“A MURAME-based technology for bank decision support in credit worthiness assessment”


| | |
|---|---|
| **AUTHORS** | Marco Corazza<br>Stefania Funari<br>Federico Siviero |
| **ARTICLE INFO** | Marco Corazza, Stefania Funari and Federico Siviero (2014). A MURAME-based technology for bank decision support in credit worthiness assessment. *Banks and Bank Systems*, *9*(3) |
| **RELEASED ON** | Tuesday, 30 September 2014 |
| **JOURNAL** | "Banks and Bank Systems" |
| **FOUNDER** | LLC “Consulting Publishing Company “Business Perspectives” |

Marco Corazza (Italy), Stefania Funari (Italy), Federico Siviero (UK)

# A MURAME-based technology for bank decision support in creditworthiness assessment


## Abstract

In the last 30-40 years multicriteria-based approaches have been strongly applied to financial decision-making problems. The contributions that use multicriteria methodologies in order to evaluate the creditworthiness of debtors and loan applicants belong to this research field and mainly refer to applications of UTADIS and ELECTRE methods. In this paper authors themselves propose a deterministic approach for creditworthiness evaluation based on the multicriteria method known as MURAME, which has been applied for the first time to creditworthiness assessment problems by the authors themselves. The MURAME-based approach proposed by the authors is articulated in two phases: a rating assignment phase, and a rating quantification phase. The first one is mainly devoted to determine the debtors' rating classes, whereas the second phase allows to obtain estimates of the probabilities of default and of transition, and permits to calculate quantities providing further information about the (cardinal) rating and the (ordinal) ranking of the debtors' credit quality. The proposed approach is applied to an important northeastern Italian bank, the Banca Popolare di Vicenza, in order to evaluate the creditworthiness in a real case.


**Keywords:** multicriteria analysis, MUlticriteria RAnking MEthod (MURAME), credit risk assessment.
**JEL Classification:** C02, G20, G32.

## Introduction

Creditworthiness assessment of debtors and loan applicants is one of the main activities of financial institutions like banks and regulatory authorities. In short, it provides quantities for measuring credit features like the rating of obligor quality, the probability that a debtor does not fulfil her/his obligations in accordance with agreed terms, and so on (for more details see Bielecki, Rutkowski, 2001; Schmid, 2004). Because of that, creditworthiness assessment constitutes the first step of any credit risk analysis, and credit risk in its turn is considered «*one of the fundamental factors of financial risk*» (Bielecki, Rutkowski, 2001).

In this paper we propose a new deterministic approach for creditworthiness assessment based on the multicriteria method known as MUlticriteria RAnking MEthod (MURAME).

As known, multicriteria methods provide support to various kind of decisions concerning a discrete set of alternatives when the multidimensional nature of real world problems and the preferences of the decision maker (DM) have to be taken into account. Their use in economic and financial contexts is fully appropriate. Indeed, multicriteria methods have been applied to financial decision-making problems since the late 1970 and they tackled questions regarding portfolio management, asset analysis, bond and loan rating, assessment of various type of risks (for details see Zopoudinis, 1999; Zopoudinis, Doumpos, 2002; Spronk et al., 2005 and the references therein).

However, the contributions that use multicriteria methods in order to evaluate the creditworthiness of firms are not so numerous as the importance of such a subject could suggest. Among the significant ones we mention Khalil et al. (2000), Doumpos et al. (2002), Kosmidou et al. (2002), Doumpos, Pasiouras (2005), Baourakis et al. (2009), Doumpos, Zopoudinis (2011), and Angilella, Manzù (2013).

In Khalil et al. (2000), the authors propose a system based on the multicriteria method ELECTRE TRI to rate firms' credit quality and to sort them into homogeneous creditworthiness groups. In the context of credit risk rating models, the ELECTRE TRI method has been used also in Doumpos, Zopoudinis (2011), although combined with an evolutionary optimization approach, and in Angilella, Manzù (2013) for evaluating the credit riskiness of Small and Medium Enterprises (SMEs).

In Doumpos et al. (2002) and Kosmidou et al. (2002), the authors investigate the applicability of the multicriteria method MHDIS in classifying firms applying loan into homogeneous groups of creditworthiness, and compare the obtained clustering with some standard classification technique results.

In Doumpos, Pasiouras (2005), the authors propose a model based on UTADIS in order to replicate the assessment of firms' creditworthiness assigned by regional rating agencies. Baourakis et al. (2009) explore the potential of the multicriteria method UTADIS in developing systems for assessing credit risk of financial institutions by using publicly available data.

All the quoted papers present real world financial applications whose results are satisfactory. With respect to the mentioned methodologies, our approach is characterized by at least three novelty elements: first, it is based on MURAME, which has been applied for the first time to creditworthiness assessment problems by the authors of this paper; second, it allows to estimate ex post probabilities of default and probabilities of transition[1], third, it permits to calculate some quantities which provide further information about the (cardinal) rating and the (ordinal) ranking of the firms' credit quality.

Anyway, the utilization of creditworthiness assessment models based on multicriteria methodologies leaves open a question that has not been discussed so far in the literature. As known, the standard theoretical framework for creditworthiness assessment models is the stochastic one of modern quantitative finance. Therefore, why adopt creditworthiness assessment models based on multicriteria methods? In the final part of this paper we discuss this issue.

The remainder of the paper is organized as follows. In Section 1 we summarize the MURAME method. In Section 2 we present our methodology, based on MURAME, which allows to evaluate the creditworthiness of firms. In Section 3 we check its capabilities when applied to a real case. In particular, we use data provided by an important northeastern Italian bank, the Banca Popolare di Vicenza. In Section 4 we propose some metholodological reasons for the development of creditworthiness assessment models in the frame of multicriteria methodologies. We give some final remarks in conclusions.

## 1. MURAME

The MURAME is a multicriteria method that has been originally proposed in Goletsis et al. (2003) as a method for project ranking. It takes inspiration from two well known methods: ELECTRE III, proposed by Roy (1968; 1991) and PROMETHEE II, proposed by Brans and Vincke (1985).

In order to describe the MURAME, in this section we first present the double threshold preference structure adopted by the MURAME itself, and then we illustrate the main steps of the methodology that, through the computation of some indexes, lead to a complete ranking of the considered alternatives.

**1.1. The double threshold preference structure.** Let us consider a set of $m$ alternatives $\{a_1, \ldots a_i, \ldots a_m\}$ to be evaluated, and a set of $n$ criteria $\{c_1, \ldots c_j, \ldots c_n\}$ by which to evaluate the alternatives. Further, let us assume to compare two generic alternatives $a_i$ and $a_k$, with $a_i \neq a_k$, according to the criterion $c_j$, and let us denote by $g_{i,j}$ the score of the alternative $a_i$ evaluated by the criterion $c_j$. The MURAME considers the following "double threshold preference" structure:

$$\begin{array}{lcl} a_i \; P \; a_k & \Leftrightarrow & g_{i,j} > g_{k,j} + p_j \\ a_i \; Q \; a_k & \Leftrightarrow & g_{k,j} + q_j \leq g_{i,j} \leq g_{k,j} + p_j \, , \\ a_i \; I \; a_k & \Leftrightarrow & \left| g_{i,j} - g_{k,j} \right| \leq q_j \end{array} \quad (1)$$

where: $P$ and $I$ indicate the preference and the indifference relations, respectively; $Q$ denotes a weak preference relation; $p_j$ and $q_j$ indicate a preference and an indifference thresholds, respectively, with $q_j \leq p_j$. Such thresholds allow to take into consideration, beyond the case in which the DM is perfectly sure to prefer a given alternative with respect to another one and the case in which the DM is indifferent between two given alternatives, also a hesitation area in which the DM is not completely sure to prefer a given alternative with respect to another one.

**1.2. The outranking index.** A crucial step in the MURAME is to build an outranking relation in order to evaluate the strength of the assertion "the alternative $a_i$ is at least as good as the alternative $a_k$", for each pair of alternatives ($a_i$, $a_k$). The outranking relation is obtained through the calculation of proper indexes, known as concordance and discordance indexes.

For each pair of alternatives ($a_i$, $a_k$), the local concordance index $C_j(a_i, a_k)$ defined as in (2) specifies the dominance of $a_i$, over $a_k$ according to a given criterion $C_j$:

$$C_j(a_i, a_k) = \begin{cases} 1 & if \;\; g_{k,j} \leq g_{i,j} + q_j \\ 0 & if \;\; g_{k,j} \geq g_{i,j} + p_j \, . \\ \dfrac{g_{i,j} - g_{k,j} + p_j}{p_j - q_j} & otherwise \end{cases} \quad (2)$$

Notice that if $g_{k,j} \geq g_{i,j} + p_j$, then the DM prefers alternative $a_k$ to alternative $a_i$ (strict preference). This entails that the local concordance index of the pair ($a_i$, $a_k$) reaches its minimum value as $a_i$ is dominated by $a_k$. On the contrary, if $g_{k,j} \leq g_{i,j} + q_j$, then $a_k$ is not preferred to $a_i$ and this implies that the local concordance index of ($a_i$, $a_k$) reaches its maximum value. In the intermediate preference region, where the DM is not sure to prefer $a_k$ or $a_i$ (weak preference), therefore, the local concordance index takes values in the interval (0, 1).

The discordance index $D_j(a_i, a_k)$ is also constructed in order to measure how much the hypothesis that $a_i$ dominates $a_k$ according to $c_j$ is not satisfied:

$$D_j(a_i, a_k) = \begin{cases} 0 & if \;\; g_{k,j} \leq g_{i,j} + p_j \\ 1 & if \;\; g_{k,j} \geq g_{i,j} + v_j \, , \\ \dfrac{g_{k,j} - g_{i,j} - p_j}{v_j - p_j} & otherwise \end{cases} \quad (3)$$

[1] The probability of default is the probability that a debtor does not fulfil her/his obligation in accordance with agreed terms. The probability of transition is the probability that an obligor migrates over time from a creditworthiness group to another one.

where $v_j$, with $v_j \geq p_j$, is the veto threshold that is used to reject the fact that the alternative $a_i$ is at least as good as the alternative $a_k$. In this case the discordance index reaches its maximum value.

By taking into account both concordance and discordance information, an outranking index $O(a_i, a_k)$ is built, which indicates how much the alternative $a_i$ outranks the alternative $a_k$ jointly considering all the criteria:

$$O(a_i, a_k) = \begin{cases} C(a_i, a_k) \\ \text{if } D_j(a_i, a_k) \leq C(a_i, a_k) \forall_j, \\ C(a_i, a_k) \prod_{j \in K} \dfrac{1 - D_j(a_i, a_k)}{1 - C(a_i, a_k)} \\ \quad \text{otherwise} \end{cases} \tag{4}$$

where $C(a_i, a_k) = \sum_{j=1}^{n} w_j C_j(a_i, a_k) \Big/ \sum_{j=1}^{n} w_j$ represents the weighted mean of the local concordance indexes, in which $w_j$ indicates the weight associated to criterion $c_j$ and $K$ denotes the subset of deponents of the criteria for which $D_j(a_i, a_k) > C(a_i, a_k)$. Notice that, if there exists even only one criterion for which there is maximum discordance (i.e. $D_j(a_i, a_k) = 1$), then the outranking index is equal to 0. This entails that, if for one of the given criteria the alternative $a_i$ is "worse" than the alternative $a_k$, then it is not more possible consider $a_i$ at least as "good" as $a_k$ although this was true for all the remaining criteria. Under this point of view, the above described outranking approach is considered prudential.

**1.3. The final net flow.** For each alternative $a_i$, a leaving flow $\varphi^+(a_i)$ and an entering flow $\varphi^-(a_i)$ are computed as follows in order to calculate the strength and the weakness of $a_i$ over the remaining alternatives:

$$\varphi^+(a_i) = \sum_{k \neq i} O(a_i, a_k), \varphi^-(a_i) = \sum_{k \neq i} O(a_k, a_i). \tag{5}$$

Then, in order to obtain a total preorder of the alternatives and not only partial ones (see Goletsis et al., 2001), for each alternative $a_i$ a final net flow is computed as the difference between the leaving and the entering flows:

$$\varphi(a_i) = \varphi^+(a_i) - \varphi^-(a_i). \tag{6}$$

In such a way the alternatives are ranked in a descending order, according to their net flows (6).

## 2. The methodology for creditworthiness evaluation

In this section we present the method based on the MURAME that allows to rate the debtors' credit quality and to sort them into a prefixed number of homogeneous creditworthiness groups. The approach is articulated in two phases which are described in the following Sections 2.1 and 2.2. We remark that in problems of creditworthiness evaluation the alternatives considered in Section 1 are the debtors on loan applicants (individuals or firms) and the criteria are the various features according to which the credit risk may be evaluated.

This MURAME-based method is applied to a real case in the next section.

**2.1. Phase 1: the rating assignment phase.** This phase is mainly addressed to determine the rating classes to which the credit applicants have to be assigned and can be articulated in the following steps.

*2.1.1. Typology of credit risk.* At the beginning of the assessment process it is important to carefully define what kind of credit risk "idea" has to be considered in the analysis. In our approach we propose to divide the investigated firms in 10 rating classes. Within this frame, we consider as bankrupt firms those whose respective debts are "bad" or doubtful. Moreover, we also estimate ex post probabilities of default and probabilities of transition.

*2.1.2. Subdivision of the firms.* The firms are subdivided in classes by using the definition of SMEs provided by the Basel capital framework, following which the SMEs can be classified on the basis of their sales revenues, of their number of employees and so on.

*2.1.3. Evaluation criteria.* A very crucial role is devoted to choose the evaluation criteria which can really affect the firms' credit risk conditions. It is known that many factors can determine the bankruptcy of a firm. Therefore, in order to perform an assessment process as better as possible, it is suitable to take into account all possible informative criteria: beyond the usual quantitative accounting ones (related, for instance, to the profitability, to the funding, to the liquidity of the firms), also qualitative criteria (related, for instance, to the management quality, to the operational efficiency, to the generational replacement of the firms), and also macroeconomic quantities (related, for instance, to the economic sector, to the regional areas in which act the firms).

*2.1.4. Reference profiles.* A reference profile is a fictitious firm which is used as benchmark and, consequently, separates contiguous creditworthiness classes. The reference profiles can be determined by a panel of experts and/or professionals. However, the approach we propose here does not need external assistance and determine the reference profiles by using only actual data. In short: first, we determine for each alternative its sample deciles[1]; then, we construct

[1] Notice that we use deciles since we consider 10 creditworthiness classes. Of course, one has to use as many quantiles as the homogeneous rating classes are.

the generic $l$-th reference profile by putting togheter the $l$-th deciles of all the alternatives.

*2.1.5. Thresholds and weight values.* A crucial role in multicriteria methods is played by the thresholds – $p_j$, $q_j$ and $v_j$ – which indicate preference, indifference and veto thresholds (see Section 1). Their values can either be determined in advance by a panel of experts and/or professionals, or they can be computed by the analysts. In credit risk evaluation problems, the presence of these threshold levels introduces some kind of flexibility in incorporating the preference structure of the credit risk managers.

*2.1.6. Application of the MURAME.* In this last step the MURAME methodology is implemented by considering as alternatives the firms applying for credit and the reference profiles identified in the previous steps. This allows us to assign to each firm a final score, which is computed by means of the net flow through formula (6) of Section 1. The ranking of the investigated firms is finally obtained by using the net flow. Notice that the use of the reference profiles allows to separate contiguous creditworthiness classes and consequently to classify the alternatives in homogeneous rating classes.

**2.2. Phase 2: the rating quantification phase.** The second phase concerns the estimation of the probabilities of default and of the probabilities of transition. It is evident that, in general, the estimates have mainly descriptive importance, but in the long run they can be used for making ex ante forecasts once a suitably sized time series of such estimates has been collected. Moreover, this phase aims at providing further information about the ordinal ranking of the firms' credit quality.

*2.2.1. Estimates of the probabilities of default.* For each homogeneous rating class it is possible to consider the probability of default, i.e. the probability that a debtor does not fulfil his obligation in accordance with agreed terms. The estimation of the probability of default of the $l$-th creditworthiness class can be computed in terms of relative frequency through the ratio:

$$\frac{\#f_{l,B}}{\#f_l}, \tag{7}$$

where $\#f_{l,B}$ indicates the number of bankrupt firms belonging to the $l$-th creditworthiness class and $\#f_l$ indicates the number of firms belonging to the $l$-th creditworthiness class.

*2.2.2. Estimates of the probabilities of transition.* The probabilities of transition are the probabilities that an obligor migrates over time from a creditworthiness group to another one. In order to estimate these probabilities, we have to consider only the not bankrupt firms concerning two consecutive business period; the estimation of probabilities of transition from the $l$-th creditworthiness group to the $m$-th creditworthiness group are computed in terms of the relative frequency:

$$\frac{\#f_{l,\overline{B},t}}{\#f_{m,t+1}}, \tag{8}$$

where $\#f_{l,\overline{B},t}$ indicates the number of not bankrupt firms belonging to the $l$-th creditworthiness class in the first period $t$ and $\#f_{m,t+1}$ indicates the number of firms belonging to the $m$-th creditworthiness class in the second period $t + 1$ ($l$ may, or may not, be equal to $m$). Notice that, by construction,

$$\sum_{l=1}^{10} \#f_{l,\overline{B},t} = \sum_{m=1}^{10} \#f_{m,t+1} . \tag{9}$$

*2.2.3. Cardinal rating vs. ordinal ranking.* By construction, the ranking of the obligors is ordinal, so it represents the obligors themselves in an equally-distributed fashion over the related support set {1-*st*, …, *N*-th}, where $N$ is the number of the alternatives. However, taking into account the ranking alone, it is not possible to claim whether a firm ranked in a given position is "good" or "bad", whether its rating is "high" or "low", and so on. Therefore, in order to make more informative the results available to the DM, it is possible to perform a simple comparison between the rating and the ranking based distribution. First, the ranking distribution is normalized over the support set of the rating distribution, with the aim to make comparable the two distributions. Secondly, for each homogeneous creditworthiness group of both distributions we compute some indicators, such as *minimum*, *maximum*, mean and standard deviation of the score and of the normalized score. Thirdly, the indicators related to the two distributions are used to make some kind of comparison, in order to check if the creditworthiness groups evaluated on the basis of the rating are better, equivalent, or worse than the analogous evaluation based on the ranking.

*2.2.4. Time worsening/improvement of the credit quality.* Passing from a running year to the next one, the obligors which are gone bankrupt in the first year are replaced by new obligors in the second one. An important question is related to the time worsening/improvement of the credit quality of the new set of alternatives and of the new homogeneous creditworthiness groups. In order to provide an estimate of this worsening/improvement it is possible to perform an analysis similar to the one proposed in the point *2.2.3.* with the only difference that the analysis is applied only to the rating distributions of the consecutive running years.

## 3. A real case analysis

In this section we apply the two phases of the evaluation methodology developed in Section 2 to evaluate the creditworthiness of firms in a real case.

We utilize data provided by the Banca Popolare di Vicenza, an important northeastern Italian bank, concerning 1000 firms which have obtained funding from the bank in the period of 2001 and 2003[1].

In Table 1 we report the number of firms that have been "healthy" (marked by a flag equal to 0), and the number of firms whose debts have been "bad" or doubtful or non-performing (marked by a flag equal to 1). Notice that, given this kind of labeling, it is not possible to distinguish the firms which have been strictly bankrupt (i.e. those whose debts have "bad" or doubtful) from the firms which have simply gone through a difficult period (i.e. those whose debts have been non-performing).

Table 1. Number of firms which have obtained funding from the bank in 2001 and in 2002, classified in "healthy" (flag equals to 0) and "not healthy" (flag equals to 1).

| Flag | 2001 | 2002 |
|---|---|---|
| 0 | 856 | 839 |
| 1 | 144 | 161 |
| Tot. | 1000 | 1000 |

The bank provided us information about the dimension of the firms: it considers as *small business* a firm whose sales revenues belong to [0.5, 3] millions of euros, and as *middle business* a firm whose sales revenues belong to [3,100] millions of euros.

**3.1. The criteria.** The criteria we adopt in the evaluation process are the set of balance indicators used by the bank in its credit risk analysis. This choice has been intended to compare our results with the ones of the bank. The balance indicators are mainly related to the aspects of profitability, funding and leverage, liquidity, growth and size.

Profitability related criteria:

$I_1$: ratio between *interest expenses* and *total debt*. This indicator measures the firm's capability to remunerate the external capital.
$I_2$: ratio between *turnover* and *total debt*. This indicator evaluates the capability of the external capital to produce sales revenues.
$I_3$: ratio between *operating income* and *total assets*. This indicator is a measure of the auto-financing capacity of the firm, and is a reflection of the probability of the capital invested into the business.

Funding and leverage related criteria:

$I_4$: ratio between *permanent capital* and *fixed assets*. This indicator measures the firm's capability to cope with the financial requirement resulting from immobilization investments through the use of permanent capital.
$I_5$: ratio between *permanent capital* and *total assets*. This indicator evaluates the portion of total assets financed through long term funding, shareholders' funds included. To be significative, it should take values around 80% (for details see Mella, 1991).
$I_6$: ratio between *debt* and *shareholders funds*. This indicator is a measure of the firm's financial leverage.
$I_7$: ratio between *capital-intangibles* and *assets-intangibles*. This indicator evaluates the degree of financial autonomy of the firm. To be significative, it should take values around 50% .

Liquidity related criteria:

$I_8$: ratio between *cash* and *total assets*. This indicator is a measure of the firm's liquidity. To be significative, it should take values around 20% (for details see Mella, 1991).
$I_9$: ratio between *cash* and *current liabilities*. This indicator is another measure of the firm's liquidity. If it assumes values greater than 1, it indicates a good liquidity level.
$I_{10}$: ratio between *net financial flow* and *cash flow*. This indicator indicates how much the financial management affects the total cash flow produced by the firm. Its ideal values should be 0 since it should imply that the firm is able to cover its financial expenses.

Growth related criteria:

$I_{11}$: *year-over-year total assets growth*. This indicator allows of monitoring over time the variation of the invested capital.
$I_{12}$: *year-over-year net sales growth*. This indicator allows of monitoring over time the variation of the sales revenues.

Size related criterion:

$I_{13}$: firm's size. This indicator classifies the firm as small business or as middle business.

Generally speaking, "healthy" firm should be characterized by high values of the indicators $I_2$, $I_3$, $I_4$, $I_9$, $I_{11}$, $I_{12}$, and $I_{13}$, and by small values of the indicator $I_1$. Therefore, in our analysis we use these indicators as criteria, maximizing the former ones and minimizing the latter one.

As regards the indicators $I_5$, $I_7$, $I_8$ and $I_{10}$, it is preferable that they take values around predetermined optimal values. So, in order to penalize their divergences from their respective optimal levels, in our testing analysis we minimize the following related criteria:

[1] The fact that the data are not particularly recent is due to discretional reasons requested by the bank in order to ensure the maximum possible privacy to its customers. However, notice that 2001-2002 is the biennium in which the so called Dot-Com (or New Economy) bubble began to deflate. Therefore, being a period characterized by strong structural changes in the economic dynamics, anyway it constitutes a valued test bench for the proposed approach.

$$\tilde{I}_5 = |I_5 - 0.8|, \tilde{I}_7 = |I_7 - 0.5|, \tilde{I}_8 = |I_8 - 0.2|, \text{and } \tilde{I}_{10} = |I_{10}|. \quad (10)$$

Finally, a further different situation regards the indicator $I_6$. In fact, since it is computed as the ratio between the firm's total liabilities – shareholders' equity excluded – and the shareholders' equity explaining the funding composition, considering the cost of funding, the firm amplifies the potential return on an investment or project, but also increases the potential loss. In such a context, it is fundamental to maintain a sound equilibrium among funding sources. In particular, if $I_6$ takes values around 2 then the firm's financial equilibrium is generally assured, whereas, if $I_6$ assumes values meaningfully lower and higher than 2 then the firm's financial equilibrium could become critical, although not in a symmetric way. So, following the relevant literature, in our analysis we minimize the following related criterion:

$$\tilde{I}_6 = f(I_6) = \begin{cases} 0 & if \quad I_6 < 1.0 \\ 125 \times I_6 - 125 & if \quad 1.0 \le I_6 < 1.8 \\ 100 & if \quad 1.8 \le I_6 < 2.2 \\ -50 \times I_6 + 210 & if \quad 2.2 \le I_6 < 4.0 \\ 10 & if \quad 4.0 \le I_6 \end{cases} \quad (11)$$

Notice also that this criterion penalizes the departure from its optimal value.

**3.2. Determination of the parameters and reference profiles.** Generally speaking, in multicriteria-based methods a tricky point concerns the choice of the values of the thresholds, of the veto parameters, and of the weights. The problem of determining appropriate values for all these parameters is investigated for example in Rogers, Bruen (1998) within ELECTRE method.

In our analysis, in order to determine the values of parameters $p_j$, $q_j$ and $v_j$ for each criterion, first we compute the range $s_j = |max\ (I_j) - min\ (I_j)|$; then, we determine the values of the parameters in the following way:

$p_j = \frac{2}{3} s_j$, $q_j = \frac{1}{6} s_j$ and $v_j = \frac{5}{6} s_j$ so that the inequalities $q_j < p_j < v_j$ are satisfied for all $j$.

As regards the values of the weights, we set them all equal since the bank has not revealed preference with regard to any of them.

The reference profiles, determined as described in Section 2.1. are shown in Table 2.

Table 2. The reference profiles

| | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $\tilde{I}_5$ | $\tilde{I}_6$ | $\tilde{I}_7$ | $\tilde{I}_8$ | $I_9$ | $\tilde{I}_{10}$ | $I_{11}$ | $I_{12}$ | $I_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r_1$ | 0.02 | 3.66 | 0.42 | 0.43 | 0.20 | 2.74 | 0.21 | 0.10 | 86.00 | 0.20 | 1.20 | 0.07 | 1.00 |
| $r_2$ | 0.02 | 2.43 | 0.25 | 0.27 | 0.10 | 2.23 | 0.14 | 0.14 | 55.10 | 0.29 | 2.00 | 0.14 | 1.00 |
| $r_3$ | 0.03 | 1.83 | 0.18 | 0.18 | 0.05 | 1.97 | 0.11 | 0.17 | 21.30 | 0.35 | 2.83 | 0.22 | 1.00 |
| $r_4$ | 0.04 | 1.48 | 0.12 | 0.12 | 0.03 | 1.77 | 0.08 | 0.18 | 10.00 | 0.41 | 3.85 | 0.35 | 1.00 |
| $r_5$ | 0.04 | 1.25 | 0.07 | 0.07 | 0.02 | 1.59 | 0.06 | 0.19 | 10.00 | 0.46 | 5.14 | 0.49 | 1.00 |
| $r_6$ | 0.05 | 1.09 | 0.03 | 0.03 | 0.01 | 1.45 | 0.05 | 0.19 | 10.00 | 0.51 | 7.30 | 0.70 | 2.00 |
| $r_7$ | 0.05 | 0.93 | 0.00 | 0.00 | 0.01 | 1.29 | 0.04 | 0.20 | 10.00 | 0.57 | 9.50 | 0.99 | 2.00 |
| $r_8$ | 0.05 | 0.78 | –0.10 | –0.10 | 0.00 | 1.14 | 0.04 | 0.20 | 10.00 | 0.62 | 13.50 | 1.50 | 2.00 |
| $r_9$ | 0.07 | 0.58 | –0.10 | –0.10 | 0.00 | 0.89 | 0.02 | 0.20 | 10.00 | 0.68 | 25.20 | 3.36 | 2.00 |

**3.3. The results.** We have performed five experiments which differ for the criteria taken into account in each of them. In Table 3 we synthetically illustrate such experiments.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Experiment 1 | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | 13 |
| Experiment 2 | ● | ● | ● | ● | ● | | ● | ● | ● | ● | ● | ● | ● | 12 |
| Experiment 3 | ● | ● | ● | ● | | | | | ● | | ● | ● | ● | 8 |
| Experiment 4 | ● | ● | ● | ● | | | | | ● | | ● | ● | | 7 |
| Experiment 5 | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● | | 12 |

In the first experiment we have considered all the criteria. In the second, the third and the fourth experiment we have progressively excluded all the criteria coming from some kind of transformation of an underlying balance indicators (namely: $\tilde{I}_5$, $\tilde{I}_6$, $\tilde{I}_7$, $\tilde{I}_8$ and $\tilde{I}_{10}$) in order to avoid possible biasing effects due to the transformations themselves. Finally, in the fifth and last experiment we have excluded only the last criteria (namely: $I_{13}$), i.e. the one concerning with the firms' size.

Although the results related to each experiments are satisfying, in this section the authors report only the results of the last experiment since it is the unique that is "coherent" with the acting of the bank as it takes into account exactly the same criteria considered by bank's experts. In particular, it is important to notice that these results well agree with those of the bank, meaning that the MURAME-based methodology for creditworthiness evaluation is able to incorporate in a satisfactory fashion the preference structure of the DM.

In Table 4 the authors report the estimates of the probabilities of default. As one should expect, in each year the default probabilities related to the first rating classes are small, and the default probabilities related to the last rating classes are considerably high. Further, one should also expect that these probabilities strictly increase as the rating classes increase. This is basically true with the exception of some intermediate classes. Reasonably, it should depend on the fact that the bank does not distinguish between the strictly bankrupt firms (that generally belong to the last rating classes) and the firms whose debt are "only" non-performing (that generally can belong also to intermediate rating classes). This shows the good classifying capabilities of our methodology.

Table 4. Vectors of the values (multiplied by 100) of the estimates of the default probabilities. The columns represent the rating classes. The second row of each year reports the same values (not multiplied by 100) reported in the corresponding first row but expressed in rational form, in accordance with the ratio (7)

| | $rc_1$ | $rc_2$ | $rc_3$ | $rc_4$ | $rc_5$ | $rc_6$ | $rc_7$ | $rc_8$ | $rc_9$ | $rc_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2001 | 0.00 | 2.99 | 3.49 | 2.06 | 7.07 | 19.64 | 14.74 | 27.72 | 30.39 | 36.90 |
| | 0/23 | 6/201 | 3/86 | 2/97 | 7/99 | 22/112 | 14/95 | 28/101 | 31/102 | 31/84 |
| 2002 | 0.00 | 1.58 | 2.11 | 8.89 | 7.84 | 10.42 | 24.74 | 29.31 | 29.00 | 44.79 |
| | 0/18 | 3/190 | 2/95 | 8/90 | 8/102 | 10/96 | 24/97 | 34/116 | 29/100 | 43/96 |

In Table 5 the authors report the estimates of the probabilities of transition. Recalling that the values along the principal diagonal represent the firms' probabilities of permanence in the same rating classes for both the years, one notices that, with the exception of the "extremal" rating classes (namely: $rc_1$, $rc_2$, $rc_9$ and $rc_{10}$), all the other ones are characterized by similar probabilities (about 40%) that from 2001 to 2002 a firm stays in the same rating classes. Anyway, considering all these classes, one notices also that there exists a meaningful average probability (about 55%) that from 2001 to 2002 a firm moves from a rating class to the contiguous ones. These remarks can be of some usefulness to the bank in order to improve its credit risk management policy.

Table 5. Matrix of the values (multiplied by 100) of the estimates of the transition probabilities from the year 2001 to the year 2002. The rows and the columns represent the rating classes

| | $rc_1$ | $rc_2$ | $rc_3$ | $rc_4$ | $rc_5$ | $rc_6$ | $rc_7$ | $rc_8$ | $rc_9$ | $rc_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $rc_1$ | 30.43 | 65.22 | 4.35 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $rc_2$ | 4.84 | 75.27 | 11.82 | 3.22 | 2.15 | 0.54 | 1.08 | 0.54 | 0.54 | 0.00 |
| $rc_3$ | 1.22 | 24.39 | 43.90 | 19.51 | 4.88 | 3.66 | 1.22 | 0.00 | 0.00 | 1.22 |
| $rc_4$ | 1.09 | 7.61 | 21.74 | 34.78 | 26.09 | 2.17 | 4.35 | 2.17 | 0.00 | 0.00 |
| $rc_5$ | 0.00 | 3.45 | 12.64 | 20.69 | 40.23 | 17.24 | 3.45 | 1.15 | 1.15 | 0.00 |
| $rc_6$ | 0.00 | 1.14 | 2.27 | 6.82 | 20.45 | 39.77 | 15.92 | 9.09 | 2.27 | 2.27 |
| $rc_7$ | 0.00 | 0.00 | 1.28 | 2.56 | 5.13 | 20.51 | 37.18 | 19.23 | 10.26 | 3.85 |
| $rc_8$ | 0.00 | 0.00 | 0.00 | 1.43 | 2.86 | 15.71 | 11.43 | 41.43 | 17.14 | 10.00 |
| $rc_9$ | 0.00 | 0.00 | 0.00 | 1.43 | 2.85 | 1.43 | 8.57 | 24.29 | 50.00 | 11.43 |
| $rc_{10}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.96 | 5.88 | 13.73 | 17.65 | 60.78 |

Finally, in Tables 6 and 7 we report the estimates of the *minimum*, the *maximum*, the mean and the standard deviation of the rating-based score and of the normalized ranking-based one, respectively. Notice that we normalize both the scores in the range [-100, 100]. In order to perform the analyses proposed in the points 2.2.3. and 2.2.4. (see Section 2.2) we utilize two classical decision making criteria: the mean-variance-based one and the expected value-based one[1]. The choice to use the former criterion is dictated by its importance in the field of the financial decision making. Nevertheless, observing that one of the primary tasks of any credit risk analysis consists in classifying the investigated firms into homogeneous creditworthiness groups (more than to evaluate the variability within the groups themselves), the authors choose to use the latter one.

As far as the analysis proposed in the point 2.2.3. is concerned, recalling that greater the score greater the credit quality, the results coming from both the decision making criteria indicate that the (cardinal) rating-based distribution of the firms is better than the (ordinal) ranking-based one.

In particular, the expected value-based criterion shows that in both the years the average rating-based score of each rating class but $rc_1$ is greater than the average normalized ranking-based score (in fact, only for the rating class $rc_1$ one has both $\mu_{2001\text{-}rt} = 83.90 < \mu_{2001\text{-}rk} = 85.00$ and $\mu_{2002\text{-}rt} = 84.68 < \mu_{2002\text{-}rk} = 85.13$).

[1] We recall that a random variable (r.v.) $X_1$ is preferred in the mean-variance sense to another r.v. $X_2$ if and only if $E(X_1) \geq E(X_2)$ and $Var(X_1) \leq Var(X_2)$ and at least one of these inequalities holds in the strict way, where $E(.)$ indicates the expectation operator and $Var(.)$ indicates the variance operator. We also recall that a r.v. $X_1$ is preferred in the expected value sense to another r.v. $X_2$ if and only if $E(X_1) > E(X_2)$. Notice that in our real case analysis the roles of r.v.s are played by the rating-based and the ranking-based scores.

Table 6. Vectors of the values (multiplied by 100) of the estimates of the minimum, the maximum, the mean and the standard deviation of the rating-based score for the years 2001 and 2002. The columns represent the rating classes

| | | $rc_1$ | $rc_2$ | $rc_3$ | $rc_4$ | $rc_5$ | $rc_6$ | $rc_7$ | $rc_8$ | $rc_9$ | $rc_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2001 | $min_{2001\text{-}rt}$ | 82.60 | 59.38 | 38.15 | 19.35 | –0.49 | –23.29 | –40.78 | –61.85 | –81.91 | –99.56 |
| | $max_{2001\text{-}rt}$ | 87.04 | 82.54 | 58.68 | 38.08 | 19.03 | –0.89 | –23.40 | –40.82 | –62.11 | –82.31 |
| | $\mu_{2001\text{-}rt}$ | 83.90 | 73.85 | 46.96 | 28.41 | 9.91 | –11.82 | –32.96 | –51.10 | –71.65 | –90.64 |
| | $\sigma_{2001\text{-}rt}$ | 1.17 | 5.97 | 5.83 | 4.83 | 5.76 | 6.11 | 6.11 | 6.67 | 6.11 | 5.27 |
| 2002 | $min_{2002\text{-}rt}$ | 83.14 | 59.28 | 40.64 | 21.91 | 2.34 | –17.34 | –36.45 | –58.89 | –78.21 | –99.72 |
| | $max_{2002\text{-}rt}$ | 86.71 | 82.88 | 58.76 | 40.36 | 21.37 | 1.63 | –17.56 | –36.48 | –59.45 | –78.73 |
| | $\mu_{2002\text{-}rt}$ | 84.68 | 73.30 | 48.37 | 30.93 | 11.88 | –7.98 | –28.90 | –46.79 | –69.36 | –88.58 |
| | $\sigma_{2002\text{-}rt}$ | 1.08 | 6.02 | 4.56 | 5.32 | 5.45 | 5.95 | 6.63 | 7.88 | 5.45 | 5.70 |

In other terms, the ranking-based standard descriptive indicators provide a picture of the homogeneous creditworthiness groups which is worse than the rating-based ones. In lack of an analysis like the one proposed in the point 2.2.3., such a biased picture could induce the bank to undertake unnecessary, if not dangerous, new credit risk management policies for improving the groups' credit quality.

Table 7. Vectors of the values (multiplied by 100) of the estimates of the minimum, the maximum, the mean and the standard deviation of the normalized ranking-based score for the years 2001 and 2002. The columns represent the rating classes

| | | $rc_1$ | $rc_2$ | $rc_3$ | $rc_4$ | $rc_5$ | $rc_6$ | $rc_7$ | $rc_8$ | $rc_9$ | $rc_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2001 | $min_{2001\text{-}rt}$ | 82.96 | 45.57 | 29.47 | 11.32 | –7.19 | –28.11 | –45.88 | –64.76 | –83.83 | –99.56 |
| | $max_{2001\text{-}rt}$ | 87.04 | 82.59 | 45.20 | 29.09 | 10.95 | –7.56 | –28.48 | –46.25 | –65.13 | –84.20 |
| | $\mu_{2001\text{-}rt}$ | 85.00 | 64.09 | 37.33 | 20.21 | 1.88 | –17.83 | –37.18 | –55.50 | –74.48 | –91.88 |
| | $\sigma_{2001\text{-}rt}$ | 1.26 | 10.77 | 4.62 | 5.21 | 5.32 | 6.01 | 5.10 | 5.42 | 5.48 | 4.52 |
| 2002 | $min_{2002\text{-}rt}$ | 83.56 | 48.24 | 30.48 | 13.65 | –0.05 | –23.34 | –41.46 | –63.10 | –81.78 | –99.72 |
| | $max_{2002\text{-}rt}$ | 86.71 | 83.19 | 47.87 | 30.11 | 13.28 | –5.77 | –23.71 | –41.83 | –63.47 | –82.15 |
| | $\mu_{2002\text{-}rt}$ | 85.13 | 65.71 | 39.17 | 21.88 | 3.94 | –14.55 | –32.58 | –52.47 | –72.63 | –90.94 |
| | $\sigma_{2002\text{-}rt}$ | 0.99 | 10.17 | 5.10 | 4.83 | 5.47 | 5.15 | 5.21 | 6.22 | 5.37 | 5.15 |

As far as the analysis proposed in the point 2.2.4. is concerned, again recalling that greater the score greater the credit quality, the results provided by both the decision making criteria strongly show a time improvement of the credit quality of all the homogeneous creditworthiness groups passing from the 2001 to the 2002. Reasonably, it means the general capability of the bank to effectively cope with new microeconomic and macroeconomic situations when they happen.

## 4. Why adopt creditworthiness assessment models based on multicriteria methods?

As premised in Introduction, the utilization of creditworthiness assessment models based on multicriteria methodologies leaves open a question that, to the best of our knowledge, has not been discussed before. As well-known, the theoretical framework of reference for credit risk modeling is the stochastic one of modern quantitative finance (for instance see Bielecki, Rutkowski, 2001; Schmid, 2004). Therefore, why adopt and utilize credit risk modeling based on methods like those taken into account in this paper? In our opinion there exist at least the following four reasons.

♦ Within the framework of modern quantitative finance, the most important approaches to credit risk modeling, namely the structural and the reduced-form ones, are based on hypotheses about the financial markets which are, although classical, sometimes unrealistic. Typically, the structural approach «*assumes a Black-Scholes type frictionless market*» (Schmid, 2004, p. 50), and the reduced-form approach assumes «*the framework of an arbitrage-free financial market model*» (Bielecki, Rutkowski, 2001, p. 223). Therefore, credit risk modeling based on multicriteria methods shows some usefulness in case of real financial market that do not satisfy these assumptions. Moreover, with specific reference to the structural approach, it is particularly based on «*the evolution of the firm's value and of the firm's capital structure*» (Bielecki, Rutkowski, 2001, p. 26). In general, the firm's value is not directly observable, and it has be «*derived from observable equity value* [...]» (Schmid, 2004, p. 53). In other words, firm has to to be listed. But, credit is very often applied by firms that, given their small or medium size, are not listed[1]. Therefore, credit risk modeling based on multicriteria methods shows significant usefulness also in case of no listed and, more generally, small and medium sized firms.

- Even if real financial markets satisfy the posed classical assumptions, and if the firms applying loan were listed, the DM could have at her/his disposal information which are precluded to the financial markets themselves, and consequently which could not be used by the above-mentioned approaches (in other terms, the considered financial markets could be characterized by some kind of inefficiency). On the contrary, credit risk modeling based on multicriteria methods can easily manage this kind of information as a criterion in order to provide better credit risk analysis results or, when already existing, to improve them.
- Credit risk modeling based on multicriteria methods is compliant with the Basel capital framework (for instance see Altman, Sabato, 2005). In particular, the Basel II capital accord introduced the possibility for the banks to develop owner internal credit risk models to be used along with those developed by external specialized agencies. This allows the DM to use in his credit risk analyses information which is not available to the external agencies themselves.
- The use of credit risk modeling based on multicriteria methods can be particularly fruitful since it is the easiest to understand by professional operators than the corresponding modeling based on refined concepts like stochastic process or martingale measure.

## Conclusions

In this paper we have proposed a methodology for credit risk assessment which is based on a deterministic multicriteria approach known as MURAME. It is possible to apply it also in case of no listed firms, and in case of small and medium sized ones. In particular, it allows to rank the firms according to their credit risk characteristics and to sort them into a prefixed number of homogeneous creditworthiness groups. Moreover, by means of the use of suitable reference profiles, this methodology allows to estimate the probabilities of default and of transition.

The use of a multicriteria-based approach for creditworthiness assessment provides several advantages and we have shown the usability of our evaluation methodology when applied to a meaningful real case.

Recalling that the goodness of the results of the evaluation methodology proposed depends also on a suitable choice of the parameters and weights, it should be interesting to perform sensitivity analyses with respect to these quantities.

Further, since other multicriteria methods for credit risk assessment have been proposed in the literature, it should be interesting to compare the performances obtained by using the MURAME-based evaluation methodology with those achieved by adopting others multicriteria methods.

It might be also interesting to investigate non-multicriteria based techniques. As possible proposal we mention the second order stochastic dominance criterion of which, as example, we graphically give the results of an application to the same analysis considered in the point 2.2.4.[2] In particular, in Figure 1 we represent the cumulative distribution functions of the rating-based scores of the two investigated consecutive running years. Generally speaking, similarly to what presented in Section 3.3. a general time improvement of the credit quality passing from 2001 to 2002 is confirmed but, contrary to what presented in the same section, this time improvement is not verified for all the rating classes.

## Acknowledgements

We thank for their helpful comments some members of the A.M.A.S.E.S., the Italian Association for Mathematics Applied to Social and Economic Sciences, and the participants to some past editions of the international conference on Computational Management Science.

[1] Notice that the wide majority of enterprises are small or medium sized. For example, they widely represent more than the 90% of the firms of the OECD member countries (for details see OECD, 2011).

[2] We recall that a r.v. $X_1$ is preferred in the second order stochastic dominance sense to another r.v. $X_2$ if and only if $\int_{-\infty}^{x} F_1(t)\,dt \le \int_{-\infty}^{x} F_2(t)dt$, or equivalently $\int_{-\infty}^{x} \left[F_1(t) - F_2(t)\right]dt \le 0$, for all the values of $x$ and this inequality holds in the strict way for at least a value of $x$, where $F_*(.)$ indicates the cumulative distribution function of the r.v. $X_*$. The choices of the analysis to perform taken into account in this application has been made at random among the possible ones.

## References


1. Altman, E.I., Sabato, G. (2005). Effects of the new Basel capital accord on bank capital requirements for SMEs, J*ournal of Financial Services Research*, 28, pp. 15-42.
2. Angilella, S., Mazzù, S. (2013). The financing of innovative SMEs: A multicriteria credit rating model, arXiv:1308.0889v1 [q-fin.RM], pp. 1-34.
3. Baourakis, G., Coniescu, M., Van Dijk, G., Pardalos, P.M., Zopounidis, C. (2009). A multicriteria approach for rating the credit risk of financial institutions, *Computational Management Science*, 6, pp. 347-356.
4. Bielecki, T.R., Rutkowski, M. (2001). *Credit Risk: Modeling, Valuation and Hedging*, Springer, Berlin.
5. Brans, J.P., Vincke, P. (1985). A preference ranking organization method: (the Promethee method for multiple criteria decision-making). *Management Science*, 31, pp. 647-656.
6. 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 Reseach*, 138, pp. 392-412.
7. Doumpos, M., Pasiouras, F. (2005). Developing and testing models for replicating credit ratings: A multicriteria approach, *Computational Economics*, 25, pp. 327-341.
8. Doumpos, M., Zopounidis, C. (2011). A multicriteria outranking modeling approach for credit rating, *Decision Sciences*, 42, pp. 721-742.
9. Goletsis, Y., Askounis, D.T., Psarras, J. (2001). Multicriteria judgements for project ranking: an integrated methodology. *Economic & Financial Modelling*, 8, pp. 127-148.
10. Goletsis, Y., Psarras, J., Samouilidis, J.M. (2003). Project ranking in the Armenian energy sector using a multicriteria methods for groups, *Annals of Operations Research*, 120, pp. 135-157.
11. Khalil, J., Martel, J.M., Jutras, P. (2000). A multicriterion system for credit risk rating. *Gestion 2000: Belgian Management Magazine*, 15, pp. 125-146.
12. Kosmidou, K., Papadimitriou, G., Doumpos, M., Zopounidis, C. (2002). Credit risk assessment using a multicriteria hierarchical discrimination approach. In: Kontoghiorghes, E.J., Rustem, B., Siokos, S. (eds.), *Computational Methods in Decision-Making, Economics and Finance*, Kluwer Academic Publishers, Dordrecht, pp. 605-622.
13. Mella, P. (1991). *Indici di Bilancio*, Pirola editore, Milan [into Italian].
14. OECD (2011). Small and medium-sized enterprises. In: *OECD Factbook 2011-2012: Economic, Environmental and Social Statistics*, OECD Publishing.
15. Rogers, M., Bruen, M. (1998). Choosing realistic values of indifference, preference and veto thresholds for use with environmental criteria within ELECTRE, *European Journal of Operational Reseach*, 107, pp. 542-551.
16. Roy, B. (1968). Classement et choix en présence de critères multiples (la méthode ELECTRE), *RIRO*, 8, pp. 57-75.
17. Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods, *Theory and Decision*, 31, pp. 49-73.
18. Schmid, B. (2004). *Credit Risk Pricing Models. Theory and Practise*, Springer, Berlin.
19. Spronk, J., Steuer, R.E., Zopounidis, C. (2005). Multicriteria decision aid/analysis in finance. In: Figueira, J., Greco, S., Ehrgott, M. (eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer, New York, pp. 799-848.
20. Zopounidis, C. (1999). Multicriteria decision aid in financial management, *European Journal of Operational Research*, 119, pp. 404-415.
21. Zopounidis, C., Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review, *European Journal of Operational Research*, 138, pp. 229-246.

## Appendix

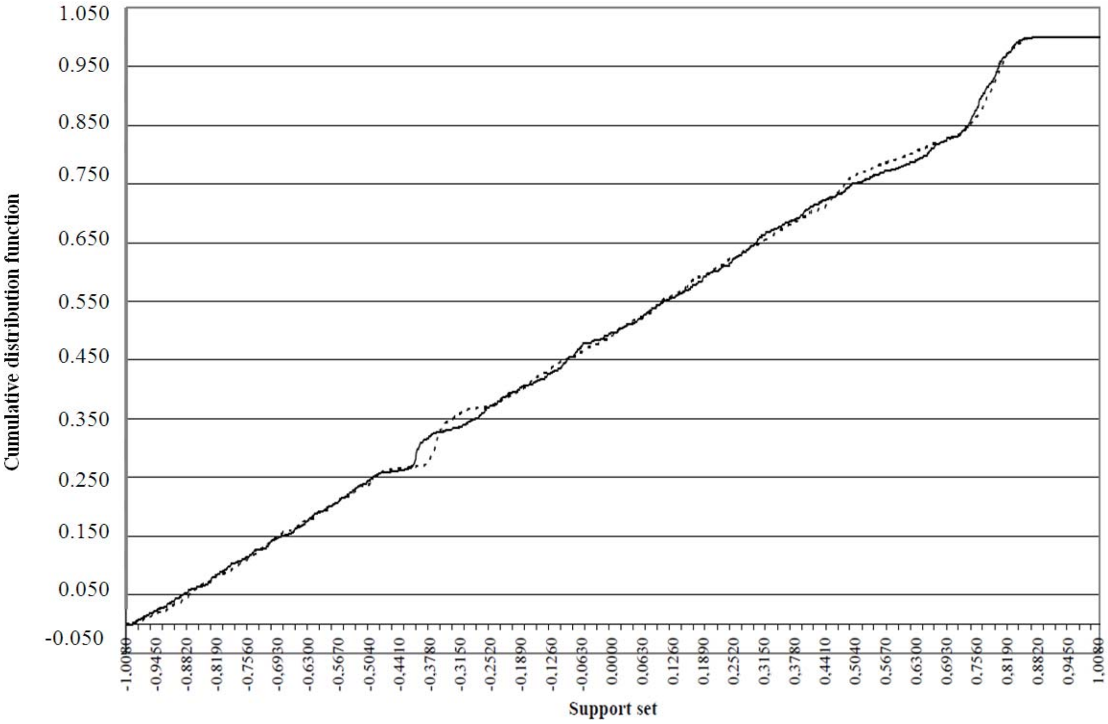


**Fig. 1. Results given by the second order stochastic dominance-based decision making criterion. The continuous and the dotted lines represent the cumulative distribution functions of the rating-based score of the 2001 and the 2002, respectively**