, 2015), which could serve as starting points, both to boost the development of new methods and for the improvement of the existing ones.

A common limitation credit analysts face is bound with the size and complexity of the databases required to test a given model: *"one of the major problems for applying ML algorithms in credit risk prediction is the unavailability, scarcity and incompleteness* […] *of credit data"* (Twala, 2010: 3326). This is reinforced by Lopez and Saidenberg (2000: 152), who note that *"due to the nature of credit risk data, only a limited amount of historical data on credit losses is available and certainly not enough to span several macroeconomic or credit cycles. These data limitations create a serious difficulty for users' own validation of credit risk models".*

In addition, the methods themselves also present challenges. Classic statistical methods commonly used to develop credit scoring models or systems, such as linear regression, logit, probit, tobit, binary tree and the minimum method, have limitations associated to the non-linearity between variables and the sensitivity to deviations from initial assumptions (*cf.* Šušteršic *et al.*, 2009). Neural network models also have limitations. Šušteršic *et al.* (2009: 4738), for instance, note that *"one of the drawbacks of EBP ANN* [Evidence Based Practice Artificial Neural Networks] *is that it can be easily over-trained. Over-training appears if after a number of iterations, that are improving predictions on the training subset, the network starts yielding worse and worse predictions".*

Models incorporating or analyzing qualitative variables or weighting criteria, in turn, have limitations associated with the subjectivity of the data. However, when combined to quantitative weights, they can be important for minimizing default risk (see Costa, 2004). Another important limitation is the bias of the data when it is not predefined, but rather defined in terms of a goal, which influences the results. As Jacobson and Roszbach (2003: 633) argue, *"the choice of default definition matters for the VaR* [Value at Risk]-*measure".* In addition, the complexity of models can also be an important limitation, insofar as it can hinder the interpretation of results (*cf.* Avery *et al.*, 2004; Thomas, 2009 and 2010; Wang *et al.*, 2010).

Indeed, no approach is exempt from limitations. Notwithstanding, given that the two main methodological limitations of the more commonly used methods are bound with the identification of the assessment criteria, and the lack of transparency in obtaining weights, there seems to be room for the combined application of cognitive mapping and multiple criteria methodologies in order to make consumer credit assessment systems more robust.

METHODOLOGICAL BACKGROUND

The current study is based on the multiple criteria decision analysis (MCDA) approach (see Roy, 1985; Belton and Stewart, 2002), and the strategic options development and analysis (SODA) methodology (Ackermann and Eden, 2001). Bana e Costa *et al.* (1997: 34) propose that the decision aid process should be divided into three main stages: (1) structuring; (2) evaluation; and (3) making recommendations. According to the authors, structuring *"is an essential phase of MCDA, as it provides the actors involved in a problematic situation with a common language for debate and learning, and with clear information about the plausible impacts of potential actions on the different points of view, thus serving to make explicit the actors' value systems".*

The structuring of a problem can be focused on a decision maker's value system (*i.e.* value-focused thinking) or on the characteristics of choice alternatives (alternative-focused thinking) (see Keeney, 1994; 1996). Multiple criteria decision aids are humanistic in character, and see decision processes as based on the consequences of actions and the preferences of actors (Roy, 1985). As Bana e Costa *et al.* (1997) note, the interaction between these two subsystems and their complementarity allow the overall results of the decision-making process to be achieved in the evaluation phase. The interactive process between the actors allows for advances in solving decision problems which are based on their own objectives and value systems.

MCDA approaches thus present themselves as a relevant methodological option for the evaluation of credit requests; in particular due to their ability to reconcile objectivity and subjectivity, in some ways materializing the complexity of the current business context (Zavadskas and Turskis, 2011; Ferreira *et al.*, 2014a; Zopounidis *et al.*, 2015). Such an approach is likely to be particularly advantageous for the assessment of a complex issue like consumer credit risk, which deals with a wide range of intangible variables. Because it allows very complete and detailed analyses, it can also be expected to lead to more coherent and robust decisions. Thus, it becomes important to understand and clarify the contribution of SODA and MCDA in the process of building an evaluation system for the analysis of consumer credit risk.

Cognitive Mapping

To support the structuring of the decision problem, we used the SODA methodology. SODA, also called JOURNEY Making, was originally conceived, by Colin Eden, for the structuring of complex decision problems (Belton and Hodgkin, 1999; Tegarden and Sheetz, 2003; Eden and Ackermann, 2004; Ackermann, 2012). Based on cognitive mapping techniques, the approach allows ideas to be structured, and that structure to be visualized, which facilitates dialogue and collaboration between decision agents, as well as the reorganization of different ideas and/or perspectives.

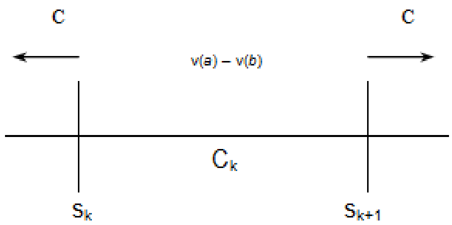
The cognitive maps used in the SODA methodology employ different techniques, including brainstorming and the listing of decision makers' goals and values (Ferreira, 2011). Cognitive maps are seen as structuring tools, whose main advantages stem from the fact that they allow a reduction in the rate of omitted criteria, and promote learning through discussion and analysis of how these same evaluation criteria relate to each other (Ferreira *et al.*, 2014b; Filipe *et al.*, 2015). Indeed, cognitive maps' ability to deal with the complexity and inherent subjectivity in decision making is widely recognized. Human cognition is defined as *"a complex process that results from the interaction between the motor-system and the neurological structures responsible for individuals' cognitive systems* [...]*"* (Ferreira, 2011: 123). In line with this, Eden (2004: 673) argues that *"a cognitive map is the representation of thinking about a problem that follows from the process of mapping"*. However, what truly characterizes a cognitive map is the aggregation of ideas, the ability to materialize thought and allow cause-and-effect relationships to be identified. In practice, the problem description is a consequence of the coming together of the decision makers, and of the information sharing which results from the development of cognitive maps.

The MACBETH Approach

Developed in the early 1990s by Carlos Bana e Costa and Jean-Claude Vansnick (see Bana e Costa and Vansnick, 1995; Bana e Costa *et al.*, 2012), the MACBETH approach is part of the MCDA domain. It is characterized as an interactive support tool for the construction of numerical interval scales, traditionally considered useful for defining weights among the criteria in a given valuation model. This approach has been gaining impact in relation to other multiple criteria methods (Ferreira *et al.*, 2015) because of the simplicity of its application to prioritization problems, which has also led to its increasing use in different areas of knowledge, such as health, economics, finance and operations management, among others (see Bana e Costa and Oliveira, 2002; Belton and Stewart, 2002; Bana e Costa *et al.*, 2006; Bana e Costa *et al.*, 2012; Ferreira *et al.*, 2012; Ferreira *et al.*, 2014b; Filipe *et al.*, 2015).

As Ferreira *et al.* (2014b: 9) refer, the MACBETH methodology is framed within the mathematical principles of Doignon and reports to the question of *"numerical representations of semi-orders for multiple thresholds"*. That is, given a point of view PV, the numerical representation of preferences is possible in a structure of m binary relations $[P^{(1)}, \ldots, P^{(k)}, \ldots, P^{(m)}]$ (where $P^{(k)}$ is a preference which is stronger the higher the value of k). The MACBETH procedure consists in associating to each element of X (where $X = \{a, b, \ldots, n\}$ is a finite set of n actions) a value x (resulting from v(.): $X \rightarrow R$), such that differences such as $v(a) - v(b)$ (where a is more attractive than b (*i.e.* a P b)) are as compatible as possible with the decision makers' judgments. As shown in *Figure 1*, for every pair of actions (a, b) allocated to a given category of difference of attractiveness C_k, the difference $v(a) - v(b)$ belong to the same interval, without overlapping (see Bana e Costa and Vansnick, 1995).

Figure 1 – Allocation of the difference of attractiveness $v(a) - v(b)$ to the category C_k



Source: Bana e Costa and Vansnick (1995).

In association with *Figure 1*, and in order to proceed with setting the intervals between consecutive difference of attractiveness categories, the next step consists of calculating the limits s_k, which can be interpreted as transition thresholds (Ferreira, 2011). Based on this assumption, and recalling the numerical representation of multiple semi-orders by constant thresholds, multiple semi-orders can easily be introduced, as long as preferences are represented by the values of function v and the s_k thresholds are considered in accordance with formulation (1):

$$a\ P^{(k)}\ b: \quad s_k < v(a) - v(b) < s_{k+1} \tag{1}$$

Given that the s_k thresholds are real positive values, the definition of the intervals between the semantic categories of attractiveness is made easier; because in fact, between the origin (*i.e.* $s_1 = 0$) and s_m, an infinite number of categories and limits could be set. It should be noted that, since a $\mathrm{P}^{(m)}$ b, it is always possible to add an additional preference level for the introduction of an action c, be it real or fictitious, as long as c is more attractive than b to a greater extent than a is more attractive than b (Ferreira, 2011).

Furthermore, as explained by Bana e Costa *et al.* (2005), the limits of the intervals should not be defined beforehand, but simultaneously with numerical value scores for the elements of X. In practice, the emphasis is on the idea that the whole process of building cardinal value scales should be interactive and developed in a simple and natural way, using the semantic categories of difference of attractiveness as shown in *Table 1*.

Table 1 – Semantic categories of difference of attractiveness

Category	Difference of Attractiveness
C_0	**Null** Difference of Attractiveness
C_1	**Very Weak** Difference of Attractiveness
C_2	**Weak** Difference of Attractiveness
C_3	**Moderate** Difference of Attractiveness
C_4	**Strong** Difference of Attractiveness
C_5	**Very Strong** Difference of Attractiveness
C_6	**Extreme** Difference of Attractiveness

As an example, if a decision maker considers an action *a* more attractive than *b*, and the difference between the two actions is weak, then $(a, b) \in C_2$. The design of an evaluation system should therefore be based on these semantic categories and, for consistency, formulations (2) and (3) (*cf.* Junior, 2008) should be analyzed taking into account the decision makers' value judgments:

$$\forall a,b \in X : v(a) > v(b) \Leftrightarrow aPb \tag{2}$$

$$\begin{aligned} &\forall k,k^* \in \{1,2,3,4,5,6\}, \forall a,b,c,d \in X \text{ com } (a,b) \in C_k \\ &\text{e} (c,d) \in C_{k^*} : k \geq k^* + 1 \Rightarrow v(a) - v(b) \geq v(c) - v(d) \end{aligned} \tag{3}$$

Then, linear programming is applied according to formulation (4) (*cf.* Junior, 2008) in order to generate an initial scale to be presented to the decision makers for discussion.

$$\begin{aligned}
&Min\, v(n) \\
&S.T.: \forall a,b \in X : aPb \Rightarrow v(a) \geq v(b) + 1 \\
&\qquad \forall a,b \in X : aIb \Rightarrow v(a) = v(b) \\
&\qquad \forall (a,b),(c,d) \in X, \text{ if the difference of attractiveness between} \\
&\qquad\qquad a \text{ and } b \text{ is bigger than between } c \text{ and } d \text{, then :} \\
&\qquad\qquad v(a) - v(b) \geq v(c) - v(d) + 1 + \delta(a,b,c,d) \\
&v(a^-) = 0 \\
&\text{where :} \\
&n \text{ is an element of } X \text{ so that } \forall a,b,c,... \in X : n\,(P \cup I)\, a,b,c,... \\
&a^- \text{ is an element of } X \text{ so that } \forall a,b,c,... \in X : a,b,c,...(P \cup I)\, a^- \\
&\delta(a,b,c,d) \text{ is the minimal number of categories of difference of attractiveness} \\
&\qquad \text{between the difference of attractiveness between } a \text{ and } b \text{ and the} \\
&\qquad \text{difference of attractiveness between } c \text{ and } d.
\end{aligned} \tag{4}$$

In practice, then, MACBETH is based on a direct question-answer logic, where pairs of actions are compared by panel members, and given a qualitative assessment of the difference of attractiveness between them. To this end, various value judgments matrices are filled in; and this process continues until a local preference scale is determined for each of the descriptors included in the model.

IMPLEMENTATION

Structuring Phase

According to Kim and Lee (1998: 303), *"knowledge engineering is one of the most important tasks in developing expert systems. One of the primary objectives* […] *is to develop a complete, consistent and unambiguous description of the knowledge base"*. From this premise, and in order to start the operational procedures of the structuring phase, a five-person panel was established with banking employees, who in their daily lives deal with assessments of consumer credit applications. The literature does not define an ideal number for decision maker panels in these methods, and the number can typically fluctuate between 5 and 12 individuals (*cf.* Belton and Stewart, 2002; Ferreira, 2011; Filipe *et al.*, 2015). In addition, it is worth bearing in mind that our study is process-oriented. This means that, with due adjustments, the process followed can work well with any group of (expert) participants. The group sessions were conducted by two facilitators (*i.e.* researchers), who also recorded the results. The structuring phase took place during two group work sessions, totaling 8 hours (*i.e.* four hours per session).

The first session began with a brief explanation of the methodology and the presentation of the trigger question. This kicked off the group discussion and allowed the use of the "post-its technique" (Ackermann and Eden, 2001), which identified the evaluation criteria and laid the foundation for the construction of a group cognitive map. The technique has simple rules: a post-it is used for each criterion, and where it has a negative cause-and-effect relationship, a (–) negative sign is placed in the upper right corner (*cf.* Ferreira, 2011). The decision makers, through discussion, and according to their own values and knowledge, filled in the post-its; and in the next phase, organized them into clusters, with the possibility of adding new criteria or eliminating no longer relevant ones, always left open. Once the criteria had been defined and grouped into clusters, the decision makers were asked to focus on each one of those clusters in turn. The aim was to analyze relations of causality or influence among the criteria, such that given each criterion's level of importance within a given cluster, hierarchies could be established among them.

Once the "post-its technique" phase had been completed, a group cognitive map was built with recourse to the *Decision Explorer* software. This served to aid the discussion on how the decision problem had been structured, and left open the possibility of changing the criteria and/or clusters, or even of restarting the entire process. Once the relevant criteria had been identified, hierarchically arranged within the clusters and the collective cognitive map had been obtained, the next stage of the structuring process consisted in analyzing the cognitive lines of the map, so that criteria which might be considered Fundamental Points of View (FPVs) could be identified.

Based on Keeney's (1996) methodological guidelines, specific areas of interest were identified, underlying the selection of the FPVs, namely: Bank's Credit Policy; Environmental Circumstances; Commercial Evaluation of the Loan Applicant; Personal Evaluation of the Loan Applicant; Application or Proposal; and Deal Breakers (see *Figure 2*).

Figure 2 – Collective cognitive map

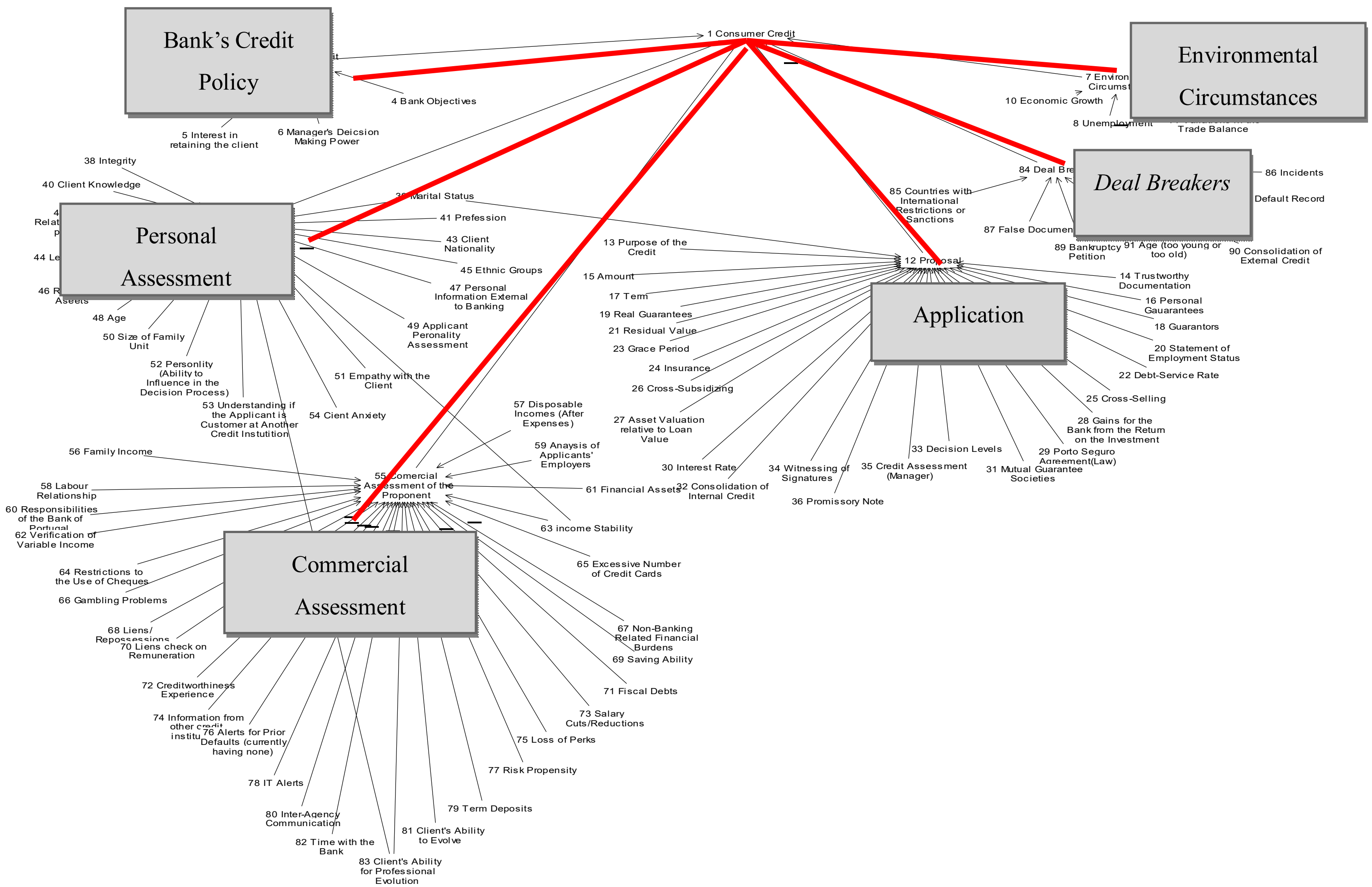
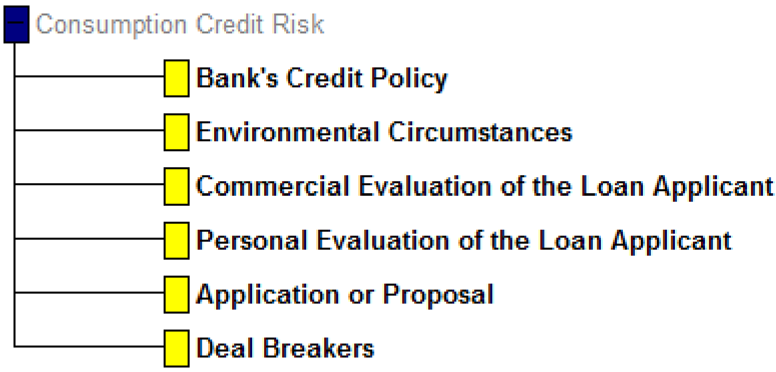


Figure 2 is representative of the most significant and/or relevant aspects for the assessment of the consumer credit process, and led to the tree of criteria presented in *Figure 3*, which was designed using the *M-MACBETH* software. Although the transition from the cognitive map to the value tree was not smooth due to the context-dependence of the process, it is worth noting that this dependence is arguably more than compensated by the amount of information discussed, the iterative nature of the procedures, and the direct involvement of the participants, which make cognitive mapping very valuable for the structuring and understanding of multi-faceted decision situations (*cf.* Ferreira *et al.*, 2015). It should be noted, in addition, that various tests were carried out in order to ensure mutual preferential independence between the comparison reference points identified in *Figure 3*.

Figure 3 – Tree of criteria



The subsequent step took place during the second group meeting and consisted in defining, for each FPV, a descriptor and its respective impact levels. This required a thorough analysis of the map and of the tree of criteria. For each cluster, the decision makers pointed out the criteria they considered most relevant for the assessment of consumer credit applications and, through an adapted version of Fiedler's (1965; 1967) least preferred co-worker (LPC) scale, proceeded to define partial performance levels, as well as the reference levels to be used for each descriptor. For instance, impact level L_1 expresses a partial performance considered "excellent", while L_n represents the worst performance level possible. *Figure 4* shows one of the descriptors and the impact levels identified for it.

Figure 4 – Descriptor and impact levels for FPV1

Descriptor PVF01 - Bank's Credit Policy (BCP)			Level	Description
Overly ambitious goals	1 2 3 4 5 6 7 8	Unambitious goals	L1	BCP Index ∈ [28-32]
Manager with decisive decision-making	1 2 3 4 5 6 7 8	Manager without any decision-making power	Good	BCP Index ∈ [21-27]
Absence of interest in customer retention	1 2 3 4 5 6 7 8	Extreme interest in customer retention	Neutral	BCP Index ∈ [16-20]
Absence of customer segmentation	1 2 3 4 5 6 7 8	Very marked customer segmentation	L4	BCP Index ∈ [4-15]

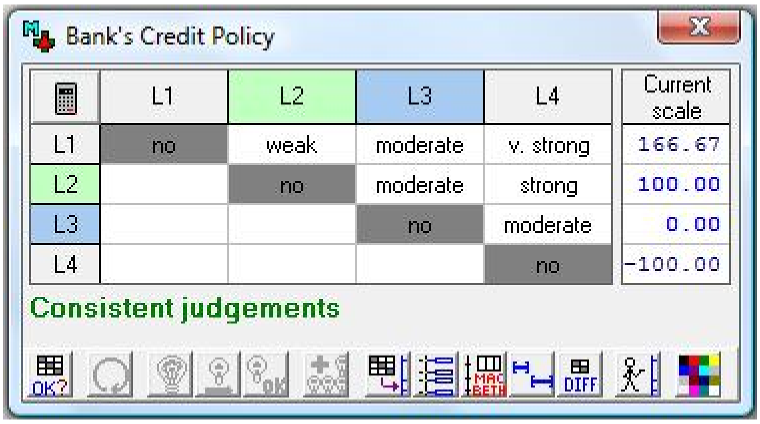
As *Figure 4* shows, the descriptor for FPV_1 allows for the analysis of the goals initially set by the bank in terms of loans, segmentation, customers and managerial decision-making power. According to the logic of the descriptors, the L_1 impact level represents an excellent performance, with the highest possible values for the assessment and approval of the credit. The L_4 impact level, on the other hand, reflects a negative assessment of the loan application, obtaining only the minimum possible values in the application assessment. This procedure was repeated for the remaining five FPVs.

Evaluation Phase

In the evaluation phase, judgment matrices are filled in, with the aim of obtaining local preference scales for each of the descriptors developed. In the current study, defining the local preference scales was of extreme importance, because it made it possible to estimate the partial performance of consumer credit applications according to each FPV. Underlying the construction of these preference scales was the MACBETH methodology, applied during the third group work session. The main advantage of using the MACBETH method came from the possibility of creating numerical scales based on semantic judgments (see Ferreira, 2011). The matrices of value judgments were filled in based on the semantic categories of differences of attractiveness shown in *Table 1*, and this gave rise to the different levels of impact for each FPV.

As *Figure 5* shows, FPV_1 was operationalized using four reference levels; and the application of the MACBETH methodology yielded a value function that assigned a classification of 166.67 points to the highest level (L_1), and -100 points to the worst level (L_4). L_2 was assessed by the decision makers as the level "Good" with a rating of 100 points; and L_3 was considered as the level "Neutral", having obtained a rating of 0 (zero) points. It should be noted that these two levels ("Good" and "Neutral") served as "anchors" to facilitate cognitive comparisons (see Filipe *et al.*, 2015). The procedure was repeated for the remaining five FPVs.

Figure 5 – Judgment values and proposed value scales for FPV_1



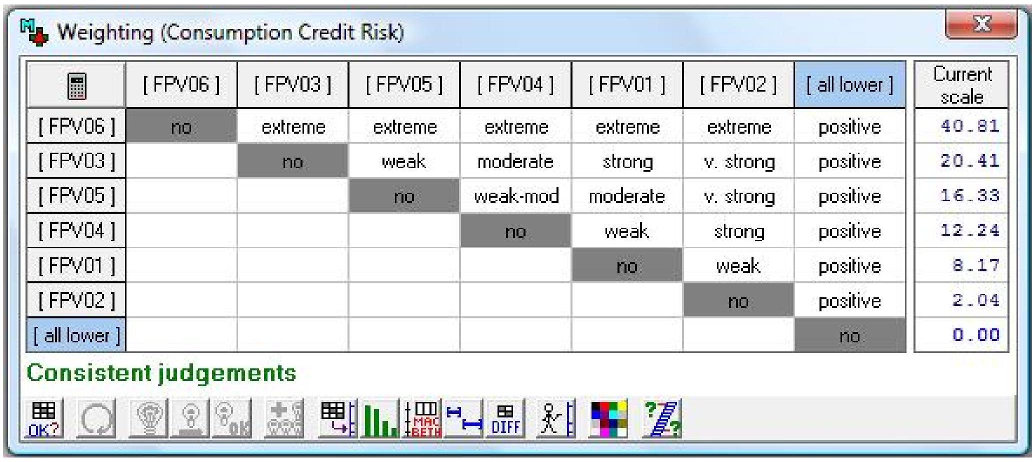
Having completed the local performance matrices, and ordered them based on their degree of relevance, another matrix was completed, in which the decision makers assigned a value of "1" to a FPV when it was considered preferentially more relevant than the others (and "0" otherwise). *Figure 6* presents the outcome of this sorting process.

Figure 6 – FPV ordering matrix

		FPV01	FPV02	FPV03	FPV04	FPV05	FPV06	Total	R
Bank's Credit Policy	FPV01		1	0	0	0	0	1	5
Environmental Circumstances	FPV02	0		0	0	0	0	0	6
Commercial Evaluation of the Loan Applicant	FPV03	1	1		1	1	0	4	2
Personal Evaluation of the Loan Applicant	FPV04	1	1	0		0	0	2	4
Application or Proposal	FPV05	1	1	0	1		0	3	3
Deal Breakers	FPV06	1	1	1	1	1		5	1

Once the FPVs had been ordered, the next step was to complete a trade-offs matrix, which allowed the differences in overall attractiveness between FPVs in the model to be circumscribed, using the *M-MACBETH* software. *Figure 7* shows this matrix and the normalized weights obtained for the FPVs in the model under study.

Figure 7 – Matrix of value judgments for trade-offs calculation



	[FPV06]	[FPV03]	[FPV05]	[FPV04]	[FPV01]	[FPV02]	[all lower]	Current scale
[FPV06]	no	extreme	extreme	extreme	extreme	extreme	positive	40.81
[FPV03]		no	weak	moderate	strong	v. strong	positive	20.41
[FPV05]			no	weak-mod	moderate	v. strong	positive	16.33
[FPV04]				no	weak	strong	positive	12.24
[FPV01]					no	weak	positive	8.17
[FPV02]						no	positive	2.04
[all lower]							no	0.00

It should be highlighted that calculating the trade-offs is critical to the implementation of the additive model presented in formulation (5), allowing an overall score for each process being evaluated to be obtained:

$$V(a) = \sum_{i=1}^{n} x_i v_i(a) \quad with \quad \sum_{i=1}^{n} x_i = 1 \quad and \quad x_i > 0 \quad and \quad \begin{cases} v_i(good_i) = 100 \\ v_i(neutral_i) = 0 \end{cases} \qquad (5)$$

This additive model aggregates the partial scores $v_i(a)$, considering the respective weights x_i, and allowing an overall score $V(a)$ for each credit application to be calculated. Technically, $v_i(good_i)$ and $v_i(neutral_i)$ stand for the partial scores of two specific impact levels (*i.e.* good and neutral), included in the model to facilitate cognitive comparisons (see Ferreira *et al.*, 2012 and 2014b; Filipe *et al.*, 2015).

Results and Recommendations

To validate the model developed, a practical application was required, as well as additional analyses to strengthen the results obtained. To this end, information on consumer credit applications (henceforth "Alphas") was needed, and was obtained by a bank employee, under conditions of absolute anonymity and confidentiality. Partial evaluations were obtained for each of the loan application; and these partial assessments were then aggregated using the additive aggregation model presented in formulation (5). *Figure 8* shows the partial and overall scores for the 10 consumer credit applications assessed.

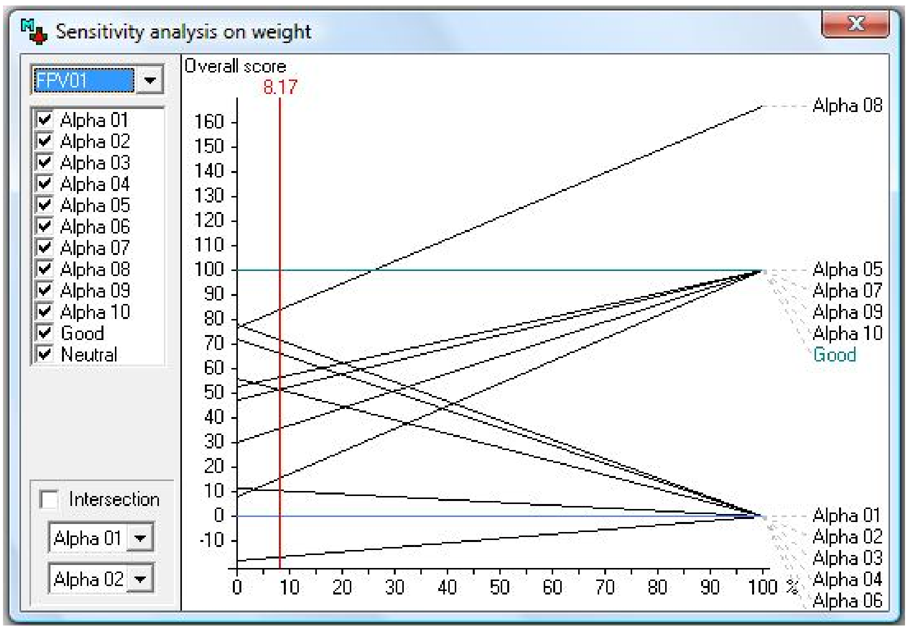
Figure 8 – Partial and overall attractiveness values for the consumer credit applications evaluated

Table of scores

Options	Overall	FPV01	FPV02	FPV03	FPV04	FPV05	FPV06
Alpha 01	66.32	0.00	-150.00	0.00	100.00	100.00	100.00
Alpha 02	71.42	0.00	100.00	0.00	100.00	100.00	100.00
Alpha 03	51.36	0.00	-150.00	66.67	0.00	0.00	100.00
Alpha 04	10.55	0.00	-150.00	66.67	0.00	0.00	0.00
Alpha 05	15.31	100.00	-150.00	-50.00	-100.00	-50.00	100.00
Alpha 06	-16.67	0.00	-150.00	-66.67	0.00	0.00	0.00
Alpha 07	35.72	100.00	-150.00	-50.00	0.00	0.00	100.00
Alpha 08	84.02	166.67	-150.00	100.00	100.00	0.00	100.00
Alpha 09	51.36	100.00	-150.00	66.67	0.00	-50.00	100.00
Alpha 10	56.46	100.00	100.00	66.67	0.00	-50.00	100.00
Good	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Neutral	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Weights :		0.0817	0.0204	0.2041	0.1224	0.1633	0.4081

According to Ferreira (2011), for the model to be validated by the panel of decision makers, its sensitivity to possible changes in the weights of the observed FPVs needs to be tested. Thus, using the *M-MACBETH* software, we carried out a number of sensitivity analyses, which allowed the overall performance of consumer credit applications in the model to be determined, in light of variations in the weights of FPVs. For illustrative purposes, *Figure 9* displays the sensitivity analysis performed for FPV_1, whose weight was defined at 8.17%.

Figure 9 – Sensitivity analysis for FPV_1



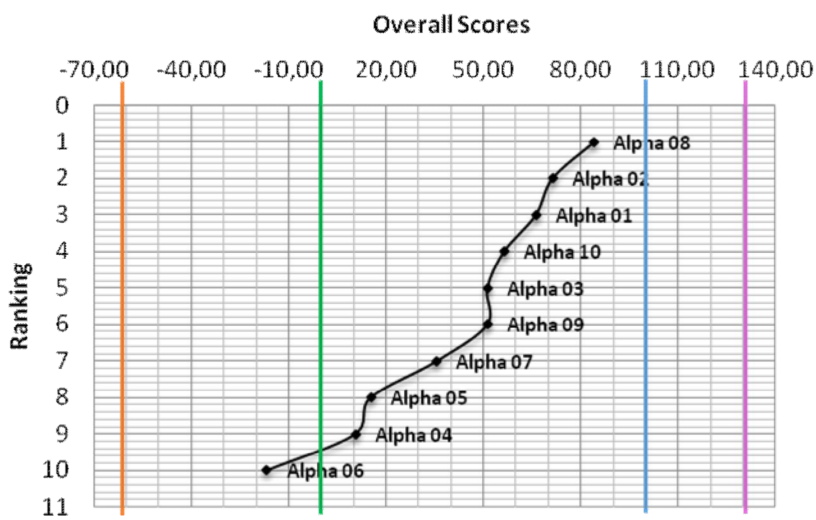
At the end of the final group session, spreads intervals were defined by the decision makers, on the basis of the default risk of each of the consumer credit applications in the study. The spread applicable to a given loan was determined according to the risk of default, and taking into account the results of the global assessments according to the four cognitive references (*i.e.* "Great", which brings together the best levels of FPVs; "Good", "Neutral", and "Terrible"). *Table 2* shows the spread thresholds.

Table 2 – Spread thresholds

Spread Projections		
Alpha ID	**Overall Score**	**Spreads**
Great	131.236	[2%; 4%]
Good	100.000	]4%; 9%]
Neutral	0.000	]9%; 16%[
Terrible	-61.566	≥ 16%

Figure 10 shows the position occupied by each of the personal consumption credit applications, within the ranges defined by the decision makers in *Table 2*. Those in the area on the right of the pink line (1st bound on the right) are in the "zone of excellence"; the area between the pink and the blue lines (1st and 2nd bounds on the right) is considered a "good zone"; the area between the blue line and the green line (2nd and 3rd bounds on the right) is considered "acceptable"; while the area between the green line and the red line (1st and 2nd bounds on the left) is "acceptable under certain conditions".

Figure 10 – Ranking of the consumer credit applications



The model developed thus allows personal consumption credit risk to be analyzed, based on the convictions and expert knowledge of specialists within banking and risk assessment in consumer credit. However, despite these very promising results, it is worth noting that, as in all models of evaluation, limitations have been identified, associated to the subjectiveness and context-dependence of the methodologies used (*i.e.* SODA and MACBETH). In this sense, the work in this paper should be considered as an instrument of negotiation and learning, rather than a quest for optimal decisions. The system developed in this study takes on contextual characteristics, and as such, generalizations from it, its implementation or extrapolation should be carried out with the necessary adaptations.

CONCLUSION

This paper makes clear that the integrated use of cognitive maps with the MACBETH approach allows for the design of multiple criteria models for more informed and transparent personal consumption credit risk assessment processes. In order to achieve this broad objective, a constructivism-based logic was followed.

Given that there are no perfect methods in risk assessment, the design of a multiple criteria system in this field, to complement and add to existing models, proved of great importance in contributing to minimize some of the gaps present in current approaches. With the application of the chosen methodology, a multiple criteria expert system was designed for the support of decision making in personal consumption credit risk assessments, which is relatively well-informed and robust. It seems clear that multiple criteria methodologies, especially those within the MCDA approach, by aggregating the experiences of decision makers, present great potential in the pursuit of more transparent and realistic valuation models. This, in turn, provides returns resulting from more accurate credit risk assessments, with benefits for bank branches, customers and, as a result, the economy a whole.

Because of the constructivist nature of our study, however, the framework developed is endowed with idiosyncratic characteristics, such that its results should not be extrapolated without caution. In this sense, the aim was not to achieve a unique optimization model, but rather to adopt a complementary (more so than comparative) perspective, contributing to the promotion of new methodologies, which, on the basis of discussion and negotiation, might improve consumption credit allocation decisions.

As for future research, it would be of interest to perform comparisons using different consumption credit-scoring systems; and include sensitivity and robustness analyses to identify which model or technique can provide more reliable risk assessments. Because the focus of our study was on the integration of cognitive mapping with the MACBETH approach in order to design robust consumer credit assessment systems, detailed comparisons with other evaluation systems were beyond the scope of this paper. Improvements to the expert system developed and presented here would also be of great interest, as would be the development of a software application that might facilitate the implementation of the proposed procedures. Given the largely unexplored potential in this field, any advance should be welcomed.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the superb contribution and infinite willingness of the panel of credit analysts: Ana Cunha, António Neves, Nuno Torres, Patrícia Oliveira and Paula Pereira.

REFERENCES

Ackermann, F. & Eden, C. 2001. SODA – journey making and mapping in practice. In: Rosenhead, J. & Mingers, J. (eds.), *Rational analysis for a problematic world revisited: problem structuring methods for complexity, uncertainty and conflict* (2nd Ed.), John Wiley & Sons, Chichester, 43-60.

Ackermann, F. 2012. Problem structuring methods ‘in the dock’: Arguing the case for soft OR. *European Journal of Operational Research*, 219(3), 652-658.

Alcarva, P. 2011. A Banca e as PME: Como gerir com eficácia o relacionamento entre as PME e a Banca. *Vida Económica Editorial*, Porto.

Altman, E. & Saunders, A. 1998. Credit risk measurement: Developments over the last 20 years, *Journal of Banking & Finance*, 21(11/12), 1721-1742.

Avery, R.; Calem, P. & Canner, G. 2004. Consumer credit scoring: Do situational circumstances matter? *Journal of Banking & Finance*, 28(4), 835-856.

Bana e Costa, C. & Oliveira, R. 2002. Assigning priorities for maintenance, repair and refurbishment in managing a municipal housing stock. *European Journal of Operational Research*, 138(2), 380-391.

Bana e Costa, C. & Vansnick, J. 1995. Uma nova abordagem ao problema da construção de uma função de valor cardinal: MACBETH. *Investigação Operacional*, 15(1), 15-35.

Bana e Costa, C., Stewart, T. & Vansnick, J. 1997. Multicriteria decision analysis: Some thoughts based on the tutorial and discussion sessions of the ESIGMA meetings. *European Journal of Operational Research*, 99(1), 28-37.

Bana e Costa, C.; Corte, J. & Vansnick, J. 2012. MACBETH, *International Journal of Information Technology & Decision Making*, 11(2), 359-387.

Bana e Costa, C.; De Corte, J. & Vansnick, J. 2005. On the mathematical foundations of MACBETH. In Figueira, J.; Greco, S. & Ehrgott, M. (Eds.), *Multiple Criteria Decision Analysis: The State of the Art Surveys*, Springer, New York, 409-442.

Bana e Costa, C.; Fernandes, T. & Correia, P. 2006. Prioritisation of public investments in social infrastructures using multicriteria value analysis and decision conferencing: A case study. *International Transactions in Operational Research*, 36(3), 279-297.

Belton, V. & Stewart, T. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*, Dordrecht: Kluwer Academic Publishers.

Belton, V. & Hodgkin, J. 1999. Facilitators, decision makers, D.I.Y. users: Is intelligent multicriteria decision support for all feasible or desirable? *European Journal of Operational Research*, 113(2), 247-260.

Carvalho, P. 2009. *Fundamentos da Gestão de Crédito*, Lisbon: Edições Sílado.

Costa, C. 2004. Estratégias bancárias e a avaliação do risco de crédito, *Proceedings of the Portuguese-Spanish Conference on Scientific Management*, Azores, 4-7 February, 1090-1097.

Crook, J.; Edelman, D. & Thomas, L. 2007. Recent developments in consumer credit risk assessment. *European Journal of Operational Research*, 183(3), 1447-1465.

Doumpos, M. & Zopounidis, C. 2001. Assessing financial risks using a multicriteria sorting procedure: The case of country risk assessment. *Omega – The International Journal of Management Sciences*, 29(1), 97-109.

Doumpos, M.; Kosmidou, K.; Baourakis, G. & Zopounidis, C. 2002. Credit risk assessment using a multicriteria hierarchical discrimination approach: A comparative analysis. *European Journal of Operational Research*, 138(2), 392-412.

Eden, C. & Ackermann, F. 2004. Cognitive mapping expert views for policy analysis in the public sector. *European Journal of Operational Research*, 152, 615-630.

Eden, C. 2004. Analyzing cognitive maps to help structure issues or problems. *European Journal of Operational Research*, 159(3), 673-686.

Ferreira, F. 2011. *Multiple Criteria Evaluation of Bank Branches: Decision Analysis Models and Applications*, Faro: University of Algarve Press.

Ferreira, F.; Jalali, M.; Meidutė-Kavaliauskienė, I. & Viana, B. 2015. A metacognitive decision making based-approach to bank customer loyalty measurement and management. *Technological and Economic Development of Economy*, 21(2), 279-299.

Ferreira, F.; Santos, S. & Dias, V. 2014a. An AHP-based approach to credit risk evaluation of mortgage loans. *International Journal of Strategic Property Management*, 18(1), 38-55.

Ferreira, F.; Santos, S. & Rodrigues, P. 2011. From traditional operational research to multiple criteria decision analysis: Basic ideas on an evolving field. *Problems and Perspectives in Management*, 9(3), 114-121.

Ferreira, F.; Santos, S.; Rodrigues, P. & Spahr, R. 2014b. Evaluating retail banking service quality and convenience with MCDA techniques: A case study at the bank branch level. *Journal of Business Economics and Management*, 15(1), 1-21.

Ferreira, F.; Spahr, R.; Santos, S. & Rodrigues, P. 2012. A multiple criteria framework to evaluate bank branch potential attractiveness. *International Journal of Strategic Property Management*, 16(3), 254-276.

Fiedler, F. 1967. *A Theory of Leadership Effectiveness*, New York: McGraw-Hill.

Fiedler, F. 1965. Engineer the job to fit the manager. *Harvard Business Review*, 43(5), 115-122.

Filipe, M.; Ferreira, F. & Santos, S. 2015. A multiple criteria information system for pedagogical evaluation and professional development of teachers. *Journal of the Operational Research Society*, 66(11), 1769-1782.

Grunert, J.; Norden, L. & Weber, M. 2005. The role of non-financial factors in internal credit ratings. *Journal of Banking & Finance*, 29(2), 509-531.

Jacobson, T. & Roszbach, K. 2003. Bank lending policy, credit scoring and value-at-risk. *Journal of Banking & Finance*, 27(4), 615-633.

Junior, H. 2008. Multicriteria approach to data envelopment analysis. *Pesquisa Operacional*, 28(2), 231-242.

Keeney, R. 1994. Creativity in decision making with value-focused thinking. *MIT Sloan Management Review*, 35(4), 33-41.

Keeney, R. 1996. Value-focused thinking: Identifying decision opportunities and creating alternatives. *European Journal of Operational Research*, 92(3), 537-549.

Kim, H. & Lee, K. 1998. Fuzzy implications of fuzzy cognitive map with emphasis on fuzzy causal relationship and fuzzy partially causal relationship. *Fuzzy Sets and Systems*, 97(3), 303-313.

Lopez, J. & Saidenberg, M. 2000. Evaluating credit risk models. *Journal of Banking & Finance*, 24(1), 151-165.

Mari, C. & Renò, R. 2005. Credit risk analysis of mortgage loans: An application to the Italian market. *European Journal of Operational Research*, 163(1), 83-93.

Roy, B. 1985. *Méthodologie Multicritère d'Aide à la Décision*, Paris: Economica.

Scarpel, R. & Milioni, A. 2002. Utilização conjunta de modelagem econométrica e optimização em decisões de concessão de crédito. *Pesquisa Operacional*, 22(11), 61-72.

Šušteršic, M.; Mramor, D. & Zupan, J. 2009. Consumer credit scoring models with limited data. *Expert Systems with Applications*, 36(3), 4736-4744.

Tegarden, D. & Sheetz, S. 2003. Group cognitive mapping: A methodology and system for capturing and evaluating managerial and organizational cognition. *Omega – The International Journal of Management Sciences*, 31(2), 113-125.

Thomas, L. 2009. Modelling the credit risk for portfolios of consumer loans: Analogies with corporate loan models. *Mathematics and Computers in Simulation*, 79(8), 2525-2534.

Thomas, L. 2010. Consumer finance: Challenges for operational research. *Journal of Operational Research*, 61(1), 41-52.

Twala, B. 2010. Multiple classifier application to credit risk assessment. *Expert Systems with Applications*, 37(4), 3326-3336.

Wang, G.; Hao, J.; Ma, J. & Jiang, H. 2010. A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, 8(1), 223-230.

Yu, L.; Wang, S. & Lai, K. 2007. An intelligent-agent-based fuzzy group decision making model for financial multicriteria decision support: The case of credit scoring. *European Journal of Operational Research*, 195(3), 942-959.

Zavadskas, E. & Turskis, Z. 2011. Multiple criteria decision making (MCDM) methods in economics: An overview. *Technological and Economic Development of Economy*, 17(2), 397-427.

Zopounidis, C.; Galariotis, E.; Doumpos, M.; Sarri, S. & Andriosopoulos, K. 2015. Multiple criteria decision aiding for finance: An updated bibliographic survey. *European Journal of Operational Research*, 247(1), 339-348.