

Document downloaded from:

<http://hdl.handle.net/10251/168598>

This paper must be cited as:

Pla Santamaría, D.; Bravo Selles, M.; Reig-Mullor, J.; Salas-Molina, F. (2021). A multicriteria approach to manage credit risk under strict uncertainty. *Top.* 29(2):494-523.
<https://doi.org/10.1007/s11750-020-00571-0>



The final publication is available at

<https://doi.org/10.1007/s11750-020-00571-0>

Copyright Springer-Verlag

Additional Information

A multicriteria approach to manage credit risk under strict uncertainty

David Pla-Santamaria · Mila Bravo · Javier Reig-Mullor · Francisco Salas-Molina

the date of receipt and acceptance should be inserted later

Abstract Assessing the ability of applicants to repay their loans is generally recognized as a critical task in credit risk management. Credit managers rely on financial and market information, usually in the form of ratios, to estimate the quality of credit applicants. However, there's no guarantee that a given set of ratios contains the information needed for credit classification. Decision rules under strict uncertainty aim to mitigate this drawback. In this paper, we propose the use of a moderate pessimism decision rule combined with dimensionality reduction techniques and compromise programming. Moderate pessimism ensures that neither extreme optimistic nor pessimistic decisions are taken. Dimensionality reduction from a set of ratios facilitates the extraction of the relevant information. Compromise programming allows to find a balance between quality of debt and risk concentration. Our model produces two critical outputs: a quality assessment and the optimum allocation of funds. To illustrate our multicriteria approach, we include a case study on 29 firms listed in the Spanish stock market. Our results show that dimensionality reduction contributes to avoid redundancy and that quality-diversification optimization is able to produce budget allocations with a reduced number of firms.

David Pla-Santamaria
Universitat Politècnica de València, Ferrándiz y Carbonell, 03801, Alcoy, Spain
E-mail: dplasan@upv.es

Mila Bravo
CEGEA-Centre of Business Management Research. Universitat Politècnica de València,
Ferrándiz y Carbonell, 03801, Alcoy, Spain
E-mail: mibrasel@cegea.upv.es

Javier Reig-Mullor
Universitat Miguel Hernandez, Elche, Spain
E-mail: javier.reig@umh.es

Francisco Salas-Molina
Universitat de València, Av. Tarongers s/n, 46022 València, Spain
E-mail: francisco.salas-molina@uv.es

Keywords: Moderate pessimism decision making; finance and banking; financial modeling; compromise programming.

MSC: 90B50, 90C29, 91B30.

1 Introduction

This paper addresses the problem of credit risk management for a set of applicants according to their creditworthiness. More precisely, we deal with the problem of scoring applicants, determining interest rates and allocating funds. To this end, we rely on financial and non-financial data, and we follow a moderate pessimism decision rule under strict uncertainty, meaning that the decision maker can estimate the consequence of decisions but not their probabilities. A key element of our novel multicriteria approach is the optimal allocation of funds in terms of quality and diversification of risks.

The concept of credit scoring refers to methods used to predict the probability that a loan applicant will default. A major revision on this concept, its beginnings and the difficulties arising when developing accurate methods of scoring is made by Mester (1997). Credit scoring models are usually based on classification models rely on information from applicants to separate good and bad credit risks (Falangis, 2007; Shi and Xu, 2016). A closely related problem is loan optimization. Banks and other credit institutions hold portfolios of loans from mortgages, car loans, credit cards that need to be optimized. However, managing risky loans has not received as much attention as the equity portfolio selection problem in spite of the fact that one of the main commercial activities of banks is credit granting. Some relevant previous works on this topic are the following. Saunders et al. (2007) and Mencía (2012) focused on optimal portfolio selection based on the trade-off of expected return and credit risk. Sirignano et al. (2016) proposed computational methods for solving the approximate optimization problem for actual loan data and compared the performance to integer program solvers. More recently, Sirignano and Giesecke (2019) developed efficient numerical methods for the analysis of large pools of loans for financial institutions.

Credit scoring and loan optimization usually implies solving multidimensional a problem. Doumpos and Zopounidis (2011) proposed the use of evolutionary algorithms to fit a credit rating model based on the ELECTRE-TRI method. Bravo and Pla-Santamaria (2012) developed a multicriteria decision support model given uncertainty to evaluate loan performance of a set of bank offices belonging to a Spanish bank. Angilella and Mazzù (2015) proposed a multicriteria credit rating model to the aim of financing innovative small and medium companies. However, none of these works addresses two critical issues: 1) the curse of dimensionality to deal with possibly correlated features used to evaluate loan risk; and 2) the use of decision rules to determine recommended interest rates and optimal loan allocations.

The decision of granting a credit is usually based on the characteristics of the borrower. However, one may wonder how many features are relevant to make the right decision and what is the relative importance of different features. The curse of dimensionality refers to several issues introduced in decision making when considering multiple features or dimensions (Bellman, 1957). One of these issues is the possibility of correlation among features used to characterize applicants. To solve this problem, we use principal component analysis (PCA) to transform initial data into a set of linearly uncorrelated features called principal components. More precisely, PCA extracts the important information from data with observations described by several dependent variables and expresses this information as a set of new orthogonal variables (Abdi and Williams, 2010). Thus, a first contribution of this work is the analysis of the impact of PCA on the outcome of our decision making model.

Once the set of relevant criteria is established, we are in position to evaluate credit applicants. Multiple criteria decision making (MCDM) provides sound methods to rank a set of alternatives from a multiple criteria perspective. Moderate pessimism (MP) is a decision making rule proposed by Ballestero (2002) to rank alternatives from multiple states of nature under strict uncertainty. This rule presents advantages with respect to other decision rules such as the principle of insufficient reason by Laplace (1825) or the maximin rule by Wald (1950). This approach is empirically applied by Ballestero (2006), where a set of 132 upholstery/curtain fabrics from a real world textile firm catalog is ranked, and by Ballestero et al. (2007) in a portfolio selection context. The outcome of a MP decision rule is a scalar evaluation of alternatives, namely, a map from alternatives to the degree in which an alternative is better than another.

Loan terms and conditions are relevant for companies in all sectors, but they are particularly important during periods of crisis. This deep concern is shared with lenders who look for analytic tools to evaluate the risk related to each firm. In our context, we use the MP evaluation as a surrogate for the ability of credit applicants to repay their debts. We use this evaluation as a measure of the quality of debt. Then, we establish interest rates inversely proportional to the score obtained for each firm. A further contribution of this work is the use of compromise programming (CP) to find a balance between quality of debt and diversification of risks, since we argue that both criteria represent the main concerns of bank managers. CP is a MCDM technique to deal with the joint achievement of two criteria (Zeleny, 1982; Yu, 1985). The use of CP allows us to elicit a set of efficient solutions ready to be selected by bank managers according to their particular preferences for quality and diversification.

To illustrate our novel approach, we apply the MP decision rule and the CP model to a real world case. We use information on criteria mainly from firms' income statements and balance sheets. We apply two variants of the MP decision rule (with and without dominated alternatives) and we also study the impact of the use of PCA on the recommendations provided by our model. The main potential users of our model are bank managers and financial consultants.

Summarizing, this paper contributes to enhance the ability of bank managers to establish interest rates among credit applicants by means of the following tools:

1. A MP decision rule to evaluate the applicants debt quality that determines a recommended interest rate.
2. A CP model to find a balance between debt quality and risk diversification in the allocation of funds.
3. An analysis to estimate the impact of dimensionality reduction and domination analysis.

The paper is organized as follows. In Section 2, we formulate the proposed method and we describe its main elements. In Section 3, we illustrate the method through a numerical example. The paper ends with some concluding remarks in Section 4.

2 Methodology

In this section, we describe the proposed methodology to evaluate companies in terms of quality of debt under a context of strict uncertainty in which the available information is limited to a set of key variables. The main elements of this methodology are: 1) a dimensionality reduction technique; 2) a decision making rule to establish interest rates; and 3) a multicriteria model to find a balance between quality of debt and diversification of risks in the allocation of funds. The rationale behind the selection of each of the previous elements is expressed in the following subsections.

2.1 Assumptions and a previous definition

Let us assume that a bank manager (the decision maker) pursues to evaluate loan applicants from multiple criteria to establish interest rates to be applied. He/she is considering a set \mathcal{A} of credit applicants, who require financing funds for their long-term investments on an imperfect competition market with multiple providers. The evaluation of applicants will consider n different criteria in set \mathcal{C} . For convenience, all criteria must be rearranged to be “the more, the better”. All applicants in \mathcal{A} are evaluated in terms of criteria in \mathcal{C} . The extreme values of such an evaluation play a critical role in the proposed methodology since we assume a decision making context characterized by strict uncertainty. Next, we adapt the definition of strict uncertainty provided by Ballestero (2002) to our context.

Definition 1 Strict uncertainty. A decision maker is said to rank alternatives under strict uncertainty when the available information is limited to:

1. A finite set of firms \mathcal{A} .
2. A finite set of criteria \mathcal{C} .

3. A scalar evaluation $V : \mathcal{A} \times \mathcal{C} \rightarrow \mathbb{R}$ of every alternative for each criteria.

This definition is critical in our paper. If every criteria or set of criteria, led to the same loan applicants evaluation, then the proposed method would be meaningless. However, several studies from the literature show this is not the case, see for example Avery et al. (2000); Weber (2012), or Eriksson et al. (2014). Therefore, an arbitrary choice of the set of criteria would lead to arbitrary results. To face this problem, a feasible path for the analyst is to search for a solution in the framework of multicriteria decision theory. Thus, we propose a solution in which particular preferences of the decision maker do not influence the evaluation of applicants, hence ensuring objectivity.

2.2 Dimensionality reduction

Set \mathcal{C} from Definition 1 is characterized by its cardinality or dimension $|\mathcal{C}| = n$, denoting the number of different criteria under study to evaluate applicants. To solve issues related with the presence of correlation among a possibly larger set of criteria, we propose the use of PCA to transform initial data into a set of linearly uncorrelated features called principal components. Although PCA is applied in nearly all scientific disciplines, we next highlight its use in efficiency assessment and financial enterprise performance. Li et al. (2015) evaluate four modelling means under the frame of forecasting business risk with support vector machines. In accordance with their conclusions, hybrid SVM models combined with PCA are more efficient than pure support vector machines models. In order to predict financial time series, Wang and Wang (2015) present a stochastic time effective function neural network with PCA. Chen (2014) focuses on the prediction of financial failures. We use PCA since it is one of the methods most commonly used to identify latent constructs, together with maximum likelihood estimation, although PCA is highly recommended (Armor, 1973; Carmines and Zeller, 1979; Spector, 1992; Nunnally and Bernstein, 1994).

In this work, the initial dataset for 29 different companies includes 55 ratios including: a) financial ratios; b) cash flow analysis; c) bankruptcy predictive models; d) market ratios and finally; and e) non-financial data. To analyze the impact of dimensionality reduction on the evaluation of credit applicants, we explore two logical ways of action:

1. We apply no dimensionality reduction technique.
2. We apply PCA to reduce the number of relevant criteria.

We expect to reduce the set of possibly correlated variables since PCA orders components according to its relative importance. The first factor or component is the one that explains most of the total variance, the second factor, the one explaining most of the left variance and so on (Kantardzic, 2011). Finally, we use varimax rotation since this scheme is the most used of the orthogonal rotation criteria (Ford et al., 1986; Fabrigar et al., 1999; Conway and Huffcutt, 2003).

2.3 Scoring firms by the MP rule

Given a set of alternatives \mathcal{A} , a set of criteria \mathcal{C} and a scalar evaluation V as described in Definition 1, a critical task is to define a decision rule that outputs a score of alternatives. Formally, a score is a map $S : \mathcal{A} \rightarrow \mathbb{R}$ that associates a positive value to each element in \mathcal{A} such that the larger value the better the alternative. A key issue in the definition of scoring function S is the selection of aggregation weights for criteria. Two main procedures have been developed in the literature:

1. Consistent weighting. In this procedure, one and only one weight corresponds to each criteria. Laplace (1825) principle of insufficient reason is the classical paradigm and assumes equal weights ($1/n$) for every criterion.
2. Flexible weighting. This method uses different weights for criteria depending on the alternative under consideration. There is an ongoing issue with this procedure due to its lack of consistency. The *maximin* rule by Wald (1950), which assumes extreme pessimism, is a representative example.

The MP decision rule (Ballester, 2002) is a contribution to this field combining the advantages of previous weighting procedures. Like Laplace's decision rule, MP relies on consistent weights. Like Wald's decision rule, MP assumes pessimism, but not extreme pessimism. Extremely pessimistic decision makers think that the worst will always occur. However, moderately pessimistic decision makers think that it is the best that will never occur. We argue that this fact is a fundamental difference and a critical advantage for decision making. A further advantage of the MP decision rule is that it uses all the available information in decision Table 1 given by the combination of sets \mathcal{A} (applicants) and \mathcal{C} (criteria).

Table 1 Decision table.

Applicants	Criteria					
	r_1	r_2	\dots	r_j	\dots	r_n
a_1	V_{11}	V_{12}	\dots	V_{1j}	\dots	V_{1n}
a_2	V_{21}	V_{22}	\dots	V_{2j}	\dots	V_{2n}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
a_i	V_{i1}	V_{i2}	\dots	V_{ij}	\dots	V_{in}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
a_m	V_{m1}	V_{m2}	\dots	V_{mj}	\dots	V_{mn}

In our context, we consider set $\mathcal{A} = \{a_1, a_2, \dots, a_m\}$ of credit applicants, indexed by $i = 1, 2, \dots, m$, and set $\mathcal{C} = \{r_1, r_2, \dots, r_n\}$ of n different criteria. For convenience, all criteria r_j , for $j = 1, 2, \dots, n$, must be rearranged to be the more the better. The combination of applicant a_i and criterion r_j results in evaluation V_{ij} . In addition, MP requires the following steps:

1. Domination analysis (DA) to classify applicants in non-dominated and dominated by convex combinations of other alternatives.

2. Veto to remove the dominated applicants from the evaluation.
3. Determine aggregation weights w_j for each criterion.

From decision Table 1, we say that an applicant a_k is dominated by a convex combination of applicants if the following relationship is satisfied:

$$\sum_{i=1}^m \varphi_j V_{ij} \geq V_{kj}, \quad \forall j = 1, 2, \dots, n, \quad (1)$$

where φ_j is a positive coefficient for the j -th criterion. One may argue that DA and veto may result in a subset of applicants that are not further considered as eligible for obtaining funds. However, bank managers may be interested in reducing credit risk by ruling out applicants below a threshold of required performance measured in terms of the set of criteria under consideration. Within a MP context, this threshold is given by DA. In an attempt to analyze the impact of domination analysis in the context of credit risk management, we evaluate applicants with and without veto after DA in the case study of Section 3.

The principle of MP leads to the following aggregation weights for the set of criteria:

$$w_j = \frac{1}{V_j^* - V_{j*}} \quad (2)$$

where w_j is the weight for the j -th criterion, $V_j^* = \max(V_{ij})$, $V_{j*} = \min(V_{ij})$, and $V_j^* \neq V_{j*}$. Weights in equation (2) are proven to be consistent to solve a decision matrix under strict uncertainty and satisfy several significant properties such as consistency and objectivity (Ballester, 2002). The rationale behind the selection of weights inversely proportional to the range of evaluations among alternatives is as follows. Assume a decision maker facing a performance problem in the context of strict uncertainty on a selected criterion. Suppose that the range of evaluations for the j -th criterion leads to a very high value of $(V_j^* - V_{j*})$. In this case, the decision maker may think "I am rather pessimistic about the soundness of this criterion since I fear that a high range of values is not realistic". The larger the difference between maximum and minimum values, the higher the distrust of the decision maker towards the j -th criterion. The decision maker fears that the maximum is overestimated with respect to the minimum. Accordingly, the higher this range, the lower the weight, as equation (2) states.

This weighting scheme has the advantage of controlling both optimistic and pessimistic evaluations. However, it may be affected by the presence of outliers within the data used for evaluation purposes. In order to control outliers, we next refine equation (2) to limit the impact of extreme values through winsorization:

$$w_j^\alpha = \frac{1}{V_j^\alpha - V_j^{1-\alpha}} \quad (3)$$

where V_j^α is the α -percentile value of the data for the j -th criterion, and $V_j^{1-\alpha}$ is the reciprocal $(1 - \alpha)$ -percentile value. Typical values for α are 0.90, 0.95 or 0.975. As a result, we control outliers by replacing extreme values with a given percentile value of the data. More precisely, in equation (3) we replace values above the α -percentile with V_j^α and values below the $(1 - \alpha)$ -percentile with $V_j^{1-\alpha}$. Finally, from aggregation weights w_j (or w_j^α), we get score S_i for each firm:

$$S_i = \sum_{j=1}^n w_j V_{ij}, \quad i = 1, 2, \dots, m. \quad (4)$$

Since we aligned criteria under the rule the more-the better, we take the view of considering each score as a measure of the ability of credit applicants to pay their debts. In other words, we assume that score S_i is a measure of the quality of the debt granted to applicant i . If this quality measure does not reach a minimum threshold set by the decision maker, then the applicant is taken out of the sample of companies. This additional veto allows to ensure that the set of companies under consideration for optimization purposes are creditworthy. From this quality score we derive a normalized risk premium π_i as follows:

$$\pi_i = \frac{S_{max} - S_i}{S_{max} - S_{min}} \quad (5)$$

where S_{max} and S_{min} are, respectively, the maximum and minimum scores among the whole set of applicants. Finally, we compute interest rates by means of the following expression:

$$R_i = R_{min} + \pi_i(R_{max} - R_{min}) \quad (6)$$

where R_i is the interest rate to be paid by the i -th firm, R_{min} is the minimum interest rate at which the bank would extend loans, and R_{max} is the maximum interest rate the firm is willing to accept. Notice that risk-based equation (6) assumes risk neutrality due to the linear relationship established between the risk premiums and interest rates. Thomas (2009) defines risk-based pricing within the context of consumer credit as the possibility to adjust the interest rate charged to borrowers to reflect the risks involved. The practice of risk-based pricing started in mortgages loans and its use was later extended to other type of loans (Freeman and Hamilton, 2002; Edelberg, 2006). Risk is usually assessed by means of different econometric models from a set of features of the borrower. We here follow the approach of setting a linear model of the features of the companies as summarized in a set of ratios. The rationale behind our choice is as follows. First, its simplicity. In general, linear models are simpler than non-linear models and this choice helps us diminish complexity. Second, linear models are widely used in risk-based in a consumer credit context. Some examples are Edelberg (2006), where a linear combination of the elements of a vector of the borrower's characteristics is used to predict the interest rate, and

the works by Adams et al. (2009), Einav et al. (2012), and Einav et al. (2013), where an almost linear relationship is empirically established between risk and the probability of payment. However, a linear model may lead to some restrictions in practice when absence of risk neutrality recommends a non-linear model. The use of FICO scores and more sophisticated models (Sengupta and Bhardwaj, 2015; Breeden, 2016) may help overcome this limitation. In this sense, a further advantage of our approach is that equation (6) can be replaced by a non-linear risk-based interest rate model without compromising the rest of our approach.

2.4 A compromise between quality and diversification

An important characteristic of our approach is that it produces two relevant outputs. The first one is a set of interest rates and the second one is a recommendation on the credit amount for each applicant. To this end, we aim to find a balance between the quality of debt, measured by the score computed in Section 2.3, and the diversification of risk among applicants, measured by the counterpart of the Herfindahl index (Woerheide and Persson, 1992). The Herfindahl index is a measure of concentration within an industry computed as the sum of squares of the market shares of the firms expressed as fractions:

$$H = \mathbf{x}^T \mathbf{x} = \sum_{i=1}^m x_i^2. \quad (7)$$

The higher the value of H , the higher the concentration of the market. In our context, x_i denotes the weight of the i -th firm in a given budget of available funds for credit applicants. In order to reduce risk, a desirable goal for bank managers is to minimize concentration or, changing the point of view, to maximize diversification.

Since we are dealing with a bicriteria problem, we obtain solutions by means of CP, which represents a sound technique to deal with the joint achievement of two criteria (Zeleny, 1982; Yu, 1985; Ballesteros and Romero, 1998). The main advantage of CP with respect to other multiple criteria decision techniques is the possibility to split the solution selection process in two parts: first, a Paretian efficient frontier in a normalized space of criteria is obtained; second, a point in this frontier is chosen according the particular preferences of decision makers through a family of distance functions. No re-run of optimization algorithms is required in this second stage, hence facilitating the sensitivity analysis of solutions due to changes in preferences and distance functions.

In our context, a CP model is based on the concept of ideal point within a normalized quality-diversification space. This ideal point is characterized by both the maximum quality and the maximum diversification attainable. Due to the usual conflict between objectives this ideal point is usually infeasible.

However, it plays a key role as a reference point in CP. Indeed, the Zeleny's axiom of choice that states that alternatives that are closer to the ideal point are preferred to those that are further (Zeleny, 1973). Then, we can formally express a CP model as method to evaluate any pair of feasible alternatives, denoted by \mathbf{x}_1 and \mathbf{x}_2 , by means of the next preference relations (Ballestero and Romero, 1998):

$$\mathbf{x}_1 \succ \mathbf{x}_2 \iff D(\mathbf{x}_1) < D(\mathbf{x}_2) \quad (8)$$

$$\mathbf{x}_2 \succ \mathbf{x}_1 \iff D(\mathbf{x}_2) < D(\mathbf{x}_1) \quad (9)$$

$$\mathbf{x}_1 \sim \mathbf{x}_2 \iff D(\mathbf{x}_1) = D(\mathbf{x}_2) \quad (10)$$

where \succ means "is preferred to", \sim means "is indifferent to", and $D(\mathbf{x})$ is a distance function between any feasible solution \mathbf{x} and the ideal point.

The main input of our CP model is the quality score for a set of firms that we obtained from the application of the MP rule. In addition, we assume that bank managers are provided with a budget to allocate among the set of firms under consideration. Let \mathbf{x} be a an m -dimensional vector with positive real values that add up to one representing the fraction of budget allocated to each of the m firms. According to the scores computed in Section 2.3, we measure global quality of the debt as the product $\mathbf{s}^T \mathbf{x}$, where the i -th element of vector \mathbf{s} contains the score obtained for the i -th. A drastic rule to ensure the highest debt quality would be allocating the entire budget to the firm with the highest score. However, risk would be extremely concentrated. To ensure diversification of risk, we minimize concentration by means of the Herfindhal index in equation (7). Within a normalized quality-diversification space, we can build a non-dominated frontier of loan portfolios by solving the next family of quadratic programs with parameter S_0 (Ballestero and Pla-Santamaria, 2004):

$$\min \mathbf{x}^T \mathbf{x} \quad (11)$$

subject to:

$$\mathbf{s}^T \mathbf{x} = S_0 \quad (12)$$

$$\mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \quad (13)$$

$$\mathbf{1}^T \mathbf{x} = 1 \quad (14)$$

where \mathbf{x} is an $m \times 1$ vector of weights (parts per unit) of a given budget that is allocated to each firm; \mathbf{s} is an $m \times 1$ vector of scores for each firm, and S_0 is a positive parameter denoting a weighted average score; \mathbf{l} and \mathbf{u} are $m \times 1$ vectors with lower and upper bounds for weights in \mathbf{x} ; and $\mathbf{1}$ is an $m \times 1$ vector of ones. Once we have obtained the set of non-dominated loan portfolios, we are in a position to select the best one according to the axioms encoded in equations (8), (9) and (10) by computing distances of each loan portfolio to the ideal point.

The model encoded in equations (11) to (14) assumes that applicants that were unable to reach a minimum quality threshold in score described in equation (4) are taken out from the set of candidates. It also assumes that there is a known budget that can be continuously divided with the restriction that some lower and upper bounds in constraint (13) are satisfied. This constraint allows the possibility to set a minimum allocation (if any) for applicants and also setting an upper bound to avoid that allocations exceed the requested amounts.

3 Case study

In order to illustrate our proposal, in this section we elaborate a case study for large companies in Spain. Furthermore, we aim to evaluate the impact of dimensionality reduction and DA in the process of eliciting both interest rates (derived from the MP rule) and amounts granted to a set of credit applicants expressed in fractions of a given budget. We also aim to evaluate the effect of outliers in data and winsorization. Summarizing, we compute rates and budget allocations in five different cases:

- **Case 1: Control.** Using the MP rule without DA and without PCA.
- **Case 2: DA.** Using the MP rule with DA and without PCA.
- **Case 3: PCA.** Using the MP rule without DA and with PCA.
- **Case 4: DA-PCA.** Using the MP rule with DA and with PCA.
- **Case 5: Winsor.** Using the MP rule without DA and without PCA after winsorization.

In all cases, we consider set \mathcal{A} with 29 Spanish firms gathered in Table 2. All these companies meet the following requirements: 1) non-bank character; 2) listed in the Spanish stock exchange market in year 2013; and 3) market capitalization equal or superior to 1000 millions of euros. The motivation behind this selection is both the availability and the homogeneity of data to perform the analysis.

Each company is classified according to its economic activity as follows: services sector (10 firms); construction/materials (5); financial services (5); food (4); transport (1); automobile (1); chemical/pharma (1); electronics (1); steel (1). Once the universe of companies is defined, the next step consists of collecting financial, non-financial and market data on each firm as summarized in Appendix A. This kind of information is widely referenced in literature when dealing with a broad spectrum of problems (Altman et al., 2016). More precisely, we consider the following five groups of criteria:

- a) Financial ratios (FR). These 29 indicators shown in Table 3 help us assess the economic and financial performance of a company from the relationships between the different accounting indicators obtained from the balance sheet and the income statement. These 29 variables are grouped in: (i) profitability ratios; (ii) liquidity ratios and (iii) solvency ratios.

Table 2 Set of credit applicants.

Id	Company	Id	Company
1	EDP RENOVAVEIS	16	ENAGAS
2	CATALANA	17	ZARDOYA
3	MAPFRE	18	ALMIRALL
4	IAG	19	ABERTIS
5	EBRO FOODS	20	A3MEDIA
6	ACERINOX	21	ACS
7	NH HOTELES	22	REPSOL
8	COLONIAL	23	CAF
9	VISCOFAN	24	OHL
10	PROSEGUR	25	ENDESA
11	AMADEUS	26	DIA
12	VIDRALA	27	IBERDROLA
13	RED ELECTRICA	28	GAS NATURAL
14	MELIA	29	TELEFONICA
15	FERROVIAL		

- b) Cash flow (CF) analysis. It involves references shown in Table 4 which allow us to assess the cash-flow resources and needs in the company to meet its payment obligations. These references are obtained from relationships between cash-flow statement indicators and accounting indicators. Seven criteria are included in this group.
- c) Bankruptcy predictive models (BM). These formulations attempt to determine the company's financial health, i.e., foresee a possible bankruptcy in the company's future. Here, we follow the approach of selecting models that have been successfully used in practice for a long period of time. These models are classified into two groups:
- The first group is obtained from combinations of various quantitative variables extracted from the financial statements by applying different statistical techniques. Five models are here considered. Altman (1968) was the first that successfully used step-wise multiple discriminant analysis to develop a prediction model obtaining a high degree of accuracy. Among the sample of 66 companies, 50% failed and 50% were successful. The accuracy rate achieved by Altman's model was 95.0%. Legault and Score (1987) analyzed thirty financial ratios in a sample of 173 Quebec manufacturing businesses having annual sales ranging between 1-20 million USD by using step-wise multiple discriminant analysis. Springate (1978) developed a model following the recommendations in Altman (1968). Step-wise multiple discriminant analysis was used to select four out of 19 popular financial ratios that best distinguished between sound business and those that actually failed. Kanitz (1974)

and Elisabetsky (1976) bankruptcy predictive models are also applied in our work.

- The second group includes the following market-based indicators: Merton’s distance to default (Merton, 1974) and credit ratings established by major rating agencies. Unlike the models specified in the first group, Merton’s distance main source of information is the market value. We here used KMV model (Kealhofer, 1993; McQuown, 1993; Vasicek, 1984), which is founded on the assumptions of Merton’s bond pricing model. Magee (2013) applies this methodology with the aim of identifying the effect of foreign currency hedging with derivatives on the probability of financial distress. Regarding credit ratings, Moody’s, Standard & Poor’s, Fitch, Axesor and Morningstar are the credit rating agencies taken into account. Credit rating agencies provide investors and analysts with objective and independent assessments of companies on its securities. In Jory et al. (2016), the presence of ratings is related to premiums paid in mergers and acquisitions.
- d) Market ratios (MR). We use ten indicators focusing mainly on the valuation of listed companies, where stock market capitalization and number of shares play a key role as reference variables, in relation to various accounting indicators.
- e) Non-financial (NF) data. They refer to non-financial information such as company’s age and expectations on the economic sector in which the company operates. Two indicators are taken into account in this group.

A summary of statistics for the 55 criteria and the 29 companies used in this paper is included in Appendix B. A few criteria present large ranges that may be caused by either the particular features of some of the companies in the sample or the presence of outliers. We here limit the impact of extreme values by three different ways: 1) the weighting scheme of the MP rule, which is specially designed to control extreme values; 2) the use of winsorization to remove possible outliers; and 3) the large number of criteria used, hence reducing the importance of a single criterion.

Once we have the input data in a decision table such as Table 1, we are in position to apply the MP rule to establish interest rates and amounts granted to the set of applicants in Table 2 by means of CP. In order to analyze the impact of DA and PCA in the results, we next consider the following cases.

3.1 Case 1. Control: without DA and without PCA

In the first case, we apply the MP rule to establish the quality of debt for a set of applicants with neither applying PCA nor DA. We label this case as

Table 3 Financial ratios: profitability (PR), liquidity (LR), and solvency ratios (SR).

Id	Ratio	Set
1	Return on invested capital (ROIC)	PR
2	Return on assets (ROA)	
3	Assets turnover	
4	Return on equity (ROE)	
5	Pretax income-to-equity ratio	
6	Ebit-to-sales ratio	
7	Net value added-to-sales ratio	
8	Quick ratio	LR
9	Current ratio	
10	Accounts receivable net turnover	
11	Total accounts receivable turnover	
12	Working capital-to-assets ratio	
13	Cash ratio	
14	Current liabilities-to-assets ratio	
15	Trade payable turnover	
16	Suppliers turnover	
17	Financial leverage	SR
18	Assets-to-liabilities ratio	
19	Liabilities-to-(liabilities+equity) ratio	
20	Liabilities-to-equity ratio	
21	Equity-to-assets ratio	
22	Retained earnings-to-equity ratio	
23	Retained earnings-to-assets ratio	
24	(Non-current liabilities+equity)-to-assets ratio	
25	Non-current liabilities-to-liabilities ratio	
26	Debt-to-liabilities ratio	
27	Non-current liabilities-to-assets ratio	
28	Leverage ratio	
29	Debt average cost	

”Control” since we later aim to analyze the impact (if any) of these procedures in the recommendations given by our model. Given a decision table with 29 applicants in rows and 55 criteria in columns, we first compute weights for each criterion using equation (2) and we next derive normalized quality scores using equation (4). We summarize the scores for this control experiment in the second column of Table 5.

A list of interest rates is then established by equation (6). However, bank managers may also be interested in eliciting funds to be allocated to each applicant. In order to provide managers with a wide range of efficient solutions, we solve the quadratic program encoded from equation (11) to equation (14) for 50 evenly spaced values of S_0 over the interval $[S_{min}, S_{max}]$, where S_{min}

Table 4 Cash flow (CF), bankruptcy models (BM), market ratios (MR) and non-financial (NF) data.

Id	Ratio	Set
30	Short term debt-to-net operating cash flow	CF
31	Debt-to-net operating cash flow	
32	Current liabilities-to-net operating cash flow	
33	Liabilities -to- net operating cash flow	
34	EBIT-to-net operating cash flow before interest and taxes	
35	NOPAT-to-net operating cash flow before interest	
36	Net income-to-net operating cash flow	
37	Legault and Score (1987)	BM
38	Elisabetsky (1976)	
39	Kanitz (1974)	
40	Springate (1978)	
41	Altman (1968)	
42	Merton (1974)	
43	Credit rating	
44	Earnings per share	MR
45	Sales per share	
46	Book value per share	
47	Price-to-sales ratio	
48	Price-to-cash flow ratio	
49	Price-to-book value ratio	
50	Price-to-earnings ratio (PER)	
51	Enterprise Value (EV)-to-Ebitda ratio	
52	Cash flow per share	
53	Dividends per share	
54	Years from foundation	NF
55	Expectations on the economic sector	

and S_{max} are, respectively, the minimum and the maximum of the scores in each of the columns of Table 5.

The results for these 50 different optimization problems are represented in Figure 1. Each point of this figure is an allocation \mathbf{x} of funds in a given budget to each of the companies under consideration. Interestingly, this bullet-shaped set of solutions of budget allocations under a quality-concentration context resembles the typical mean-variance efficient frontier shape in portfolio selection problems (Ballesteros and Pla-Santamaria, 2004). For obvious reasons, bank managers are only interested in solutions located in the upper part of the bullet shape since they can achieve a higher quality for a similar degree of concentration.

Moreover, solutions in Figure 1 are not aligned, meaning that managers aim to maximize quality but to minimize concentration of risks. To facilitate

Table 5 Moderate pessimism scores (D=Dominated).

Applicant	Control	DA	PCA	DA-PCA	Winsor
EDP RENOVAVEIS	0,642	0,642	0,396	0,396	0,597
CATALANA	0,586	0,586	0,668	0,668	0,590
MAPFRE	0,520	0,520	0,498	0,498	0,518
IAG	0,313	0,313	0,272	D	0,295
EBRO FOODS	0,570	0,570	0,668	0,668	0,571
ACERINOX	0,428	0,428	0,389	0,389	0,414
NH HOTELES	0,303	0,303	0,277	0,277	0,279
COLONIAL	0,293	0,293	0,277	0,277	0,279
VISCOFAN	0,708	0,708	0,641	0,641	0,745
PROSEGUR	0,411	0,411	0,448	D	0,436
AMADEUS	0,555	0,555	0,507	0,507	0,598
VIDRALA	0,621	0,621	0,621	0,621	0,636
RED ELECTRICA	0,522	0,522	0,459	0,459	0,527
MELIA	0,315	0,315	0,394	D	0,313
FERROVIAL	0,452	0,452	0,345	0,345	0,444
ENAGAS	0,561	0,561	0,467	0,467	0,572
ZARDOYA	0,542	0,542	0,575	0,575	0,584
ALMIRALL	0,519	0,519	0,403	0,403	0,497
ABERTIS	0,401	0,401	0,384	0,384	0,401
A3MEDIA	0,280	0,280	0,428	0,428	0,285
ACS	0,214	0,214	0,344	0,344	0,240
REPSOL	0,478	0,478	0,404	0,404	0,470
CAF	0,401	0,401	0,486	0,486	0,405
OHL	0,316	0,316	0,316	0,316	0,328
ENDESA	0,455	0,455	0,494	0,494	0,468
DIA	0,449	0,449	0,517	0,517	0,485
IBERDROLA	0,475	0,475	0,468	0,468	0,469
GAS NATURAL	0,473	0,473	0,467	0,467	0,477
TELEFONICA	0,426	0,426	0,466	0,466	0,439

the selection of the best budget allocation, we transform these solutions into a normalized quality-diversification space. Normalization is achieved by means of two indexes ranging in $[0, 1]$ so that the higher the better. First, a quality index:

$$\text{Quality index} = \theta_1 = \frac{Q_k - Q_{min}}{Q_{max} - Q_{min}} \quad (15)$$

where $Q_k = s^T \mathbf{x}_k$ is the global quality obtained by budget allocation \mathbf{x}_k , and Q_{max} and Q_{min} are, respectively, the maximum and minimum quality among all 50 optimal allocations derived from solving the quadratic program encoded from equation (11) to equation (14). Second, a diversification index:

$$\text{Diversification index} = \theta_2 = \frac{H_{max} - H_k}{H_{max} - H_{min}} \quad (16)$$

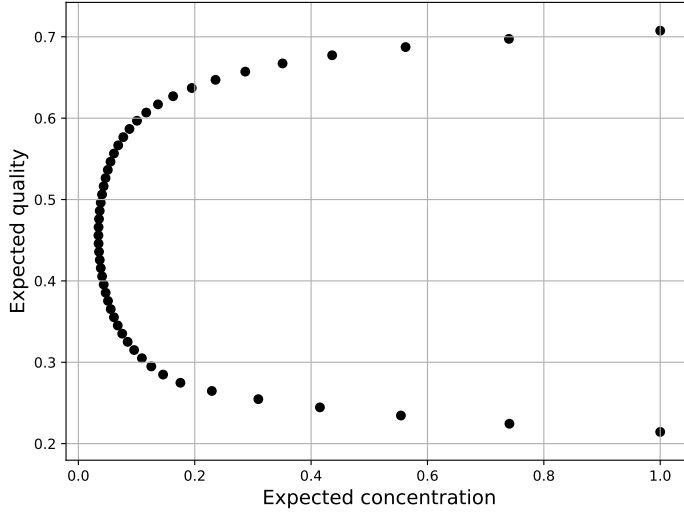


Fig. 1 Bullet-shaped quality-concentration set of solutions.

where $H_k = \mathbf{x}_k^T \mathbf{x}_k$ is the Herfindhal index obtained by budget allocation \mathbf{x}_k , and H_{max} and H_{min} are, respectively, the maximum and minimum Herfindhal indexes among all 50 optimal allocations.

The main advantage of the use of indexes θ_1 and θ_2 is that we can graphically compare solutions in a normalized quality-diversification space through the construction of an efficient frontier with non-dominated budget allocations. The efficient frontier for the first control case without the application of PCA and DA is shown in Figure 2. Within this normalized quality-diversification space, alternative solutions are compared in terms of the distance of each of them to the ideal point (1,1) where the maximum attainable quality and diversification are obtained. Finally, CP is based on the Zeleny's axiom of choice that states that alternatives that are closer to the ideal point are preferred to those that are further (Zeleny, 1973). In order to compute the distance to the ideal point when considering particular preferences, the following parametric family of weighted distances is commonly used in CP:

$$\mathcal{L}_h = [w^h \cdot (1 - \theta_1)^h + (1 - w)^h \cdot (1 - \theta_2)^h]^{1/h} \quad (17)$$

where h is a positive integer parameter and weight $w \in [0, 1]$ denotes the particular preference of the bank manager for quality and diversification. Note that \mathcal{L}_1 is the weighted Manhattan distance; \mathcal{L}_2 is the weighted Euclidean distance, and \mathcal{L}_∞ is the weighted Chebyshev distance. In what follows, we use the Euclidean distance ($h = 2$) to select the best budget allocation for a neutral bank manager, $w = 0.5$ i.e., without any bias to either quality or

diversification. We use the Euclidean distance since we think that it is a more widely known concept among bank managers.

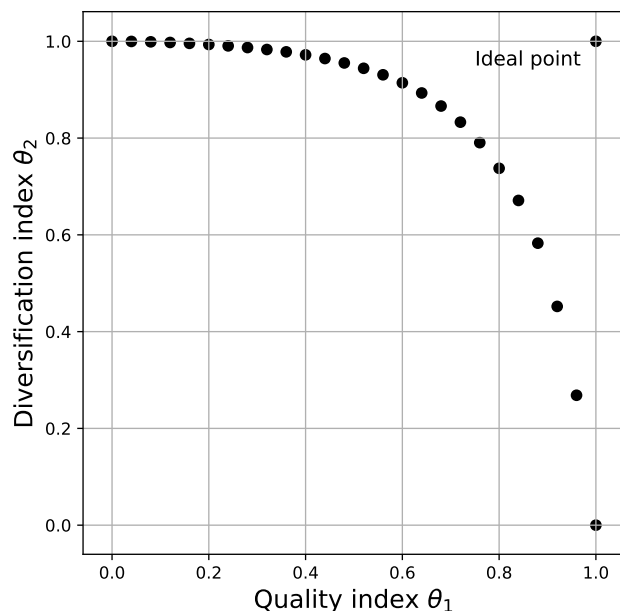


Fig. 2 A normalized quality-diversification efficient frontier.

Since we do not apply DA in this case, no company is vetoed and, consequently, all of them are initially allowed to apply for funds. However, the recommended budget allocation derived from our bicriteria quality-diversification approach only contains seven firms with a fraction higher than 1% as shown in the second column of Table 7. These results show that the mean budget is 14% but with a remarkable variation given by the standard deviation of firm's allocations and the range between the minimum and maximum budget. The larger budget percentage was allocated to VISCOFAN and the smallest to AMADEUS with almost no difference with respect to ENAGAS.

3.2 Case 2. DA: with DA and without PCA

Recall from Section 2.3, that an applicant is dominated if a convex combination of the rest of applicants is at least as good as the applicant. To determine if applicant a_k is dominated, we start from decision Table 1 and we solve the

following linear program for all $k = 1, 2, \dots, m$ (Ballester, 2002):

$$\min \varphi_k \quad (18)$$

subject to:

$$\sum_{i=1}^m \varphi_i V_{ij} \geq V_{kj}, \quad \forall j = 1, 2, \dots, n \quad (19)$$

$$\sum_{i=1}^m \varphi_i = 1. \quad (20)$$

$$\varphi_i \geq 0, \quad \forall i = 1, 2, \dots, m. \quad (21)$$

If the solution for the previous linear program is $\varphi_k = 1$ and the remainder φ_i ($i = 1, 2, \dots, m; i \neq k$) are equal to zero, then applicant a_k is non-dominated. On the contrary, if $\varphi_k = 0$ and there is at least one φ_i ($i \neq k$) greater than zero, then applicant a_k is dominated. By solving the program encoded from equation (18) to equation (21) for all the applicants in Table 2 and all the criteria in Tables 3 and 4, we found that no firm was dominated by a convex combination of the rest of applicants. Consequently, the results obtained for Case 1: Control remain valid for the Case 2: DA, as summarized in Table 7. It is important to highlight that although DA has no influence in our case study for 29 applicants and 55 different criteria, this fact does not imply that bank managers should not routinely rely on DA to make important decisions. This kind of analysis helps identify applicants whose performance is below a given threshold, which we think it is a critical task to manage risk. The method described in this section represents a suitable tool to do it.

3.3 Case 3. PCA: without DA and with PCA

Although the total set of criteria amounts to 55, when analyzing the interdependencies between criteria included in each set gathered in Tables 3 and 4, PCA yields a reduced set of 24 criteria. In our context, ratios within the same set may be correlated. Then, by applying PCA to criteria that may contain similar information (profitability, liquidity, solvency, cash-flow and market), we aim to reduce the impact of correlation within our initial data set. To this end, we use SPSS Statistics software package to apply PCA to every set of ratios in Tables 3 and 4, except for the bankruptcy predictive models (BM) and non-financial data (NF), since we do not expect these ratios to be correlated. More precisely, we use varimax rotation with Kaiser normalization (Kaiser, 1958) and select those components that explain at least 80% of the variance as shown in Table 6. In order to measure the adequacy of the PCA dimensionality reduction, we use the Kaiser-Meyer-Olkin (KMO) test (Kaiser, 1974; Cerny and Kaiser, 1977) and the sphericity Bartlett's test (Bartlett, 1937; Snedecor, 1989). In our experiments, the KMO coefficient is above 0.6 and the p-value in the sphericity Bartlett's test is below 0.001. Once the PCA

transformation is made, the dataset is no longer interpretable in terms of criteria. It represents a new data set that we use to assess the quality of credit applicants.

Table 6 Sets, components and total variance explained by PCA (Comp=Component; Var=Variance; Cum=Cumulative).

Set	Ratio Id	Comp	Criteria's loading	Total variance	
				% Var	Cum %
Profitability	1-5	PR 1	Returns	50,76%	50,76%
	6-7	PR 2	Margin of sales	32,05%	82,81%
Liquidity	8-11	LR 1	Accounts receivable	47,51%	47,51%
	12-14	LR 2	Current liabilities	21,44%	68,95%
	15-16	LR 3	Suppliers	15,48%	84,42%
Solvency	17-20	SR 1	Liabilities	41,27%	41,27%
	21-23	SR 2	Non-current liabilities	22,56%	63,83%
	24-27	SR 3	Finance autonomy	14,50%	78,33%
	28-29	SR 4	Financial leverage	8,30%	86,63%
Cash-flow	30-33	CF 1	Operating cash-flow	55,93%	55,93%
	34-36	CF 2	Earning and cash-flow	30,63%	86,56%
Market	44-46	MR 1	Value per share	34,30%	34,30%
	47-49	MR 2	Price of market	20,77%	55,07%
	50-51	MR 3	Business value	13,80%	68,87%
	52-53	MR 4	Value per share	11,37%	80,25%

The output of PCA is a decision table with 29 firms and 24 different components playing the role of new criteria to apply the MP decision rule. Notice that criteria groups BM and NF maintain its initial number of criteria after applying PCA. As a result, we obtain a second evaluation of firms in terms of quality of debt that we use to establish interest rates. The correlation between the evaluation derived from the Control case and this PCA case is depicted in Figure 3. As expected, we observe a strong correlation between evaluations represented by the trend line in the figure.

The next step in our case study is the use of this new evaluation to determine an additional optimal budget allocation. The results for the PCA case are summarized in the fourth column of Table 7. Interestingly, both the quality index and the diversification index increased with respect to the control case, moving closer to the ideal point (1, 1). We can reasonably think that this behavior is produced by the ability of PCA to reduce, to some extent, the influence of correlated (redundant) criteria. Furthermore, this performance in terms of quality and diversification is achieved by a reduced number of firms in the final budget with a more balanced allocation as the mean, the standard deviation, and the maximum and minimum budget allocation results show. The presence of a 37% allocation for VISCOFAN in the Control case penalizes

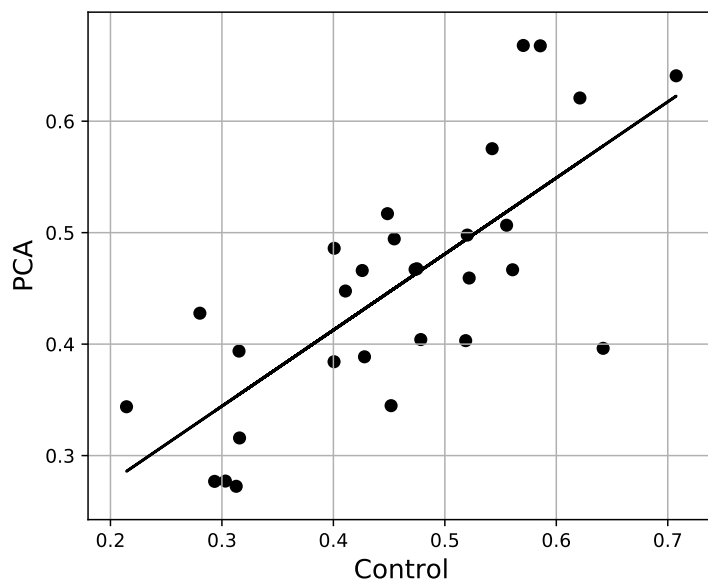


Fig. 3 Correlation between Control and PCA evaluations.

more the diversification index than the reduced number of firms in the PCA case.

It is important to highlight the fact that a new firm (ZARDOYA) is included in the optimal budget allocation in the PCA case. Moreover, two companies (EDP RENOVAVEIS and AMADEUS) are excluded from the optimal allocation. This point may be caused by the impact that PCA has on redundancy reduction. By applying PCA, the influence of redundant criteria is reduced and a new quality index is obtained. Firms that reduced their quality index in Table 5 are either excluded (EDP RENOVAVEIS and AMADEUS) or reduced (VISCOFAN) in terms of budget allocation. Firms that increased their quality index are either promoted (CATALANA and EBRO FOODS) or included (ZARDOYA). Finally, firms experimenting no change in their evaluation (VIDRALA), remain with the same budget allocation. All these statements hold when the quality index is enough to enter in the optimal budget allocation. In other words, in spite of experimenting a remarkable improvement in their quality indexes, some firms (A3MEDIA or ACS) do not enter in the optimal budget allocation due to a low initial quality index value.

3.4 Case 4. DA-PCA: with DA and with PCA

The fourth case that we analyze in this paper is the application of DA to a reduced decision table with 24 criteria after PCA. As shown in Table 5, three firms (IAG, PROSEGUR and MELIA) are dominated when using the DA procedure described in Section 3.2. The reason for the existence of dominated firms in this case should be found in the fact that PCA produces dimensionality reduction. From an initial set of 55 criteria, we obtain a smaller set of 24 components. A reduced number of criteria implies a higher probability that DA may influence in the analysis. Indeed, in the extreme case when only one criterion is considered, all applicants except one would be dominated.

However, the results derived from the CP model for quality-diversification optimization, produced no change in the optimal budget allocation with respect to Case 3: PCA. The reason for these results is that the quality indexes for the dominated firms were not high enough to enter in the optimal budget allocation. Consequently, DA after PCA had no influence in this case.

3.5 Case 5. Control with winsorization

To study the effect of outliers in data, we here solve Case 1 after winsorizing data according to equation (3) with $\alpha = 0.975$. We replace values above the 0.975-percentile with exactly the 0.975-percentile value and values below the 0.025-percentile with the 0.025-percentile value from Appendix B. Later, we get score S_i for each firm by means of equation (4) and aggregation weights $w_j^{0.975}$. Scores after winsorizing extreme values gathered in Table 5 (column Winsor) show a slight difference with respect to the control case. However, the results derived from the optimization procedure summarized in Table 7, show several adjustments in the percentage allocation for each firm and the inclusion of an additional firm (ZARDOYA). These results suggest that management of outliers in data may have an impact in budget allocation.

3.6 Summary of results and analysis

From the results summarized in Table 7, we next further elaborate on some interesting findings. Recall that we aimed to apply a MP decision rule to establish interest rates for a set of applicants and CP to determine the optimal budget allocation among them in terms of quality and diversification. In order to analyze the impact of DA, PCA and outlier management, we considered five different cases: a control case, DA without PCA, PCA without DA, the joint effect of DA and PCA, and winsorization.

Disregarding the application of DA and PCA, the first interesting finding is that a reduced number of companies is enough to reach optimal quality-diversification budget allocations. Indeed, when using the MP decision rule to establish quality from a set of criteria, and when measuring diversification by

Table 7 Summary of best results for the case study

Case	Control	DA	PCA	DA-PCA	Winsor
Quality index θ_1	0,75	0,75	0,88	0,88	0,70
Diversification index θ_2	0,79	0,79	0,81	0,82	0,79
Mean budget	0,14	0,14	0,20	0,20	0,13
Standard deviation budget	0,12	0,12	0,06	0,06	0,12
Max budget	0,37	0,37	0,25	0,25	0,42
Min budget	0,03	0,03	0,10	0,10	0,04
Firms with budget > 1%	7	7	5	5	8
Firms budget for case	Control	DA	PCA	DA-PCA	Winsor
EDP RENOVAVEIS	0,22	0,22	-	-	0,09
CATALANA	0,10	0,10	0,25	0,25	0,08
EBRO FOODS	0,06	0,06	0,25	0,25	0,04
VISCOFAN	0,37	0,37	0,21	0,21	0,42
AMADEUS	0,03	0,03	-	-	0,10
VIDRALA	0,18	0,18	0,18	0,18	0,18
ENAGAS	0,04	0,04	-	-	0,04
ZARDOYA	-	-	0,10	0,10	0,06

means of the complement of the Herfindhal concentration index, only a small fraction from 17 to 24% of the initial set of 29 applicants is enough to ensure an optimal allocation of funds. This fact does not mean that the rest of the applicants do not deserve credit. Our approach is a decision support tool to evaluate applicants in terms of quality and also to allocate a credit budget considering diversification. By using this tool, bank managers are empowered to manage risk by allocating funds to those firms with more capability to pay their debts. This management process is a rational one, i.e., a process that allows bank managers to choose between alternatives by means of an optimization process. As long as the credit market is large enough, applicants not included in the optimal credit allocation may find credit in other banks. Furthermore, if the initial set of applicants had not been formed by 29 but for only 2 applicants, any bank manager would have been unable to achieve enough risk diversification. As a result, we claim that our method is a tool for decision making, not a method to ensure credit for all applicants.

A similar reasoning fits well in the explanation of our second finding: the reduced impact that DA had on the final results. While the procedure of DA and veto within a MP decision rule may seem quite strong at first glance, credit applicants are economic agents competing for funds. It is reasonable to accept that better qualified applicants obtain a larger fraction of available funds. DA plays the role of a threshold of required performance measured in terms of the set of criteria under consideration. However, when considering the optimal allocation of funds in terms of quality and diversification, even a remarkable subset of companies are excluded from the best alternative in spite of being accepted for consideration by the MP rule. As a result, quality-diversification

optimization is a stronger rule than DA. The influence of DA is limited by the fact that it is likely that dominated companies present lower quality results than the rest, hence preventing them to form part of the optimal allocation.

Note, however, that a lower number of criteria implies a higher probability that DA may influence in the analysis since the comparison has lower dimensions and the probability that an applicant performs worse than the rest of applicants for all the criteria increases. The extreme case would be a single criterion, with an applicant dominating the rest. In this sense, dimensionality reduction techniques such as PCA may influence the final allocation of credit as described in Case 3. The ability of PCA to reduce the impact of redundant criteria may benefit some companies at the expense of others. However, we think this redundancy reduction is a clear advantage of the use of PCA to improve credit allocation. Only uncorrelated and relevant criteria should be considered to evaluate credit applicants. To this end, PCA is a sound tool to better estimate the relevant set of criteria, even when an arbitrary initial set is used. Furthermore, we observe better performance in terms of quality and diversification when PCA is applied with more balanced results.

A final comment must be done about outlier treatment. Our results show that procedures to winsorize data modify quality scores and ultimately the optimal budget allocation. In this sense, adapting the initial quality evaluation by means of winsorization to deal with outliers may be regarded as a suitable method in the deployment of the MP decision rule in practice. To this end, our MP methodology can be adjusted to the needs of practitioners by means of equation (3).

4 Concluding remarks

In this paper, we propose the use a MP decision rule to establish interest rates for a set of applicants and CP to determine the optimal budget allocation among them in terms of quality and diversification. On the one hand, the use of MP ensures that neither extreme optimistic and pessimistic decisions are taken. On the other hand, CP allows to find a balance between quality of credit applicants and diversification of risk. As a result, our model concatenates moderate pessimism by Ballestero (2002) and the compromise programming axiom of choice by Zeleny (1973) to produce two critical outputs: 1) a quality evaluation determining the recommended interest rates for each applicant within a given interval; and 2) the optimum allocation of credit funds among applicants in terms of quality and diversification.

To assess the ability of applicants to repay their loans, we rely on financial and market information, usually in the form of a set ratios. Our results show that dimensionality reduction techniques such as PCA help improve credit allocation by reducing redundancy in the set of criteria under consideration. Furthermore, we observe better performance in terms of quality and diversification when PCA is applied. We do not claim that these results can be generalized to every set of applicants. However, we firmly believe that PCA

represents a recommended preliminary step to better estimate quality among applicants. In this sense, the CP model described in this paper is also a useful tool for evaluation purposes since a range of preferences for quality and diversification ensures the flexibility to adjust decisions to different profiles.

The use of a MP decision rule implies the application of DA as a required step. However, we find that DA has a low impact on quality-diversification optimization results. The impact of DA is limited by the likely fact that dominated firms present low quality criteria. This fact leads us to conclude that quality-diversification optimization is a stronger restriction than DA.

One of the limitations of our approach is its application in terms of time scheduling. In practice, banks receive credit applications that have to be accepted or rejected on a continuous basis. As a result, our approach can be applied when the rate of applications is enough to use the model with a sufficient number of applicants. In this sense, the insight derived from the case study on the reduced number of companies that is able to produce optimal CP results in terms of quality of debt and diversification of risk is a remarkable finding. Provided that a large set of firms apply for credit in a single bank, only a fraction of them is enough to achieve an optimal allocation of funds. However, it is also important to say that our method is a tool for decision making, and that many other commercial and technical questions should be considered to apply this method in practice. Some of these questions can be implemented by means of some lower and upper bounds for loan allocations in our CP model. These constraints allow the possibility to set a minimum allocation for applicants and also setting an upper bound to avoid allocations in excess of the requested loan amount.

Finally, since criteria may vary with time, a natural extension of this work would be the analysis of optimal budget allocations when changes in criteria realizations result in a different quality evaluation. An additional interesting future line of work would be a comparative analysis of MP and other decision rules such as Wald's or Laplace's criteria. On the bank management practical side, the study of correlations between loans, the incorporation of portfolio risk measures as additional goals, and further empirical validation of the model presented in this paper and other possible variants represent interesting topics for future research.

References

- Abdi, H. and Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459.
- Adams, W., Einav, L., and Levin, J. (2009). Liquidity constraints and imperfect information in subprime lending. *American Economic Review*, 99(1):49–84.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609.

- Altman, E. I., Iwanicz-Drozdzowska, M., Laitinen, E. K., and Suvas, A. (2016). Financial and non-financial variables as long-horizon predictors of bankruptcy. *The Journal of Credit Risk*, 12(4):49–78.
- Angilella, S. and Mazzù, S. (2015). The financing of innovative smes: A multicriteria credit rating model. *European Journal of Operational Research*, 244(2):540–554.
- Armor, D. J. (1973). Theta reliability and factor scaling. *Sociological methodology*, 5:17–50.
- Avery, R. B., Bostic, R. W., Calem, P. S., and Canner, G. B. (2000). Credit scoring: Statistical issues and evidence from credit-bureau files. *Real Estate Economics*, 28(3):523–547.
- Ballester, E. (2002). Strict uncertainty: A criterion for moderately pessimistic decision makers. *Decision Sciences*, 33(1):87–108.
- Ballester, E. (2006). Ranking alternatives from the decision maker's preferences: An approach based on utility and the notion of marginal action. *Journal of the Operations Research Society of Japan*, 49(1):49–65.
- Ballester, E., Günther, M., Pla-Santamaria, D., and Stummer, C. (2007). Portfolio selection under strict uncertainty: A multi-criteria methodology and its application to the frankfurt and vienna stock exchanges. *European Journal of Operational Research*, 181(3):1476–1487.
- Ballester, E. and Pla-Santamaria, D. (2004). Selecting portfolios for mutual funds. *Omega*, 32(5):385–394.
- Ballester, E. and Romero, C. (1998). *Multiple criteria decision making and its applications to economic problems*. Kluwer Academic Publishers, Dordrecht.
- Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 160(901):268–282.
- Bellman, R. (1957). *Dynamic programming*. Princeton University Press, New Jersey.
- Bravo, M. and Pla-Santamaria, D. (2012). Evaluating loan performance for bank offices: a multicriteria decision-making approach. *INFOR: Information Systems and Operational Research*, 50(3):127–133.
- Breeden, J. L. (2016). Incorporating lifecycle and environment in loan-level forecasts and stress tests. *European Journal of Operational Research*, 255(2):649–658.
- Carmines, E. G. and Zeller, R. A. (1979). *Reliability and validity assessment*. Quantitative applications in the social sciences. SAGE.
- Cerny, B. A. and Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate behavioral research*, 12(1):43–47.
- Chen, M.-Y. (2014). Using a hybrid evolution approach to forecast financial failures for taiwan-listed companies. *Quantitative Finance*, 14(6):1047–1058.
- Conway, J. M. and Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational research methods*, 6(2):147–168.

- Doumpos, M. and Zopounidis, C. (2011). A multicriteria outranking modeling approach for credit rating. *Decision Sciences*, 42(3):721–742.
- Edelberg, W. (2006). Risk-based pricing of interest rates for consumer loans. *Journal of Monetary Economics*, 53(8):2283–2298.
- Einav, L., Jenkins, M., and Levin, J. (2012). Contract pricing in consumer credit markets. *Econometrica*, 80(4):1387–1432.
- Einav, L., Jenkins, M., and Levin, J. (2013). The impact of credit scoring on consumer lending. *The RAND Journal of Economics*, 44(2):249–274.
- Elisabetsky, R. (1976). *Um modelo matemático para decisões de crédito no banco comercial*. PhD thesis, Dissertação (Mestrado)–Escola Politécnica, Universidade de São Paulo, São Paulo.
- Eriksson, K., Jonsson, S., Lindbergh, J., and Lindstrand, A. (2014). Modeling firm specific internationalization risk: An application to banks' risk assessment in lending to firms that do international business. *International Business Review*, 23(6):1074–1085.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4(3):272.
- Falangis, K. (2007). The use of MSD model in credit scoring. *Operational Research*, 7(3):481–503.
- Ford, J. K., MacCallum, R. C., and Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel psychology*, 39(2):291–314.
- Freeman, L. and Hamilton, D. (2002). A dream deferred or realized: The impact of public policy on fostering black homeownership in new york city throughout the 1990's. *American Economic Review*, 92(2):320–324.
- Jory, S. R., Ngo, T. N., and Wang, D. (2016). Credit ratings and the premiums paid in mergers and acquisitions. *Journal of Empirical Finance*, 39:93–104.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3):187–200.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1):31–36.
- Kanitz, S. (1974). Como prever falencias de empresas. *Revista de Negocios Em Exame, Dezembro*, pages 95—102.
- Kantardzic, M. (2011). *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, New York.
- Kealhofer, S. (1993). Portfolio management of default risk. *KMV Corporation*.
- Laplace, P.-S. (1825). *Essai philosophique sur les probabilités*. H. Remy, Paris.
- Legault, J. and Score, A. (1987). Ca score. warning system for small business failures. *Bilanas, June*, pages 29–31.
- Li, H., Hong, L.-Y., Zhou, Q., and Yu, H.-J. (2015). The assisted prediction modelling frame with hybridisation and ensemble for business risk forecasting and an implementation. *International Journal of Systems Science*, 46(11):2072–2086.
- Magee, S. (2013). The effect of foreign currency hedging on the probability of financial distress. *Accounting & Finance*, 53(4):1107–1127.

- McQuown, J. (1993). Market versus accounting-based measures of default risk. *KMV Corporation*.
- Mencía, J. (2012). Assessing the risk-return trade-off in loan portfolios. *Journal of Banking & Finance*, 36(6):1665–1677.
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470.
- Mester, L. J. (1997). What’s the point of credit scoring? *Business review*, 3(Sep/Oct):3–16.
- Nunnally, J. C. and Bernstein, I. (1994). *Psychometric Theory (McGraw-Hill Series in Psychology)*, volume 3. McGraw-Hill New York.
- Saunders, D., Xiouros, C., and Zenios, S. A. (2007). Credit risk optimization using factor models. *Annals of Operations Research*, 152(1):49–77.
- Sengupta, R. and Bhardwaj, G. (2015). Credit scoring and loan default. *International Review of Finance*, 15(2):139–167.
- Shi, J. and Xu, B. (2016). Credit scoring by fuzzy support vector machines with a novel membership function. *Journal of Risk and Financial Management*, 9(4):13.
- Sirignano, J. and Giesecke, K. (2019). Risk analysis for large pools of loans. *Management Science*, 65(1):107–121.
- Sirignano, J. A., Tsoukalas, G., and Giesecke, K. (2016). Large-scale loan portfolio selection. *Operations Research*, 64(6):1239–1255.
- Snedecor, George W, C. W. G. (1989). *Statistical methods*. Iowa State University Press, Iowa.
- Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Number 82. Sage, Newbury Park.
- Springate, G. L. (1978). *Predicting the possibility of failure in a Canadian firm: a discriminant analysis*. PhD thesis, Simon Fraser University.
- Thomas, L. C. (2009). *Consumer credit models: pricing, profit and portfolios: pricing, profit and portfolios*. Oxford University Press, Oxford.
- Vasicek, O. (1984). The philosophy of credit valuation: the credit valuation model. *KMV Corporation*.
- Wald, A. (1950). *Statistical decision functions*. Wiley, New York.
- Wang, J. and Wang, J. (2015). Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. *Neurocomputing*, 156:68–78.
- Weber, O. (2012). Environmental credit risk management in banks and financial service institutions. *Business Strategy and the Environment*, 21(4):248–263.
- Woerheide, W. and Persson, D. (1992). An index of portfolio diversification. *Financial services review*, 2(2):73–85.
- Yu, P.-L. (1985). *Multiple-criteria decision making: concepts, techniques, and extensions*. Plenum Press, New York.
- Zeleny, M. (1973). Compromise programming. In *Multiple criteria decision making (Cochrane, JL, Zeleny, M, Editors)*, pages 262–301. University of South Carolina Press, Columbia.
- Zeleny, M. (1982). *Multiple criteria decision making*. McGraw-Hill, New York.

A Appendix: Definition of criteria

	Criterion	Definitions
1	Return on invested capital (ROIC)	$\frac{\text{NOPAT}}{\text{EQUITY} + \text{NET DEBT}}$
2	Return on assets (ROA)	$\frac{\text{EBIT}}{\text{ASSETS}}$
3	Assets turnover	$\frac{\text{SALES}}{\text{ASSETS}}$
4	Return on equity (ROE)	$\frac{\text{NET INCOME}}{\text{EQUITY}}$
5	Pretax income-to-equity ratio	$\frac{\text{PRE-TAX INCOME}}{\text{EQUITY}}$
6	Ebit-to-sales ratio	$\frac{\text{EBIT}}{\text{SALES}}$
7	Net value added-to-sales ratio	$\frac{\text{NET VALUE ADDED}}{\text{SALES}}$
8	Quick ratio	$\frac{\text{ASSETS} - \text{INVENTORY}}{\text{LIABILITIES}}$
9	Current ratio	$\frac{\text{CURRENT ASSETS}}{\text{CURRENT LIABILITIES}}$
10	Accounts receivable net turnover	$\frac{\text{SALES}}{\text{AVERAGE RECEIVABLES}}$
11	Total accounts receivable turnover	$\frac{\text{REVENUE}}{\text{AVERAGE RECEIVABLES}}$
12	Working capital-to-assets ratio	$\frac{\text{WORKING CAPITAL}}{\text{ASSETS}}$
13	Cash ratio	$\frac{\text{CASH}}{\text{LIABILITIES}}$
14	Current liabilities-to-assets ratio	$\frac{\text{LIABILITIES}}{\text{ASSETS}}$
15	Trade payable turnover	$\frac{\text{REVENUE} - \text{COST} + \text{OTHER EXPENSES}}{\text{AVERAGE TRADE PAYABLE}}$
16	Suppliers turnover	$\frac{\text{COST OF SALES}}{\text{AVERAGE SUPPLIERS}}$
17	Financial leverage	$\frac{\text{PRE-TAX INCOME}}{\text{EQUITY}} \times \frac{\text{EBIT}}{\text{ASSETS}}$
18	Assets-to-liabilities ratio	$\frac{\text{ASSETS}}{\text{LIABILITIES}}$
19	Liabilities-to-(Liabilities+equity) ratio	$\frac{\text{LIABILITIES}}{\text{LIABILITIES} + \text{EQUITY}}$
20	Liabilities-to-equity ratio	$\frac{\text{LIABILITIES}}{\text{EQUITY}}$
21	Equity-to-assets ratio	$\frac{\text{EQUITY}}{\text{ASSETS}}$
22	Retained earnings-to-equity ratio	$\frac{\text{RETAINED EARNINGS}}{\text{EQUITY}}$
23	Retained earnings-to-assets ratio	$\frac{\text{RETAINED EARNINGS}}{\text{ASSETS}}$
24	(Non-current liabilities + equity)-to-assets ratio	$\frac{\text{NON-CURRENT LIABILITIES} + \text{EQUITY}}{\text{ASSETS}}$
25	Non-current liabilities-to-liabilities ratio	$\frac{\text{NON-CURRENT LIABILITIES}}{\text{LIABILITIES}}$
26	Debt-to-liabilities ratio	$\frac{\text{DEBT}}{\text{LIABILITIES}}$
27	Non-current liabilities-to-assets ratio	$\frac{\text{NON-CURRENT LIABILITIES}}{\text{ASSETS}}$
28	Leverage ratio	$\frac{\text{DEBT}}{\text{EQUITY}}$

29	Debt average cost	$\frac{\text{INTEREST EXPENSE}}{\text{AVERAGE DEBT}}$
30	Short term debt-to-net operating cash flow ratio	$\frac{\text{SHORT TERM DEBT}}{\text{NET OPERATING CASH FLOW}}$
31	Debt-to-net operating cash flow ratio	$\frac{\text{DEBT}}{\text{NET OPERATING CASH FLOW}}$
32	Current liabilities-to-net operating cash flow ratio	$\frac{\text{CURRENT LIABILITIES}}{\text{NET OPERATING CASH FLOW}}$
33	Liabilities -to- net operating cash flow ratio	$\frac{\text{LIABILITIES}}{\text{NET OPERATING CASH FLOW}}$
34	EBIT-to-net operating cash flow before interest and taxes ratio	$\frac{\text{EBIT}}{\text{NET OPERATING CASH FLOW BIT}}$
35	NOPAT-to-net operating cash flow before interest ratio	$\frac{\text{NOPAT}}{\text{NET OPERATING CASH FLOW BI}}$
36	Net income-to-net operating cash flow ratio	$\frac{\text{NET INCOME}}{\text{NET OPERATING CASH FLOW}}$
37	Legault and Score (1987)	Bankruptcy predictive model
38	Elisabetsky (1976)	Bankruptcy predictive model
39	Kanitz (1974)	Bankruptcy predictive model
40	Springate (1978)	Bankruptcy predictive model
41	Altman (1968)	Bankruptcy predictive model
42	Merton (1974)	Bankruptcy predictive model
43	Credit rating	Bankruptcy predictive model
44	Earnings per share	$\frac{\text{NET INCOME}}{\text{NUMBER OF SHARES}}$
45	Sales per share	$\frac{\text{SALES}}{\text{NUMBER OF SHARES}}$
46	Book value per share	$\frac{\text{EQUITY}}{\text{NUMBER OF SHARES}}$
47	Price-to-sales ratio	$\frac{\text{PRICE}}{\text{SALES}}$
48	Price-to-cash flow ratio	$\frac{\text{PRICE}}{\text{CASH FLOW}}$
49	Price-to-book value ratio	$\frac{\text{PRICE}}{\text{EQUITY}}$
50	Price-to-earnings ratio (PER)	$\frac{\text{PRICE}}{\text{NET INCOME}}$
51	Enterprise Value (EV)-to-Ebitda ratio	$\frac{\text{ENTERPRISE VALUE}}{\text{EBITDA}}$
52	Cash flow per share	$\frac{\text{CASH FLOW}}{\text{NUMBER OF SHARES}}$
53	Dividends per share	$\frac{\text{DIVIDENDS}}{\text{NUMBER OF SHARES}}$
54	Years from foundation	(Self-explanatory)
55	Expectations on the economic sector	(Self-explanatory)

B Appendix: Summary of criteria statistics

Key:

- **Mean:** Average value.
- **Std:** Standard deviation.
- **Max:** Maximum value.
- **Min:** Minimum value.
- **0,025-Pct:** 0.025-Percentile value.
- **0,975-Pct:** 0.975-Percentile value.

Criterion	Mean	Std	Max	Min	0,025-Pct	0,975-Pct
1	0,08	0,07	0,36	-0,03	-0,01	0,25
2	0,07	0,06	0,32	-0,05	-0,02	0,23
3	0,60	0,55	2,92	0,02	0,11	1,93
4	0,13	0,23	1,07	-0,50	-0,24	0,66
5	0,20	0,30	1,56	-0,10	-0,06	0,95
6	0,25	0,28	0,95	-0,12	-0,05	0,63
7	0,42	0,25	1,03	0,08	0,10	0,75
8	0,99	0,78	4,68	0,08	0,25	1,46
9	1,17	0,80	4,68	0,08	0,39	1,82
10	918,72	4.792,10	26.275,89	1,59	1,81	72,03
11	23,08	84,50	466,61	0,79	0,92	36,41
12	-0,01	0,15	0,26	-0,43	-0,39	0,17
13	0,29	0,22	1,00	0,00	0,01	0,78
14	0,28	0,16	0,70	0,02	0,11	0,62
15	2,06	2,22	10,17	0,00	0,00	7,33
16	3,93	2,62	11,53	0,28	1,33	10,07
17	0,37	0,24	0,80	-0,09	-0,06	0,80
18	0,38	0,19	0,89	0,05	0,11	0,64
19	2,00	1,53	9,35	1,06	1,12	2,79
20	0,62	0,19	0,95	0,11	0,36	0,89
21	0,38	0,84	1,20	-3,59	-1,43	1,05
22	0,20	0,18	0,56	-0,20	-0,07	0,56
23	2,65	3,15	17,35	0,12	0,56	10,65
24	0,72	0,16	0,98	0,30	0,38	0,89
25	0,45	0,21	1,00	0,07	0,15	0,75
26	0,57	0,22	0,99	0,09	0,20	0,82
27	0,34	0,16	0,66	0,00	0,11	0,62
28	2,09	4,08	16,15	-12,79	-5,75	11,08
29	0,05	0,02	0,09	0,02	0,03	0,09
30	2,73	6,93	35,40	-6,26	-2,47	23,05
31	13,57	42,76	236,48	-18,95	-7,49	110,44
32	7,47	20,50	105,49	-30,45	-11,50	58,31
33	21,15	64,39	353,46	-47,85	-18,15	165,36
34	1,18	2,64	8,97	-9,63	-4,34	6,09

Criterion	Mean	Std	Max	Min	0,025-Pct	0,975-Pct
35	1,43	3,30	12,28	-10,10	-5,31	8,91
36	2,86	5,71	27,05	-1,99	-1,81	21,11
37	-0,44	0,93	1,74	-1,98	-1,73	1,54
38	-0,13	1,02	1,43	-4,85	-2,48	0,51
39	7,23	4,98	28,28	-2,28	-0,21	9,62
40	0,78	2,73	12,95	-0,97	-0,82	1,24
41	2,52	2,41	12,95	-0,36	0,37	9,36
42	7,36	3,38	16,69	0,14	1,43	15,10
43	0,59	0,13	0,85	0,27	0,30	0,84
44	25,81	51,14	248,08	-60,73	-44,99	160,68
45	8,35	32,01	55,39	-150,65	-58,05	28,12
46	3,32	4,54	23,05	0,25	0,59	16,94
47	4,40	8,11	36,32	0,21	0,34	5,47
48	30,30	82,10	447,84	0,12	1,47	252,31
49	16,26	37,96	212,76	0,28	0,70	105,17
50	2,00	2,89	11,28	-5,14	-3,67	8,14
51	1,34	35,54	30,00	-162,17	-32,60	26,83
52	1,87	4,86	26,56	-2,06	-1,05	13,06
53	0,10	0,25	0,90	0,00	0,00	0,83
54	0,52	0,34	1,49	0,02	0,15	1,06
55	0,28	0,22	0,75	0,00	0,00	0,75