

# **A model for selecting and tranching pools of small and medium size enterprise loans**

By

Pascal Terborg

May 2010

Supervision:



Dennis Roubos  
Rene Bekker



Jorn de Vries  
Paul Wessels

*Business Mathematics and Informatics*

## Preface

In order to complete the masters degree in Business Mathematics and Informatics at the VU University Amsterdam, an internship had to be completed. This internship took 6 months, during which a challenge at a company had to be studied, after which a solution is provided. I got the opportunity to do this internship at the ABN AMRO bank.

In this section I would like to acknowledge some people that played an essential role during this internship.

First of all I would like to thank Jorn de Vries, who was the team-leader of Functional Management Financial Accounting (the department I did the internship at), for creating the opportunity for me to do this internship. Next, I would like to thank Paul Wessels, who worked at CMG for taking time out of his busy schedule, to brainstorm on ideas, and to provide guidance when it came to the quantitative fundamentals of portfolio risk assessment.

I would also like to thank Dennis Roubos from the university for his help, guidance and patience throughout this project. Many thanks to Rene Bekker from the university, for reviewing, and providing feedback and comments on the documents I created. I also received lots of help from Stanislav Zakovic from the Risk department of ABN AMRO, in getting past some difficult practical modeling issues during this internship.

Last but not least, I would like to thank my family and girlfriend for the support and encouragement during these last months of this study.

This internship was a great learning experience for me. It was great to see the statistical, and other mathematical theory I learned during the study actually being put into practice.

## Summary

This thesis is setup in to two phases, were in the first phase we build a model to improve the current selection process of small and medium enterprise (SME) loans that are to be included in a portfolio to be securitized. This new selection method reduces the duration and error proneness of the selection process. To accomplish this, the extract, load and transform (ETL) method is used. In this method we first extract the data from the individual source systems. Next the data is cleansed, transformed and combined, after which it is loaded into the data mart area. At this point we have a complete set of the loan portfolio.

After the ETL process we impose selection criteria on the data, where each loan get a 1 or 0 indicator for inclusion or exclusion respectively. We also add variables, with which we indicate which of the selection criteria causes the loan to be excluded from the portfolio. These variables can be used to analyze the data using techniques such as what if scenarios. With this result we conclude the first phase.

After concluding the first phase, we move on to discuss the mathematical background and notation that is needed in the second phase.

In the second phase we build a model to create tranches (slices) of a portfolio, where each tranche has a given credit rating<sup>1</sup>. We accomplish this by using Monte Carlo techniques, to create a portfolio loss distribution, which can be used to determine portfolio risk measures necessary for determining the tranches. The underlying model for this simulation is based upon the single factor model in which the the single factor is the industry sector the obligor participates in. These industry sectors are correlated according to a given correlation matrix which will be inputed into this model.

The portfolio selected in the the first phase is then inputed into the simulation model build in the second phase. The output gives a loss distribution, which is then tranced according to the default rates of the given rating levels.

---

<sup>1</sup> for example, Standard & Poor's uses the AAA, AA, BBB letters.

## Chapter index

Preface.....	2
Summary.....	3
1 Introduction.....	5
1.1 ABN AMRO .....	5
1.2 Overview of securitisation.....	6
1.3 Problem statement.....	7
1.4 Phases of the project.....	8
1.5 Technical aspects.....	9
2 Pool Selection environment.....	10
2.1 Extract Transform and Load .....	10
2.2 Check.....	11
2.3 Transformations.....	11
2.4 Combining.....	12
2.5 Loading.....	12
2.6 Selection.....	13
2.7 Requirements.....	14
2.8 Asset selection.....	14
3 Mathematical background.....	15
3.1 Notation.....	15
3.2 Credit Loss distribution.....	15
3.3 Default Correlation.....	16
3.4 Simplified firm's value model.....	17
3.5 The single factor model.....	18
3.6 Cholesky factorization.....	20
4 Simulation.....	22
4.1 Portfolio credit risk model.....	22
4.2 Monte Carlo.....	22
4.3 Setting up the simulation.....	23
4.4 Modeling the correlation.....	23
4.5 Simulation steps.....	24
4.6 The limiting distribution of losses.....	24
4.7 Calibrating the parameters .....	26
4.8 Testing the model.....	26
4.9 The effect of the correlations.....	28
4.10 Tranching.....	28
5 Conclusions.....	32

# 1 Introduction

First we start with a general introduction of the context of this study. Here an introduction of the bank, departments (the ones relevant for this thesis), and a high-level overview of securitisation is given. In chapter 2 the details of the datamart and the selection process are covered. The mathematical background, and models that are used in this paper are discussed in chapter 3. In chapter 4, we move to the simulation, testing of the model, and the results. We conclude this thesis with the conclusions in chapter 5.

## 1.1 ABN AMRO

ABN AMRO Bank N.V. is a Dutch bank, currently owned by the government of the Netherlands. With 20.000 employees, ABN AMRO has over 3 million private clients, and also serves a substantial amount of small and medium enterprises and large corporates.

### Data Management

Data Management is part of the ABN AMRO/Finance/Accounting & Reporting department. Data Management is responsible for the financial and management accounting systems and for the processing and reporting of financial measures. Data Management enables the management of ABN AMRO to manage the company by delivering accurate and complete financial information.

Data Management is divided in 4 sub-departments; Functional Management, Data providing, Master data Management and the Expertise Center. This internship took place within Functional Management. Functional management is responsible for creating new data processes in order to satisfy the needs of the different departments within the finance department.

### Capital Management Group

One of the businesses within Finance, to which Data Management delivers information, is Capital Management Group (CMG). One of the objectives of CMG is to manage and optimize the balance sheet of the bank in terms of working capital, credit risk and liquidity on a portfolio level. CMG also aims to maintain a capital structure that is consistent with the bank's rating targets. One of the techniques that CMG uses to accomplish its objectives, is by doing structured financial transactions, such as securitisations.

## **1.2 Overview of securitisation**

Securitisation is a popular financial technique, that was developed to enable an entity (originator) to trade a collection of assets, that are not trade-able by nature, for cash. In the process of securitisation the assets are transformed into bonds. The central element of the securitisation is that the repayment of the obligation depends on the cashflows of the underlying assets, and not on the owner of the asset. Thus, the rating of the bonds is independent, and not necessarily equal to the rating of the originator.

In the process of securitisation, a portfolio of assets is refinanced on the capital market by a Special Purpose Vehicle (SPV) which sells bonds to investors, in order to pay for the assets. The assets are removed from the balance sheet of the originator, and sold to the SPV, who then pays the cash received from the investors. The main purpose of the SPV is to guide the transaction and to detach the portfolio of assets from the original owner. The concept dis-attachment is called bankruptcy remoteness and is one of the elements that makes securitisations more attractive than alternative types of financing (Kenneth Ayotte and Stav Gaon 2005). In case the originator goes bankrupt, the creditors do not have any claim on the assets in the portfolio. Therefore the bondholders will still receive the repayments according to the original contract, even if the originator goes bankrupt.

The bonds are created by cutting up the portfolio in different slices, which are called tranches. Each tranche carries a certain level of risk of the portfolio and is sold off separately. The cashflows and losses of the assets, are assigned to the different tranches according to seniority. Creditors with a higher seniority will receive less interest than those with lower seniority but will receive priority over the repayments in the case of default of the underlying assets. This method of repayment is also called the waterfall or cascade cashflow method, because the repayments starts at the higher levels and then falls down to the lower levels.

There are many different forms of securitisations, of which three large categories are the ABS (asset backed securities), MBS (mortgage backed securities) and CDO (collateralized debt obligation). During this project the focus was on the ABS transaction, where the underlying assets are SME (small medium enterprise) loans.

### **Steps in securitisation the process**

The essence of the securitisation process consists of three steps (Vink 2009) and is shown in figure 1.

In the first step the sponsor uses specific selection criteria to identify the assets suited to be included in a securitisation transaction. These selection criteria are imposed on a set of assets to form a portfolio to be securitized.

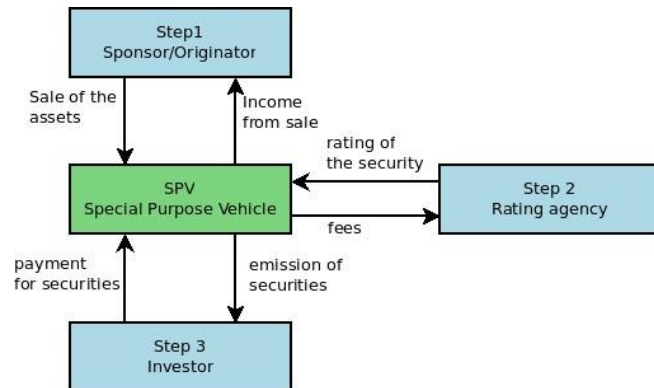


Figure 1: The securitisation process.

In the second step the quality of the portfolio is measured by different rating agencies. The rating agency will assign a rating to the securities. Based upon this credit rating, the interest payments for the bonds can be decided. The rating agency receives a fee for its services. These fees can reach up to 10 basis points of the face value, with minimum amounts of \$30.000, and can go up to \$2.4 million (Partnoy 2006).

In step three the new securities are sold to the investors by the SPV. From the moment the deal is completed until maturity of the transaction, the investors will receive the interest that is generated by the cashflows of the assets in the portfolio. These cashflows are collected by the originator (gets a fee for servicing the assets), who pays the SPV, who then pays the investors the interest on the bonds.

### **1.3 Problem statement**

There may be several reasons for initiating a securitisation process. Experience has shown, that regardless of the reason for the securitisation, the response times of delivering data to CMG need to be short. One of the first deliverables after initiation of such a securitisation process, is always the data. This data does not only need to be delivered quickly but should also be reliable. The whole data-delivery process also needs to be transparent so that any question concerning the source of the data and the calculation of the measures can be answered with little effort.

In the past the data-delivery has been done by some ad-hoc programming. Such type programming is very error prone because the application itself does not go through the usual development design and testing phases. Since important decisions with huge financial impact are made based upon the

supplied information, this has to be very reliable. In order to ensure high data quality the process has to be monitored very closely, which leads to a very time-consuming and labor intensive process that can take up to several days.

Another issue with the current process is that it lacks flexibility, in terms of doing “what-if” analysis. “What if analysis” are analysis that show the effect on certain measures of the portfolio if the selection criteria are changed. For example: “what happens to the total exposure of the portfolio if a certain asset type is removed from the portfolio?” Each time one of the criteria is changed, a bulk of the software had to be rerun, then delivered to CMG again, until the desired portfolio was found. Needless to say, this again was a very time consuming process.

Next, given a set of selected assets, the tranche structure of the selected portfolio is not known until this is sent to the rating agencies. As stated in the previous section, the rating agencies receive a substantial fee for there services. It is therefore very useful to gain more insight in the portfolio. In this project we study a model for assessing the credit risk of a portfolio and provide a method for tranching the selected portfolios.

This model may then be used by the bank as a tool to provide more insight in the credit structure of the portfolio, which can enable CMG to engage in discussion with rating agencies.

In short, the process of selecting the assets that are to participate in a securitisation pool (deal) takes too long. This process does not allow for data analysis in terms of scenario analysis. There also is no in-house method to measure the credit quality of the portfolio after the assets have been selected.

## ***1.4 Phases of the project***

As can be read in the problem statement, the challenge contains two parts. For this reason this project was also split up into 2 phases, due to the different set of skills needed for the challenge.

The first phase covers the business and computer science aspect of the project. In this phase the goal is to improve the efficiency, reduce the error proneness of the asset selection process, and to set up a more flexible environment to do scenario (what-if) analysis. The solution here was to set up the process in a more structured manner. Here the ETL (extract, transform and load) model was used to create a data mart containing all the measures needed for the securitisation process. With this database as the base, different selection criteria can be imposed to create a set of eligible loans, which can be analyzed for possible inclusion.



Phase two covers the mathematical and computer science aspect of the project. Here a model for rating a newly selected portfolio is studied and used to assess the credit risk of a portfolio of selected assets. Using simulation techniques similar to the ones the rating agencies use, portfolios are split up into several rating categories. With this tool in hand, CMG is able to get an indication of the size of these rating categories, even before the rating agencies get involved. If necessary, the composition of assets in the pool can be adjusted to get a more desired tranche structure.

## **1.5 Technical aspects**

Throughout the project the SAS programming language was used to build all software. SAS is very diverse and is the main data-processing language used within Data Management. Besides data-processing, SAS is also very suited for statistical analysis. This programming language was used in both phases of the project. The platforms (mainframe and NT network), however, do differ between phase one and phase two.

In the first phase all the data-processing takes place on the mainframe environment. This environment was chosen due to the fact that most of the source data is available in this environment. For mainframe development a version of SAS, named SAS mainframe, is used.

For phase two the software is built on the windows NT platform. The choice to switch to the NT platform was because of two reasons. The first one is that most of the data needed for this part is sourced on the windows platform. The second reason for this choice is that SAS mainframe has limited graphing abilities. The SAS PC version (as the windows version is called) of SAS contains many graphing functions that can not be utilized on the mainframe version. Since there is a need to visualize the results, the choice for SAS PC is obvious.

## 2 Pool Selection environment

As the solution to the problem of the timely and labor intensive data-delivery process as described in the problem statement earlier, the decision was made to create a data mart, which we will name the PSE (pool selection environment). The PSE will contain the data on an asset level.

The PSE software will be scheduled to run on a monthly basis. This way the PSE will contain the end of month positions of the assets.

One of the requirements of the PSE is that it needs to be self-sufficient throughout the securitisation process. This means that all the data needed during this process should be present in the database. For this reason a broad selection of asset attributes are included in this data mart.

The design and implementation of the data mart is according to the extract transform and load (ETL) model.

### 2.1 Extract Transform and Load

Extracting data from the systems.

In data warehousing terms, collecting all the data from the information systems is called extraction of the data. In this step data from different information systems (some of these are listed in appendix A) are consolidated into the PSE environment.

One of the challenges in this step is that information systems store data in different formats, such as DB2 (an IBM database), flat files, SAS or Microsoft Access.

The SAS programming language has capabilities to read all of these types of files and to store the data in its own, SAS format. To import the data into the the PSE environment data-connectors were made to read these data into the

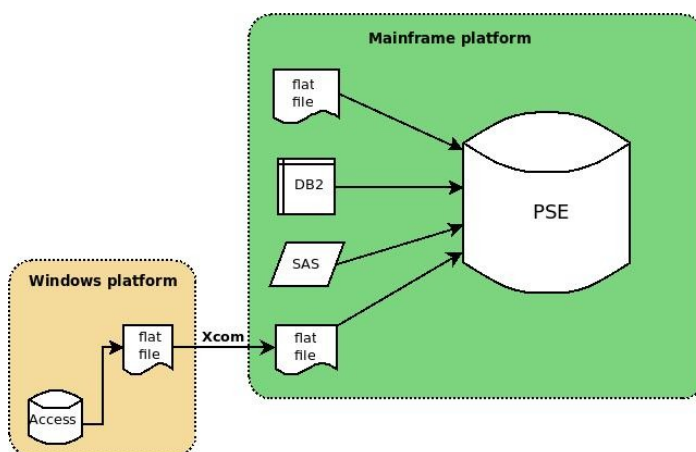


Figure 2: Data extraction

system. Since the DB2, flat files and SAS files are located on the mainframe platform this could be done relatively simple. To collect the Microsoft Access data, some extra steps had to be included, since Microsoft Access is a pure Windows application. Thus, somehow the data had to be delivered from the Access database on Windows to a SAS datafile on the mainframe system. There are several ways

to accomplish this, and the following solution<sup>2</sup> was chosen; First the data is exported out of Access as a CSV file, after which a data transport tool named XCOM is used to upload this to the mainframe environment. After this is done, the data can be extracted into the PSE environment, as shown in figure 2.

## 2.2 Check

Not all data sources are ready to have data extracted from them at the same time, because some processes take longer than others. The ready date of the first data source can be up to 2 weeks before the ready date of the last data source. Before the extraction we first need to make sure that all the source data is ready, therefore a check object is placed at the beginning that checks to make sure all the data is available. This object returns a 1 if all the data is present and a 0 otherwise. Such a check object is using the Job Control Language (JCL), which is used to control and schedule software in the mainframe environment.

## 2.3 Transformations

When all the data is collected in the PSE, the data is not ready to be used yet. This data still has to go through a certain transformation process. These transformations can range from changing the name of an attribute or summing up amounts to computing complex statistical measures. Almost all variables undergo a name change during the process to conform to naming conventions used by the CMG and Data Management. Three examples of transformations in the PSE include

- Data cleansing:  
Some values that are collected in the extraction process contain faulty data. Some of these can be “fixed”. For example; a certain identification code contained leading zeros. Since for this code the leading zeros are not desired, these are removed.
- Calculating the coverage percentage:  
This measure is a function of the total collateral, total exposure and possible future exposures, which indicates a percentage of the clients exposure that definitely will be recovered in case of default. Each of the underlying values are also values determined in the transformation process.
- Currency conversion:  
Some loans are in a foreign currency. All these have to be converted into euro using the actual foreign currency exchange table.

---

<sup>2</sup> Reasons for choosing this method are due to the fact that it is better maintainable than others.

## 2.4 Combining

The data extracted from each system needs to be combined in order to get one record of data per loan contract. This is done by using different keys that are present within each system, and by making use of assistance-tables. For example: say we want to add the postal code of a client to a loan contract. We need to combine a table from the loan system (containing only loan information) with a table from the client database (containing only client information). In order to perform this merge, a common key needs to be found. These two tables do not have such a common key, i.e, the client system does not have a reference key to the contract data and the loan system does not have a reference key to the client system. There is, however, a table that is created and updated by a department within Data Management that contains two attributes; a client number and a contract number. We can use this assistance-table to perform the match between the loan and client data. This is done by merging the loan with the assistance-table by using the contract number, to retrieve the client number. After this is done we can merge the resulting table with the client table by using the client number as displayed in figure 3.

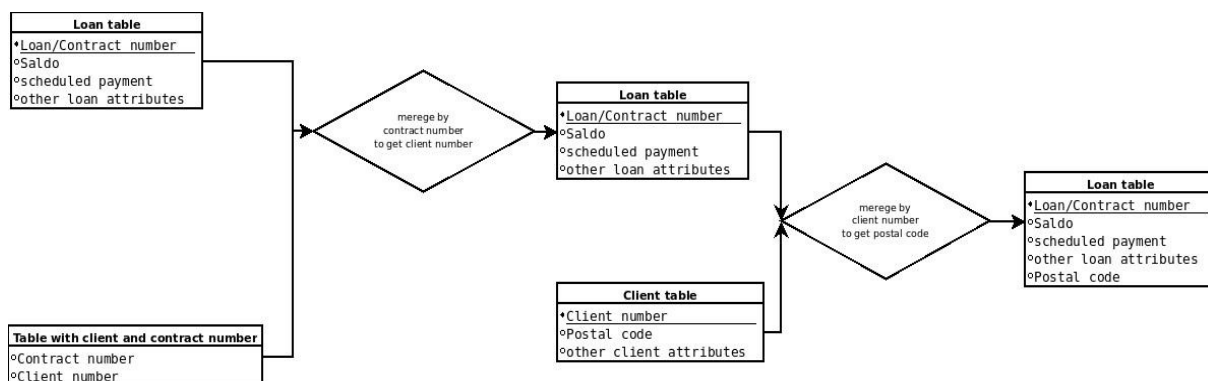
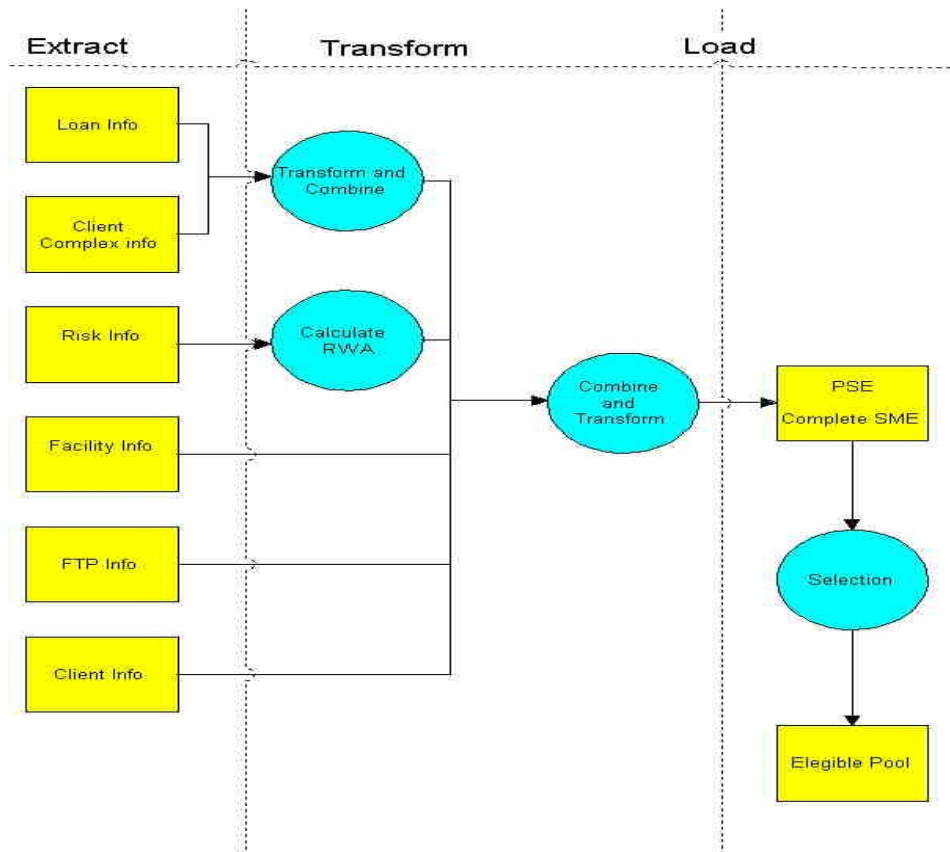


Figure 3: Example of a combination.

## 2.5 Loading

After all the transformations have been done, all the data is combined and placed in a data-mart where analysis can be done. This step is also referred to as loading the data. In the PSE this data-mart contains one table with all the necessary information that is needed for the securitisation at an asset level.

The above process getting data from the source systems into the datamart, is called the ETL process and the data flow is shown in the diagram below.



Pool Selection Environment Data Flow

Figure 4: ETL process.

## 2.6 Selection

For the past securitisations, CMG decided on a set of requirements for loans in a securitisation portfolio. Data management takes these requirements, selects all the loans that satisfy the given requirements, and delivers them to CMG.

These requirements cover many different aspects of the client and the loan itself. All this data is now present in the PSE, and is ready to be analyzed. Up until now this has been a very manual process. Due to the fact that the selection requirements can change from time to time, some of the processes must be run by hand. The pool selection environment is set up in such a way that the process up until the selection itself, is automated and the selection process semi-automated. This enables flexible analysis of the system by changing the selection parameters.

## 2.7 Requirements

There are more than 20 constraints applied to the loans during the selection process. Such constraints are, for example:

- duration of the loan must be greater than some constant period D (loan attribute)
- the credit rating of the client must be within a given range (client attribute)
- the exposure of the loan has to be decreasing in size over time (loan attribute)
- the loan must not be in a previous securitisation (loan attribute)
- clients that have had delinquencies in the past may not be included (client attribute)

These are just a couple of constraints and, depending upon the aim of the transaction, these may be made more tight or loose. Thus, this part of the software needs to be flexible.

## 2.8 Asset selection

After the ETL process is completed, all the data that is needed to perform the selections is available. First an automated selection is performed, which is a broad selection. This means that the selection criteria are automatically imposed on all the SME loans in the PSE, but this is a wide<sup>3</sup> selection, to an intermediate pool (contains more loans than due to the loose criteria). For each criteria an indicator variable is created, which can be either a 1 or 0 to indicate “meets criteria” or “does not meet criteria”, respectively. Let us take two criteria, for example:

- duration of a loan  $> d$
- credit rating between A and C

We then create two variables “crit<sub>1</sub>” and “crit<sub>2</sub>”<sup>4</sup> to indicate whether the criteria 1 or 2 is met or not. Since there are over 20 criteria, more than 20 of these variables will be created. In order to easily select the loans that meet all the criteria, a variable name “totcrit”<sup>5</sup> is calculated by multiplying all the

criteria indicators. For  $n$  the number of criteria, this will be  $totcrit = \prod_{i=1}^n crit_i$ . If

“totcrit” is 1 then the asset is selected.

Next, all the selected loans are exported into an Microsoft Access database, where further analysis can be done using simple SQL queries in order to further narrow down the assets to be included in the portfolio.

---

<sup>3</sup> These are very loose selections, that should always be used.

<sup>4</sup> Crit<sub>1</sub>= the first criteria, and crit<sub>2</sub> is the second criteria.

<sup>5</sup> Totcrit= 1 if all the criteria are met.

### 3 Mathematical background

In this chapter the mathematical concepts and underlying formulas that are used in the following chapters are described. First we start with the notation that is used. Next the credit loss distribution, which is the backbone of portfolio risk assessment, is described. After that, the correlations between industry sectors, and obligors are described, followed by a description of the firm's value model which is the bases of the single factor model. In order to correlate a vector of variables, we first need to perform a Cholesky factorization to the matrix containing the correlations. We conclude this chapter with a discussion about this technique.

#### 3.1 Notation

The notation that will be used throughout this chapter is as follows<sup>6</sup>:

- $i$  is the  $i^{\text{th}}$  asset in the portfolio
- $a_i$  is the sector to which obligor  $i$  belongs
- $r_i$  the value of obligor  $i$ 's asset
- $\alpha_a^i$  is the asset  $i$ 's correlation to the sector  $a$ . In essence, this is the portion of the asset's value explained by the common factor 'a'
- $R_a$  is the return of the common factor 'a', drawn from the standard normal distribution.  $R_a$  and  $R_b$  are correlated with  $\rho_{ab}$
- $\alpha_a^2$  is the correlation within the sectors (intra-sector correlation), and is in the interval  $[0,1]$
- $\rho_{ab}$  is the correlation between the sectors (inter-sector correlation), and is in the interval  $[0,1]$
- assets  $i$  and  $j$  in sectors  $a$  and  $b$ , respectively, are correlated according to  $\alpha_a \rho_{ab} \alpha_b = \Sigma_{ij}$ . Thus,  $\Sigma_{ij}$  is the correlation between assets  $i$  and  $j$
- $\Phi(0,1)$  is the cumulative distribution function for a normal distribution with mean 0 and standard deviation 1.

#### 3.2 Credit Loss distribution

The credit loss distribution of a portfolio is used for determining several useful measures of portfolio risk, such as the expected and unexpected loss. This portfolio loss distribution is also central to understanding the value and risk inherent to an asset backed security (ABS). In essence the loss distribution represents the probability of experiencing a given level of loss over a given period<sup>7</sup>.

<sup>6</sup> The currency is always Euro.

<sup>7</sup> In this paper the period is 1 year.



The cumulative loss distribution function  $F(x)=P(X\leq x)$  gives us the probability that the portfolio losses  $X$  are less than a certain value  $x$ . The probability of a certain loss in the interval  $[a,b]$  occurring is expressed by  $P(a\leq x\leq b)$ .

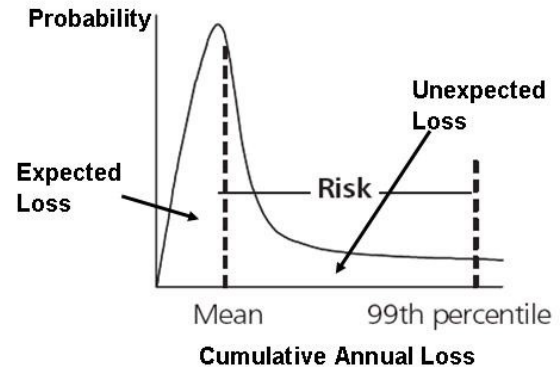


Figure 5: Example of the loss distribution.

To create the loss distribution, the losses within the portfolio can be simulated using the Monte Carlo simulation approach. Using this approach a number of independent trials are simulated. For each trial, a vector of size  $N$  (size of the portfolio) containing standard normal and correlated loss events  $X_i$  is generated. If the event  $X_i$  is less than some threshold  $y_i$ , the corresponding asset is defaulted<sup>8</sup> for that trial. For each run, the amount that is lost is summed and divided by the total portfolio principal balance, to get the loss percentage for that trial, which represents the default rate. Collecting all the observed default rates from all the simulation scenarios, generates a probability distribution for the default rate, which is the loss distribution.

In the estimation of the credit loss distribution, the correlation between two counterparties has a significant impact on the shape of the distribution (this is shown in chapter 4). The correlation therefore has to be included in the model, which will be described next.

### 3.3 Default Correlation

Default correlation is a measure of dependency between two counterparties. Such a dependency can have several causes; for example one counterparty could be a customer of the other one, or two counterparties operate in the same sector or market. Default correlation with respect to multiple counterparties is highly relevant for the pricing and credit risk management of collateralized debt obligations, for portfolio management of loan or bond portfolios and bank-wide risk management: Diversification of credit risk can only be successful if an adequate portfolio credit risk model is in place that can quantify the risks in the portfolio (Schönbucher 2000).

The default correlation  $\Sigma_{ij}$  between counterparty  $i$  and counterparty  $j$  gives us a measure of linear correlation between counterparty  $i$  and  $j$ . If  $\Sigma_{ij}=0$ , the

<sup>8</sup> Default occurs when a debtor has not met his or her legal obligations according to the debt contract.



counterparties default independently. For  $\Sigma_{ij}=1$  they are totally dependent and move in the same direction, and for  $\Sigma_{ij}=-1$  they are totally dependent and move in opposite direction. Such correlations can be found by historical analyses of default data.

Ideally the correlation between each counterparty would be used. Obtaining all the data to calculate these correlations. This is, however, not possible. Even if the data to calculate the correlation for each individual obligor is available, it would be almost impossible to use them due to the fact that you would need very large matrices, and therefore lead to very long computing times or “out of memory” messages. For this reason the counterparties are grouped together into clusters.

When using clustering, the following assumptions are made:

- all clients within the same cluster (intra-cluster) are similarly correlated
- the movements between two clusters (inter-cluster) are equal

In this study the clusters will be based upon the industry sectors that the obligors operate in.

As input we use two types of correlations:

1. Intra-sector correlations: These are the correlations between any two counterparties within the same sector. This is calculated by taking the average of all pair-wise correlations within an sector. These correlations are denoted by a vector  $\alpha^2$ .
2. Inter-sector correlations: These are the correlations between any two counterparties in different sectors. This is calculated by taking the average of all pair-wise correlations between the two sectors the counterparties operate in. These correlations are stored in a matrix  $\rho$  where the correlation between sector a and b is  $\rho_{ab}$ .

### **3.4 Simplified firm's value model**

In this study we use the single factor model to represent the (correlated) defaults. The single factor model (Vasicek 2002) is a simplified version of the firm's value model (Merton 1974).

The basic idea of the firm's value model, is that an obligor default occurs when the asset value of a firm<sup>9</sup> falls below the debt level. When this happens, selling the assets to repay the debt is not possible (Cespuedes 2002). Hence, the obligor is not able to fulfill its commitments and defaults.

---

9 A firm is an obligor.

In the firm's value model it is assumed that a default of an obligor occurs due to a change of its asset value  $r_i$ , where  $r_i \sim \Phi(0,1)$ <sup>10</sup>. In this model obligor  $i$  defaults if its value  $r_i$  drops below some threshold  $z_i$ .

### 3.5 The single factor model

Factor models (also known as conditionally independent credit risk models) are among the few models that can replicate a realistic correlated default behavior while still retaining a certain degree of analytical tractability<sup>11</sup>. Many models that are used in practice are based on this approach (Schönbucher 2000).

In the single factor model it is assumed that the assets of the obligor are composed of a standard normally distributed common part  $R$  and a standard normally distributed idiosyncratic part  $\varepsilon_i$ .  $R$  represents the common factor, which in our case is the sector specific and  $\varepsilon_i$  represents the company specific value.  $R$  and  $\varepsilon_i$  are uncorrelated.

The single factor model is stated as follows:

$$r_i^a = \alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i$$

#### Claim:

If  $\alpha^2$  is the intra-correlation (correlation between obligors) within sectors,  $R \sim \Phi(0,1)$  has a correlation structure according to the inter-correlation  $\rho$ , and  $\varepsilon_i \sim \Phi(0,1)$ , then  $r$  is correlated according to  $\Sigma = \text{Corr}(r_i^a, r_j^b) = \alpha_a \rho_{ab} \alpha_b$ .

This is shown as follows:

- First we show that the expectation of the asset return is zero, because both the stochastic variables have an expectation of zero.

$$E(r_i^a) = \alpha_a E(R_a) + \sqrt{1 - \alpha_a^2} E(\varepsilon_i) = 0.$$

- The variation of the  $r_i$  should be 1

$$\begin{aligned} \text{Var}(r_i^a) &= \text{Var}(\alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i) \\ &= \alpha_a^2 \text{Var}(R_a) + (\sqrt{1 - \alpha_a^2})^2 \text{Var}(\varepsilon_i) + 2\alpha_a \sqrt{1 - \alpha_a^2} \text{Cov}(R_a, \varepsilon_i) \\ &= \alpha_a^2 + 1 - \alpha_a^2 + 0 = 1. \end{aligned}$$

<sup>10</sup>  $\Phi(0,1)$  Signifies a normal distribution with a mean of zero and a variance of 1.

<sup>11</sup> Easily managed or controlled.

- To show that the correlation between the returns is as desired we can derive

$$\text{Corr}(r_i^a, r_j^b) = \frac{\text{Cov}(r_i^a, r_j^b)}{\sqrt{\text{Var}(r_i^a)\text{Var}(r_j^b)}} = \text{Cov}(r_i^a, r_j^b).$$

$$\begin{aligned} \text{Cov}(r_i^a, r_j^b) &= E(r_i^a r_j^b) + E(r_i^a)E(r_j^b) = E(r_i^a r_j^b) \\ &= E[(\alpha_a R_a + \sqrt{1-\alpha_a^2}\varepsilon_i)(\alpha_b R_b + \sqrt{1-\alpha_b^2}\varepsilon_j)] \\ &= E[\alpha_a \alpha_b R_a R_b] + E[\alpha_a R_a \sqrt{1-\alpha_b^2}\varepsilon_j] + E[\alpha_b R_b \sqrt{1-\alpha_a^2}\varepsilon_i] \\ &\quad + E[\sqrt{1-\alpha_a^2}\varepsilon_i \sqrt{1-\alpha_b^2}\varepsilon_j]. \end{aligned}$$

Since  $E(\varepsilon_i \varepsilon_j) = 0$  and  $E(\varepsilon_i^a R_b) = 0$ , the terms

$$E[\alpha_a R_a \sqrt{1-\alpha_b^2}\varepsilon_j] = \alpha_a \sqrt{1-\alpha_b^2} E[R_a \varepsilon_j],$$

$$E[\alpha_b R_b \sqrt{1-\alpha_a^2}\varepsilon_i] = \alpha_b \sqrt{1-\alpha_a^2} E[R_b \varepsilon_i],$$

and

$$E[\sqrt{1-\alpha_a^2}\varepsilon_i \sqrt{1-\alpha_b^2}\varepsilon_j] = \sqrt{1-\alpha_a^2} \sqrt{1-\alpha_b^2} E[\varepsilon_i \varepsilon_j] \text{ are } 0,$$

so we are left with:

$$\text{Cov}(r_i^a, r_j^b) = E[\alpha_a \alpha_b R_a R_b] = \alpha_a \alpha_b E[R_a R_b] = \alpha_a \alpha_b \rho_{ab} = \text{Corr}(r_i^a, r_j^b).$$

- If obligor i and j are in the same sector then  $a=b$  and  $\rho_{ab}=1$  so,

$$\Sigma_{ij} = \text{Corr}(r_i^a, r_j^b) = \begin{cases} \alpha_a^2 & , \text{if } a=b / \text{obligors} \in \text{same sector}, \\ \alpha_a \rho_{ab} \alpha_b & , \text{if } a \neq b / \text{obligors} \in \text{different sectors}. \end{cases}$$

Figure 6 shows the correlation relationships between two sectors, two obligors within a sector, two obligors from different sectors and between an obligor and its sector.

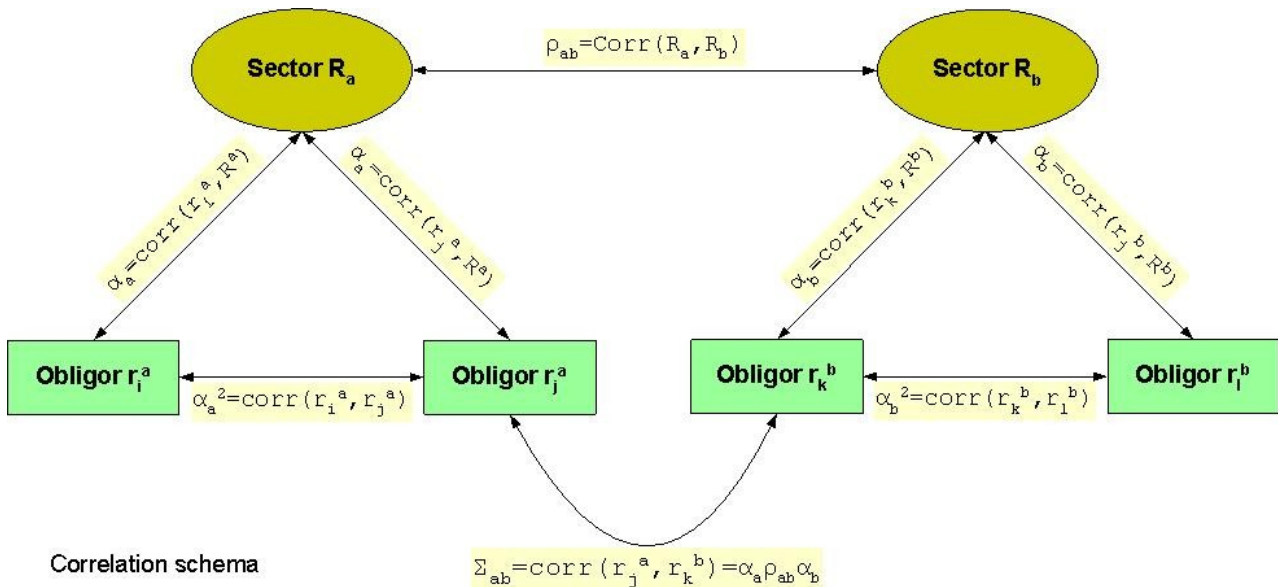


Figure 6: Correlation between sectors and obligors.

### 3.6 Cholesky factorization

The market return vector  $R$  is drawn from a standard normal distribution and needs to have some correlation structure according to a correlation matrix  $\rho$ . In order to get the desired dependency, we need to perform a linear transformation to the vector  $R'$  as  $R = AR'$  such that,  $\text{Cov}(R') = \rho = A^T A$ . Where  $A^T$  is the transpose of matrix  $A$ . To find this matrix  $A$ , we need to perform a Cholesky factorization on  $\rho$ .

So, starting with a given correlation matrix  $\rho$ , we can use the Cholesky factorization to find a matrix  $A$  such that  $\rho = A^T A$ . If we then multiply  $A$  with  $R$  we obtain the vector  $R'$  with the desired correlation structure according to a given correlation  $\rho$ .

If the entries of the vector  $R$

$$R = \begin{bmatrix} R_1 \text{ Agriculture} \\ R_2 \text{ Mining} \\ R_3 \text{ Auto industry} \\ R_4 \text{ Steel industry} \\ \vdots \\ \vdots \\ \vdots \\ R_n \text{ Clothing industry} \end{bmatrix}$$

are the random variables with a variance of 1 and mean 0, then the covariance and correlation matrix whose (a, b) is the correlation

$$\rho_{ab} = \text{Cov}(R_a, R_b) = E[(R_a - E[R_a])(R_b - E[R_b])^T]$$

Because  $E[R_a]$  and  $E[R_b] = 0$ , this leaves us with

$$\text{Cov}(R_a, R_b) = E[(R_a)(R_b)^T] = E[(AR'_a)(AR'_b)^T] = AA^T E(R'R'^T)$$

$E(R'R'^T)$  is the identity matrix  $I$ ,

so,

$$\rho_{ab} = \text{Cov}(R_a, R_b) = AA^T$$

### **Cholesky factorization formula**

If we have

$$\rho = A^T A$$

$$\begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ 0 & a_{22} & a_{32} \\ 0 & 0 & a_{33} \end{bmatrix}$$

then matrix A can be obtained using the following equations:

$$a_{ii} = \sqrt{s_{ii} - \sum_{k=1}^{i-1} a_{ik}^2} \text{ for the diagonal}$$

and

$$a_{ij} = \frac{1}{a_{ii}} \sqrt{(s_{ij} - \sum_{k=1}^{i-1} a_{ik} a_{jk})} \text{ for } i < j.$$

## 4 Simulation

In this chapter the simulation is discussed. We first discuss the Monte Carlo simulation technique. Then we move on to talk about the input parameters of the simulation, and the modeling of the correlations between assets within the portfolio. Fitting the correlations within this model leads to some technical difficulties. We also discuss a limiting distribution, which will be used to test our model. In the last paragraph the results are shown.

### 4.1 Portfolio credit risk model

Models for ABS risk analysis are based on the estimation of the underlying asset default probabilities and recoveries, and the linkage of these through a dependency model, which specifies the joint default behavior. With such a model the loss distribution of a portfolio, different risk measures can be determined.

To estimate the portfolio credit risk, a Monte Carlo approach is used.

### 4.2 Monte Carlo

Monte Carlo simulations relate to or involve “the use of random sampling techniques and often the use of computer simulation to obtain approximate solutions to mathematical or physical problems especially in terms of a range of values each of which has a calculated probability being the solution” (Merriam-Webster 2010). This definition from Webster's dictionary is a good description of Monte Carlo studies.

#### **Why Monte Carlo?**

Monte Carlo techniques are very useful in cases when it comes to estimation of the probability distribution of outcomes for a system whose properties make analytic or numeric solutions impractical, as is the case here. Since there are no tailor-made mathematical tools (i.e. no closed form expression of the portfolio loss distribution) for evaluating the credit risk of the portfolio, Monte Carlo simulations can be used. By characterizing the distribution of the portfolio losses, through Monte Carlo, we can obtain the values that are necessary to tranche the portfolio.

This approach for estimating the loss distribution, is similar to the approach used by the credit rating agencies, and is described in the next sections.

### 4.3 Setting up the simulation

#### **Simulation inputs**

The model is based upon the following inputs being available:

1. EAD (exposure at default): amount of risk of the obligor at the point of default. This is the maximal amount of loss that can occur if the obligor defaults
2. Sector id: indicator for sector the obligor operates in
3. mapping of scenario default rates (SDR: loss-percentiles) to ratings
4. intra-sector correlation vector  $\alpha$
5. inter-sector correlation matrix  $\rho$
6. PD: probability of default within 1 year
7. LGD (loss given default): expected percentage of the EAD that will be lost, given a default occurs

#### **Basis of the simulation**

Assume the portfolio consists of N assets and let  $X_i$  denote the performance of the  $i^{\text{th}}$  asset, with poor performance corresponding to small values of  $X_i$ . The event that the  $i^{\text{th}}$  asset defaults is equivalent to some quantity  $z_i$ , such that  $P(X_i < z_i) = PD_i$ . Thus  $z_i = \Phi^{-1}(PD_i)$ .  $X_i$  is assumed to be normally distributed with a mean of 0 and a variation of 1, and must have a correlation structure according to a given correlation matrix  $\Sigma$ . For the simulation a vector  $r$ , containing the asset values of the obligors is constructed by applying the single factor model:

$$r_i^a = \alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i^a.$$

### 4.4 Modeling the correlation

At first the idea was to model the asset returns independently by imposing the industry correlation matrix  $\Sigma$  directly. To accomplish this, however, the correlation matrix had to be expanded to a matrix A of size NxN for each obligor combination. Since the portfolio can get quite large (in a realistic case about 13000 assets can be marked as eligible), this results in a very large matrix with 13000 rows and 13000 columns, which caused significant data processing and computational problems.

The second modeling approach was the factor approach, in which individual asset returns are modeled as a combination of responses to a set of correlated factors and a set of independent idiosyncratic returns. Here the return of an asset is described by a single-factor model, where a portion of the return is explained by the common industry factor and the other portion by an asset specific factor (idiosyncratic). This model is discussed in section 3.5 and is

defined as  $r_i^a = \alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i^a$ , where the  $\alpha$  can be interpreted as the portion of asset return variation explained by the common industry factor  $R$ , and  $\varepsilon_i$  is the idiosyncratic return, where  $r$  has a correlation structure according to the industry correlation matrix  $\Sigma$ . Hence, all correlation in asset returns are then derived from their dependence on the correlated factors, which reduces the dimensions of the correlation matrix to the number of industry factors.

#### 4.5 Simulation steps

In this section the simulation steps are described. Suppose the portfolio is  $N$  obligors large, the number of sectors is  $M$  and the number of simulations is  $S$ :

1. decompose the inter-correlation matrix  $\rho$  using Cholesky decomposition, and obtain the lower triangular matrix  $A$
2. create a vector  $R' \sim \Phi(0,1)$  of size  $M$ , and impose the inter-correlation on this vector by performing  $R = R' \cdot A$
3. step 2 is repeated  $S$  times, to get a matrix containing a set of correlated normally distributed random sector returns for each iteration of the simulation
4. calculate the threshold value  $z_i = \Phi^{-1}(PD_i)$  for each obligor  $i$ , where  $\Phi^{-1}$  is the inverse of the cumulative standard normal distribution function
5. calculate  $r_i^a = \alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i^a$  where  $\alpha$  is the intra-correlation of the sector of the obligor and  $\varepsilon_i$  is drawn from the standard normal distribution
6. determine if obligor  $i$  defaults by evaluating  $r_i < z_i$ , and calculate
 
$$Loss = \sum_{k=1}^N LGD_i \times EAD_i \times 1_{(r_i < z_i)}$$
 where  $1_{(r_i < z_i)}$  is the default indicator for the  $i^{\text{th}}$  obligor.
7. calculate the loss-percentage as loss/(total portfolio exposure)
8. step 5-8 is repeated  $S$  times for each iteration

The above simulation results in a vector containing loss-percentages for each simulation run. If the number of simulations  $S$  is large enough, the distribution of the portfolio losses can be determined at a high level of accuracy. For an Asset Backed Security (ABS) linked to a single portfolio of assets (as in this case), the portfolio loss distribution contains all of the information to determine the performance of each ABS tranche (Gilkes 2005).

#### 4.6 The limiting distribution of losses

In order to verify that our simulation works we need to find a scenario that gives us a closed form-expression of a loss distribution. We can then use our to



replicate the closed-form distribution using our simulation model. If the results from the simulation approach the distribution described by the closed-form expression, we can assume that our model works.

Given the following assumptions on the loan portfolio, a closed form expression can be found for the loss distribution (Vasicek 1991).

1. The number obligors,  $N$ , goes to infinity (very large number of obligors).
2. The exposure size is  $1/n$  for the obligors, so as  $n$  goes to infinity, the exposure tends to go to zero.
3. All obligors have the same probability of default  $PD$ .
4. The asset value for all obligors follows an i.i.d Gaussian process:

$$x_i = \alpha Y + \sqrt{1 - \alpha^2} \varepsilon_i.$$

Where  $Y$  is a common factor to all obligors (single factor), and  $\varepsilon_i$  is an idiosyncratic shock.

5. The asset correlation between all obligors is  $\alpha^2$ .
6. The loss given default (LGD) of each obligor is 1.

Under these assumptions the closed formula for the distribution of defaults is given by:

- The cumulative distribution:

$$F(x, \alpha^2 PD) = \Phi\left(\frac{\sqrt{1 - \alpha^2} \Phi^{-1}(x) - \Phi(PD)}{\alpha}\right).$$

- The density of the loss distribution:

$$f(x, \alpha^2 PD) = \sqrt{\frac{1 - \alpha^2}{\alpha^2}} \exp\left(-\frac{1}{2\alpha^2} (\sqrt{1 - \alpha^2} \Phi^{-1}(x) - \Phi^{-1}(PD))^2 + \frac{1}{2} (\Phi^{-1}(x))^2\right).$$

where:

- $\Phi(\cdot)$  is the standard normal cumulative distribution.
- $\Phi^{-1}(\cdot)$  is the inverse normal cumulative distribution.
- $\alpha^2$  is the asset correlation.
- $PD$  is the individual probability of default.

The difference between this model by Vasicek (1991) and our single factor model used in the simulation is in the assumptions. If there is no clustering then all assets have a single default correlation factor.

Given this closed formula for the distribution of defaults, we can calibrate our single factor model (in the simulation)  $r_i^a = \alpha_a R_a + \sqrt{1 - \alpha_a^2} \varepsilon_i$  in such a way that its loss distribution can be described by the closed-form expression above. This way we can run several simulations for different correlation values, and show

that the simulation gives the desired results.

## 4.7 Calibrating the parameters

To show that the model used in the simulation gives the desired results, we first need to calibrate the input parameters of the simulation in such a way that the assumptions made in deriving the Vasicek model are met.

### Correlation

As shown in section 3.5 the correlation of our model is given as follows:

$$\text{Corr}(r_i^a, r_j^b) = \begin{cases} \alpha_a^2 & , \text{if } a=b / \text{obligors} \in \text{same sector}, \\ \alpha_a \rho_{ab} \alpha_b & , \text{if } a \neq b / \text{obligors} \in \text{different sectors}. \end{cases}$$

If we set the inter-correlation  $\rho$  between all sectors to 1, then the correlation between all obligors in the pool will be equal to  $\alpha^2$ . By doing this we force all the assets within the pool into one single industry sector, and set the intra-correlation of that industry sector to  $\alpha^2$ . This way all the obligors within the pool have a correlation factor of  $\alpha^2$ .

The other input parameters (individual exposure, PD and LGD) in the simulation are easily set to the values implied by the assumptions.

## 4.8 Testing the model

Now we can apply the assumptions, to the simulation and plot the resulting distribution of defaults against the limiting density distribution  $f(x, \zeta, pd)$ .

For a probability of default (PD) of 5% the plots for different correlation values  $\zeta = \{1\%, 5\%, 10\%, 20\%, 30\%, 50\%, 70\%\}$  are given in figure 7. The results are achieved with 3000 iterations and a portfolio of size  $N=1000$  assets. The solid line is the distribution obtained with simulation, and the dotted line resembles the analytical distribution.

As can be seen in the plots, the density distribution derived by the Monte Carlo simulation approaches the analytical (closed) formula for the distribution of defaults. As the number of simulation iterations get larger, the error will get even smaller.

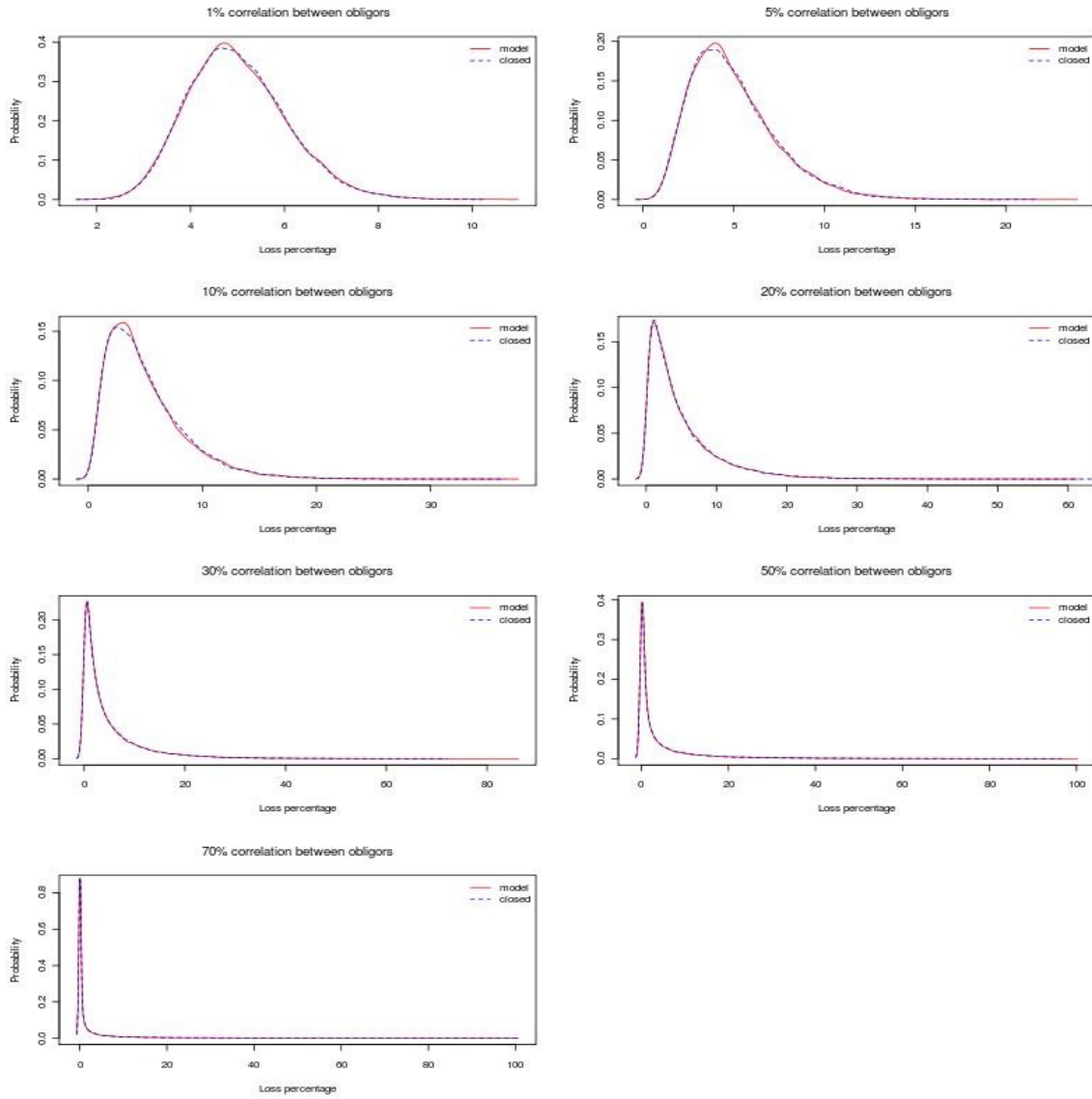


Figure 7: Density plots for different correlation: analytical outputs vs Monte Carlo simulation.

## 4.9 The effect of the correlations

Figure 8 shows the simulated distribution of losses for previous portfolio (PD=5%, N=1000 assets) for different correlation values  $\zeta = [1\%, 5\%, 10\%, 20\%, 30\%, 50\%, 70\%]$ . If the asset correlation increases, the peak (probability weight) shifts to the left, and the tails get longer.

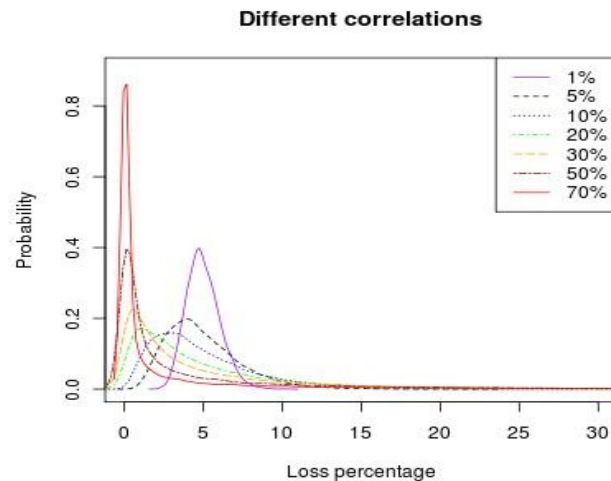


Figure 8: Loss distribution for different correlation values with a probability of default of 5%.

This correlation effect can be seen even more clearly if we look at the percentiles of the loss distribution. Table 1 gives the 99.9% and 95% percentiles for each of the correlations.

From the 95 and 99.9% percentiles, it is clear that the tail

probabilities increase as the asset correlations increase. If we look at the 99.9 percentile, a 10% increase in correlation (from 20 to 30%) results in a (53.02 - 38.75) 14.27% loss percentage increase.

		Percentiles	
		95,0%	99,9%
Correlations	1%	6,87	9,03
	5%	9,52	16,34
	10%	11,78	24,2
	20%	15,47	38,75
	30%	18,58	53,02
	50%	24,61	78,76
	70%	30,91	96,21

Table 1: 95% and 99.9% percentiles for a one factor model. Parameters: PD 5%, #assets N=1000, asset correlations 1, 5, 10, 20, 30, 50 and 70%

## 4.10 Tranching

After the loss distribution is determined using the simulations, the portfolio can be split up in different tranches. In this section the method for determining the tranches of the portfolio is described.

### Scenario default rates

The first measure that needs to be determined is the scenario default rate

(SDR), which is a percentile of the portfolios loss distribution consistent with a given rating.

The SDR is also described as the portfolio default rate, such that the probability of defaults in the portfolio exceeding this portfolio default rate is no greater than the probability of default of a corporate bond with that rating (Bergman 2001). Given the loss distribution of a portfolio and the default rates, the scenario default rates (SDR) can be found with little effort. Let us say  $h$  is the default rate of a given rating, and  $X$  is the portfolio default rate. Then  $P(X > SDR) = h \Rightarrow P(X \leq SDR) = 1 - h$ . Thus we need to look for the portfolio default rate at the  $(1-h)$  percentile of the portfolio loss distribution. This SDR is the attachment point (AP) of the tranche, and the detachment point (DP) of the previous tranche.

If we do this for each rating the whole portfolio can be tranching and the portfolio losses can be allocated to each tranche, starting with the equity tranche. When the losses get passed a certain percentage AP, the losses from AP to DP will be allocated to the investors who hold notes within that tranche. In this study, the one year default rates (table 3) are retrieved from (Gilkes 2005).

AAA	AA	A	BBB	BB	B	CCC
0,0001	0,004	0,012	0,16	1,722	3,971	34,17

*Table 3: Rating one year default rates*

Figure 10 shows the loss distribution of the portfolio from the previous section (N=1000 assets, correlation = 10%) with its tranches according to the default rates above. Figure 9 gives a bar-chart with the sizes of each tranche.

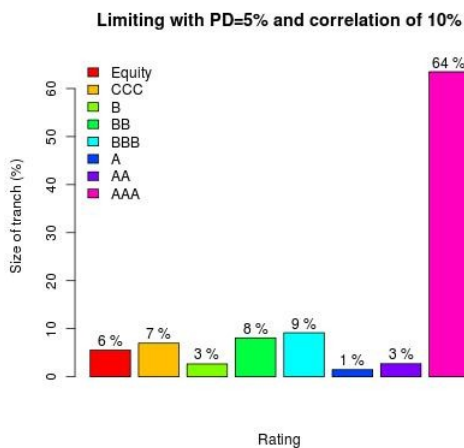


Figure 9: Size of the tranches for pool with PD=5% and a correlation of 10%

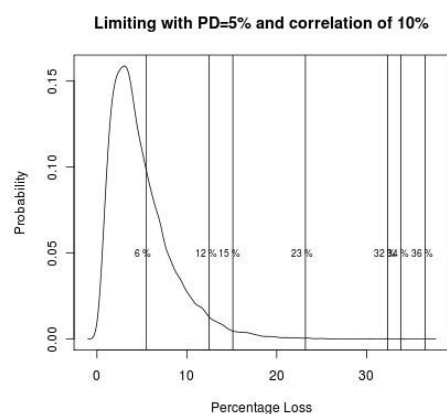


Figure 10: Loss ditribution of the Vasicek distribution with PD=5% and a correlation of 10%

### **Realistic pool**

Let us randomly generate a realistic pool (by realistic we mean a diversified pool with different probabilities of default, exposures and clusters). Here we create a pool containing assets with a PD between .5% and 15%, and exposures between 10.000 and 150.000 euros. We use the correlations as supplied. We can then create the loss distribution and tranches as given in figure 11 and 12 respectively.

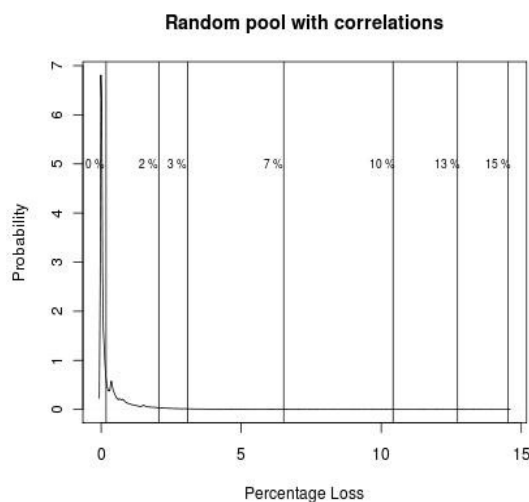


Figure 12: Loss distribution of a pool with random attributes

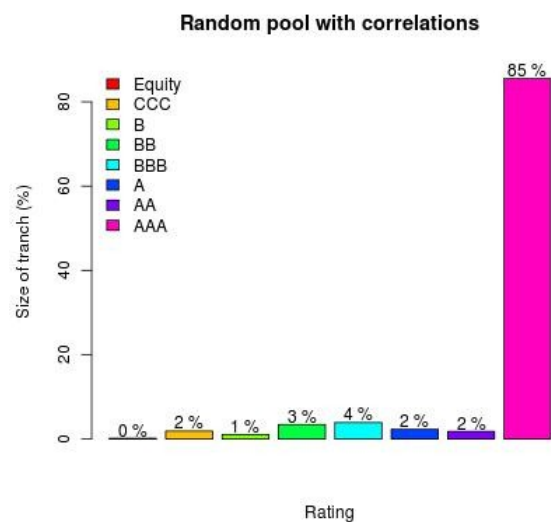


Figure 11: Tranches of a pool with random attributes

In figure 12 the first vertical line is the attachment and detachment point of the equity tranche at 0%, so there is no equity tranche. The second vertical line is at 2% which marks the detachment point for the CCC tranche. This means that the portfolio losses from 0-2% go to the CCC note holders. When the total losses of the portfolio pass 2%, the losses are assigned to the B tranche. This B tranche has a detachment point of 3%, thus has a size of 1%. The attachment point of the AAA tranche is at 15% loss. The default rate for the AAA tranche is 0,0001%. Thus, a attachment point of 15% means that we can say with 99.999% (100%-0,0001%) confidence, that the portfolio losses will not be more than 15%. The detachment point for this tranche is at 100%, so the AAA tranche has a size of 85%.

### **Increase the correlations**

It is interesting to see whether the changing the correlations also has such a huge impact if the portfolio is more diversified.

We therefore increase supplied correlations with 5.9%. Given that the correlation is:

$$\text{Corr}(r_i^a, r_j^b) = \begin{cases} \alpha_a^2 & , \text{if } a=b / \text{obligors} \in \text{same sector,} \\ \alpha_a \rho_{ab} \alpha_b & , \text{if } a \neq b / \text{obligors} \in \text{different sectors.} \end{cases}$$

we can easily accomplish this by increasing  $\alpha$  by a factor  $\sqrt{5.9\%}$ . The results after running the simulation are given in figures 13 and 14.

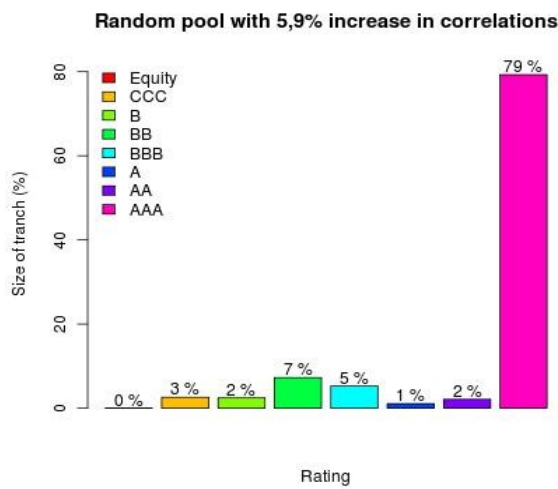


Figure 13: Tranch distribution of a pool with random attributes and correlations increased by 5.9%

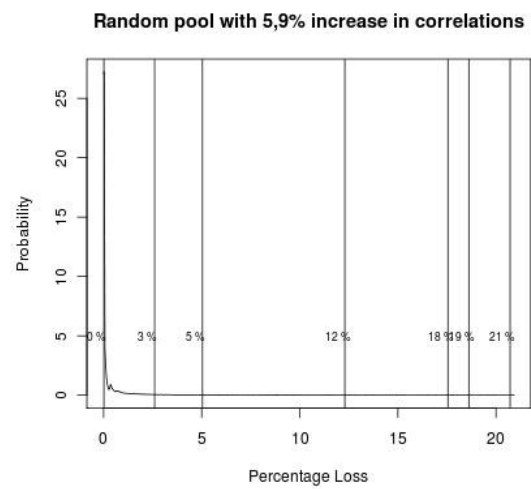


Figure 14: Loss distribution of a pool with random attributes and correlations increased by 5.9%

If we compare the tranches of the loss distributions in Figures 13 and 11 we see that increasing the correlations 5.9% has a significant impact on the tranche distribution of the portfolio. The AAA tranche decreases from a size of 85% to 79% of the portfolio.



## 5 Conclusions

With a structured solution as the Pool Selection Environment (PSE) all the data needed for a planned securitisation transaction, can now be extracted out of the data-mart with little effort. Therefore the data delivery times are reduced significantly. Where it could take up to 4 days to select a pool of eligible assets, this can now be reduced to less than 1 hour. A query tool (for these purposes, Microsoft Access) can now be used to perform further analysis to select the final portfolio to be securitized. This pool can then be input into our simulation model, which returns a portfolio loss distribution. This loss distribution is then analyzed to determine the tranches. A schematic overview of this whole process is given in figure 14 below.

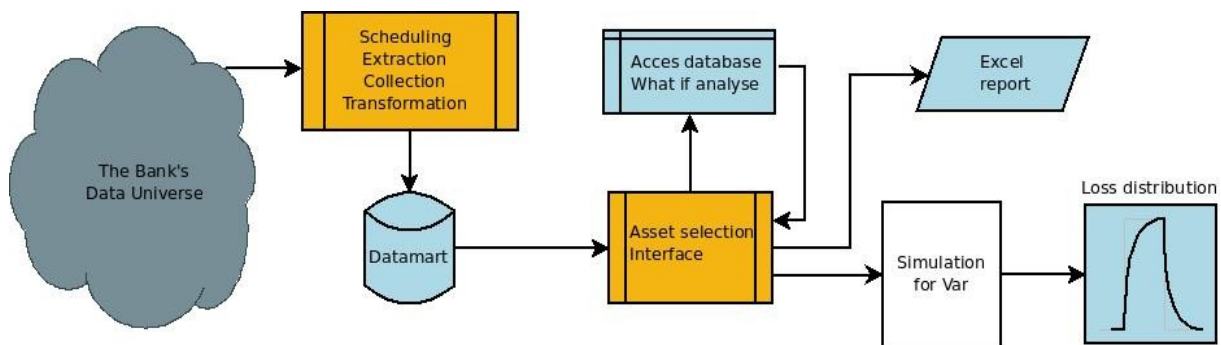


Figure 15: Flow from asset selection to a loss distribution

In a further study, it would be worth looking into a more flexible tool for analysis. Though Microsoft Access is a versatile tool, it does require some technical expertise (which many business users do not have) to use.

As shown in chapter 4, even a 5.9% increase in correlation has a significant impact on the loss distribution of a portfolio. Thus, if the aim is to generate tranches that are comparable (to tranches generated by a rating agency) tranches, it is very important to use similar correlation factors. Another important point to note is that the rating agencies also calculate the probability of default (PD) for the obligors, and therefore these do not necessarily equal the PDs calculated by the bank. The rating agencies also apply a adjustment factor on the tranches, which should be applied after the tranches are determined (Bergman 2001, 2). Keeping these in mind the model can be used to enter gain more insight into the portfolio, which enables CMG to enter into discussions with the rating agencies.



## Appendix

This appendix gives an overview of what types of source systems are used during this project.

**Loans database:** This is the information system that is used by the center for credit administration to manage credit information for medium term loans. The data from this information system is pulled out from the production environment into a back-office environment regularly. The back-office environment is the backbone of the PSE.

**Client facility system:** In this information system all the facilities information of the client is administered.

**Client database:** In this system the client information such as name, address and KVK number are administered

**Risk system:** This is a system containing credit risk information of a client. Out of this information system, risk specific data such as Loss given default and probability default are gathered.

**RWA (Risk weighted assets):** This is a tool that contains the risk weight measure for each asset, which is the minimum amount of capital the bank is required to keep to cover the risk of specific asset.

**Funds transfer pricing system:** A system that contains the funds transfer pricing information. This is where the data regarding the internal funding rate of the contract is derived from.

**Several mapping tables:** These tables are used to determine industry sector of the obligor.

## Bibliography

- Bergman, Sten. 2001. CDO Evaluator Applies Correlation and Monte Carlo Simulation to the Art of Determining Portfolio Quality. Standard & Poors, November 12.
- Cespuedes, Juan Carlos Garcia. 2002. Credit Risk Modelling and Basel II. *Algo Research Quaterly* 5, no. 1.
- Gilkes, Kai. 2005. CDO Evaluator Version 3.0: Technical Document. Standard & Poors, December 19.
- Merriam-Webster. 2010. Monte carlo - Definition and More from the Free Merriam-Webster Dictionary. April 10. [http://www.merriam-webster.com/dictionary/monte carlo](http://www.merriam-webster.com/dictionary/monte%20carlo).
- Merton, Robers. 1974. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29, no. 2 (May): 449-470.
- Partnoy, Frank. 2006. SSRN-How and Why Credit Rating Agencies are Not Like Other Gatekeepers by Frank Partnoy. May. [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=900257&download=yes](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=900257&download=yes).
- Raeves, Gert. 2008. Bare necessity. *The Markit Magazine*.
- Schönbucher, Philipp J. 2000. Factor Models for Portfolio Credit Risk (December).
- Vasicek, Oldrich Alfons. 1991. Limiting Loan Loss Probability Distribution (August 9).
- . 2002. The Distribution of Loan Portfolio Value. *Risk* 15, no. 12 (December): 160-162. doi:10.1.1.139.5752.
- Vink, Dennis. 2009. SECURITISATIE: HOE NU VERDER? <http://www.dennisvink.nl/>.