COMISEF Computational Optimization Methods in Statistics, Econometrics and Finance

- Marie Curie Research and Training Network funded by the EU Commission through MRTN-CT-2006-034270 -

COMISEF WORKING PAPERS SERIES

WPS-039 25/05/2010

# Threshold Accepting for Credit Risk Assessment and Validation

M. Lyra
A. Onwunta
P. Winker

# Threshold Accepting for Credit Risk Assessment and Validation*

Marianna Lyra[†] Akwum Onwunta[‡] Peter Winker[§]

May 25, 2010

**Abstract**

According to the latest Basel framework of Banking Supervision, financial institutions should internally assign their borrowers into a number of homogeneous groups. Each group is assigned a probability of default which distinguishes it from other groups. This study aims at determining the optimal number and size of groups that allow for statistical ex post validation of the efficiency of the credit risk assignment system. Our credit risk assignment approach is based on Threshold Accepting, a local search optimization technique, which has recently performed reliably in credit risk clustering especially when considering several realistic constraints. Using a relatively large real-world retail credit portfolio, we propose a new technique to validate ex post the precision of the grading system.

**Keywords:** credit risk assignment, Threshold Accepting, statistical validation.

---

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of Deutsche Bank AG.

[†]Department of Economics, University of Giessen, Germany.

[‡]Deutsche Bank AG, Frankfurt, Germany.

[§]Department of Economics, University of Giessen, Germany, and Center for European Economic Research (ZEW), Mannheim, Germany.

# 1 Introduction

It is evident from the current financial and credit market crisis that credit institutions should pursue a more valid approach to credit risk assignment based on realistic assumptions. A core input for assigning bank clients into credit classes are their default probabilities; that is, the probability that the clients will not fulfill their credit obligations during the next, say, 1 year. Having at hand these estimates, homogeneous clients are assigned into the same class.[1] The precise estimation, assignment and validation of credit risk permit banks to retain an adequate capital level. This will ensure their solvency not only under normal economic conditions, but also under more extreme negative economic conditions with a given confidence level. In this paper we deal with credit risk assignment and ex post validation.

The main contribution of this article is the determination of the optimal number and size of groups necessary for the ex post validation of the efficiency of a credit risk classification system. First, we grade borrowers using as an indicator their default probability. The grading is made in such a way as to minimize the *within* and maximize the *between* classes distance and satisfy the constraints imposed by Basel Committee on Banking Supervision (2006). Next, two validation techniques are compared that ensure the precision of the grading system. It is shown, that the precision level of the ex post validation techniques is dependent on the portfolio size. The first validation technique relies on the precise statement of the number of defaults. A lower and upper number of defaults in a bucket are defined using statistical benchmarks. The second is based on the correct statement of the unexpected losses, and hence the regulatory capital, as required by Basel Committee on Banking Supervision (2006).

We validate our rating system by making the simplifying assumption that default risks and actual defaults are independently and identically distributed. The independence assumption, to a reasonable extent, makes sense particularly for retail credit portfolios (in the context of which we are doing our empirical analysis) but does not hold in general. Obviously, defaults are a two-state event and we therefore assume that they are binomially distributed.[2]

To classify borrowers into homogeneous buckets, we rely on Threshold

[1] We will be using the terms 'class', 'grade' and 'bucket', interchangeably.

[2] To lessen the computational time though, in the empirical application we used the normal approximation to the binomial distribution since our data set is relatively large.

Accepting (TA) - an optimization technique suitable for the clustering of credit risk. It originated in the early 1990s (Dueck and Scheuer (1990)) and was established a few years later for applications in economics and finance by Winker (2001). Under real world conditions, TA, as shown in Lyra *et al.* (2010), grades borrowers more efficiently and effectively than other optimization techniques such as differential evolution and genetic algorithm.

The paper proceeds as follows. Section 2 reports on credit risk assessment under the framework of Basel II Capital Accord. Section 3 introduces the TA heuristic optimization technique and its basic properties. Two techniques on how to deal with Basel II constraints are also presented. Section 4 describes the two validation techniques. Section 5 presents the data structure and discusses the major results. Finally, Section 6 concludes and suggests further possible improvements.

# 2 Basel II and Credit Risk Assignment

## 2.1 General Framework

The latest framework of Banking Supervision - Basel II (Basel Committee on Banking Supervision, 2006) aims at maintaining the financial stability of credit institutions. It allows for a new form of credit risk management by providing the framework for internally estimating the capital level that ensures bank solvency. The Basel II framework comprises three parts: minimum capital requirements, authorities' supervision of capital adequacy and banks' disclosures for market supervision. Of particular importance is the adequate calculation of banks' capital to account for credit, operational and market risk. Our focus is on credit risk.

Under the new framework, there are two alternative approaches that a bank can apply to determine the adequate capital level for credit risk, namely, the standard approach and the internal rating-based approach (IRB). Institutions applying either approach should retain a minimum solvency ratio greater than 8%. The solvency ratio is defined as the capital requirement for assets' credit risk divided by the 'risk-weighted assets':

$$\text{Solvency ratio} = \frac{\text{capital requirement}}{\text{risk-weighted assets}} \geq 8\%\,. \tag{1}$$

The standard approach requires credit risk ratings or risk weights for asset values to be determined by an external credit rating agency. The IRB approach allows the internal estimation of credit risk. This new risk assessment practice allows for grading of small and medium bank clients, which are not graded by (authorized) credit rating agencies. The IRB approach classifies a bank's assets into sovereign, bank, corporate, retail and equity. For each asset category, a different capital is required. In our application, we consider retail borrowers.

Banks are required to retain a minimum capital level to cover unexpected losses from defaulted borrowers. Unexpected losses might result, for example, from economic depression. While banks can cover expected losses through pricing, provisions and write-offs, unexpected losses have to be covered with capital charges on credit instruments (Basel Committee on Banking Supervision, 2005). Here, we assume that a bank's stability is conditional only on one factor, e.g. the general economic condition.

Next, to motivate the concept of conditional default probability which we will often use in the rest of the paper, we adopt the structural (firm-value) credit risk modeling approach based on a one-factor model. In a structural model, a firm is said to default if its asset value (hereafter called the ability-to-pay) falls below a certain threshold called the default threshold (Merton (1974)). Let $n$ be the number of borrowers in a given loan portfolio. For a given borrower $k \in \{1, \cdots, n\}$ the expected loss amount, $L_k$, equals the product of the exposure at default, $EAD_k \in \mathbb{R}$, the loss given default, $LGD_k \in [0, 1]$, and the default probability, $p_k \in [0, 1]$; that is,

$$L_k = L(p_k) = EAD_k \cdot LGD_k \cdot p_k.$$

Borrower $k$ defaults when its ability-to-pay variable $A_k$ has fallen below a pre-specified threshold value, say, $c_k$. More formally,

$$I_k = 1 \Leftrightarrow A_k \leq c_k, \tag{2}$$

where $I_k$ is the default indicator:

$$I_k = \mathbf{1}_{\{A_k \leq c_k\}}. \tag{3}$$

Here, we make the simplifying assumption that $A_k$ is standardized and normally distributed. Thus, the probability of default $p_k \in [0, 1]$ of the $k$th obligor in the portfolio is given by

$$p_k = \mathbb{P}(A_k \leq c_k) = N(c_k),$$

where $N$ is the cumulative standard normal distribution function. The default threshold $c_k$ is therefore linked to the default probability via

$$c_k = N^{-1}(p_k). \tag{4}$$

In order to aggregate the loss variables $L_k$ of the individual obligors to a loss variable $L$ of the entire portfolio, the dependence structure of $A_k$ needs to be specified. Of course, obligors default and therefore the Bernoulli variables $I_k$ are generally not independent.[3]

Firms' financial well-being depends on macroeconomic factors. This induces cyclical default dependence. Consequently, a firm's default probability is a function of the realization of the state of the economy and is thus random. This leads to the introduction of a factor model consisting of systematic and idiosyncratic factors (see e.g. Bluhm *et al.* (2002)). More precisely, each ability-to-pay variable $A_k$ is decomposed into a sum of systematic factors $\psi_1, \ldots, \psi_m$, with weights $w_{k1}, \ldots, w_{km}, m < n$ and an idiosyncratic (or specific) factor $\varepsilon_k$, that is

$$A_k = \sqrt{\rho_k} \sum_{i=1}^{m} w_{ki}\psi_i + \varepsilon_k \sqrt{1-\rho_k}. \tag{5}$$

Each systematic factor $\psi_i$ is a centered random variable and it is assumed that the vector $\psi = (\psi_1, \ldots, \psi_m)$ follows a $m$-dimensional normal distribution with mean $\mathbf{0} = (0, \ldots, 0)$ and covariance matrix $\Sigma$. The systematic weights $w_{k1}, \ldots, w_{km} \in \mathbb{R}$ determine the impact of each systematic factor on the ability-to-pay variable $A_k$. The systematic weights are scaled such that the systematic component

$$\phi_k := \sum_{i=1}^{m} w_{ki}\psi_i, \tag{6}$$

as well as each idiosyncratic factor $\varepsilon_k$, are standardized normally distributed variables. The idiosyncratic factors $\varepsilon_1, \ldots, \varepsilon_n$ are independent of each other as well as independent of the systematic factors.

An important instance of this model class is the homogeneous one-factor model. One-factor models have the advantage of incorporating simple dependence structures which are to a certain degree sufficient for a rudimentary

[3]In this analysis actual defaults are assumed to be independent. This is a rather sensible assumption for retail loans. Further work will be focused on relaxing the independence assumption and using a portfolio of corporate loans.

form of credit risk management.[4] In these models, obligors depend on a factor representing the general state of the economy. In particular, (5) specializes to

$$A_k = \psi\sqrt{\rho_k} + \varepsilon_k\sqrt{1-\rho_k}. \tag{7}$$

The dependency (or the asset correlation[5]) equals $\rho_k$. It represents the sensitivity of a firm with respect to the systematic factor $\psi$. For retail portfolios, it is given by (Basel Committee on Banking Supervision (2006), § 283)

$$\rho_k = 0.03\left[\frac{1-\exp(-35\cdot p_k)}{1-\exp(-35)}\right] + 0.16\left[1 - \frac{1-\exp(-35\cdot p_k)}{1-\exp(-35)}\right]. \tag{8}$$

In the asset correlation function (8) above, the figures 0.03 and 0.16 are the limit asset correlations for very high and very low default probabilities, 100% and 0%, respectively. Empirical findings in the literature, however, reveal that there is sufficient variation in asset correlations as to make a constant assumption inaccurate (Lopez (2002) and Roesch and Scheule (2010), Chapter 10). Correlations between these limits are modeled by a decaying exponential weighting function that displays the dependency on default probability. The rate of decay of the exponential function is determined by the so-called 'k-factor' which is 35 for retail exposures. This line of argument is motivated by empirical findings that there exists a negative relationship between asset correlation and default probability.

However, the relationship between asset correlation and default probability still remains a controversial issue since authors come to rather different conclusions. For instance, Lopez (2002) finds that asset correlation is a decreasing function of default probability and an increasing function of the firm's asset size. The former result seems quite intuitive: the higher the default probability, the higher the idiosyncratic (individual) risk components of a borrower. The default risk depends less on the overall state of the economy and more on the individual risk drivers. Nevertheless, the empirical evidence in Roesch and Scheule (2010), Chapter 10 and Dietsch and Petey (2004)

[4] However, it should be pointed out that portfolio models that use a single systematic factor across all industries do not provide sufficient flexibility to capture the complex dependence structure exhibited by the portfolio, see e.g. Roesch and Scheule (2010), Chapter 10.

[5] A formal definition of asset correlation and its relationship with default correlation are given in the Appendix.

show that there is no systematic dependence between the asset correlation and default probability.

Besides, Lee *et al.* (2009) find that the stylized relationship for retail exposures is an increasing one instead and may underestimate portfolio risk for high default probability portfolios. These findings therefore call for a review of the use of this relationship by the Basel Committee in the calculation of banks' capital requirement. Observe that (8) does not take into account the firm's size although Lopez (2002) argues that the larger a firm, the higher its dependency upon the overall state of the economy, and vice versa; smaller firms are more likely to default for idiosyncratic reasons. The paper further shows that the decline in $\rho_k$ due to an increase in default probability is faster in (8) than suggested by calibrated asset correlation. It finds, however, that for well-diversified portfolios, $\rho_k$ and calibrated asset correlation are reasonably comparable.

The variable $\psi$ can be interpreted as a portfolio common factor, such as an economic index. The term $\psi\sqrt{\rho_k}$ is the obligor's exposure to the common factor and the term $\varepsilon_k\sqrt{1-\rho_k}$ represents the obligor's specific risk. Conditional on a given realization of $\psi$, one finds from (2), (4) and (7), that the default probability of the $k$-th obligor is

$$p_{c,k} = \mathbb{P}(I_k = 1|\psi) = N\left(\frac{N^{-1}(p_k) - \psi\sqrt{\rho_k}}{\sqrt{1-\rho_k}}\right). \tag{9}$$

Observe from (9) that the conditional probability $p_{c,k}$ depends on both the asset correlation and the default probability. In particular, $p_{c,k}$ is a decreasing function of $\psi$ : positive realizations of the macroeconomic factor correspond to a healthy economy whereas negative values imply a distressed economy. Basel II requires banks to hold a regulatory capital in order to ensure with confidence 99.9% that the total actual losses from extraordinary events, in the subsequent year, will not exceed the provisions given by the sum of expected loss among borrowers, $\sum_k L_k$. Thus, in the context of Basel II, $\psi$ is specified by $N^{-1}(0.001)$ so that Equation (9) becomes

$$p_{c,k} = N\left(\frac{N^{-1}(p_k) - N^{-1}(0.001)\sqrt{\rho_k}}{\sqrt{1-\rho_k}}\right). \tag{10}$$

## 2.2 Credit Risk Assignment

Banks can either estimate only the default probabilities using the foundation IRB approach, or estimate internally the default probabilities and the LGDs using the advanced IRB approach (Basel Committee on Banking Supervision, 2006). A qualifying IRB approach requires the design of a risk rating system for credit risk assignment and validation, that fulfills specific minimum requirements. Banks' borrowers should be assigned to a number of grades according to their default probabilities ($p_k$). We seek the optimum number of homogeneous but distinguishable grades with sufficient size to permit the precise ex post validation of the grading system. The optimal bucket setting is the one that minimizes the regulatory capital and satisfies the other constraints imposed by Basel II.

In particular, for a retail portfolio, all borrowers in each grade $g$ are assigned the same default probability $\overline{p}_g$ and loss-given-default $LGD_g$. To optimize the number of homogenous grades and to ensure that no structurally similar buckets exist, alternative objective functions are minimized. Equation (11) transforms mathematically paragraph 401 of Basel II, where the within grades' variance resulting from the deviation of $p_k$ from $\overline{p}_g$ must be minimized:

$$min \sum_g \sum_{k \in g} \left(p_k - \overline{p}_g\right)^2 . \tag{11}$$

Alternatively, a bank can assess the accuracy of the credit risk assignment system by minimizing the sum of the squared distance of individual conditional probabilities $p_{c,k}$ from pooled ones:[6]

$$min \sum_g \sum_{k \in g} \left(p_{c,k} - \overline{p}_{c,g}\right)^2 . \tag{12}$$

Another way to evaluate the efficiency of a grading system is to make sure that portfolio obligors with similar loss structure are placed in the same grade. Thus, we can minimize the total absolute deviation of $L(p_k)$ from $L(\overline{p}_g)$.

$$min \sum_g \sum_{k \in g} \left| L\left(p_k\right) - L\left(\overline{p}_g\right) \right| . \tag{13}$$

[6] We point out here that we can get a trivial solution by equating the number of buckets to the number of obligors in a portfolio. However, this practice is not admissible by Basel II and does not satisfy the constraints it imposes (Section 3.2).

In order to comply with Basel II regulatory capital requirements, a bank should be concerned about keeping the adequate regulatory capital. Risk or regulatory capital $RC$ is the one that ensures solvency for credit institutions in case the cumulative realized losses from extraordinary events $\sum_k RL_k$ exceed the provisions $\sum_k L_k$. That is,

$$RC(p_k) = 1.06 \cdot \left[ RL(p_{c,k}) - L(p_k) \right],$$

where[7]

$$RL_k = RL(p_{c,k}) = EAD_k \cdot LGD_k \cdot p_{c,k}, \ ,$$

and adjusting Equation 12 to comply with capital requirements

$$min \sum_g \sum_{k \in g} \left| RC(p_k) - RC(\overline{p}_g) \right|. \tag{14}$$

To construct optimal classes (in terms of number and width) of homogeneous obligors, (14) implies that banks can grade borrowers to minimize the distance between individual regulatory capital $RC(p_k)$ and regulatory capital based on mean $p_g$, $RC(\overline{p}_g)$. To calculate $RC(\overline{p}_g)$ we substitute $p_k$ with $\overline{p}_g$ in (10). Thus, (9) becomes,

$$p_{c,g} = N \left( \frac{N^{-1}(\overline{p}_g) - N^{-1}(0.001)\sqrt{\rho_k}}{\sqrt{1 - \rho_k}} \right). \tag{15}$$

# 3 An Optimization Heuristic for Credit Risk Assignment

In application to financial problems, heuristics are proven to be a reliable tool not only for binary classification (Varetto (1998) and Shin and Lee (2002)), but also for credit risk rating (Krink *et al.* (2007), Krink *et al.* (2008) and Lyra *et al.* (2010)), portfolio performance improvement (Dueck and Winker (1992) and Zhang and Maringer (2009)).

A diverse range of optimization heuristics have been used for credit risk assignment. Some of them are Genetic Algorithms (GA), Differential Evolution (DE) and Threshold Accepting (TA), introduced by Holland (1975),

[7] The multiplier 1.06 controls possible underestimation of $RC(p_k)$ in the advanced IRB approach (Basel Committee on Banking Supervision (2006), § 14).

Storn and Price (1997) and Dueck and Scheuer (1990), respectively. While DE, has already shown better performance than GA and PSO in tackling the credit risk bucketing problem (Krink *et al.* (2007)), TA outperforms DE in terms of precision and computational time (Lyra *et al.* (2010)).

Our discussion in this paper is focused on TA. TA is part of a broad class of optimization heuristics, called local search heuristics. Their basic property is that they iteratively update a set of initial candidate solutions that improve the objective function.

## 3.1 Threshold Accepting

A key feature of TA is that it enables the search to escape local minima by accepting not only an improvement, but also an impairment of the objective function value, as long as it does not exceed a certain threshold $\tau$. Until the stopping criterion $I$ is met, the current candidate solution $\chi^0$, for a given objective function $f$, is compared with a neighboring solution $\chi^1$, the condition $\Delta = (f(\chi^1) - f(\chi^0)) < \tau$ is checked, where the threshold value $\tau$ determines to what extent not only local improvements, but also local impairments are accepted. A general outline of the TA implementation is presented in Algorithm 1. For a comprehensive overview of TA see Winker (2001).

**Algorithm 1** General description of TA algorithm.

Initialize $I$ and $\tau_{iter}, iter = 1, 2, \ldots, I$
Select at random a set of initial cluster thresholds $\chi^0 \in p_n$
**for** $iter = 1$ to $I$ **do**
    Generate neighbor at random, $\chi^1 \in \mathcal{N}(\chi^0)$ (neighborhood of $\chi^0$)
    **if** $f(\chi^1) - f(\chi^0) < \tau_{iter}$ **then**
        $\chi^0 = \chi^1$
    **end if**
**end for**

Winker and Fang (1997) suggest a data driven approach for the threshold sequence $\tau$ that is constructed ex-ante based on the problem's search space. Local differences of the fitness function are sorted in a descending order representing a diminishing threshold sequence. The reducing neighborhood structure allows a wider search space at the beginning of the search and a greedy search towards the end (see Gilli and Winker (2009) for a detail review).

Alternatively, Lyra *et al.* (2010) propose a threshold sequence based on the differences in the fitness of candidate solutions that are found in the area of the search space currently under consideration. Local differences actually calculated during the optimization run are considered. As a result, the threshold sequence adapts to the region of the search space to which the current solution belongs and to the objective function used. By using a moving average of 100 local differences, a smooth threshold sequence is obtained. In addition, the threshold values are downweighted as a linear function of the current number of iterations ($iter$), namely $1 - iter/I$. In our empirical study we apply this threshold sequence generation approach.

## 3.2 Constraint Handling Techniques for TA

The design of a risk rating system for the assignment of borrowers into homogenous grades is subject to a number of constraints imposed by Basel II. Apart from having at least seven clusters for non-defaulted borrowers (§ 404 of Basel Committee on Banking Supervision (2006)), the EAD in each bucket shall be no higher than 35% of the total borrowers' exposure in a given portfolio (Krink *et al.* (2007)). Thus, we avoid having high concentration of exposure in a given grade. Further, $\overline{p}_g$ for each bucket $g$ should exceed 0.03%, (§ 285 of Basel Committee on Banking Supervision (2006)). An additional constraint regarding a lower bound on the number of borrowers in each bucket is necessary to ensure that no bucket is empty and that the number of borrowers in each bucket is adequate to make statistical inferences.[8] In this contribution, this lower bound is set using statistical benchmarks presented in Section 4.

Constraint handling techniques are applied to handle the constraints imposed by Basel II. Here we consider two techniques: the first technique accepts a new candidate solution considering first the number of constraint violations and then the objective function value. More precisely, in every iteration the algorithm selects between the old and the new solution the one that violates fewer number of constraints and/or the one that satisfies $\Delta < \tau$. Once a feasible solution has been found, this technique limits the search to the feasible region and avoids any infeasible ones.

The next technique is the so-called penalty technique, which is introduced

[8] Previous literature (Krink *et al.* (2007)) specifies that it should exceed a given percentage, e.g. 1%, of the total number of borrowers in a given portfolio.

when TA with the previous constraint handling technique fails to reach the feasible region. This approach allows for penalized infeasible solutions in order for the algorithm to enter the unconstrained search space. For that a penalty term is multiplied by the objective function (so as to 'force' the TA algorithm to enter eventually the feasible region again). Here, the candidate solutions get punished according to the number of constraints they violate ($\beta$). The exact derivation of $\beta$ is given in the Appendix. The magnitude of the penalty term increases as the number of iterations ($iter$) increases to ensure that feasible solutions are obtained at the end ($I$).[9]

$$f_{c,g} = f_{n,g} \cdot (1 + \exp(iter/I))^{\beta}. \tag{16}$$

# 4 Optimal Number of Homogenous Grades

The central aim of an efficient grading system is to obtain an optimal number of rather homogeneous grades. In what follows, we statistically verify that our classification algorithm does exactly that. First, in Section 4.1, we propose a validation approach based on the actual number of defaults. In this context, one is basically interested in knowing whether, at a given confidence level, the observed actual number of defaults matches or is quite close to the one predicted by the grading system. The confidence level must be chosen meaningfully by regulation authorities and / or banks based on their objectives. In the case of ex post validation of the actual number of defaults, not all desired confidence levels are feasible for a given sample size. Moreover, in the interest of supervisory authorities and banks, this verification can further be extended to check whether or not the unexpected loss, and hence the regulatory capital is adequately calculated. In this spirit, following Lyra *et al.* (2010) we also validate our classification system based on unexpected loss. This approach is discussed in Section 4.2.

## 4.1 Actual Number of Defaults Verification

To analyze the ex post validation of the actual number of defaults in a given grade $g$, we compare the actual number of defaults $M_g^a$ with the forecast $M_g^f$

[9]We have chosen an exponential function of *iter* to ensure that the penalty increases very fast with the number of iterations so that the candidate solutions get punished more as the number of iterations increases. In comparison to alterative functional forms, the exponential function resulted more frequently in feasible solutions.

based on the mean default probability $\overline{p}_g$ and the number of borrowers in $g$ grades, $n_g$:

$$M_g^f = n_g \overline{p}_g \,.$$

To judge whether a deviation of $M_g^a$ from $M_g^f$ should be considered as being significant, i.e. challenging the credit classification system, the distribution of $M_g^f$ has to be analyzed under the null hypothesis that $\overline{p}_g$ is an unbiased estimator. Furthermore, to statistically verify the grading system we shall make sure, that the actual number of defaulted obligors in a grade does not exceed the lower bound $M_{g,l}^f$ and the upper bound $M_{g,u}^f$ of the number of forecasted defaulted obligors with more than a percentage $\varepsilon$. Thus, $\varepsilon$ indicates the precision level for which the grading system can be ex post verified. Thus with a given probability $1-\alpha$, $\alpha \in [0,1]$,

$$\mathbb{P}_{int} = \mathbb{P}\left(M_{g,l}^f \leq M_g^a \leq M_{g,u}^f\right) \geq 1-\alpha \,. \tag{17}$$

The corresponding confidence interval for the default rates, i.e. $M_g^a/n_g$ will shrink with a growing number of borrowers $n_g$ in bucket $g$, all other conditions being equal. Thus, any requirement on the size of the confidence interval will impose a lower bound on $n_g$. We consider symmetric confidence intervals around $M_g^f$ of size $2\varepsilon$ as long as the confidence interval falls in the interval $[0,1]$, otherwise, the confidence interval is censored, i.e.,

$$M_{g,l}^f = n_g \cdot \max(\overline{p}_g - \varepsilon, 0), \tag{18}$$

$$M_{g,u}^f = n_g \cdot \min(\overline{p}_g + \varepsilon, 1). \tag{19}$$

The choice of an absolute definition of approximation errors rather than imposing a relative error margin is motivated by its economic impact. In fact, any deviation of the actual ex post default rates from the estimated ones by, e.g., one percentage point will have the same effect on actual defaults independent from the level of the estimated default rate ceteris paribus.

Given that the actual default on a given loan is a binary variable, the number of actual defaults within a bucket can be modeled by the binomial distribution.[10] Hence, for a given $\alpha$, a $(1-\alpha)$ confidence interval for $M_g^a$ is modeled by:

[10] Here, we made the simplifying assumption that that actual defaults are independent, which might be a sensible assumption for retail loans, but might be challenged for other segments of the loan market. In such a case, the statistical model would have to be adjusted.

$$\mathbb{P}_{int} \quad = \quad \sum_{k=M_{g,l}^f}^{M_{g,u}^f} \binom{n_g}{k} \cdot \overline{p}_g^k \cdot \left(1 - \overline{p}_g\right)^{n_g - k} \geq 1 - \alpha. \tag{20}$$

Given that we impose a minimum constraint on the precision, we do not have to solve Equation (20) for the number of elements in the grade $n_g$. Instead, we need to confirm that $\mathbb{P}_{int} \geq 1 - \alpha$ for all grades. Thus, the grading mechanism is efficient if we can verify with accuracy $\alpha$ and $\varepsilon$ that (17) is satisfied.

However, not all combinations of $\alpha$ and $\varepsilon$ will be feasible for a given total number of loans to be considered, taking into account the other constraints imposed by the Basel II framework. In fact, a rough calculation shows that for our data with a mean $p$ of around 6% in the last bucket, values of $\alpha = 5\%$ and $\varepsilon = 1\%$ would require around 3 000 observations in that bucket. With a small sample size this is not feasible given the constraints imposed, in particular the constraint that no more than 35% of total exposure at default should belong to one bucket. Yet, our sample of roughly 95 000 observations is sufficient for reporting feasible results even for $\alpha = 1.5\%$ and $\varepsilon = 1\%$.

## 4.2 Unexpected Losses Verification

As an alternative to our new approach, we also validate our model based on the correct statement of unexpected loss $UL$, and hence, regulatory capital $RC$, since $RC = 1.06 \cdot UL$.[11] We use the unexpected loss constraint because Basel II emphasizes it. Moreover, we would like to compare this validation approach with the validation of the actual number of defaults.

For validation based on unexpected losses, however, the lower bound $M_{g,l}^f$ and the upper bound $M_{g,u}^f$ of the number of forecasted defaulted obligors from (17) are transformed in such a way that we obtain

$$n_g \cdot \overline{p}_g \cdot \left(1 - \varepsilon \cdot \frac{\overline{UL}}{\overline{UL_g}}\right) \leq M_g^a \leq n_g \cdot \overline{p}_g \cdot \left(1 + \varepsilon \cdot \frac{\overline{UL}}{\overline{UL_g}}\right), \tag{21}$$

[11] UL equals $[RL(p_{c,k}) - L(p_k)]$

where

$$UL = \sum_k (EAD_k \cdot LGD_k \cdot (p_{c,k} - p_k)), \tag{22}$$

$$UL_g = \sum_{k \in g} (EAD_k \cdot LGD_k \cdot (p_{c,g} - p_g)) \tag{23}$$

for some predefined $\varepsilon$ (see Appendix). $\overline{UL}$ is the mean unexpected loss of the entire portfolio whereas, $\overline{UL_g}$ is the mean value of the grade unexpected loss $UL_g$. Since defaults are binomially distributed, the above idea can be expressed in terms of (20) but with $M^f_{g,l}$ and $M^f_{g,u}$ replaced by

$$M^f_{g,l} = n_g \cdot \overline{p}_g \cdot \left(1 - \varepsilon \cdot \frac{\overline{UL}}{\overline{UL_g}}\right) \tag{24}$$

$$M^f_{g,u} = n_g \cdot \overline{p}_g \cdot \left(1 + \varepsilon \cdot \frac{\overline{UL}}{\overline{UL_g}}\right), \tag{25}$$

respectively. Equations (17) and (21) simply guarantee a sufficient number of borrowers in each bucket $g$ so that, with a certain confidence level $1-\alpha$, one can ex ante state that the actual number of defaults lies within the interval $[M^f_{g,l}; M^f_{g,u}]$. Observe that this interval increases with the number of obligors and the pooled probability of the bucket, as well as $\varepsilon$.[12]

Since a high number of buckets reduces the precision error resulting from substituting the individual default probabilities by the 'pooled' ones, we choose an optimum number of buckets that is consistent with some predefined values for $\alpha$ and $\varepsilon$. In practice both $\alpha$ and $\varepsilon$ are chosen meaningfully by regulation authorities and / or banks based on their objectives. As noted before in the case of ex post validation of the actual number of defaults, not all combinations of $\alpha$ and $\varepsilon$ are feasible for a given sample size.

# 5 Empirical Analysis

## 5.1 Data

We use a real-world test portfolio consisting of 93 580 retail borrowers. The LGDs range between 0.17 and 1. The default probabilities vary between

[12] The above specification replaces the constraint to have at least 1% of all borrowers in each bucket.

0.000001% and 30%. Figure 1 shows the distribution of the default probabilities. The distribution is highly skewed to the right (as expected), with mean and variance of 2.1% and 0.14%, respectively. The expected loss of the portfolio amounts to 0.877% of the total exposure.

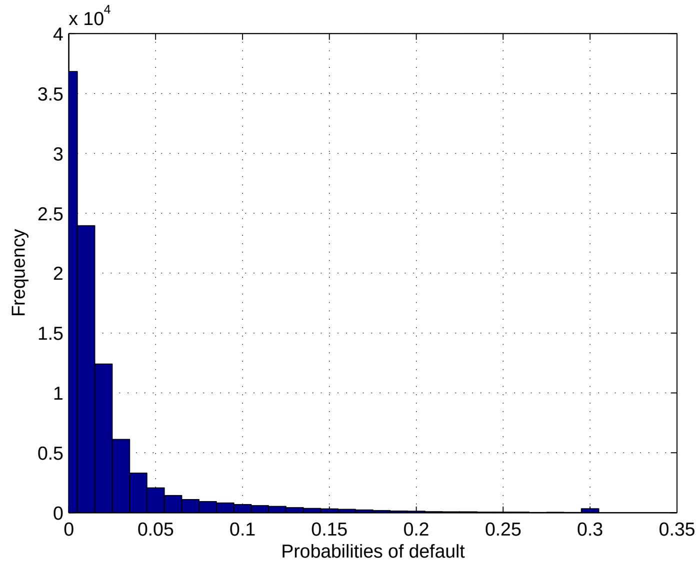


Figure 1: The distribution of the PDs

## 5.2 Discussion of Empirical Results

In this section we report the ex post validity of the credit risk grading system. Figure 2 shows, for over 10 restarts of the algorithm, the evolution of the mean values of the objective function minimizing the within grades variance (12) as the number of grades $g$ changes, whereas Figure 3 depicts the same for the objective function value minimizing RC (14). In both cases, the left graph represents the mean value using the validation approach discussed in Section 4.1 (see equations (18) and (19)) while the right one is for the approach in Section 4.2 (see equations (24) and (25)). Detailed results are reported in the Appendix, Tables 1 and 2. In these experiments, the grading system is considered valid if it ensures, at a given confidence level[13], that the actual

[13]In the Appendix a grid search for determining the optimal precision level is provided.

number of defaults neither falls below the lower bound nor exceeds the upper bound specified in Equations (18) and (19), (24) and (25).

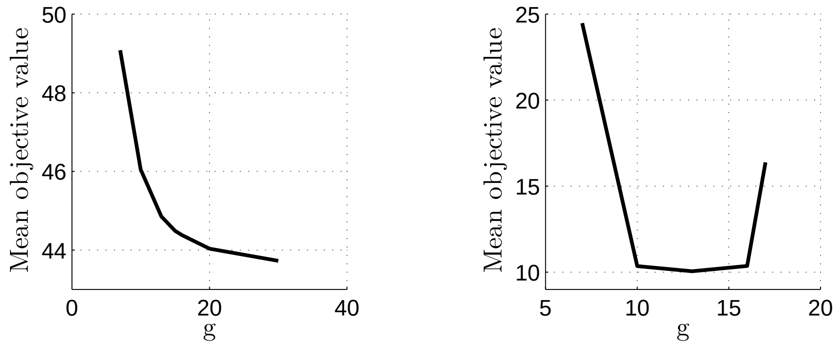


Figure 2: Mean value for objective (12) as a function of grade number (g)

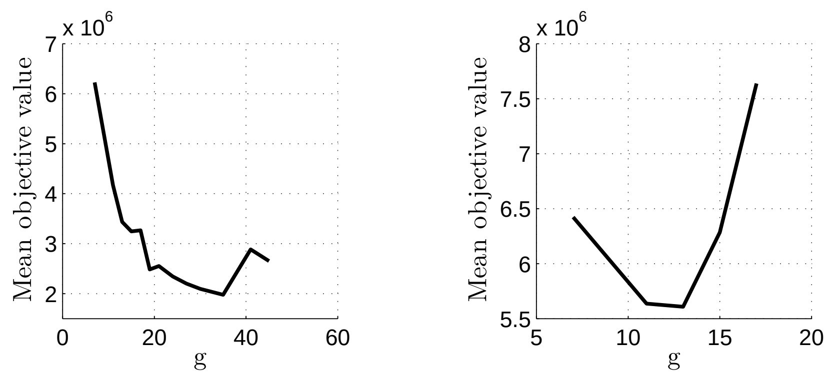


Figure 3: Mean value for objective (14) as a function of grade number (g)

For the objective function minimizing the within grades variance (12), a general finding is that our proposed validation approach using the number of defaults (Section 4.1) results, for most bucket settings, in better mean objective function values with lower variation for 10 restarts than the unexpected loss approach (Section 4.2). For this specific objective function and the proposed technique, the optimum bucket setting seemed to be at g = 30 for $\alpha = 1.5\%$ and $\varepsilon = 1\%$.[14] For the unexpected loss validation technique

[14]When $\alpha = 2.5\%$ and $\varepsilon = 3\%$ the grading system resulted in an optimum bucket setting

(Section 4.2), the optimal bucket number is found at g = 17.[15]

In Table 2, for the objective function value minimizing RC (14), the new simple validation technique based on actual number of defaults (Section 4.1) results in an optimum bucket number of g = 35. For the unexpected loss validation (Section 4.2) approach the optimum number of grades is 17. After this number of buckets no feasible solutions could be found.

The two validation techniques result in different optimum number of homogeneous buckets. This is not surprising since the unexpected loss validation imposes stricter constraints and hence only a smaller number of larger buckets satisfies them. By applying a simple, computationally less demanding, validation technique based on the number of borrowers in the buckets, we can validate ex post the precision of a grading system. However, the precision level is a function of the portfolio size. Hence, the bucket size is a determinant factor for the precise validation of a credit risk grading technique. According to the objectives of each credit institution either validation approaches can be applied to determine the optimal number of buckets. This contradicts the usual practices of many credit institutions, where the number of buckets and hence their size is fixed. This number varies from bank to bank; for instance, Deutsche Bank uses 26 buckets. Although by fixing the number and size of buckets makes the clustering simpler and consistent through time, it endangers possible under or over estimation of RC.

# 6 Conclusion

We have studied the use of Threshold Accepting (TA) for an optimal clustering of banks' borrowers based on their default probabilities. We have also proposed a new computationally tractable technique based on the actual number of defaults in each homogeneous bucket to validate ex post the precision of the grading system. In particular, this is achieved by making the simplifying assumption that the defaults are independent and identically distributed. This assumption makes sense in the case of retail portfolios and yields reasonable results. The validation approach based on the unexpected

above 55. This is a result of less strict constraints. Since, such relaxed bounds do not result in a reasonable bucket setting, we do not further discuss these findings.

[15]In all results no feasible solutions could be found after the reported bucket number even after applying the penalty technique (Section 3.2) and after increasing the iteration number to 1 000 000 and the restarts to 20.

loss constraint is computationally more intensive and time-consuming than the approach we have proposed.

Depending on the objectives of the credit institution, either of the two validation approaches are applied. Furthermore, the size of the portfolio, the parameters (i.e. $\alpha$ and $\varepsilon$), and hence the bucket size, are determining factors for the precise validation of a credit risk grading technique. In particular, with a relatively large data set of a real-world credit portfolio, we find that the validation approach based on the actual number of defaults generally yields better mean values of the objective functions considered in this paper than its unexpected loss counterpart. However, for small portfolios the ex post validation based on the actual number of defaults approach is not feasible at a reasonable level of precision.

The optimal number of grades, suggested by the TA classification algorithm, is only applicable for one period window (here for one year). It also depends on the objective functions and the associated constraints faced by the bank. In a dynamic framework, TA might result in a different number of grades for optimal classification. However, different number of grades over two periods can make the calculation of credit migration matrices very difficult, if not impossible. As it is often desirable by credit institutions to calculate how the obligors 'migrate' over time with respect to their PDs, it is an open challenge to apply TA for several time periods for the calculation of credit migration matrices.

Another challenge is the relaxation of the independence assumption among actual defaults. It is indeed a rather simplistic assumption for modeling defaults in general. Thus, a more complex distributional assumption (such as heavy tailed multivariate t-distributions) may be required to model the dependence structure of defaults, especially for other loan types such as corporate. Finally, one can further investigate the possible relationship between sample size and confidence and precision levels.

# Acknowledgements

Financial support from the EU Commission through MRTN-CT-2006-034270 COMISEF is gratefully acknowledged.

# Appendix

## Asset and Default Correlation

Since correlation is our measure of dependence, we proceed to discuss default and asset correlations. The dependence structure of the default indicators as defined in (2) and (3) is specified through the factor model (5).

The default or event correlation $\rho_{jk}^D$ of obligors $j$ and $k$ is defined as the correlation of the respective default indicators. Since

$$\mathrm{Var}(I_k) = \mathrm{E}(I_k^2) - (\mathrm{E}(I_k))^2 = \mathrm{E}(I_k) - (\mathrm{E}(I_k))^2 = p_k - p_k^2,$$

we have that the default correlation equals

$$\rho_{jk}^D = \mathrm{Corr}(I_j, I_k) = \frac{\mathrm{E}(I_j I_k) - p_j p_k}{\sqrt{(p_j - p_k^2)(p_j - p_k^2)}}. \tag{26}$$

There exists an obvious link between default correlation $\rho_{jk}^D$ and asset correlation $\rho_{jk}$. For given default probabilities, the default correlations $\rho_{jk}^D$ are determined by joint default probability $\mathrm{E}(I_j I_k)$ according to (26). It follows from (2) and (3) that

$$\mathrm{E}(I_j I_k) = \mathbb{P}(A_j \leq c_j, A_k \leq c_k) = \int_{-\infty}^{c_j} \int_{-\infty}^{c_k} f_{jk}(u, v) du dv,$$

where $f_{jk}(u, v)$ is the joint density function of $A_j$ and $A_k$. Hence, default correlations depend on the joint distribution of $A_j$ and $A_k$. If $(A_j, A_k)$ is bivariate normal the correlation of $A_j$ and $A_k$ determines the copula of their joint distribution and hence the default correlation:

$$\mathrm{E}(I_j I_k) = \frac{1}{2\pi\sqrt{1 - {\rho_{jk}}^2}} \int_{-\infty}^{c_j} \int_{-\infty}^{c_k} \exp\left(-\frac{u^2 - 2\rho_{jk} uv + v^2}{2(1 - {\rho_{jk}}^2)}\right) du dv. \tag{27}$$

Using historical ratings data one can compute default correlation $\rho_{jk}^D$ and asset correlation $\rho_{jk}$ from (26) and (27), see, e.g., in Roesch and Scheule (2010), Chapter 10, the contribution on validating structural credit risk portfolio models. Note that for general ability-to-pay variables outside the multivariate normal class, the asset correlations do not fully determine the default correlations, McNeil *et al.* (2005).

## Penalty Function

To calculate the exponent $\beta$ in (16) the candidate solutions get punished according to the number of constraints they violate.

$$\begin{aligned}\beta &= \left(0.5 \cdot \sum_g D_{EAD,g} \cdot \frac{\sum_{k\in g} EAD_{k,g} - 35\% \cdot \sum_g \sum_{k\in g} EAD_{k,g}}{65\% \sum_g \sum_{k\in g} EAD_{k,g}}\right) \\ &+ \left(0.5 \cdot \frac{\sum_g D_{n,g} \cdot \frac{1-\alpha-\mathbb{P}_{int}}{1-\alpha}}{\sum_g D_{n,g}}\right),\end{aligned} \tag{28}$$

and $f_{c,g}$ is the current value of the objective function in a specific grade. The probability $\mathbb{P}_{int}$ is as defined in Equation (20). If no constraints are violated, then $\beta = 0$; if the number and magnitude of constraints violated is maximum, then $\beta = 1$. The first component of the exponent $\beta$ refers to the EAD constraint, whereas the second component constrains the actual default number so as to maintain the statistical validity of the classification system. It checks whether the actual default number in each grade stays inside a predefine lower and upper bound. A detailed explanation and derivation of $\mathbb{P}_{int}$ are given in Section 4.1. Both components of the exponent $\beta$ are equally penalized. More precisely, the indicator variables $D_{EAD,g}$ and $D_{n,g}$ are given, respectively, by

$$D_{EAD,g} = \begin{cases} 1, & \text{if } \sum_{k\in g} EAD_{k,g} > 35\% \cdot \sum_g \sum_{k\in g} EAD_{k,g} \\ 0, & \text{otherwise} \end{cases}$$

and

$$D_{n,g} = \begin{cases} 1, & \text{if } \mathbb{P}_{int} < 1-\alpha \\ 0, & \text{otherwise.} \end{cases}$$

## Grid Search

We extend the previous literature by investigating what levels of precision can be achieved with a relatively large dataset. In an attempt to determine the optimal precision level we run the algorithm for 7, 11 and 13 bucket settings, changing every time the precision parameters $\alpha$ and $\varepsilon$. The results depict the feasible region as we alter the two parameters.

Figures 4 and 5 illustrate the boundaries of the feasible search space for objectives (12) and (14) when considering the actual number of defaults constraint (left hand side) and the unexpected loss constraint (right hand side), explained in Sections 4.1 and 4.2, respectively. With about 95 000 observations the proposed classification system could be validated with the following precision levels: For objective (12) and the actual number of defaults constraint (equations (18) and (19)) with precision close to $\alpha = 2.5\%$ and $\varepsilon = 3\%$. In this application we test the possible ex post validation of the grading system using also $\alpha = 1.5\%$ and $\varepsilon = 1\%$.[16]

The optimal $\alpha$ and $\varepsilon$ for the unexpected loss constraint constraint (equations (24) and (25)) are $\alpha = 5\%$ and $\varepsilon = 20\%$. These are the minimum accuracy levels for which a feasible solution can be found. For objective (14) the minimum precision levels are $\alpha = 1.5\%$ and $\varepsilon = 1\%$ for the actual number of defaults constraint and $\alpha = 5\%$ and $\varepsilon = 22\%$ for unexpected loss constraint, respectively. It should be specified that these values are optimal only for the specific dataset length.

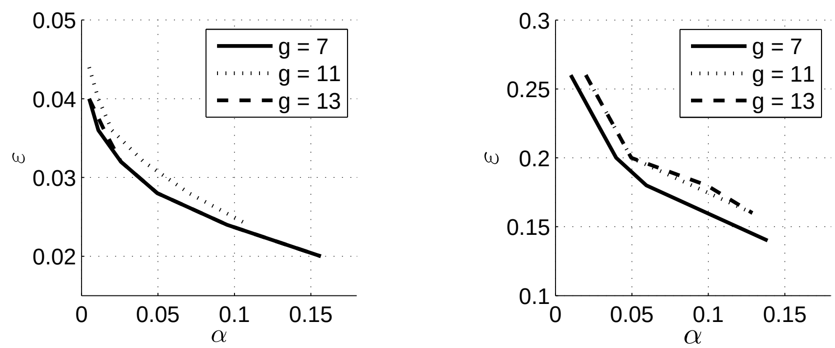


Figure 4: Feasible region for different confidence and precision levels for objective 12

## Results for Optimal Number of Grades

Table 1 presents results for the objective function minimizing within grades variance (12) and Table 2 for the objective function minimizing RC (14),

[16]The precision level of $\alpha = 2.5\%$ and $\varepsilon = 3\%$ imposes a less strict constraint. Such relax bounds are may not be desirable by a credit institution.

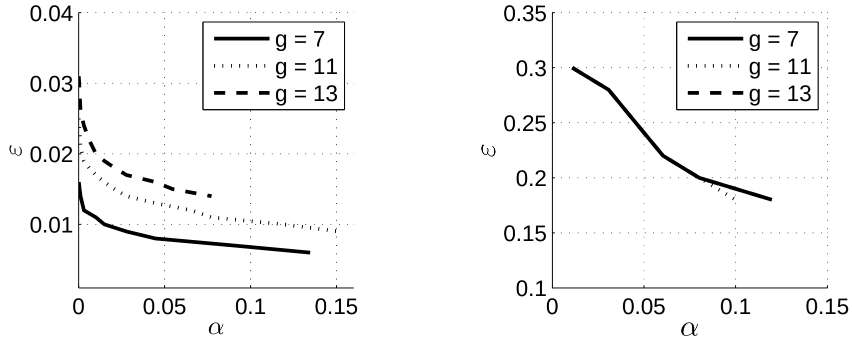


Figure 5: Feasible region for different confidence and precision levels for objective 14

both for the two validation techniques. TA is run for 1 000 000 iterations for both objective functions. (12), whereas for objective function (14) only 500 000 iterations are run. Since objective function (14) is computationally more demanding, we decided to limit its iterations to keep the computational time at the same level. TA was restarted 10 times for each objective function and some descriptive results are presented in Tables 1 and 2. These include the best objective value obtained over 10 restarts of the algorithm, the mean, the worst value, the standard deviation, the 80th and 90th percentiles and the frequency the best value appears over the 10 restarts.

Table 1: Objective function for minimizing within grades variance (12) with $it = 1\,000\,000$

| | Best | Mean | Worst | s.d. | q80% | q90% | Freq |
|---|---|---|---|---|---|---|---|
| | | | | $g = 7$ | | | |
| TA[a] | 18.6836 | 18.6836 | 18.6836 | $3.6731 \cdot 10^{-8}$ | 18.6836 | 18.6836 | 8/10 |
| TA[b] | 18.6552 | 24.4809 | 46.2984 | 8.2478 | 24.8204 | 24.8221 | 1/10 |
| TA[c] | 49.0847 | 49.0849 | 49.0852 | 0.0002 | 49.0852 | 49.0852 | 1/10 |
| | | | | $g = 10$ | | | |
| TA[a] | 9.7293 | 9.7293 | 9.7293 | $5.3490 \cdot 10^{-7}$ | 9.7293 | 9.7293 | 1/10 |
| TA[b] | 9.1118 | 10.3545 | 10.9233 | 0.8520 | 10.8863 | 10.9108 | 1/10 |
| TA[c] | 45.9783 | 46.0478 | 46.1033 | 0.0408 | 46.0769 | 46.0862 | 1/8 |
| | | | | $g = 13$ | | | |
| TA[a] | 6.6716 | 6.6716 | 6.6716 | $2.9353 \cdot 10^{-6}$ | 6.6716 | 6.6716 | 1/10 |
| TA[b] | 6.5974 | 10.0515 | 14.5469 | 2.7151 | 10.4310 | 12.4890 | 1/6 |
| TA[c] | 44.7501 | 44.8534 | 44.9603 | 0.0749 | 44.9027 | 44.9315 | 1/4 |
| | | | | $g = 16$ | | | |
| TA[a] | 5.2454 | 5.2454 | 5.2454 | $1.9032 \cdot 10^{-6}$ | 5.2454 | 5.2454 | 1/10 |
| TA[b] | 10.3647 | 10.3647 | 10.3647 | 0.0000 | 10.3647 | 10.3647 | 1/1 |
| TA[c] | 44.3419 | 44.3771 | 44.4123 | 0.0352 | 44.3982 | 44.4052 | 1/2 |
| | | | | $g = 17$ | | | |
| TA[a] | 4.9268 | 4.9268 | 4.9268 | $2.3253 \cdot 10^{-6}$ | 4.9268 | 4.9268 | 1/10 |
| TA[b] | 12.6106 | 16.3882 | 20.1657 | 3.7775 | 18.6547 | 19.4102 | 1/2 |
| TA[c] | 44.2395 | 44.2461 | 44.2527 | 0.0066 | 44.2500 | 44.2513 | 1/2 |
| | | | | $g = 20$ | | | |
| TA[a] | 4.2859 | 4.2871 | 4.2921 | 0.0024 | 4.2859 | 4.2912 | 1/10 |
| TA[c] | 44.0334 | 44.0334 | 44.0334 | 0.00 | 44.0334 | 44.0334 | 1/1 |
| | | | | $g = 30$ | | | |
| TA[a] | 3.4613 | 3.4617 | 3.4623 | $3.3069 \cdot 10^{-4}$ | 3.4613 | 3.4614 | 1/10 |
| TA[c] | 43.7280 | 43.7280 | 43.7280 | 0.00 | 43.7280 | 43.7280 | 1/1 |
| | | | | $g = 35$ | | | |
| TA[a] | 3.3043 | 3.3111 | 3.3170 | 0.0038 | 3.3128 | 3.3149 | 1/6 |
| | | | | $g = 40$ | | | |
| TA[a] | 3.2202 | 3.2210 | 3.2239 | 0.0016 | 3.2218 | 3.2228 | 1/5 |
| | | | | $g = 45$ | | | |
| TA[a] | 3.1456 | 3.1501 | 3.1523 | 0.0024 | 3.1521 | 3.1522 | 1/5 |
| | | | | $g = 50$ | | | |
| TA[a] | 3.0979 | 3.1095 | 3.1196 | 0.0070 | 3.1150 | 3.1177 | 1/8 |
| | | | | $g = 55$ | | | |
| TA[a] | 3.0707 | 3.0824 | 3.0771 | 0.0044 | 3.0808 | 3.0816 | 1/6 |

[a] Actual number of defaults constraint, $\alpha = 2.5\%$ and $\varepsilon = 3\%$

[b] Unexpected loss constraint, $\alpha = 5\%$ and $\varepsilon = 20\%$. After g = 17 no feasible solution could be found.

[c] Actual number of defaults constraint, $\alpha = 1.5\%$ and $\varepsilon = 1\%$. After g = 30 no feasible solution could be found.

Table 2: Objective function for minimizing RC (14) with $it = 500\,000$

| | **Best** | **Mean** | **Worst** | **s.d.** | **q80%** | **q90%** | **Freq** |
|---|---|---|---|---|---|---|---|
| | | | $g = 7$ | | | | |
| TA[a] | 6,228,874 | 6,228,874 | 6,228,874 | $9.82 \cdot 10^{-10}$ | 6,228,874 | 6,228,874 | 10/10 |
| TA[b] | 6,419,727 | 6,423,788 | 6,426,403 | 2,052.62 | 6,419,727 | 6,420,826 | 1/10 |
| | | | $g = 11$ | | | | |
| TA[a] | 4,165,257 | 4,167,952 | 4,182,902 | 5,998.45 | 4,165,257 | 4,165,257 | 7/10 |
| TA[b] | 5,534,072 | 5,636,388 | 5,814,094 | 101,283.20 | 5,534,072 | 5,538,839 | 1/10 |
| | | | $g = 13$ | | | | |
| TA[a] | 3,425,092 | 3,435,627 | 3,436,798 | 3,701.71 | 3,425,092 | 3,436,798 | 1/10 |
| TA[b] | 5,192,944 | 5,608,280 | 5,929,156 | 230,629.77 | 5,809,236 | 5,846,709 | 1/9 |
| | | | $g = 15$ | | | | |
| TA[a] | 3,245,441 | 3,245,636 | 3,247,260 | 571.05 | 3,245,441 | 3,245,445 | 1/10 |
| TA[b] | 5,627,306 | 6,285,472 | 7,166,148 | 647,632 | 6,724,873 | 6,945,510 | 1/3 |
| | | | $g = 17$ | | | | |
| TA[a] | 3,183,949 | 3,268,412 | 3,471,793 | 76,331.08 | 3,292,805 | 3,316,323 | 1/10 |
| TA[b] | 5,564,546 | 7,639,295 | 9,055,551 | 1,499,331 | 8,752,446 | 8,903,999 | 1/3 |
| | | | $g = 19$ | | | | |
| TA[a] | 2,331,271 | 2,484,098 | 2,713,301 | 197,258.90 | 2,331,271 | 2,331,273 | 1/10 |
| | | | $g = 21$ | | | | |
| TA[a] | 2,541,485 | 2,551,267 | 2,561,056 | 9,764.05 | 2,561,027 | 2,561,042 | 1/4 |
| | | | $g = 24$ | | | | |
| TA[a] | 2,343,282 | 2,343,283 | 2,343,283 | 0.66 | 2,343,283 | 2,343,283 | 1/2 |
| | | | $g = 27$ | | | | |
| TA[a] | 2,201,570 | 2,201,575 | 2,201,579 | 4.66 | 2,201,577 | 2,201,578 | 1/2 |
| | | | $g = 30$ | | | | |
| TA[a] | 2,094,821 | 2,095,378 | 2,095,934 | 556.07 | 2,095,711 | 2,095,822 | 1/2 |
| | | | $g = 35$ | | | | |
| TA[a] | 1,977,254 | 1,977,425 | 1,977,596 | 171.08 | 1,977,528 | 1,977,562 | 1/2 |
| | | | $g = 41$ | | | | |
| TA[a] | 2,653,114 | 2,885,592 | 3,118,070 | 232,478.22 | 3,025,079 | 3,071,575 | 1/2 |
| | | | $g = 45$ | | | | |
| TA[a] | 2,651,825 | 2,651,825 | 2,651,825 | 0.00 | 2,651,825 | 2,651,825 | 1/1 |

[a]Actual number of defaults constraint, $\alpha = 1.5\%$ and $\varepsilon = 1\%$

[b]Unexpected loss constraint, $\alpha = 5\%$ and $\varepsilon = 22\%$. After g = 17 no feasible solution could be found.

# References

Basel Committee on Banking Supervision (2005). An explanatory note on the Basel II IRB Risk Weight Functions. Technical report. Bank for International Settlements.

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards a revised framework comprehensive version. Technical report. Bank for International Settlements.

Bluhm, C., L. Overbeck and C. Wagner (2002). *An Introduction to Credit Risk Modeling*. CRC Press/Chapman & Hall.

Dietsch, M. and J. Petey (2004). Should SME exposures be treated as retail or corporate exposures? a comparative analysis of default probabilities and asset correlations in French and German SMEs. *Journal of Banking and Finance* **28**, 773–788.

Dueck, G. and P. Winker (1992). New concepts and algorithms for portfolio choice. *Applied Stochastic Models and Data Analysis* **8**, 159–178.

Dueck, G. and T. Scheuer (1990). Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *J. Comp. Phys* **90**, 161–175.

Gilli, M. and P. Winker (2009). Heuristic optimization methods in econometrics. In: *Handbook on Computational Econometrics* (E. Kontoghiorghes et al., Ed.). pp. 81–119. Elsevier. Amsterdam.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Harbor.

Krink, T., S. Paterlini and A. Resti (2008). The optimal structure of PD buckets. *Journal of Banking and Finance* **32**(10), 2275–2286.

Krink, T., S. Paterlini and A. Restic (2007). Using differential evolution to improve the accuracy of bank rating systems. *Computational Statistics & Data Analysis* **52**(1), 68–87.

Lee, J., J. Wang and J. Zhang (2009). The relationship between average asset correlation and default probability. *KMV corporation. Working Paper*.

Lopez, J. A. (2002). Empirical relationship between average asset correlation, firm probability of default and asset size. *Federal Reserve Bank of San Francisco. Working Paper*.

Lyra, M., J. Paha, S. Paterlini and P. Winker (2010). Optimization heuristics for determining internal rating grading scales. *Computational Statistics & Data Analysis. forthcoming*.

McNeil, A.J., R. Frey and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* **29**(2), 449–470.

Roesch, D. and H. Scheule (2010). *Model Risk. Identification, Measurement and Management*. Intensive Media, UK.

Shin, K. S. and Y. J. Lee (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications* **23**(3), 321–328.

Storn, R. and K. Price (1997). Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *J. Global Optimization* **11**, 341–359.

Varetto, F. (1998). Genetic algorithms applications in the analysis of insolvency risk. *Journal of Banking and Finance* **22**, 1421–1439.

Winker, P. (2001). *Optimization Heuristics in Econometrics: Applications of Threshold Accepting*. Wiley, New York.

Winker, P. and K.-T. Fang (1997). Application of threshold accepting to the evaluation of the discrepancy of a set of points. *SIAM Journal on Numerical Analysis* **34**(5), 2028–2042.

Zhang, J. and D. Maringer (2009). Improving sharpe ratios and stability of portfolios by using a clustering technique. In: *Proceedings of the 2009 International Conference of Computational Intelligence and Intelligent Systems*. pp. 1–6. Lecture Notes in Engineering and Computer Science. ISBN 978-988-17012-5-1. WCE.