Chapter 3
Banking Management

Abstract Banking institutions play a central role in the financial and business environment. Decision making in banking involves a wide spectrum of issues. This chapter focuses on the evaluation of the performance of banks. In this context, a multicriteria approach is presented, which is based on an outranking method. The proposed multicriteria methodology is illustrated through an application on a sample of Greek banks. The relationship with bank efficiency assessment is also discussed.

Keywords Banking · Bank risk rating · Outranking methods · Bank efficiency

3.1 The Regulatory Framework

Banks are at the heart of the worldwide financial system, acting as intermediaries by providing credit to firms and individuals using deposits and their investment activities. Over the years, the role of banks has undergone significant changes and their importance has increased. Nowadays, banks have extended their range of traditional commercial activities, through the introduction of specialized deposit, financing and investment products, providing new services to their customers, and expanding their operations in the global financial markets. Clearly, this context creates a wide range of new opportunities. At the same time, however, it also creates a plethora of challenges, as it has been clearly demonstrated by the recent credit crisis that began from the USA and later transmitted to Europe in the form of a banking and sovereign debt crisis.

As a consequence of the diverse nature of a bank's operation, the area of banking management is involved with a wide range of issues related to all types of financial risks faced by banks, their investment and financing activities, the efficiency of their operation, as well as the regulatory and supervisory framework that governs their full range of operations. The latter has been a focal point for policy makers over the past two decades.

M. Doumpos and C. Zopounidis, *Multicriteria Analysis in Finance*,
SpringerBriefs in Operations Research, DOI: 10.1007/978-3-319-05864-1_3,

The regulatory framework of Basel II [18], which is currently active, has been designed to improve the risk management practices in financial institutions and ensure the stability of the global financial system. The framework consists of three mutually related pillars.

- The first pillar (minimum capital requirements) is related to procedures required for specifying the minimum level of capital, which must be reserved by financial institutions, as a safety net against the undertaken risks.
- The second pillar (supervisory review process) defines the procedures that must be adopted (a) from the supervisors in order to evaluate how well banks are assessing their capital needs and (b) from the bank's risk managers for ensuring that the bank has reserved a sufficient capital to support its risks.
- Finally, the third pillar (market discipline) requires from banks to provide disclosures with how senior management and the board of directors assess and manage the various types of risk.

The upcoming revision of Basel III is expected to bring a more refined approach with new risk dimensions (e.g., liquidity risk). Even through it is now apparent that the existing regulatory framework failed to prevent the global credit crunch of 2007–2008, the adoption of common rules in a global context can be indeed positive for financial stability.

3.2 Bank Performance Evaluation

Obviously the implementation of successful policies at all levels of a bank's operation should lead to improved overall performance and reduced exposure to excessive risks. The evaluation of the performance and viability of banks has received much interest among researchers, bank managers, and regulators. Such evaluations are performed considering all factors that describe the activities, operations, and risks of a bank. The most popular evaluation framework is based on the consideration of multiple performance and risk attributes categorized in six major dimensions:

1. capital adequacy,
2. assets quality,
3. management competence,
4. earnings generating ability,
5. liquidity, and
6. sensitivity to market risks.

The evaluation context consisting of these dimensions is known as CAMELS (capital, assets, management, earnings, liquidity, sensitivity to market risks). Due to lack of sufficient historical data about bank defaults, bank performance evaluation systems are usually based on empirical assessment techniques (i.e., peer assessments). Sahajwala and Van den Bergh [212] present a comprehensive overview of the practices followed by supervisory authorities in G10 countries with respect to the adoption of risk assessment and early warning systems used for evaluating and monitoring the performance of banks. The overview indicates that central banks often use more than one system based on CAMELS and other similar frameworks, usually following a peer review approach combining financial and qualitative data.

The diverse nature of the evaluation criteria involved (qualitative and quantitative) as well as the importance of incorporating the judgments of expert banking analysts, makes MCDA a well-suited approach for building bank evaluation models. Several multicriteria techniques have been used in this context. For instance, Mareschal and Brans [163], Mareschal and Mertens [164] as well as Kosmidou and Zopounidis [146] used the PROMETHEE method, Zopounidis et al. [264] and Spathis et al. [223] used disaggregation techniques, Raveh [200] used the Co-plot method, whereas Ho [117] implemented the grey relational analysis. The evaluation of banking institutions has also been explored in a ranking context using goal programming formulations inspired by data envelopment analysis [11, 100, 126], which is a common technique for efficiency analysis with numerous applications in banking (see [88] for a comprehensive review).

Most of these studies have focused on the financial aspects of the performance of banks, using financial criteria mainly in the form of financial ratios. Other studies using MCDA approaches have considered additional aspects related to the regulatory and supervisory framework [95, 96, 125], customer-oriented criteria [106, 227]), while specific pillars of the Basel II capital adequacy framework have also been considered (e.g., operational risk [21].

Of course, banking management is not restricted to bank performance evaluation. Other important areas with applications of multicriteria techniques include:

- Asset-liability management [57, 145, 238].
- Bank branches network management [126].
- Evaluation of electronic banking services [122, 201].
- Customer relationship management [106].

The rest of this chapter is devoted to the presentation of a multicriteria approach for bank risk rating. The proposed methodology is based on the PROMETHEE II method and it has implemented in a decision support system currently, which is currently used by the Bank of Greece [73].

3.3 A Multicriteria Approach for Bank Risk Rating

The main output of bank rating models is an evaluation of the overall risk and performance of banks. In a supervisory context, expert analysts (supervisors of a central bank) gather detailed information that enables the evaluation of a bank's condition and the monitoring of its compliance with the regulatory framework. The result of this evaluation process is a rating (CAMELS rating), which provides a forward-looking approach of a bank's current overall condition and potential risk.

In common practice, the ratings are usually assigned in a scale of 1–5, which *resembles* an ordinal classification setting. Banks with ratings of 1 or 2 are considered to present few supervisory concerns, while banks with higher ratings present moderate to extreme degrees of supervisory concern. The definition of the grades in such a rating system, is based on the composite score of the banks obtained by aggregating their performance on all evaluation criteria. This score is expressed on a scale similar to the ratings (e.g., in [1, 5] or [0.5, 5.5]) so that each rating can be matched to a predefined score interval. Within this context, bank rating does not correspond to a "traditional" multicriteria classification problem, in the sense that the actual outcome of the evaluation process is a numerical evaluation score, which is matched to a risk grade at the final stage of the evaluation process, as a means of "defuzzification". This approach provides flexibility to the supervisory authorities, which may take similar actions for banks whose rating scores are very similar, even if they correspond to different ratings.

In accordance, with the CAMELS model which is currently in use by the Bank of Greece, a multicriteria methodology has been implemented that enables not only to define the required risk grades, but also to develop an overall performance index that permits comparisons on the relative performance of the banks. The methodology is based on the PROMETHEE II method [33]. The workflow of the methodology is given in Fig. 3.1.

The PROMETHEE method is widely used to rank a set of alternatives on the basis of pairwise comparisons. Except for this kind of analysis, the method was also used to perform an absolute evaluation in comparison to a pre-specified reference point. Thus, the use of the PROMETHEE method enables the consideration of both the relative and absolute performance of the banks in a unified context. The relative evaluation enables the consideration of the strengths and weaknesses of a bank as opposed to other banks (i.e., on the basis of the conditions that prevail in the banking sector), whereas the absolute evaluation enables the analysis of the condition of a bank compared to predefined reference points representing specific risk profiles. The combination of these approaches provides supervisors with a comprehensive view of the risks that banks face, taking into account the characteristics of each individual bank, the interrelationships between the banks, and the overall condition of the banking sector. The consideration of these two issues in other MCDA models (e.g., a value function) would require the introduction of specific criteria, which were difficult to define and measure in this case.

The subsections below provide details on the implementation of the PROMETHEE method in both these contexts. Details on the evaluation criteria and the details of the evaluation process are given in Sect. 3.4.

3.3.1 Relative Evaluation

The evaluation of the banks in the context of the PROMETHEE method is based on pairwise comparisons. In particular, for each pair of banks (i, j) the global preference

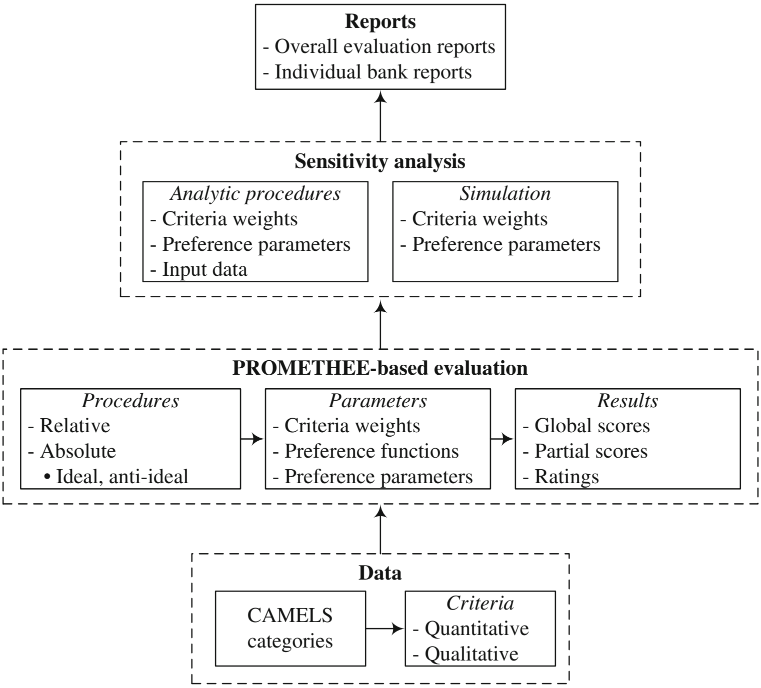

Fig. 3.1 The workflow of the multicriteria methodology

index $P(\mathbf{x}_i, \mathbf{x}_j)$ is computed, where $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ is the vector with the description of bank i on n evaluation criteria. The global preference index is defined as the weighted sum of partial preference indices:

$$P(\mathbf{x}_i, \mathbf{x}_j) = \sum_{k=1}^{n} w_k \pi_k(x_{ik}, x_{jk})$$

where w_k is the weight of criterion k and $\pi_k(x_{ik}, x_{jk})$ is the corresponding partial preference index, which measures (in a [0, 1] scale) the strength of the preference for bank i over bank j on criterion k.

The partial preference index $\pi_k(x_{ik}, x_{jk})$ is a function of the difference $x_{ik} - x_{jk}$ in the performances of the banks on criterion k. A popular choice is the Gaussian function:

$$\pi_k(x_{ik}, x_{jk}) = \begin{cases} 0 & \text{if } x_{ik} \leq x_{jk} \\ 1 - \exp\left[-\dfrac{(x_{ik} - x_{jk})^2}{2\sigma_k^2} \right] & \text{if } x_{ik} > x_{jk} \end{cases}$$

where $\sigma_k > 0$ is a user defined parameter. If a low value is used for σ_k, then even a small difference $x_{ik} - x_{jk} > 0$ may lead to a significant preference for bank i over bank j. On the contrary, for large values of σ_k, strict preference may only occur when $x_{ik} \gg x_{jk}$.

An alternative function for the definition of the partial preference index is the linear generalized criterion:

$$\pi_k(x_{ik}, x_{jk}) = \begin{cases} 0 & \text{if } x_{ik} - x_{jk} \leq 0 \\ \dfrac{x_{ik} - x_{jk}}{p_k} & \text{if } 0 < x_{ik} - x_{jk} \leq p_k \\ 1 & \text{if } x_{ik} - x_{jk} > p_k \end{cases}$$

where $p_k > 0$ is the preference threshold, which defines the minimum difference $x_{ik} - x_{jk}$ above which bank i is assumed to be strictly preferred over bank j on criterion k. Note that the above functions are only meaningful for quantitative data, but alternative options for handling qualitative criteria [33].

Assuming a set of M banks under evaluation, the results of all the pairwise comparisons are aggregated into a global performance index (net flow) as follows:

$$\Phi(\mathbf{x}_i) = \frac{1}{M-1}\left[\phi^+(\mathbf{x}_i) - \phi^-(\mathbf{x}_i)\right] \tag{3.1}$$

where $\phi^+(\mathbf{x}_i) = \sum_{j \neq i} P(\mathbf{x}_i, \mathbf{x}_j)$ is the outgoing flow representing the outranking character of bank i over all the other banks and $\phi^-(\mathbf{x}_i) = \sum_{j \neq i} P(\mathbf{x}_j, \mathbf{x}_i)$ is the incoming flow representing the outranking character of all banks in the sample over bank i. Thus, the above net flow index combines the strengths and weaknesses of a bank compared to its competitors in an overall evaluation measure. The overall net flow index $\Phi(\mathbf{x}_i)$ ranges in $[-1, 1]$, with higher values associated with low risk/high performance banks.

The net flow index (3.1) can be alternatively written in additive form as:

$$\Phi(\mathbf{x}_i) = \sum_{k=1}^{K} w_k \phi_k(\mathbf{x}_i) \tag{3.2}$$

where $\phi_k(\mathbf{x}_i) = \varphi_k^+(\mathbf{x}_i) - \varphi_k^-(\mathbf{x}_i)$ is the partial evaluation score (uni-criterion net flow), defined for criterion k, with

$$\varphi_k^+(x_{ik}) = \frac{1}{M-1}\sum_{j \neq i} \pi_k(x_{ik}, x_{jk}) \quad \text{and} \quad \varphi_k^-(x_{ik}) = \frac{1}{M-1}\sum_{j \neq i} \pi_k(x_{jk}, x_{ik})$$

representing, respectively, the strengths and weaknesses of bank i compared to the others with respect to criterion k.

The advantage of using the additive form (3.2) over (3.1) is that it provides a decomposition of the overall performance of a bank on each evaluation criterion through the corresponding uni-criterion flows. Thus, the strengths and weaknesses of the bank can be easily identified in terms of the criteria.

In order to build the required bank rating model, the evaluation scale for both the overall performance index, as well as for all the partial performance indices can be modified to enable the definition of a 5-point rating scale. In this model calibration step, the partial net flows $\phi_k(\mathbf{x}_i)$ can be used to define a modified partial evaluation function as follows:

$$v_k(x_{ik}) = \begin{cases} 0.5 & \text{if } x_{ik} \geq x_k^* \\ 0.5 + 5\dfrac{\phi_k(x_{ik}) - \phi_k(x_k^*)}{\phi_k(x_{k*}) - \phi_k(x_k^*)} & \text{if } x_{k*} < x_{ik} < x_k^* \\ 5.5 & \text{if } x_{ik} \leq x_{k*} \end{cases} \tag{3.3}$$

where x_{k*} and x_k^* are the least and most preferred values of criterion k, respectively. With this normalization, the partial evaluation of the banks on a criterion k ranges in a scale from 0.5 (best performance) to 5.5 (worst performance), and the final evaluation model is just a modified version of the net flow model (3.2):

$$V(\mathbf{x}_i) = \sum_{k=1}^{K} w_k v_k(x_{ik}) \in [0.5, 5.5] \tag{3.4}$$

This model can be used to rank the banks in terms of their relative performance, thus providing insight into the strengths and weaknesses of each bank within the competitive market and the conditions that prevail. Given the overall score defined in this way, the associated relative rating is specified by defining the intervals [0.5, 1.5] for group 1, (1.5, 2.5] for group 2, (2.5, 3.5] for group 3, (3.5, 4.5] for group 4 and (4.5, 5.5] for group 5.

It should be noted, however, that while the net flow model (3.2) is purely relational (e.g., the evaluation of a bank is expressed solely in terms of the other banks in the sample), with the introduction of the transformation (3.3), the final evaluation model (3.4) incorporates both relational and absolute aspects. This is because the least and most preferred values of the criteria are not defined on the basis of the banks under consideration. Instead, they represent reference points corresponding to high and low risk bank profiles, defined on the basis of the risk analyst's attitude towards risk. In that respect, as the banking sector is improving, the differences $\phi_k(x_{ik}) - \phi_k(x_k^*)$ will decrease, thus leading to improved ratings. Similarly, as the sector deteriorates as a whole, the differences $\phi_k(x_k^*) - \phi_k(x_{ik})$ will increase, resulting in a deterioration of the ratings. Therefore, the rating score of a bank combines its relative performance as opposed to other banks, as well as the performance of the banking sector as a whole compared to predefined risk profiles. The relative evaluation enables the consideration of the interrelationships and interactions between the banks, which is related to systematic risk.

3.3.2 Absolute Evaluation

Except for the above "hybrid" evaluation process, which combines both relative and absolute elements, a purely absolute evaluation approach can also be realized within the context of the PROMETHEE methodology. In this case the results are based only on the comparison of the banks to a pre-specified reference point, whereas the relative performance of the banks is excluded from the analysis.

In cooperation with the analysts in the Bank of Greece, two options were defined for the specification of the reference point. In the first case the banks are compared to the ideal point (ideal bank). This kind of evaluation provides an assessment of the capability of the banks to perform as good as possible. The second option uses an anti-ideal point. Both the anti-ideal and the ideal point ($\mathbf{x}_*$ and $\mathbf{x}^*$, respectively) are defined by the analysts of the Bank of Greece, each consisting of the least and most preferred values of each criterion, i.e. $\mathbf{x}_* = (x_{1*}, x_{2*}, \ldots, x_{K*})$ and $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_K^*)$.

In the case where the banks are compared to the ideal point, the partial evaluation function is adjusted as follows:

$$v_k(x_{ik}) = \begin{cases} 5.5 & \text{if } x_{ik} \leq x_{k*} \\ 0.5 + 5\dfrac{\pi_k(x_k^*, x_{ik})}{\pi_k(x_k^*, x_{k*})} & \text{if } x_{ik} > x_{k*} \end{cases}$$

On the other hand, when the anti-ideal point is used, the following partial evaluation function is used:

$$v_k(x_{ik}) = \begin{cases} 0.5 + 5\dfrac{\pi_k(x_k^*, x_{*k}) - \pi_k(x_{ik}, x_{*k})}{\pi_k(x_k^*, x_{*k})} & \text{if } x_{ik} < x_k^* \\ 0.5 & \text{if } x_{ik} \geq x_k^* \end{cases}$$

3.3.3 Analytic Sensitivity Analysis

Naturally, the multicriteria evaluations defined above incorporate some uncertainty and subjectivity, mainly with regard to the parameters of the PROMETHEE method, which include the criteria weights and the parameters σ_k and p_k of the partial preference functions. Furthermore, since banks operate in a dynamic environment, it is also important to identify changes in the input data that may lead to changes in the rating result. This analysis is performed both for the complete set of banks, as well as for each individual bank separately.

In a first stage, these issues can be addressed by analytic sensitivity procedures. For the criteria weights, the objective of the analysis is to define a range of values for the weight of each criterion k for which the rating of the banks remains unchanged. This can be easily done by imposing the condition that the global score $V(\mathbf{x}_i)$ of each bank i should remain within the score range associated with its rating, as defined with the pre-specified weights.

A similar process can also be employed for the parameters of the criteria preference functions. However, with the pairwise relative evaluation scheme of the PROMETHEE method, the partial preference indices are generally non-monotone and non-convex functions of the corresponding parameters σ and p. Thus, in this case it is not possible to define specific bounds for these parameters within which the rating of the banks does not change. On the other hand, the bounds can be explicitly defined for the absolute evaluation process. In particular, let us assume a bank i which is assigned to the rating group ℓ, defined by a range of scores $(\alpha_\ell, \beta_\ell]$ and suppose that a range $[l_k, u_k]$ should be defined for the parameters of the preference function of a criterion k, such that the rating group of the bank does not change, i.e. $\alpha_\ell < V(\mathbf{x}_i) \leq \beta_\ell$. Then:

$$V(\mathbf{x}_i) > \alpha_\ell \Leftrightarrow v_k(x_{ik}) > \max\left\{0.5, \frac{\alpha_\ell - \sum_{j\neq k} w_j v_j(x_{ij})}{w_k}\right\} \tag{3.5}$$

For illustrative purposes, it can be assumed that: (1) the Gaussian preference function is used, (2) the absolute evaluation is performed in comparison to the ideal point, and (3) $x_{k*} < x_{ik} < x_k^*$. Then, taking into account that $v_k(x_{ik})$ decreases with the preference parameter, and denoting by z_{ik} the left-hand side of (3.5), the upper bound u_k is defined as follows:

$$\begin{aligned}
0.5 + 5\frac{\pi_k(x_k^*, x_{ik})}{\pi_k(x_k^*, x_{k*})} &> z_{ik} \Rightarrow \\
\pi_k(x_k^*, x_{ik}) &> \frac{(z_{ik} - 0.5)\pi_k(x_k^*, x_{k*})}{5} \Rightarrow \\
1 - \exp\left[-\frac{(x_k^* - x_{ik})^2}{2u_k^2}\right] &> \frac{(z_{ik} - 0.5)\pi_k(x_k^*, x_{k*})}{5} \Rightarrow \\
u_k &< \sqrt{\frac{-(x_k^* - x_{ik})^2}{2\ln[1 - 0.2(z_{ik} - 0.5)\pi_k(x_k^*, x_{k*})]}}
\end{aligned}$$

Note that if $z_{ik} > 0.5 + 5/\pi_k(x_k^*, x_{k*})$, then $u_k = +\infty$. The same process is used to define the lower bound l_k:

$$\begin{aligned}
V(\mathbf{x}_i) \leq \beta_\ell \Leftrightarrow v_k(x_{ik}) \leq \min\left\{5.5, \frac{\beta_\ell - \sum_{j\neq k} w_j v_j(x_{ij})}{w_k}\right\} &= o_{ik} \Rightarrow \\
1 - \exp\left[-\frac{(x_k^* - x_{ik})^2}{2l_k^2}\right] &\leq \frac{(o_{ik} - 0.5)\pi_k(x_k^*, x_{k*})}{5} \Rightarrow \\
l_k &\geq \sqrt{\frac{-(x_k^* - x_{ik})^2}{2\ln[1 - 0.2(o_{ik} - 0.5)\pi_k(x_k^*, x_{k*})]}}
\end{aligned}$$

With $l_k = 0$ whenever $o_{ik} < 0.5$.

A similar procedure can also be applied with the linear preference function and the comparison to the anti-ideal point. In addition to the specification of bounds for the parameters of the preference functions, additional information can be obtained by observing the general impact of the preference parameters to the overall evaluation of the banks (as a whole and individually). This is done with the calculation of a sensitivity index Δ_k, which measures the mean maximum percentage change in the global evaluation of the banks due to a change in the preference parameter of criterion k. In particular, let $v_k(x_{ik}, a_k)$ denote the partial performance of bank i on criterion k, expressed as a function of x_{ik} and the criterion's preference parameter a_k. Then, two optimization problems are solved to find the parameter value a_{*ik} (a^*_{ik}) that minimize (maximize), the partial performance of bank i on criterion k, i.e.:

$$v_k^{\min}(x_{ik}, a_{*ik}) = \min_{a_{ik}>0} v_k(x_{ik}, a_{ik}) \quad \text{and} \quad v_k^{\max}(x_{ik}, a^*_{ik}) = \max_{a_{ik}>0} v_k(x_{ik}, a_{ik})$$

Then, the sensitivity index δ_{ik} measuring the impact of criterion's k preference parameter on the global performance of bank i is defined as follows:

$$\delta_{ik} = \max\left\{ w_k \frac{v_k^{\max}(x_{ik}, a^*_{ik}) - v_k(x_{ik})}{V(\mathbf{x}_i)},\ w_k \frac{v_k(x_{ik}) - v_k^{\min}(x_{ik}, a_{*ik})}{V(\mathbf{x}_i)} \right\} \tag{3.6}$$

where $V(\mathbf{x}_i)$ is the global performance of the bank obtained with criterion's k preference parameter defined by the decision-maker and $v_k(x_{ik})$ the corresponding partial score. For instance, a sensitivity index $\delta_{ik} = 0.3$ indicates that a change in the preference parameter of criterion k, may lead to a change of up to 30 % in the global performance of bank i. The direction of the change (decrease or increase) can be easily found by identifying which of the two arguments provides the maximum in (3.6).

The sensitivity index Δ_k is then calculated as:

$$\Delta_k = \frac{1}{M} \sum_{i=1}^{M} \delta_{ik}$$

In the case of absolute evaluation $v_k^{\min}(x_{ik}, a_{*ik})$ and $v_k^{\max}(x_{ik}, a^*_{ik})$ are easy to find because $v_k(x_{ik}, a_k)$ is a monotone function of a_k, and the extremes are found by imposing a range of reasonable values for a_k (e.g., between 0.001 and 100). On the other hand, in the relative evaluation process, $v_k(x_{ik}, a_k)$ is generally a non-convex function of a_k. In this case, a simple genetic algorithm is employed in order to find $v_k^{\min}(x_{ik}, a_{*ik})$ and $v_k^{\max}(x_{ik}, a^*_{ik})$.

3.3.4 Robustness Analysis Through Simulation

The analytic procedures described in the previous section, provide useful local information about the sensitivity of the rating results. Further information can be derived through simulation approaches to obtain a holistic view of the robustness of the results. Simulation is used to analyze the robustness of the ratings with respect to changes in the weights of the criteria, but the process can be easily extended to consider the parameters of the preference functions, too.

The simulation involves the generation of multiple scenarios regarding the weights of the criteria. Two options can be considered for the generation of the weights. In the first case, the weights are generated at random over the unit simplex. Alternatively, the decision maker can provide a ranking of the criteria according to their relative importance, and then random weights are generated, which are in accordance with the ordering of the criteria.

The results of the simulation can be analyzed in terms of the mean and median of the global performance scores, their standard deviation and confidence intervals. Furthermore, for each individual bank useful conclusions can be drawn on the distribution of its rating under different weighting scenarios.

3.3.5 Implementation

The proposed multicriteria methodology has been implemented in an integrated decision support system (DSS) [73]. The system enables multiple users (senior or junior level analysts) to work simultaneously on a common data base. Senior bank analysts are responsible for defining the evaluation criteria and setting the main parameters of the evaluation process (criteria weights, the type of the criteria preference functions, and preference parameters). Lower level analysts have full access to all features of the multicriteria evaluation process, but they are not allowed to perform permanent changes in the evaluation parameters.

Except for data base management and the use of the multicriteria tools, the DSS includes a user-friendly interface that facilitates the preparation of several reports in graphical and tabular format. The system also includes multivariate statistical analysis techniques such as principal components analysis (PCA) as well as some additional modules that support analysts in the specification of the criteria weights using the rank-order centroid (ROCD) and rank-sum (RS) approaches [131]. PCA is a multivariate statistical analysis tool that enables the examination of the explanatory power of the criteria from a statistical point of view. On the other hand, the ROCD and RS approaches simplify the definition of the weights of the criteria. Both techniques only require the user to define a weak-order of the criteria according to their relative importance, without asking for the specification of the exact trade-offs. The ROCD estimates are derived by the centroid of the polyhedron defined by the constraints on the criteria trade-offs, whereas the RS approach relies on the order statistics of the uniform distribution. In particular, assuming that K criteria have been ranked from the most to the least important ones (criterion 1 is assumed to be the most important and criterion K the least important one), the ROCD and RS weights for criterion k are defined as follows:

$$\text{ROCD weight: } w_k = \frac{1}{K}\sum_{\ell=k}^{K}\frac{1}{\ell} \qquad \text{RS weight: } w_k = \frac{K+1-k}{0.5K(K+1)}$$

The system runs on any MS Windows-based PC and it is currently used by the Risk Analysis and Supervisory Techniques Division of the Bank of Greece for evaluating and monitoring the strengths and weaknesses of Greek banks, on the basis of the supervisory policy defined in accordance with the international regulatory framework.

The next section presents an illustrative application of the methodology on sample data for Greek commercial banks over the period 2001–2005.

3.4 Application

3.4.1 Data and Evaluation Parameters

The data involve detailed information for all Greek banks during the period 2001-2005. Overall, 18 banks are considered. The banks are evaluated on a set of 31 criteria (Table 3.1), selected in co-operation with expert analysts of the Bank of Greece, who are responsible for monitoring and evaluating the performance of the banks. The criteria are organized into six categories (capital, assets, management, earnings, liquidity, sensitivity to market risks), in accordance with the CAMELS framework. Overall, 17 quantitative and 14 qualitative criteria are used. By "quantitative"/"qualitative" criteria, we refer to criteria used to evaluate the financial and non-financial, respectively, aspects of the operation of banks. All criteria are actually measured in numerical scales. For the qualitative criteria an interval 0.5–5.5 scale is used (with lower values indicating higher performance), in accordance with the existing practice followed by the risk analysts of the Bank of Greece, who are responsible for collecting and evaluating the corresponding information.

The weights of each category of criteria and the criteria therein have been defined by the expert analysts of the Bank of Greece. Table 3.2 presents the weights defined for each category of criteria along with the corresponding ROCD and RS estimates defined using the ordering of the criteria according to the expert's weights. It is interesting to note that the RS estimates are very close to the actual relative importance of each criteria group. The same was also observed at the individual criteria level. Overall, the quantitative criteria are assigned a weight of 70 %, with the remaining 30 % involving qualitative criteria.

Table 3.1 Evaluation criteria

Categories	Abbr.	Criteria
Capital	Cap1	Capital adequacy ratio
	Cap2	TIER II capital/TIER I
	Cap3	Qualitative
Assets	Ass1	Risk-weighted assets/assets
	Ass2	Non performing loans − provisions/Loans
	Ass3	Large exposures / (TIER I + TIER II capital)
	Ass4	[0.5(Non performing loans) – provisions]/equity
	Ass5	Qualitative
Management	Man1	Operating expenses/operating income
	Man2	Staff cost/assets
	Man3	Operating income/business units
	Man4	Top management competencies, qualifications and continuity
	Man5	Managers' experience and competence
	Man6	Resilience to change, strategy, long term horizon
	Man7	Management of information systems
	Man8	Internal control systems
	Man9	Financial risk management system
	Man10	Internal processes charter – implementation monitoring
	Man11	Timely and accurate data collection
	Man12	Information technology systems
Earnings	Ear1	Net income/assets
	Ear2	Net income/equity
	Ear3	Interest revenue/assets
	Ear4	Other operating revenue/assets
	Ear5	Qualitative
Liquidity	Liq1	Cash/assets
	Liq2	Loans – provisions/deposits
	Liq3	Real funding from credit institutions/assets
	Liq4	Qualitative
Market	Mar1	Risk-weighted assets II/Risk-weighted assets (I and II)
	Mar2	Qualitative

Table 3.2 Weights of each category of criteria

Categories	Weight	ROCD weights	RS weights
Capital	30	47.92	30.77
Assets	20	22.92	23.08
Management	15	10.42	15.38
Earnings	15	10.42	15.38
Liquidity	10	4.17	7.69
Market	10	4.17	7.69

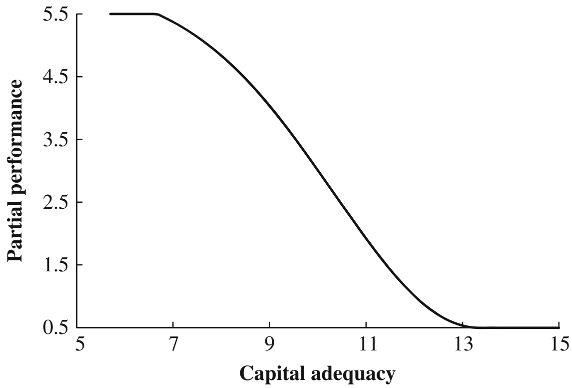

Fig. 3.2 The partial performance function for the capital adequacy ratio (absolute evaluation)

All quantitative criteria are considered through the Gaussian preference function, whereas a linear preference function is used for the qualitative criteria. Figure 3.2 illustrates the partial performance function for the capital adequacy ratio. The function decreases with the values of the criterion, thus indicating that higher capital adequacy values are associated with higher performance and lower risk. The least and most preferred values have been set by the expert analysts to 6.67 and 13.33, respectively. Thus, banks with capital adequacy ratio higher than 13.33 achieve a partial score of 0.5, whereas high risk banks with capital adequacy ratio below 6.67 have a partial score of 5.5. In all cases, the preference parameters have been set in such a way so as to ensure that the partial scores of the banks span, as much as possible, the whole range of values in the pre-specified score range [0.5, 5.5].

3.4.2 Results

Table 3.3 presents the overall evaluation results using the relative assessment procedure. Similar results are also obtained with the absolute evaluation process.[1] The results indicate that most banks achieved a rating grade of 2 or 3, each corresponding to performance scores in (1.5, 2.5] and (2.5, 3.5], respectively. There is no bank in the first (best) grade (score $\leq$1.5) nor in the highest (5th) risk grade (scores $>$4.5).

The dynamics of the performance scores of the banks, indicate that no significant changes are observed between the 5 years of the analysis. Nevertheless, 2002 appears to have been the worst year; compared to 2001 only two banks managed to improve their performance. In 2003 most banks improved their performance (compared to 2002). In 2004 and 2005 no noticeable trend is observed. The highest performance

[1] On average, the rating scores with the absolute evaluation using the ideal reference point were lower (better) compared to the relative evaluation (average difference -0.064). Throughout the 5 years, the ratings were identical in 92 % of the cases with 2 downgrades and 5 upgrades. Similarly, the rating scores with the absolute evaluation using the anti-ideal reference point, were on average higher (worse) compared to the relative evaluation (average difference 0.06). Throughout the 5 years, the ratings were identical in 87 % of the cases with 13 downgrades and none upgrade.

Table 3.3 Overall evaluation results (relative assessment)

Banks	2001	2002	2003	2004	2005
A1	2.59	3.38	2.82	2.65	2.33
A2	2.43	2.48	1.89	1.77	2.03
A3	2.70	3.26	3.35	3.21	3.04
A4	3.19	3.01	2.71	2.91	3.15
A5	2.29	2.45	2.43	2.54	2.52
A6	2.09	2.88	3.02	3.07	3.04
A7	2.03	2.18	1.70	1.63	1.61
A8	1.60	2.00	2.09	1.93	1.95
A9	–	2.10	2.31	2.32	2.59
A10	2.85	3.67	3.42	3.74	3.43
A11	2.31	2.82	2.52	2.17	2.52
A12	2.34	2.49	2.35	2.39	2.36
A13	2.16	2.20	2.26	2.75	2.13
A14	–	2.28	2.18	2.62	3.78
A15	–	2.58	2.40	2.44	2.50
A16	2.64	2.58	2.40	2.27	2.24
A17	–	2.18	2.40	1.98	2.03
A18	–	2.49	2.24	2.32	1.95

improvements have been achieved by banks A7 (20.7 % improvement in 2005 compared to 2001) and A18 (21.4 % improvement in 2005 compared to 2002). On the other hand, the highest decreases in performance involve banks A14 and A6. Bank A14 is the only bank that has been downgraded by more than one rating point during the examined time period. In 2002 (the first year being evaluated) bank A14 was assigned in the 2nd risk grade, deteriorated in the 3rd grade in 2004 and then in the 4th grade in 2005. This downgrade has been mainly due to the deterioration of the assets quality and the weakening of the earnings of the bank.

Table 3.4 provides some sensitivity analysis results for each category of criteria. The presented results involve the weights ranges within which the rating of the banks remains unchanged in each year. When compared to the pre-specified weights of each category of criteria, it becomes apparent that the rating of the banks is most sensitive to changes in the relative importance of the capital dimension. The earnings dimension also seems to be critical (mainly in 2002 and 2003). On other hand, the relative importance of the management dimension is the least likely to alter the rating of the banks. Overall, the ratings in 2002 and 2005 seem to be the most sensitive to changes in the relative importance of the criteria categories, since the obtained bounds are generally closer to the pre-specified weights. As far as the individual criteria are concerned, the most critical ones (as far as their weighting is concerned) were found to be Cap1 (capital adequacy ratio) and Mar1 (risk-weighted assets II / risk-weighted assets I and II). The same two criteria were also found to have among the highest sensitivity indices, particularly in the most recent years (2004–2005). In general, the sensitivity indices were found to be limited (lower than 4 % in all cases).

Table 3.4 Sensitivity analysis results

Categories	Weight	2001	2002	2003	2004	2005
Capital	30	[21.9, 36.5]	[29.4, 31.8]	[25.3, 33.9]	[25.4, 34.8]	[29.9, 32]
Asset	20	[11.7, 29.1]	[17.8, 23]	[4.2, 24.5]	[13.4, 34.2]	[0, 20.5]
Management	15	[0.3, 29.6]	[12.2, 16]	[0.0, 23.1]	[0.9, 22.7]	[12.3, 15.4]
Earnings	15	[7.2, 23.4]	[11.3, 15.7]	[13.4, 20.1]	[5.9, 21.2]	[13.7, 15.2]
Liquidity	10	[4.2, 22.2]	[4.3, 11.6]	[8.9, 14.1]	[6.4, 14.4]	[8.4, 10.1]
Market risk	10	[0, 18.9]	[8.3, 10.9]	[5.3, 11.9]	[4.2, 13.1]	[9.8, 11.4]

On the other hand, in the case of absolute evaluation the impact of the preference parameters was higher, with sensitivity indices up to 8.5 %.

Further results on the sensitivity of the ratings to the weighting of the criteria are obtained with Monte Carlo simulation. The simulation is based on 1,000 different weighting scenarios. In each simulation run, a weighting vector is generated at random, but taking into account the ranking of the criteria according to their importance as defined by the expert analysts. Summary results for 2005 are presented in Table 3.5. The results involve statistics on the global performance score of the banks (mean, median 95 % confidence interval), as well as the distribution of the ratings for each bank. The obtained results are in accordance with the ones given earlier in Table 3.3. In most cases, the rating of the banks is quite robust under different weighting scenarios. The most ambiguous cases involve banks A5, A9, A10, A11 and A15. Future revisions of the rating process or changes in the input data for these banks are highly likely to affect their ratings.

Banks A10 and A14 are the only ones for which a high risk rating seems quite applicable. Further analysis for these two high risk banks is performed by examining the correlations between the randomly generated criteria weights and the global performance of the banks, throughout the simulation experiment. Table 3.6 summarizes the results for the most influential criteria, i.e., the ones whose weight has the highest absolute correlation with the performance of the banks. Criteria with negative correlations are associated with the strong points of the banks, in the sense that an increase in the weight of these criteria leads to a decrease in the global performance score of the banks, thus to lower (better) rating. On the other hand, criteria with positive correlations indicate the weaknesses of the banks, in the sense that an increase in the weight of these criteria leads to an increase in the global performance score of the banks, thus to higher (worse) rating. The obtained results show that the major weaknesses of bank A10 involve its exposure to liquidity risk and its weak earnings. On the other hand, its exposure to market risk is limited, thus leading to an improvement of its overall performance. The exposure to market risk is also a strength for bank A14, which seems to suffer from poor earnings, low asset quality and low capital adequacy.

Table 3.5 Simulation results for 2005

	Statistics				Rating distribution				
Banks	Mean	Median	95 % CI		1	2	3	4	5
A1	2.36	2.37	2.05	2.62	0.0	83.2	16.8	0.0	0.0
A2	2.02	2.03	1.57	2.39	0.7	99.1	0.2	0.0	0.0
A3	3.11	3.10	2.85	3.37	0.0	0.0	99.8	0.2	0.0
A4	3.17	3.17	2.86	3.45	0.0	0.0	98.6	1.4	0.0
A5	2.55	2.56	2.26	2.80	0.0	34.8	65.2	0.0	0.0
A6	3.00	3.00	2.73	3.29	0.0	0.0	100.0	0.0	0.0
A7	1.68	1.69	1.32	2.00	20.2	79.8	0.0	0.0	0.0
A8	1.91	1.92	1.48	2.29	2.9	97.1	0.0	0.0	0.0
A9	2.55	2.56	2.23	2.84	0.0	35.1	64.9	0.0	0.0
A10	3.48	3.48	3.17	3.78	0.0	0.0	56.1	43.9	0.0
A11	2.48	2.48	2.21	2.73	0.0	55.5	44.5	0.0	0.0
A12	2.38	2.37	2.15	2.64	0.0	82.1	17.9	0.0	0.0
A13	2.08	2.08	1.77	2.38	0.1	99.7	0.2	0.0	0.0
A14	3.75	3.74	3.39	4.16	0.0	0.0	11.0	89.0	0.0
A15	2.52	2.53	2.13	2.85	0.0	43.7	56.3	0.0	0.0
A16	2.18	2.18	1.93	2.42	0.0	99.6	0.4	0.0	0.0
A17	2.01	2.01	1.77	2.24	0.0	100.0	0.0	0.0	0.0
A18	1.91	1.91	1.54	2.27	1.6	98.4	0.0	0.0	0.0

Table 3.6 Correlations between the criteria weights and the performance of banks A10, A14

A10		A14	
Mar1	−52.2	Cap2	−56.6
Ass4	−40.4	Mar1	−48.1
Cap1	−33.0	Mar2	−26.6
Mar2	−26.9	Liq3	−13.4
Liq2	20.7	Ass3	−11.5
Liq3	24.6	Ear1	18.7
Ass2	32.1	Ear2	22.7
Ear1	32.2	Ass2	22.8
Liq1	32.6	Ass4	24.0
Ear2	34.5	Cap1	54.7

3.5 Bank Efficiency Versus Bank Performance

The analysis of efficiency in the banking sector has been a major research topic in the area of banking management. According to production theory, efficiency is defined as the ratio between the outputs of a production unit over the inputs used in the production process.

In the context of banking management, the efficiency of banks can be considered under a profit or an intermediation approach [77, 195]. The profit approach focuses on the ability of a bank to control its costs in order to maximize its profits. In this setting, revenue components are the outputs and cost components define the inputs. The intermediation approach, on the other hand, focuses on the ability of a bank to produce financial services (e.g., providing loans to customers, the investment activities of the bank, etc.) using its available resources (personnel, fixed assets, loans, deposits, equity, etc.).

Efficiency assessments are based on frontier methods, with non-parametric techniques such as data envelopment analysis (DEA) being widely used [88]. DEA provides estimates of the relative efficiency for a set of decision making units (i.e., banks), based on their inputs and outputs.

In particular, let $\mathbf{X}$ be a $K \times M$ data matrix for K input variables of M DMUs and $\mathbf{Y}$ be a $O \times M$ matrix for O outputs. Then, the efficiency of the ith DMU is measured by the ratio:

$$\theta_i = \frac{\mathbf{u}_i \mathbf{y}_i}{\mathbf{y}_i \mathbf{x}_i} \in [0, 1]$$

where $\mathbf{x}_i$ and $\mathbf{y}_i$ are, respectively, the available data for the inputs and outputs of DMU i, whereas $\mathbf{u}_i, \mathbf{v}_i \geq \mathbf{0}$ are weight vectors corresponding to these input/outputs.

DEA provides an assessment of the relative efficiency of a DMU compared to a set of other DMUs. In this relative evaluation setting, each DMU is free to specify its own combination of input-output weights that maximize its performance relative to its peers (i.e., competitors). Under constant returns to scale (CRS) and assuming an input orientation, the optimal efficiency for the ith DMU can be estimated through any of the two following linear programming formulations (CCR model, [43]):

$$\begin{array}{ll} \textit{Primal}: & \textit{Dual}: \\ \max \mathbf{u}_i \mathbf{y}_i & \min \theta_i^C \\ \text{s.t. } \mathbf{v}_i \mathbf{X} - \mathbf{u}_i \mathbf{Y} \geq \mathbf{0} & \text{s.t. } \theta_i^C \mathbf{x}_i - \mathbf{X}\boldsymbol{\lambda} \geq \mathbf{0} \\ \quad \mathbf{v}_i \mathbf{x}_i = 1 & \quad \mathbf{Y}\boldsymbol{\lambda} \geq \mathbf{y}_i \\ \quad \mathbf{u}_i, \mathbf{v}_i \geq \mathbf{0} & \quad \boldsymbol{\lambda} \geq \mathbf{0}, \;\; \theta_i^C \in \mathbb{R} \end{array} \tag{3.7}$$

The estimate θ_i^C obtained from the CCR model provides a global technical efficiency measure without taking into consideration any scale effects. In that sense, it is assumed that all DMUs are operating at an optimal scale [46]. To take into account cases where this assumption is not true, variable returns to scale (VRS) can be introduced by simply adding the convexity constraint $\lambda_1 + \lambda_2 + \cdots + \lambda_N = 1$ to the dual CCR model. This constraint ensures that a DMU is benchmarked only against other units of similar size. The resulting model is known as the BCC model [17].

The combination of the results obtained from the CCR and BCC models provides a decomposition of the global efficiency as follows:

$$\theta_i^C = \theta_i^V \theta_i^S$$

where $0 \leq \theta_i^V \leq 1$ is the pure efficiency score obtained under VRS from the BCC model and $0 \leq \theta_i^S \leq 1$ is the scale efficiency factor. Thus, the inefficiency of a DMU can be attributed to inefficient operation (e.g., too small θ_i^V), disadvantageous exogenous conditions (corresponding to scale inefficiency), or both.

The methodological framework of efficiency analysis with DEA has significant similarities but also notable differences with the framework of MCDA. For instance, Joro et al. [136] focused on the connections between DEA and multiobjective optimization, and noted that both fields are interested in identifying efficient points and projecting inefficient units to the efficiency frontier. However, in DEA the projections are made with “optimally” selected weights which differ for each DMU, whereas in MCDA predefined preferential weights are used for all cases under consideration. In that regard, the authors considered multicriteria methods as *ex ante* planning tools and DEA as an *ex post* evaluation tool.

Furthermore, several authors have suggested using DEA for multicriteria evaluation purposes, given that DEA is a data-driven approach that requires minimal information [214]. However, most of such DEA-based evaluation models (e.g., cross-efficiency and super-efficiency models) have significant methodological shortcomings [32].

In the area of banking management, DEA models are useful for evaluating the relative efficiency of banks and discriminating between efficient and inefficient banks. Nevertheless, efficiency is only one aspect of the overall performance and risk of the banks. For instance, one cannot assume that all efficient banks are performing equally well. The same applies to inefficient banks as direct comparisons among such cases are generally meaningful only for those sharing similar characteristics (i.e., belonging to the same facet of the efficient frontier). Furthermore, important aspects, such as capital adequacy, risk management systems, liquidity, and external conditions, are only indirectly relevant to efficiency assessments, whereas in a performance evaluation setting they are considered as key issues. Finally, a bank performance assessment model should be transparent and allow the analysis of all banks in a common setting. In DEA, however, the assessment model does not have a well-defined functional form. Instead, it is based on the solution of a linear program, the analysis is meaningful only for samples of adequate size (it is pointless to perform multidimensional efficiency comparisons for sample with only a few banks), and the weightings of the input/output variables differ for each bank. Even though these properties make sense in an efficiency analysis setting, from an performance evaluation point of view they are troublesome.

MCDA models, on the other hand, are appropriate for benchmarking purposes, allowing the consideration of all pertinent factors that describe (direct or indirectly) the performance of banks, and enabling comparisons to be performed over time based on well-defined functional, relational, or symbolic models, that a bank analyst can use in a straightforward manner.

With the above remarks in mind, it is meaningful to explore the synergies between efficiency analysis based on frontier methods and performance assessment models constructed with MCDA techniques.