Erasmus University Rotterdam

Erasmus School of Economics

Master Thesis

Econometrics and Management Science - Quantitative Finance

Reverse stress testing

A framework to identify stress scenarios

Author:
Bas van Oers (434545)

Supervised by:
Dick van Dijk (Erasmus University)
Steyn Verhoeven (Zanders)

Second assessor:
Chen Zhou (Erasmus University)

November 13, 2017

ABSTRACT

Reverse stress tests can be used to identify the scenarios that exhaust the capital buffer of a bank and cause the default of the bank. This report presents a framework to apply a quantitative reverse stress test on the portfolio of a bank. The framework includes the selection of appropriate risk factors, the identification of the distribution of these risk factors, including dependency, estimation of the influence of these risk factors on the occurrence of a default and a grid search to identify the set of stress scenarios. Applying this framework to a standardized portfolio leads to a set of stress scenarios, that can be used as early warning signs or as input for a traditional stress test. Furthermore, to decrease the computation time needed in the grid search used in this framework, the application of scenario reduction algorithms is analyzed. However, due to computational limits, it is not possible yet to reduce the scenario set in an efficient manner. Furthermore, the incorporation of higher frequency data for the risk factors is evaluated, using a Mixed Data Sampling (MIDAS) extension. As a result of assumptions made to prevent the scenarios set to increase even further, the more extreme scenarios are smoothed out, so no stress scenarios were identified using this extension.

Keywords: reverse stress testing, mixed data sampling, scenario reduction algorithm, scenario clustering, copula

TABLE OF CONTENTS

1. INTRODUCTION

As a result of the financial crisis that hit in the period from 2007 till 2009, stress testing became increasingly important for financial institutions. The financial crisis raised the question whether the stress testing models used before were sufficient, as also stated by the Basel Committee on Banking Supervision (2009). This is reflected by the fact that since the crisis, the rules surrounding stress testing are strengthened by the Committee of European Banking Supervision (2009) and the Financial Services Authority (2008, 2009). In stress testing, extreme scenarios are analyzed to check the influence on the viability of institutions. The results of this analysis are used, among others, as an indication for the amount of capital needed to cover losses caused by large shocks in the economy. Alternatively an institution can use the results of the stress test to take action to decrease the level of risk exposure. Although the choice of scenarios in stress testing is in general argued quite well, based on expert opinion or historical data, it always remains arbitrary. Even though also a subjective set of scenarios already gives information about the effect of changes in some economic factors, it could lead to missing an extreme, yet still plausible, scenario that could cause the inviability of an institution.

Following the regulations by the Committee of European Banking Supervision (CEBS) and the Financial Services Authority (FSA), large institutions are also required to perform a reverse stress test, in addition to traditional stress tests. In contrast to traditional stress testing, reverse stress testing specifically look for the so-called stress scenarios that cause the business plan of an institution to become unviable or scenarios where the capital requirements are not met. This way scenarios can be found that are not covered by the subjective set of stress scenarios used in standard stress testing, but can still cause an institution to become unviable. Additionally, reverse stress testing can help discover hidden and unexpected risks or relations between risk factors. When these stress scenarios are identified, the institution can use the information these scenarios provide to test whether their strategy is robust in a stressed economic environment. If not, they can adjust their strategy in such a way that the uncovered weaknesses of the current strategy are resolved. Moreover, the stress scenarios specify the economic environment in which the viability of the organization is at stake. Using this information, institutions can react quickly to changes in the economy if necessary. However, regulatory authorities do not specify how the reverse stress test should be performed. Instead, institutions have to show that their approach is appropriate. Furthermore, reverse stress testing needs to be implemented proportionally, so smaller institutions are allowed to use less complicated reverse stress tests, which can maybe be more qualitative than quantitative.

Reverse stress testing involves several challenges, which make it a nontrivial exercise. First of all, it is necessary to define the risk factors that will be investigated. When applying a traditional stress test, a few regulatory scenarios are evaluated, supplemented by some historical and narrative scenarios. In reverse stress testing, the risk factors first have to be selected and then these risk factors are used to create a set of scenarios. In this selection process various types of risk factors need to be considered, that can have influence on the viability of an institution. The most common risk types evaluated in stress testing models are market risk, credit risk and liquidity risk. Market risk can be further decomposed in interest rate risk, exchange rate risk and equity risk. In addition to these risk types also macroeconomic variables can be considered, because these variables partly represent the state of the economy and thus can affect the viability as well. Another important challenge in reverse stress testing is the fact that when the number of risk factors under investigation increases or the evaluated portfolio has a complex structure, discovering the stress scenarios can become computationally challenging. This is due to the fact that if more risk factors are investigated, the number of scenarios increases exponentially. For each scenario, the value of the portfolio has to be determined and the probability of occurrence has to be calculated, to assess the plausibility of the scenario. This means that the computation time of the reverse stress test increases sharply with each added risk factor, soon making it hardly possible to perform the test. This shows that there is a clear downside in adding risk factors to the framework, so the most efficient combination of risk factors needs to be found. This trade-off between extra information contained in a risk factor and the increasing dimensionality is one of the biggest issues regarding reverse stress testing. A third challenge in reverse stress testing is the limited availability of data. The framework makes use of default data, which is only available on an annual basis. Because of this, risk factor data also needs to be generated on yearly basis, which results in the loss of information contained in higher frequency data.

This report researches these reverse stress testing challenges and introduces methods to deal with them. By combining these methods, a framework for reverse stress testing can be created, which can be used to identify the stress scenarios and assess their plausibility. The existing literature on reverse stress testing is quite limited, due to the fact that institutions only adopted this risk measuring technique after the financial crisis. Because of this, an unified approach for reverse stress testing does not yet exist. In the past years different approaches are applied to construct a reverse stress test model. On of these approaches is the framework built by Grundke & Pliszka (2015), which is used as basis for this research. However, a few extensions to this framework are made to be able to deal with the challenges in reverse stress testing. Current literature only investigates models containing a few risk factors, so the computational challenge when the number of risk factors increases, is not fully experienced. When models containing more risk factors are used, the full extent of the dimensionality problem is experienced. The first extension applied in this research deals with both the selection process for the risk factors and this dimensionality problem. This extension is applying a specific-to-general selection approach, which selects the most informative risk factors for the model, while at the same time limiting the number of risk factors and thus the number of scenarios that has to be evaluated. Using this approach, in each step one risk factor will be added to the model, but only when the additional explanatory power is higher than the extra parameter uncertainty. This way only risk factors that add sufficient extra information will be added to the model. In this research this leads to a model containing four risk factors, representing the interest rate risk, equity risk and the macroeconomic environment. Including more risk factors leads to a sharp increase of the dimensionality, while it does not provide a lot of extra information.

The second extension in this research is a scenario reduction algorithm, as suggested by Gröwe-Kuska et al. (2003). This extension is another method to solve the dimensionality issue in reverse stress testing and can be applied after the risk factors are selected and a scenario set is created. Using the concept developed in Heitsch & Römisch (2003) and Dupačová et al. (2003), Gröwe-Kuska et al. (2003) introduce two algorithms: the so-called 'Simultaneous backward reduction' algorithm, which deletes one scenario in each step, and the 'fast forward selection' algorithm, which adds one scenario to the selection in each step. These algorithms only use the distribution of the input parameters of the scenarios in the reduction process. However, when two scenarios are close in terms of input parameters, the final output of these scenarios does not necessarily need to be close as well. To solve this, Li & Floudas (2014) extend these algorithms by including the output result in the algorithm, so also the distance between the output of the scenarios is considered. Unfortunately, this research shows that, due to computational limits, it is not yet possible to use these reduction algorithms to reduce the scenario set. These limits prevent the distances between the scenarios to be calculated efficiently, so applying the reduction algorithms does not reduce the computation time necessary for the reverse stress test.

A third extension also deals with the dimensionality problem, but then regarding the interpretability, by applying a clustering algorithm to set of stress scenarios resulting from the reverse stress test. Mirzai & Müller (2013) suggest to use a cluster analysis to search for common factors in the resulting set of stress scenarios, which also expands when the number of risk factors increases. They apply a k-means clustering algorithm, as developed by Witten et al. (2011), which creates a chosen number of clusters that can be analyzed individually. This way it is possible to identify a main driving factor for each cluster of stress scenarios. This extension addresses the dimensionality issue of the result of the reverse stress test, to increase the interpretability of the set of stress scenarios. Because of this, institutions are better able to incorporate these results in their strategy and decision-making process. This clustering algorithm leads to a number of cluster scenarios, summarizing the complete stress scenario set, that can be used as input for traditional stress testing.

A final extension applied in this research is using Mixed Data Sampling (MIDAS) to address the limited availability of data for the reverse stress testing framework. Due to the fact that default data is only available on an annual basis, the risk factors also need to be generated yearly, resulting in loss of information contained in the higher frequency data. Another issue is that this aggregation provides only a limited number of observations, so it is difficult to get significant coefficients in the model for the asset returns. To be able to capture the information contained in the risk factors better, an estimation method is necessary that can correlate data sampled at various frequencies. Using a distributed lag model, where all current and lagged values are used as regressors to explain the low frequency dependent variable, quickly leads to overfitting of the model. This problem can be solved by mixed frequency data, using the MIDAS extension, as introduced by Ghysels et al. (2004), to include quarterly values for the risk factors in the model. A big advantage is that MIDAS methods use a large part of the information contained in the data with a higher frequency, while at the same time the number of coefficients to estimate stays relatively small. This is done by using a weighting polynomial, that weights the quarterly values before using it as input for the model. Ghysels et al. (2007) extend this model by including the possibility to sample the explanatory variables at different frequencies. Foroni & Marcellino (2013) show that the MIDAS model is already used in quite some econometric applications, where different versions of the model are developed to make it applicable within these methods. Freitag (2014) specifically focuses on combining the MIDAS model with a Probit model, which can for instance be used to model the number of defaults. In this research this MIDAS-Probit model will be used in the framework of reverse stress testing, in an attempt to capture the information contained in the risk factors better. However, because not all quarterly values can be included in the scenario generation process, as this leads to an exploding number of scenarios, an autoregressive structure is needed for the quarterly structure. This way, only the first quarterly value is included in the scenario generation and the rest is estimated afterwards. This research shows that, due to this autoregressive structure and the weighting of the resulting quarterly values, the extreme effects in the scenarios are smoothed out. Because of this, the resulting risk factor values are not extreme enough to cause a default of a portfolio.

The framework presented by Grundke & Pliszka (2015) resulted from research on reverse stress testing, primarily conducted in the past ten years, following the financial crisis. However, an approach for reverse stress testing was first mentioned by Studer (1997), who presented the so-called worst case approach, later discussed by Čihák (2007). In this approach the plausibility of the scenario is given and depending on this plausibility, the scenario with the worst effect on the portfolio is identified. Here the plausibility regions are defined as regions with a given probability mass, which depends on the used risk factors. Breuer et al. (2009) state that this definition causes this worst effect to be dependent on the number of risk factors in the model. Thus also irrelevant risk factors or highly cross-correlated risk factors in the model have influence on this maximum loss, because they influence the plausibility of the scenario and assess whether it falls in the chosen plausibility region. They solve this problem by using regions where the distances are measured in standard deviations. This method is later extended to a multi-period version by Breuer et al. (2012). McNeil & Smith (2012) based on their framework to find the least solvent likely event (LSLE) on the worst case approach discussed by Studer. They came up with a different way to construct the set of plausible scenarios and defined the LSLE to be the scenario within this set with the worst effect on the portfolio.

Additional to the worst case approach, both Čihák (2007) and McNeil & Smith (2012) also described a so-called threshold approach, which searches for the scenario with the highest plausibility that reaches a given threshold, for instance on the portfolio value. A similar approach to this threshold approach is the bottom-up approach presented by Grundke (2010, 2011). The bottom-up approach models the influence of the different risk types on the organization individually using separate models for each risk type. Aggregation of the influences of these individual risk types leads to defining a stress scenario. Grundke states that this approach solves the inversion problem of reverse stress testing, while it also enables the calculation of scenario probabilities. Using this approach, it is possible to identify all scenarios breaching a threshold by applying a grid search. After this the probability of occurrence can be calculated for each of these scenarios, so the most plausible stress scenario can be found. Grundke (2012) expands this format by including a time-varying bank rating and contagion effects between single issuers. The framework of Grundke & Pliszka (2015) uses some ideas mentioned by Grundke (2011, 2012), but use a top-down approach instead of a bottom-up approach. In this top-down approach the influence of all the evaluated risk types is evaluated at the same time, using a model measuring the combined effect. Grundke & Pliszka identify the marginal distributions of the different risk factors and then link these using a copula to get the joint distribution. This way they are able to create scenarios and calculate the probability of each of these scenarios. These scenarios are then used to determine the combined effect of these risk types on the organization. Furthermore, Grundke & Pliszka use empirical data in their reverse stress testing model instead of just using simulating data. Finally, they partially deal with the dimensionality problem by using a principal components analysis on the term structure of the interest rate. This analysis aggregates the information in the term structure into only a few risk factors and therefore reduces the number of risk factors that has to be included in the model.

The remainder of this paper is structured as follows. In section 2 the reverse stress testing framework is discussed. Section 3 explains the methods to select the risk factors and identify the marginal and joint distributions of these risk factors, while section 4 presents the data used in this research. Section 5 shows the results of the calibration of the framework and the reverse stress tests following from it, using the model only based on yearly observations. Section 6 presents the results of the calibration and actual reverse stress test when the MIDAS extension is applied to the framework. Finally, section 7 concludes this research, while section 8 discusses the topics for further research.

2. REVERSE STRESS TESTING FRAMEWORK

To perform a reverse stress test on a portfolio of an institution it is important to first create the framework for this test. In this section the methods used in this framework are discussed. This framework contains methods to identify the stress scenarios and to choose the most appropriate risk factors. In section 2.1 the methods to define the value of a portfolio are discussed, followed in section 3.1 by the description of the methods to select the appropriate risk factors. Section 2.3 presents the methods used to identify stress scenarios based on the valuation of the portfolio. Finally, in section 2.4 the MIDAS model is discussed, along with the adjustments that need to be made to the model to apply this specification.

2.1 Portfolio valuation

In this research it is assumed that institutions have a stylized portfolio, only consisting of assets and liabilities structured as zero-coupon bonds, with notional 1 for all positions. Besides that, the institutions are assumed to be positive maturity-transforming, a strategy often applied by banks. This means that they have negative net cash flows in the short term and positive in the long term. A final assumption concerning the framework is that the term structure of the assets and liabilities does not change over time. Because of this, changes in the time to maturity are considered not to influence the value of the portfolio.

The stylized portfolio is assumed to consist of several asset positions $n = 1, ..., N$, issued by different obligors, and liability positions $m = 1, ..., M$. The portfolio is set up in such a way that the initial value of the liability positions and the initial capital buffer add up to the initial value of the asset positions. Following the capital rules set by CRD (2013), the total capital ratio should amount to at least 8% of all risk-weighted assets. In practice, the required capital ratio is equal to approximately 2.5% of the total asset value. This percentage will be used as required capital buffer in this report, as the simplification of the asset positions does not allow for a risk-weighted asset value to be calculated. In the initial situation, most banks often hold double the amount required, which in this case results in an initial capital buffer of approximately 5% of the total asset value. The remaining 95% of the value of the asset positions is equal to the initial value of the liability positions. The values of the asset and liability positions together give the total value of the portfolio $V_p(t)$. So at the risk horizon H the portfolio value can be calculated using

$$V_p(H) = \sum_{n=1}^{N} V_n(H) - \sum_{m=1}^{M} V_m(H), \tag{1}$$

where $V_n(H)$ is the value of the asset position issued by obligor n and $V_m(H)$ is the value of the liability position m, at the risk horizon H.

The value of a zero-coupon bond at future risk horizon H is assumed to be dependent on the credit rating of the obligor that issued the asset. This credit rating determines the credit spread, which is used together with the appropriate interest rate to calculate the position value. Because of the low number of defaults in some of the rating classes, the different ratings are aggregated into two broader classes, the Investment grade and the Speculative grade. The Investment grade contains all issuers with credit rating ranging from AAA till BBB, while the lower ratings, ranging from BB till C, are contained in the Speculative grade. Within these broader classes, the characteristics, like the credit spread, are assumed to be equal. This also implicates that issuers close to the threshold between these classes are not distinguished from issuers further from this threshold. However, due to the fact that for some credit ratings defaults are very uncommon, this simplification is necessary to be able to calibrate the framework.

Following the framework of Grundke & Pliszka (2015) the asset returns are assumed to be the driving factor for the credit quality of an obligor. Using this assumption, the value of an individual asset position in the portfolio can be calculated, based on its current credit rating i_n^0 and credit rating at the risk horizon i_n^H using

$$V_n(H) = I_{[R_n(H) > RD_i]}\ exp\{-q \cdot \big(r_q(H) + S_{i_n^H}\big)\} + \big(1 - I_{[R_n(H) > RD_i]}\big) P_r. \tag{2}$$

Here $R_n(H)$ is the return of the asset position issued by obligor n at the risk horizon H, RD_i is the rating dependent default threshold for the return and $S_{i_n^H}$ is the credit spread, depending on the credit rating of the obligor at the risk horizon. The end-of-year interest rate at the risk horizon, with the appropriate maturity q, is given by $r_q(H)$. It is assumed that in case of a default of an issuer, when the indicator function $I_{[R_n(H)>RD_i]}$ is equal to zero, the bank receives a recovery payment P_r. This recovery payment can be determined using the recovery of treasury assumption as mentioned by Huang & Huang (2012):

$$P_r = \delta_n V_{rf}, \tag{3}$$

where V_{rf} is the value of the risk-free zero coupon bond with the same maturity as the risky zero coupon bond. This value can calculated using equation (2), without the possibility of default and without credit spread. The recovery fraction δ_n is assumed to be beta-distributed, which guarantees values between 0 and 1. This fraction does not depend on the credit rating of the obligor.

The value of a liability position in the portfolio can be calculated using a similar structure as in equation (2). It is assumed that the credit rating of the institution belongs to the investment grade and will not change until the risk horizon. However, the value of a liability position still depends on the scenario, as the scenario determines the interest rate curve used to value the liability. So, the value of a liability position can be generated using

$$V_m(H) = exp\{-q \cdot \big(r_q(H) + S_{bank}\big)\}, \tag{4}$$

where S_{bank} is the credit spread for the credit rating of the institution, which is assumed to belong to the investment grade rating class.

As the asset return is the driving factor for the credit rating of the obligor, it is necessary to model the asset returns at the risk horizon. It is assumed that these asset returns are affected by a number of risk factors, so a linear model is created to model the asset returns. For each of the two rating classes a separate model is formulated, as the effects of the risk factors are expected to differ between the rating classes. Investment grade rated obligors are expected to be more stable than speculative grade rated obligors and thus less sensitive to various risk factors. The linear factor model for the asset return of obligor n with current rating grade i_n^0 is then given by

$$R_n(t) = \beta_{i_n^0,1} X_1(t) + ... + \beta_{i_n^0,K} X_K(t) + \sigma_{i_n^0} \epsilon_n(t). \tag{5}$$

Here $X_1(t), ..., X_K(t)$ represent the risk factors and $\epsilon_n(t)$ denotes the idiosyncratic risk for obligor n. The idiosyncratic risk contains the risk factors not included in the model, so the influence of this error depends on the used risk factors. This idiosyncratic risk is assumed to be normally distributed with mean zero and variance $\sigma^2_{i_n^0}$, which does not depend on the obligor itself, but only on its rating. The parameters $\beta_{i_n^0,1}, ..., \beta_{i_n^0,K}$ give the sensitivity of the asset return with respect to the specific risk factor, conditional on the credit rating of the obligor.

As the asset return is the driving factor for the credit quality of an issuer, it is important to define the influence of the asset return on the credit rating. The credit rating can be determined by comparing the asset return with the transition thresholds, which are based on migration probabilities for the two rating classes. These thresholds are used to assess whether an issuer moves to a different rating class or goes into default. The transition thresholds also depend on the initial credit rating of the issuer. To determine these thresholds a Monte Carlo simulation is applied for the asset returns, which simulates 100,000 draws for the risk factors. For each of these draws the asset return is calculated, for both rating classes. This results in an empirical distribution, which is used to determine the thresholds, in combination with the migration probabilities. The default threshold is then equal to the quantile of the empirical distribution equal to the default probability of the specific rating class. To determine the threshold for transition between the rating classes, the cumulative probabilities are used to identify which quantile needs to be used.

Applying these models and definitions, it is possible to define a stress scenario. The final step is to get a definition of a default of an institution. As stated earlier, the initial buffer B_0 of the institution is equal to approximately 5% of the initial value of the asset positions. A default is defined as a scenario where the capital buffer falls below the required 2.5% of the total asset value at the risk horizon. To identify these so-called stress scenarios the approach of Grundke & Pliszka (2015) is used. This way the definition of the subset $\mathbf{\Omega}^*$, containing the stress scenarios, is given by

$$\mathbf{\Omega}^* = \{\omega \in \mathbf{\Omega} | E[V_p(H)] - q_{1-\alpha}(V_p(H)|\omega) = B_0 - B_H\}. \tag{6}$$

Here $\mathbf{\Omega}$ gives the total set of scenarios, $E[V_p(H)]$ is the unconditional expectation of the portfolio value and B_H is the required capital buffer at the time horizon, equal to 2.5% of the asset value. Furthermore, $q_{1-\alpha}(V_p(H)|\omega)$ is the $(1-\alpha)$-percentile of the probability distribution, conditional on scenario ω. In this definition the difference between the unconditional expectation and the $(1-\alpha)$-percentile is defined as the portfolio loss in the specific scenario. In case of a stress scenario this portfolio loss is equal to the difference between the initial buffer and the required buffer. An interval is created around this difference, so the scenarios are identified that just cause a default and not those with an extreme level of stress. The unconditional expectation of the portfolio value can be calculated using the 100,000 Monte Carlo simulations of the risk factors. By generating the portfolio value for each of these simulations, the unconditional expected value can be calculated. For each scenario ω the $(1-\alpha)$-percentile of the conditional probability distribution can be calculated using a Monte Carlo simulation, where 1000 observations are drawn for the portfolio value. The $(1-\alpha)$-percentile is used instead of the mean, because then the scenario is evaluated in the circumstances where also the risk factors not included in the model have a negative influence on the portfolio value. These risk factors are contained in the error term in the model for the asset returns, as given by equation (5). Because the uncertainty is present in the asset return model, 1000 draws for each asset positions are necessary for the Monte Carlo simulation.

Finally, the most plausible stress scenario ω^* in the subset $\mathbf{\Omega}^*$ can be found using

$$\omega^* = \arg\max_{\omega \in \mathbf{\Omega}^*} P(\omega), \tag{7}$$

where $P(\omega)$ is the probability of having all risk factors in the neighborhood of scenario ω. $P(\omega)$ will be defined in section 2.3. The described framework shows that there is a trade-off between how extreme a scenario is and its plausibility. By only considering these scenarios that just decrease the initial capital buffer below the required buffer, extreme and improbable scenarios are excluded.

2.2 Model estimation

When the framework for the reverse stress test is created, the model for the asset returns, as presented in equation (5), needs to be estimated. When this model is estimated based on historical values of the asset returns, a default of an obligor results in missing values for this asset return in the years after the default. To solve this problem, the asset return is marked as a latent variable, only visible when no default occurs, while the number of defaults in each year is used as observed value. The probability of a default is equal to the probability that the asset return falls beneath this threshold, as defined in section 2.1. To estimate the parameters for the asset return model, based on the number of defaults in each year, the maximum likelihood approach is used, where the log-likelihood is based on a binomial distribution. This log-likelihood, which depends on the initial rating of the obligor, is given by

$$\begin{aligned} l_i = \sum_{t=1}^{T} log\Bigg\{ \binom{N_i(t)}{d_i(t)} P\Big(R_n(t) \leq RD_i | X_1(t) = x_1, ..., X_K(t) = x_K\Big)^{d_i(t)} \\ \cdot \Big(1 - P\Big(R_n(t) \leq RD_i | X_1(t) = x_1, ..., X_K(t) = x_K\Big)\Big)^{N_i(t) - d_i(t)} \Bigg\}, \end{aligned} \tag{8}$$

where

$$P\Big(R_n(t) \leq RD_i | X_1(t) = x_1(t), ..., X_K(t) = x_K(t)\Big) = \Phi\Big(\frac{RD_i - \beta_{i,1}x_1(t) + ... + \beta_{i,K}x_K(t)}{\sigma_i}\Big). \tag{9}$$

Here $N_i(t)$ is the total number of issuers, $d_i(t)$ is the number of defaults for each rating group i and $\Phi(\cdot)$ denotes the CDF of the standard normal distribution. The optimization algorithm chooses the parameters in such a way that the likelihood of the empirical fraction of defaults in each year is maximized. This format of the log-likelihood allows for different factor sensitivities for the various initial rating grades. The log-likelihood can be calculated as stated in equation (8), because conditional on specific values of the risk factors, the defaults are independent.

2.3 Identifying stress scenarios

After selecting the appropriate risk factors, the stress scenarios can be identified and the optimization problem, used to find the most plausible stress scenario, can be solved. This can be done by applying a grid search for the selected risk factors. For each of the risk factors the sample mean μ_k and the sample standard deviation σ_k are calculated and using these characteristics the interval $[\mu_k - 4\sigma_k, \mu_k + 4\sigma_k]$ is defined. In this interval 17 equally spaced grid points x_k are obtained, with a step size of $0.5\sigma_k$. By choosing a finer grid the results of the reverse stress test will be more accurate, but the computational effort will increase sharply. Because the probability of a specific value is zero, an interval is created around each of these grid points, for all risk factors $k = 1, ..., K$: $[x_k - 0.25\sigma_k, x_k + 0.25\sigma_k]$. The probability of having all risk factors in the neighborhood of a specific scenario $\omega = (x_1, ..., x_K)$, within the used scenario set, can then be calculated using

$$P(\omega) = P\big(x_1 - 0.25\sigma_1 < X_1 \le x_1 + 0.25\sigma_1, ..., x_K - 0.25\sigma_K < X_K \le x_K + 0.25\sigma_k\big). \tag{10}$$

It should be noted that this probability does not give the real life plausibility of the scenario, but only the probability in the model, based on the chosen risk factors. This probability $P(\omega)$ for such a K-dimensional rectangle can be written like

$$P(\omega) = \sum_{i_1=1}^{2} \cdots \sum_{i_K=1}^{2} (-1)^{i_1+\cdots+i_K} F(x_{1,i_1}, ..., x_{K,i_K}), \tag{11}$$

where $F(x_{1,i_1}, ..., x_{K,i_K})$ is the joint distribution of the risk factors. In section 3.2 the determination of this joint distribution is explained. For all $k = 1, ..., K$ risk factors the values $x_{k,1} = x_k - 0.25\sigma_k$ and $x_{k,2} = x_k + 0.25\sigma_k$ are the lower and upper bound for the respective grid point. So, equation (11) gives the probability that the realizations of each of the risk factors lie within the interval of the corresponding grid point. For each of the 17^K scenarios ω the portfolio evaluation, using equation (1), (2) and (5), will be performed. By doing so, it can be determined whether the scenario should be included in the set of stress scenarios, using equation (6).

2.3.1 Scenario reduction

When the number of risk factors increases, the number of scenarios that needs to be evaluated in the grid search increases exponentially. To deal with this problem of dimensionality Heitsch & Römisch (2003) propose a method of scenario reduction. This method uses the Kantorovich distance D to address the goodness-of-fit of the approximation of the scenario space by the reduced set of scenarios. This distance measures the probability distance between the original set of scenarios and the selected subset. To get the optimal subset of scenarios the following statement should be minimized, depending on the chosen size of the reduced scenario set:

$$D = \sum_{s \in J} p_s \min_{j \notin J} D_{sj}, \tag{12}$$

where $D_{sj} = |\omega_s - \omega_j|$ is the distance between the scenarios s and j, using a norm with the dimension equal to the number of risk factors. So, $\min\limits_{j \notin J} D_{sj}$ is the distance from scenario s to the closest scenario still in the set. Furthermore, J is the index set of deleted scenarios and p_s denotes the probability of all risk factors being in the neighborhood of scenario ω_s, which can be calculated using equation (11).

Heitsch & Römisch (2003) suggest two algorithms to minimize this statement and perform the scenario reduction. For both algorithms first of all the distances D_{sj} between the different scenarios are calculated. The first algorithm they propose is a backwards reduction algorithm, which, in each step, deletes one scenario from the original set, based on the probability of the scenarios and the distance between the scenarios. Each of these steps i contains the same optimization procedure, following algorithm 1. To decide which scenario can best be deleted from the set, the new Kantorovich distance, as given in equation (12), is calculated for each of the options. Thus, for each scenario in the reduced set of step i, the Kantorovich distance is calculated in case this scenario is deleted. The scenario with the smallest value for the Kantorovich distance is the one that will be deleted. Algorithm 1 gives the calculations that need to be made to define the optimal reduced set.

Algorithm 1:

1. For each scenario $l \notin J^{i-1}$:
   (a) For each scenario $s \in J^{i-1} \cup \{l\}$, find the distance to the closest scenario, which is still in the subset, when scenario l is removed: $D^i_{sl} = \min\limits_{j \notin J^{i-1} \cup \{l\}} D_{sj}$.
   (b) Compute the Kantorovich distance when scenario l is removed: $z^i_l = \sum_{s \in J^{i-1} \cup \{l\}} p_s D^i_{sl}$.
2. Choose the scenario l for which the Kantorovich distance is the smallest: $j^i \in \arg\min\limits_{l \notin J^{i-1}} z^i_l$.
3. Add this scenario to the set of deleted scenarios: $J^i := J^{i-1} \cup \{j^i\}$.

This way the scenario is deleted that is closest to the scenarios with the highest probability. This optimization procedure is repeated until the size of the reduced scenario set reaches the chosen size. The new probability of having all risk factors in the neighborhood of scenario j, p^*_j, can be calculated using

$$p^*_j = p_j + \sum_{s \in J(j)} p_s, \tag{13}$$

where $J(j) = \{s \in J : j = j(s)\}$ and $j(s) \in \arg\min\limits_{j \notin J} D_{sj}$. Thus $J(j)$ denotes the subset of deleted scenarios that have scenario ω_j as closest scenarios still in the scenario set. So, the new probability is equal to the old probability summed with the probabilities of the closest deleted scenarios.

The second algorithm is a forward selection algorithm, where in each step one scenario is added to the subset. In that case each step i requires the optimization procedure as given by algorithm 2. Also in this algorithm the Kantorovich distance is minimized in each step, but here one scenario is added to the subset in each iteration. For each possibility, so each scenario that can be added, the new Kantorovich distance needs to be calculated, to choose the optimal scenario to be added. To update the Kantorovich distance, the distance to the closest scenario in the subset has to be identified for the remaining scenarios in the deleted set. This can be done by checking whether the scenario that was added in the previous step is closer than the scenario that were already in the reduced set. Finally, the scenario that is added is the one that leads to the highest decrease of the Kantorovich distance.

Algorithm 2:

1. For each scenario $l \in J^{i-1}$:
   (a) For each scenario $s \in J^{i-1}$ update the distance to the closest scenario in the reduced scenario set, when scenario l is added to the set: $D^i_{sl} = min\{D^{i-1}_{sl}, D^{i-1}_{sj^{i-1}}\}$.
   (b) Compute the Kantorovich distance when scenario l is added: $z^i_l = \sum_{s \in J^{i-1} \setminus \{l\}} p_s D^i_{sl}$.
2. Choose the scenario l for which the Kantorovich distance is the smallest: $j^i \in \arg\min\limits_{l \in J^{i-1}} z^i_l$.
3. Add this scenario to the reduced scenario set: $J^i := J^{i-1} \setminus \{j^i\}$.

In the first step D^1_{sl} is equal to D_{sl}, as there is no scenario added to the reduced set in the previous step. This algorithm is repeated until the required number of scenarios is reached and then the new scenario probabilities can be calculated again with equation (13).

A disadvantage of these reduction algorithms is the fact that the resulting portfolio value in a given scenario is not included in the reduction process. When two scenarios are close, based on the value of the risk factors in these scenarios, the distance based on the portfolio value in these scenarios does not necessarily need to be small as well. By including the output of the scenario, the number of scenarios that does not lead to a default can possibly be reduced. However, when using this approach is when two different scenarios lead to stress, and thus give a similar output, both scenarios need to stay in the reduced set. When including the output in the reduction algorithm, the risk arises that one of these scenarios is removed, although this is not very likely when the input factors differ significantly. This is an issue that needs to be monitored when applying these scenario reduction algorithms, by evaluating the set of stress scenarios following from the reduced scenario set.

Li & Floudas (2014) suggest an approach to include the output of a scenario in the reduction algorithm. Because the calculation of the portfolio value, the final output of a scenario evaluation in this research, requires a lot of computational effort, the asset return is used as output instead. By considering the asset return already a good indication of the output value can be achieved. To decrease the computation time of the reduction algorithm even more only the asset return for the Investment grade will be investigated for each scenario. The minimization statement to get the optimal subset of scenarios is than equal to

$$D = \sum_{s \in J} p_s \min_{j \notin J} D_{sj} + R^{\text{error}}_{max} + R^{\text{error}}_{min} + R^{\text{error}}_{exp}, \tag{14}$$

where

$$\begin{aligned} D_{sj} &= |\omega_s - \omega_j| + |R_s - R_j| \\ R^{\text{error}}_{max} &= R^{full}_{max} - R^{new}_{max} \\ R^{\text{error}}_{min} &= R^{new}_{min} - R^{full}_{min} \\ R^{\text{error}}_{exp} &= |R^{full}_{exp} - R^{new}_{exp}|. \end{aligned}$$

The maximum, minimum and expected value of the asset return, among the scenarios, in the complete scenario set are given by R^{full}_{max}, R^{full}_{min} and R^{full}_{exp}, while in the reduced set these are equal to R^{new}_{max}, R^{new}_{min} and R^{new}_{exp}. Including these terms in the minimization statement ensures that the empirical distribution of the output of reduced set is similar to the empirical distribution when the full set is used. As can be seen the calculation of the distance between the scenarios s and j also changed compared to algorithm 1 and 2, because now also the difference in value of the asset return in the specific scenarios is taken into account.

Also in this case a backwards reduction algorithm and a forward selection algorithm can be used to perform the scenario reduction. For both of these algorithms first of all the maximum, minimum and expected value of the asset return in the complete set need to be calculated, before the algorithm can be used. Besides that, the distances between the different scenarios in the complete set need to be calculated. In the backwards reduction algorithm each step i then contains the same calculations, which are given in algorithm 3. Just as in algorithm 1, the Kantorovich distance is calculated for each scenario, in case it is removed from the set. However, the Kantorovich distance here also uses the minimum, maximum and expected value of the asset return, as showed in equation 14. Again, the scenario with the smallest value for the Kantorovich distance will be removed from the subset.

Algorithm 3:

1. For each scenario $l \notin J^{i-1}$:

   (a) For each scenario $s \in J^{i-1} \cup \{l\}$, find the distance to the closest scenario, which is still in the subset, when scenario l is removed: $D^i_{sl} = \min\limits_{j \notin J^{i-1} \cup \{l\}} D_{sj}$.

   (b) Compute the maximum, minimum and expectation of the asset return in the reduced scenario set, when scenario l is removed:
   $R^i_{max}(l) = \max\limits_{j \notin J^{i-1} \cup \{l\}} R_j$, $R^i_{min}(l) = \min\limits_{j \notin J^{i-1} \cup \{l\}} R_j$ and $R^i_{exp}(l) = \sum_{j \notin J^{i-1} \cup \{l\}} p_j R_j$.

   (c) Compute the Kantorovich distance when scenario l is removed, where the maximum, minimum and expected value of the asset returns are compared to the previous step:
   $z^i_l = \sum_{s \in J^{i-1} \cup \{l\}} p_s D^i_{sl} + \big(R^{i-1}_{max}(j^{i-1}) - R^i_{max}(l)\big) + \big(R^i_{min}(l) - R^{i-1}_{min}(j^{i-1})\big) + \big|R^{i-1}_{exp}(j^{i-1}) - R^i_{exp}(l)\big|$.

2. Choose the scenario l for which the Kantorovich distance is the smallest: $j^i \in \arg\min\limits_{l \notin J^{i-1}} z^i_l$.

3. Add this scenario to the set of deleted scenarios: $J^i := J^{i-1} \cup \{j^i\}$.

The forward selection algorithm, where in each step a scenario is added to the subset, is based on algorithm 2. In that case each step i requires the optimization procedure as given by algorithm 4. The main difference is that now the distribution of the output, in this case the asset returns, also has influence on the Kantorovich distance. Once again, in each step the scenario that leads to the highest decrease of the Kantorovich distance is added to the reduced set.

Algorithm 4:

1. For each scenario $l \in J^{i-1}$:

   (a) For each scenario $s \in J^{i-1}$, update the distance to the closest scenario in the reduced scenario set, when scenario l is added to the set: $D^i_{sl} = min\{D^{i-1}_{sl}, D^{i-1}_{sj^{i-1}}\}$.

   (b) Compute the maximum, minimum and expectation of the asset return in the reduced scenario set, when scenario l is added to the reduced set:
   $R^i_{max}(l) = \max\limits_{j \notin J^{i-1} \setminus \{l\}} R_j$, $R^i_{min}(l) = \min\limits_{j \notin J^{i-1} \setminus \{l\}} R_j$ and $R^i_{exp}(l) = \sum_{j \notin J^{i-1} \setminus \{l\}} p_j R_j$.

   (c) Compute the Kantorovich distance when scenario l is removed, where the maximum, minimum and expected value of the asset returns are compared to the full set:
   $z^i_l = \sum_{s \in J^{i-1} \setminus \{l\}} p_s D^i_{sl} + \big(R^{full}_{max} - R^i_{max}(l)\big) + \big(R^i_{min}(l) - R^{full}_{min}\big) + \big|R^{full}_{exp} - R^i_{exp}(l)\big|$.

2. Choose the scenario l for which the Kantorovich distance is the smallest: $j^i \in \arg\min\limits_{l \in J^{i-1}} z^i_l$.

3. Add this scenario to the reduced scenario set: $J^i := J^{i-1} \setminus \{j^i\}$.

In this algorithm in each step the scenario for which the new reduced set is closest to the full set, is added to the reduced set. This distance will be measured in terms of distribution of the input parameters and the minimum, maximum and expected value of the full set and the reduced set.

For algorithm 3 and 4 still equation (13) can be applied to calculate the new scenario probabilities for the scenarios in the reduced set. In this report all algorithms will be used to reduce the number of scenarios, so they can be compared based on the approximation of the initial scenario set and computation time.

Using the grid search, based on the scenarios left in the reduced scenario set, all the stress scenarios can be identified. The next step is then to find the most plausible scenario in this subset Ω^*. This can be done using the new probability of all risk factors being in the neighborhood of the scenario, assigned to each scenario in Ω^*, using equation (13).

2.3.2 Scenario clustering

To be able to interpret the set of stress scenarios, identified by the grid search, that cause a default, it would be useful to cluster the scenarios, as suggested by Mirzai & Müller (2013). Using a clustering algorithm summarizes the set of stress scenarios into a few scenarios, by searching for patterns in this set. Mirzai & Müller use a k-means clustering algorithm as described by Witten et al. (2011), to find a driving factor for each of these clusters. The following algorithm is used to identify the chosen number of clusters.

Scenario clustering algorithm:

1. Choose k random points in the scenario space as cluster center. These cluster centers do not necessarily need to be scenarios in the set of stress scenarios.
2. Assign the remaining scenarios to the closest cluster center, using a norm with appropriate dimension.
3. Calculate the mean of each of the clusters and use these means as new cluster centers.
4. Repeat step 2 and 3 for the new cluster centers and iterate until the scenarios are assigned to the same cluster in consecutive rounds.

2.4 MIDAS specification

Due to the fact that data regarding the number of defaults is only available on an annual basis, the model for the asset returns, given by equation (5), is estimated using yearly data. However, this estimation can possibly be improved by including explanatory variables into the model in a higher frequency, which ensures that more information contained in the risk factors is captured by the model. This can be done using the mixed data sampling (MIDAS) method of Ghysels et al. (2004). By applying this approach to the linear model for the asset returns, the effect of the risk factors at a higher frequency can be evaluated. This way not only the movement over the complete past year can be used in the model, but also the movements within sections of this past year. In this research the risk factor data will be used on a quarterly basis in this extension, as this is the highest frequency at which all the risk factor data can be generated. Using this approach equation (5) can be rewritten as

$$R_n(t) = \beta_{i_n^0,1} X_1^{(m)}(t,\theta) + ... + \beta_{i_n^0,K} X_K^{(m)}(t,\theta) + \sigma_{i_n^0}\epsilon_n(t), \tag{15}$$

where

$$X_k^{(m)}(t,\theta) = \sum_{j=0}^{m-1} w_j(\theta) L^{j/m} X_k^{(m)}(t). \tag{16}$$

Here, m is the ratio between the frequency of the explanatory variables and the frequency of the dependent variable, L is the lag-operator and $w_j(\theta)$ is a polynomial function used to weight the regressor data. In this research $m = 4$, because the explanatory variables will be generated on a quarterly basis. As can be seen in equation (16) the values of the risk factors on a quarterly basis will be weighted before they are included in the linear model for the asset returns. This way the influence of this risk factor will still be measured by only one coefficient, so extra information is included in the model, while it is still rather parsimonious. However, to determine the weighting of each of the lagged terms a polynomial function needs to be estimated. The most commonly used polynomial function is the Almon lag polynomial (Almon, 1965), given by

$$w_j(\theta) = \frac{e^{(j+1)\theta_1 + ... + (j+1)^Q \theta_Q}}{\sum_{j=0}^{m-1} e^{(j+1)\theta_1 + ... + (j+1)^Q \theta_Q}}. \tag{17}$$

In this research the two-parameter case is used, so the number of extra parameters that has to be estimated in this function is limited. Choosing more parameters in this case results in estimating the same number of parameters as including all the lagged terms in the model individually. This two-parameter case is given by

$$w_j(\theta) = \frac{e^{(j+1)\theta_1 + (j+1)^2 \theta_2}}{\sum_{j=0}^{m-1} e^{(j+1)\theta_1 + (j+1)^2 \theta_2}}. \tag{18}$$

To limit the number of parameters that needs to be estimated even further, it is assumed that all risk factors have the same weighting polynomial. This way only two extra parameters have to be estimated, compared to original model using yearly observations.

A big advantage of the MIDAS model is that only two assumptions are necessary in order to estimate the model, as showed by Andreou et al. (2010). First of all the error term $\epsilon_n(t)$ needs to be white noise. Besides that the weighting for each of the lagged terms of the risk factors needs to satisfy two conditions:

$$0 \leq w_j(\theta) \leq 1 \qquad \text{and} \qquad \sum_{j=0}^{q} w_j(\theta) = 1. \tag{19}$$

The MIDAS model can be estimated by integrating the MIDAS structure in the log-likelihood function in equation (8), which then transforms into

$$\begin{aligned} l_i = \sum_{t=1}^{T} log \binom{N_i(t)}{d_i(t)} P\Big(R_n(t) \leq RD_i | X_1^{(m)}(t,\theta) = x_1, ..., X_K^{(m)}(t,\theta) = x_K\Big)^{d_i(t)} \\ \cdot \Big(1 - P\Big(R_n(t) \leq RD_i | X_1^{(m)}(t,\theta) = x_1, ..., X_K^{(m)}(t,\theta) = x_K\Big)\Big)^{N_i(t)-d_i(t)}, \end{aligned} \tag{20}$$

where $X_k^{(m)}(t,\theta)$ can be calculated using equation (16) for each risk factor $k = 1, ..., K$.

When the MIDAS specification is used in the linear factor model for the asset returns, as given by equation (15), a few adjustments need to be made to the identification of stress scenarios. In case of the MIDAS specification, the risk factors each consist of four quarterly values instead of one yearly value. To limit the number of scenarios that needs to be evaluated, only the first quarterly value for each risk factor will be included in the scenario generation process. The other three quarterly values will then follow using an AR(1) model for each of the risk factors:

$$X_{k,t+1q} = \alpha_k + \phi_k X_{k,t} + u_{t+1q}, \tag{21}$$

where α_k is the constant and ϕ_k is autoregressive sensitivity coefficient. The error term u_{t+1q} follows the marginal distribution of the specific risk factor, with an adjusted mean of zero. The parameters of this model can be estimated based on historical observations, using maximum likelihood. Using this model four quarterly values are generated for each risk factor in all scenarios. Using equation (16) these quarterly values can then be included in the model.

The MIDAS specification also influences the calculation of the scenario probabilities, because now for each quarter the probability of the combination of risk factors needs to be calculated. In each quarter the dependency between the risk factors is present and needs to be taken into account. In addition, for the last three quarters also the dependency over time needs to be taken into account, following the AR(1) structure. This dependency can be captured by calculating the probability of these quarters conditional on prior observations. This can be done by using the preceding value as mean in the marginal distribution. For each of the quarters equation (11) can be used to calculate the joint probability of the risk factors. The product of the probability of the four quarters gives the total joint probability of all quarterly values. This results in the following calculation for the probability of having all quarterly risk factor values in the neighborhood of the scenario:

$$P(\omega) = \sum_{j=0}^{3} \sum_{i_1=1}^{2} \cdots \sum_{i_K=1}^{2} (-1)^{i_1+\cdots+i_K} F(x_{1,i_1}(t-jq), ..., x_{K,i_K}(t-jq)), \tag{22}$$

where $x_{k,1} = x_k - 0.25\sigma_k$ and $x_{k,2} = x_k + 0.25\sigma_k$ are still the lower and upper bound for the respective risk factor value, for each $k = 1, ..., K$. The joint distribution $F(x_{1,i_1}(t-jq), ..., x_{K,i_K})(t-jq)$ differs a little for $j = 1, 2, 3$, as here the mean of the marginal distribution is different.

3. SELECTING RISK FACTORS

In this section the procedure used to select risk factors and define the joint distribution of these risk factors is described. Section 3.1 describes the methods used to select the most informative risk factors, while dealing with the dimensionality problem arising when to many risk factors are incorporated in the model. In section 3.2 the process to identify and estimate the joint distribution of the selected risk factors is explained.

3.1 Selection methods

As mentioned before, there are different categories of risk factors that might be useful in this context. Credit risk, interest rate risk and equity are commonly used as risk types, but also exchange rate risk or a macroeconomic variable could be a source of risk. However, adding a lot of risk factors to the model sharply increases the computational effort. This computational effort mainly concerns the exponential increase of the scenario set that needs to be evaluated in the grid search when an additional risk factor is included. Furthermore, adding a risk factor to the model increases the parameter uncertainty in the model. Thus, there exists a trade-off between the extra information contained in a risk factor and the increase of the dimensionality of the scenario set and the parameter uncertainty caused by adding this factor to the model. For this reason, only interest rate risk, equity risk and macroeconomic variables are considered as risk factors in this report. For each of these risk types, at least one risk factor will be selected to be incorporated in the model.

The interest rate risk can be represented by the term structure of the interest rate. To reduce the number of risk factors added to the model, a principal component analysis is used for the interest rate term structure. This method is often used to summarize the term structure of the end-of-year interest rate over time, because the resulting principal components have a clear economic interpretation. The principal component analysis applies an orthogonal linear transformation to reduce the dimensionality. The first principal component is the linear combination of the interest rate for different maturities that has the maximum variance. This way it contains as much variation present in the data as possible. To guarantee an unique solution, the sum of squared coefficients should be equal to one. So, the coefficients $\mathbf{c_1}$ for the first principal component C_1 can be found using

$$\mathbf{c_1} = \arg\max_{\mathbf{c_1}} \{var(C_1) = \mathbf{c_1'\Sigma\ c_1}\} \qquad \text{subject to } \mathbf{c_1}'\mathbf{c_1} = 1, \tag{23}$$

where Σ is the covariance matrix of the components of the term structure. For the next principal components the same function is optimized, but in each step an additional constraint is added to the problem. This constraint ensures that the new principal components are uncorrelated with the other principal components. Thus the coefficients $\mathbf{c_j}$ for the next principal component C_j can be found using

$$\begin{aligned} \mathbf{c_j} = \arg\max_{\mathbf{c_j}} \{var(C_j) = \mathbf{c_j'\Sigma\ c_j}\} \qquad \text{subject to } \mathbf{c_j}'\mathbf{c_j} &= 1 \\ \mathbf{c_j'\Sigma c_1} &= 0 \\ &\vdots \\ \mathbf{c_j'\Sigma c_{j-1}} &= 0. \end{aligned} \tag{24}$$

The interest rates for the different maturities, $r_q(t)$, with maturity q, can be calculated again from the Z principal components using

$$r_q(t) = \sum_{j=1}^{Z} c_{j,q} C_j(t). \tag{25}$$

This way it is possible to get the term structure of the interest rates back from the principal components. Because not all the principal components will be included in the stress testing framework, the generated term structure will be an approximation. The precision of this approximation depends on the percentage of the variance explained by the principal components included in the model. Concluding, the principal component analysis summarizes the term structure of the interest rate, so the interest rate risk can be included in the framework in an efficient manner.

When the interest rate risk is included in the model in a parsimonious manner, also a measure for equity risk and the appropriate macroeconomic variables need to be chosen. Also here a trade-off exists between the information added by an additional variable and the increased dimensionality following from this addition. To choose an efficient position in this trade-off, the specific-to-general approach is applied. In this approach the model for the asset returns, as given by equation (5), is estimated for combinations of the measures for equity risk and of the macroeconomic variables. Equity risk can be measured by the Dow Jones index or the S&P 500 index. As macroeconomic variables the GDP, the house price index, the CPI and the unemployment rate are considered. These variables assess different aspects of the economic environment that might contain information about the asset returns. In the first step of the specific-to-general approach one equity risk measure and one macroeconomic variable are added to the model already containing the selected principal components. The different combinations of macroeconomic variable and equity risk measure are compared to select the best performing model. To select the best performing model, two aspects are considered. First of all, the significance of the estimated coefficients is compared between the different models. In addition, also the goodness-of-fit of the different models is evaluated using the AIC and BIC. However, only evaluating the significance and goodness-of-fit of the models can lead to a model that contradicts the existing economic theory. So, to assure that the coefficients in the asset return model do not contradict the economic theory, also the economic interpretation of these coefficients is assessed. By combining these criteria, a model can be formulated that be used to estimate the return for a position in the portfolio.

3.2 Distribution of risk factors

When the appropriate risk factors are selected, the joint distribution of these risk factors is necessary to assess the plausibility of the scenarios and to be able to perform simulations. Furthermore, in case of the MIDAS specification, the joint distribution of the risk factors is used to simulate the quarterly values for each scenario. In this report the joint distribution is determined using a copula function, which also takes the dependency between the risk factors into account. As a result, the joint distribution of the risk factors can be estimated as

$$\hat{F}(x_{1t}, ..., x_{Kt}) = C\big(\hat{F}_1(x_{1t}), ..., \hat{F}_K(x_{Kt})\big), \tag{26}$$

where $\hat{F}_1(.), ..., \hat{F}_K(.)$ denote the estimated marginal distributions of the risk factors.

Thus, to identify the most appropriate joint distribution, first the marginal distribution of each of the risk factors needs to be determined. Section 3.2.1 describes the methods to identify the marginal distribution of each risk factor. In section 3.2.2 the approach to select the most appropriate copula to combine these marginal distributions.

3.2.1 Marginal distributions

For simplicity the normal distribution will be considered as first possible marginal distribution for the risk factors. This assumption will first be checked by a visual inspection, using a QQ plot. By visualizing the empirical data, it is possible to see whether the empirical distribution distinguishes from a normal distribution. If this is the case, a QQ-plot also shows in which part of the empirical distribution this distinction occurs. After this, the null hypothesis of a normal distribution will be tested for each of the risk factors using a formal test. The most common test for checking normality is Jarque-Bera, which tests whether the kurtosis and skewness of the empirical distribution coincide with those of a normal distribution. This test can be performed using the test statistic

$$JB = \left(\sqrt{n/24}(K-3)\right)^2 + \left(\sqrt{n/6}S\right)^2. \tag{27}$$

Here n is the number of observations, while K and S represents the kurtosis and the skewness of the empirical distribution. This test statistic follows approximately a Chi-squared distribution with two degrees of freedom, which can be used to determine the critical value.

For a lot of risk factors, the empirical distribution has heavier tails than a normal distribution. So, when the null hypothesis of a normal distribution is rejected, it is necessary to find a distribution that allows for these heavy tails. A distribution that can be used to capture this excess kurtosis in the distribution is the student-t distribution. Another possible reason to reject the null hypothesis is the presence of skewness. When both excess kurtosis and skewness is present in the data, a skewed student-t distribution can be applied to include these characteristics in the distribution. The number of parameters for this distribution is still quite low, which is a motivation to use this distribution if the normal distribution is rejected. Furthermore, no visual judgment is required, which is necessary when a separate distribution is estimated for the tail section. In that case the threshold between the middle part and the tail section is often determined based on visual inspection. Following Hansen (1994) the skewed student-t distribution is given by

$$F(x) = \begin{cases} b \cdot c\Big(1 + \frac{1}{\eta-2}\big(\frac{\frac{b}{\sigma}(x-\mu)+a}{1-\lambda}\big)^2\Big)^{-(\eta+1)/2} & \frac{x-\mu}{\sigma} < -\frac{a}{b} \\ b \cdot c\Big(1 + \frac{1}{\eta-2}\big(\frac{\frac{b}{\sigma}(x-\mu)+a}{1+\lambda}\big)^2\Big)^{-(\eta+1)/2} & \frac{x-\mu}{\sigma} \geq -\frac{a}{b}, \end{cases} \tag{28}$$

where μ is the location parameter, $\sigma > 0$ is the scale parameter, $2 < \eta < \infty$ is the degrees of freedom and $-1 < \lambda < 1$ is the extra skewness parameter. The constants a, b and c can be calculated using

$$a = 4\lambda c\left(\frac{\eta-2}{\eta-1}\right), \tag{29}$$

$$b = \sqrt{1 + 3\lambda^2 - a^2}, \tag{30}$$

$$c = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma\left(\frac{\eta}{2}\right)}. \tag{31}$$

When $\lambda > 0$ the density is skewed to the right, while it is skewed in the opposite direction for $\lambda < 0$. The parameters for this distribution can be estimated using maximum likelihood. For each of the estimated marginal distributions, the Kolmogorov-Smirnov test is performed to measure the goodness-of-fit of this distribution, based on the empirical observations.

3.2.2 Joint distribution

When the marginal distributions are identified, the next step is to select the most appropriate copula to combine these distributions. There exist several copulas that can be used to model the dependencies between the risk factors, which all have different properties, especially regarding the tail dependency. The most often used copulas belong to the elliptical class or the Archimedean class. Elliptical copulas, like the normal or t-copula, have a symmetric dependence structure. Archimedean copulas, like the Gumbel or Clayton copula, do allow for asymmetric dependence structures, so the upper and lower tail dependence could differ. Another big difference between the two types of copulas is caused by the implicit assumption in Archimedean copulas that the dependence between the different risk factors is the same. This is a large drawback of the Archimedean copulas compared to elliptical copulas, as for these copula a correlation matrix is used as parameter for the dependence. This correlation matrix does allow for differences in dependence between risk factors. This second difference leads suggests that the elliptical copulas might be more suitable in this model. However, still both types of copulas are evaluated in this report. As Archimedean copulas the Gumbel and Clayton copula are considered, while the normal and t-copula are investigated as elliptical copula, where for the t-copula the degrees of freedom vary between two and five.

To check which copula is most adequate for the used risk factors, a goodness-of-fit test is applied. The statistic used for this purpose is the Cramér-von Mises test statistic, which is found most powerful in studies by Genest et al. (2006) and Genest et al. (2009). This Cramér-von Mises test statistic is given by

$$\hat{T}_{CM} = T\int_0^1 \left\{\hat{C}(u) - C_1(u)\right\}^2 dC_1(u), \tag{32}$$

where $\hat{C}(u)$ is the empirical copula, $C_1(u)$ is the estimated copula and T is the sample size. Berg (2009) compares different goodness-of-fit tests and suggests to use an approach proposed by Genest & Rémillard (2008). This approach uses the empirical copula

$$\hat{C}(u) = \frac{1}{T+1} \sum_{t=1}^{T} I_{[Z_{1t} \leq u_1, ..., Z_{Kt} \leq u_K]}, \tag{33}$$

where $\mathbf{u} = (u_1, ..., u_K) = [0,1]^K$, $Z_{kt} = \frac{W_{kt}}{T+1}$ and W_{kt} is the rank of the observation at time t for risk factor k, X_{kt}, amongst $(X_{k1}, ..., X_{kT})$. In this case the Cramér-von Mises test statistic can be written as

$$\hat{T}_{CM} = \sum_{t=1}^{T} \left\{ \hat{C}(z_t) - C_1(z_t) \right\}^2. \tag{34}$$

This method uses the estimated marginal distributions and then estimates the copula parameters using

$$\theta = \arg \max_{\theta} \sum_{t=1}^{T} ln\ C\big(\hat{F}_1(x_{1t}), ..., \hat{F}_K(x_{Kt}); \theta\big). \tag{35}$$

Here $\hat{F}_1(.), ..., \hat{F}_K(.)$ are the estimated marginal distributions of the risk factors.

In some cases the goodness-of-fit test does not give a decisive answer, because more than one copula cannot be rejected. In that case information criteria can be applied to make a choice based on the trade-off between a good approximation and the dimensionality of the copula. One of these information criteria is the Bayesian Information Criterion (BIC), given by

$$BIC = -2l_C + k_C ln(T). \tag{36}$$

Here l_C denotes the log-likelihood of the copula, k_C gives the number of parameters that need to be estimated for the copula and T represents the sample size. Another information criterion that can be used is the Akaike Information Criterion (AIC), which is given by

$$AIC = -2l_C + 2k_C. \tag{37}$$

The AIC tilts a little bit towards the good approximation side of the trade-off and thus adds on average more parameters. Because of this the AIC and BIC will be used together to find a copula that is most adequate. For both information criteria, the copula with the lowest value for the information criterion is the best choice.

4. DATA

In this section the data used in this research is described. To increase the accuracy of the model, it is necessary to use as much observations as possible. To decide on the frequency and period of the data, the input factor with the lowest frequency is leading. In reverse stress testing the number of issuers and number of defaults is required to estimate the parameters in the model for asset returns. Unfortunately, this data is only available on an annual basis, so for the standard asset return model the data of other variables is also restricted to this frequency. However, it is possible to use higher frequency data when the MIDAS specification is applied. In this report it is chosen to use quarterly data for the explanatory variables in this specification. The number of observations included in the model also depends on the variable that has the least historic data available. Data regarding the term structure of interest rates is only available from 1982 onward, so the time period is set on the period between 1982 and 2015. This means that the data contains for instance the financial crisis, which started in 2008, but also the recovery period after this crisis. This shows that the data contains periods with different states of the economy, so it can give a good indicator of the factors that cause the default of an issuer.

In the reverse stress testing framework a few potential risk factors are evaluated. These risk factors consider equity risk and interest rate risk and are supplemented by a few macroeconomic variables. These macroeconomic variables can represent the state of the economy, which is also expected to have a strong influence on the number of defaults at a given time. Possible indicators of the state of the economy are the GDP, unemployment rate, consumer price index (CPI) and the FHFA house price index. For these risk factors the log returns will be considered, except for the unemployment rate, where the net change will be used. Because of this transformation the data will be stationary, so the upward trend over time is removed. This way the risk factors are better able to explain the default rate, which is also stationary. For the unemployment rate net changes are used instead of log returns, because the unemployment rate itself is already a percentage. This data will be obtained for the U.S., because the economic state of the U.S. gives a good representation of the worldwide state of the economy. Besides that, most of the defaults happen in the U.S., so the economic environment for these defaults should be considered. The equity risk is quantified using the S&P 500 index and the Dow Jones index, also transformed into log returns. Annual and quarterly data for these variables, over the period between 1982 and 2015, is obtained via Bloomberg.

The data regarding these risk factors is gathered on a quarterly and a yearly basis, to be applied to the model with or without the MIDAS specification respectively. Table 1 depicts the characteristics of the risk factors when yearly observations are used.

Table 1: Characteristics yearly observations macroeconomic variables

	Mean	**Standard deviation**	**Skewness**	**Excess kurtosis**
Dow Jones	0.037	0.063	-1.214	2.046
S&P 500	0.035	0.072	-1.360	2.295
GDP	0.022	0.009	-0.602	2.229
Housing price	0.016	0.017	-0.941	0.764
CPI	0.012	0.005	0.180	-0.232
Unemployment rate	-0.172	0.998	1.238	2.079

The mean, standard deviation, skewness and kurtosis of the potential risk factors, on a yearly basis. As equity risk factors the log return of the Dow Jones index and S&P 500 index are considered. Besides that the GDP, the housing price index, CPI and unemployment rate for the U.S. are investigated. For the GDP, housing prices and CPI also the log returns are used, while for the unemployment rate the net change is considered.

The characteristics in table 1 show that most risk factors are skewed and have excess kurtosis. This was to be expected, as most of the distributions of log returns possess heavier tails than a normal distribution. This is something that can be solved by using a different distribution, which is better able to model these heavy tails, as discussed in section 3.2.1. The characteristics based on quarterly observations differ a little from the yearly observations and can be found in table 2. For the quarterly observations the mean and standard deviations are lower than for the yearly observations, which is to be expected. Also with quarterly observations the variables are skewed, in most cases to the left, with the skewness approximately equal to the yearly data. Notable is the fact that the kurtosis for most of the variables is much lower in case of quarterly data. Only the CPI differs from the other variables, with a large skewness and kurtosis. This is caused by one extreme observation, in the last quarter of 2008, the start of the financial crisis.

Table 2: Characteristics quarterly observations macroeconomic variables

	Mean	**Standard deviation**	**Skewness**	**Excess kurtosis**
Dow Jones	0.0093	0.0338	-1.0058	1.9438
S&P 500	0.0088	0.0349	-0.9368	1.5271
GDP	0.0055	0.0031	-0.7623	3.2067
Housing price	0.0039	0.049	-1.2552	2.6564
CPI	0.0029	0.0026	-2.2904	16.6051
Unemployment rate	-0.0430	0.2957	1.4003	5.1432

The mean, standard deviation, skewness and kurtosis of the potential risk factors, on a quarterly basis. As equity risk factors the log return of the Dow Jones index and S&P 500 index are considered. Besides that the GDP, the housing price index, CPI and unemployment rate for the U.S. are investigated. For the GDP, housing prices and CPI also the log returns are used, while for the unemployment rate the net change is considered.

The interest rate risk is included in the framework using principal component analysis, which summarizes the term structure of the interest rate. To estimate these components, the term structure of the yield is obtained on an annual and a quarterly basis via Bloomberg. This term structure contains the end-of-period U.S. treasury yield with a maturity of 3 and 6 months and 1, 2, 3, 5, 7, 10 and 30 year. The data for the term structure contains some missing values for the maturity of 30 year. For this maturity it is assumed that the percentage change for the missing values is the same as for the 20 year maturity in that period. Using this assumption also the level of the interest rate for this maturity can be determined for the missing values. Figure 1 shows the term structure of the yield, averaged over time. This shows that, as expected, the yield increases along with the maturity, but when the maturity is higher, the increase with maturity is slower.

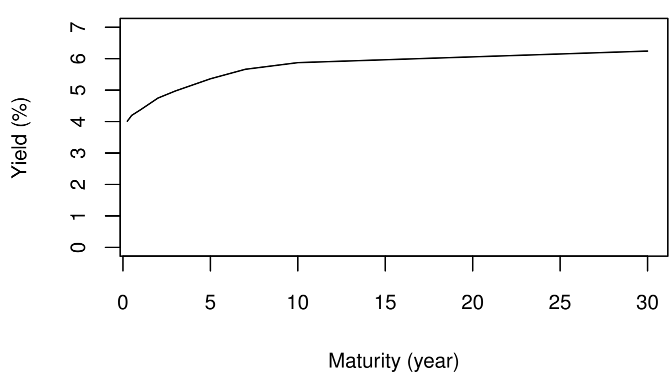

Figure 1: The term structure of the yield, averaged over the period 1982-2015.

Before the resulting principal components are added to the model, it might be necessary to take first differences of the principal components to generate a stationary time series. This transformation would result in the loss of one observation. In case this transformation is necessary for at least one of the principal components, only the period between 1983 and 2015 can be used in the model, also for the other used variables.

To perform the maximum likelihood estimation also data regarding the number of defaults is necessary, as stated earlier. This data is generated from an annual report by Standard & Poor's (2016). This report contains data regarding companies all over the world, with a large concentration in the U.S. As told in section 2, the rating grades are aggregated in two rating groups, the speculative and the investment grade. While in 1981 the number of issuers in the investment grade was a lot larger than in the speculative grade, the size of both groups was approximately equal in 2015. The default rates for both rating groups over the period between 1981 and 2015 are depicted in figure 2, as well as the default rates for the total group of issuers.

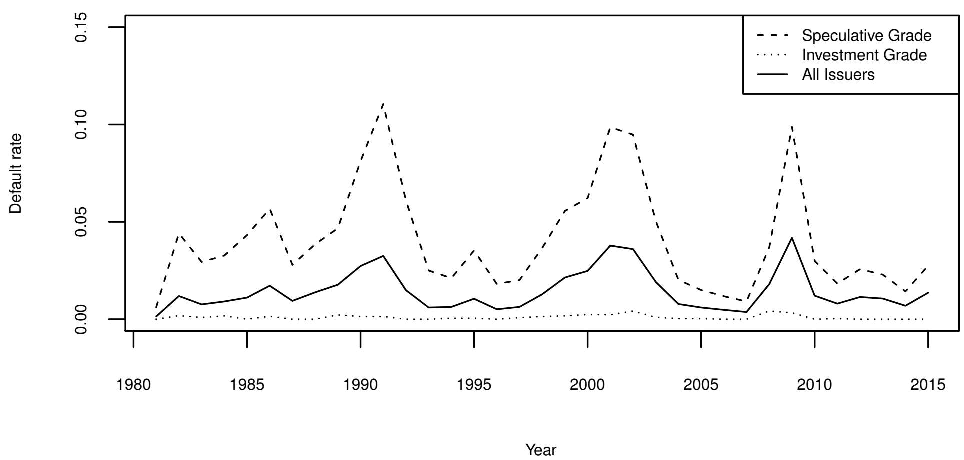

Figure 2: Default rates over the period from 1981 till 2015, based on data by Standard & Poor's (2016). The solid line represents the default rate for all the issuers listed at Standard & Poor's. The dashed line represents the issuers with a rating belonging to the Speculative Grade, while the dotted line gives the default rate of issuers contained in the Investment Grade.

Figure 2 shows that there exists a significant difference between the two rating classes, regarding the default rates. The default rate for issuers with an investment grade is low, while it is a lot higher for the speculative grade. Also a few peaks are visible, especially for the speculative grade, which coincide with economic recessions. The peak in the early 90's represents the recession, which started in the United States, following an oil price shock. The peak between 1999 and 2002 is a result of the Dotcom bubble, followed by the attack on the WTC in New York. The peak in 2009 is caused by the financial crisis, which started with the default of Lehman Brothers.

To determine which rating grade is applicable for each issuer at the risk horizon, the transition threshold is necessary. The rating grade at risk horizon decides which credit spread should be applied in the position valuation and thus influences the value of the position. The default and transition thresholds can be defined according to the historical transition probabilities, which are depicted in table 3. These transition probabilities are averaged over the period between 1981 and 2015, approximately the same period as used in the model, and obtained from Standard & Poor's (2016). Using the quantiles of the empirical distribution of the returns, determined using a Monte Carlo simulation, the thresholds can be calculated, as discussed in section 2.1.

Table 3: Transition probabilities

From/To:	**AAA**	**AA**	**A**	**BBB**	**BB**	**B**	**CCC/C**	**Default**
AAA	89.94%	9.30%	0.55%	0.05%	0.08%	0.03%	0.05%	0.00%
AA	0.55%	90.32%	8.40%	0.55%	0.06%	0.07%	0.02%	0.02%
A	0.03%	1.90%	91.86%	5.65%	0.35%	0.14%	0.02%	0.06%
BBB	0.01%	0.12%	3.79%	91.13%	4.07%	0.55%	0.13%	0.20%
BB	0.01%	0.03%	0.14%	5.62%	84.97%	7.70%	0.71%	0.81%
B	0.00%	0.03%	0.10%	0.24%	5.97%	84.39%	4.99%	4.28%
CCC/C	0.00%	0.00%	0.17%	0.24%	0.72%	15.22%	52.40%	31.25%

One-year transition rates, averaged over the period 1981-2015 (Standard & Poor's, 2016), roughly the same period as considered in the rest of the model. These rates are adjusted for rating withdrawals, which are removed from the sample. The rates thus reflect the percentage of issuers after removing the withdrawn issuers.

Table 3 shows that most of the issuers keep the same rating grade in the next year and hardly move more than two classes up or down. In this report it is assumed that the investment grade issuers have the same transition probabilities as a A rated company, while the speculative grade issuers have the properties of a B rated company. This is decided, because the average default rate over time for the investment grade is closest to the default rate of the A rating, while for the speculative grade the B rating has an approximately equal default rate. Because the issuers hardly move more than two rating classes up or down over a year, issuers are not likely to move to another rating class. Table 3 also shows that, as expected, the default probability for the speculative grade is then a lot higher than for the investment grade.

The recovery rate necessary to get the value of a position in the portfolio, as defined by equation (2), is supposed to be beta-distributed. Following Grundke & Pliszka (2015), the recovery rate in this framework is assumed to be equal to those of senior unsecured bonds. The mean and standard deviation of this recovery rate over the period from 1987 till 2011, equal to 0.518 and 0.389, are generated from Standard & Poor's (2011).

Finally, some data regarding the credit spread is necessary to calculate the value of the different positions in the portfolio, as stated in equation (2). This data is obtained for the last available point in time, 25 May 2017, for Investment grade rated issuers (BofA Merrill Lynch, 2017a) and Speculative grade rated issuers (BofA Merrill Lynch, 2017b). This results in a credit spread of 118 bps for the Investment grade and 367 bps for the Speculative grade, which is applicable to each maturity. As expected, the credit spread is higher for issuers with a lower credit rating.

5. EMPIRICAL RESULTS BASIC FRAMEWORK

This section describes the selection of the risk factors and performs the reverse stress test using the framework without the MIDAS extension. Section 5.1 describes the selection of risk factors and the calibration of the framework. In section 5.2 a reverse stress test is performed on a stylized portfolio.

5.1 Framework calibration

The first step in the calibration of the reverse stress testing framework is the selection of the risk factors to include in the model. However, due to the strongly increasing dimensionality when more risk factors are added, the number of risk factors should be limited. To lessen the number of risk factors necessary, first of all a principal component analysis is performed on the interest rate term structure. Two of the principal components resulting from this analysis, which together explain 99.8% of the total variance in the term structure, will be added to the model. Both these principal components have a clear economic interpretation. The first principal component represents the level of the interest rate term structure, while the second principal component represents the slope of the interest rate curve. Figure 3 shows the factor loadings for the two principal components for the different maturities.

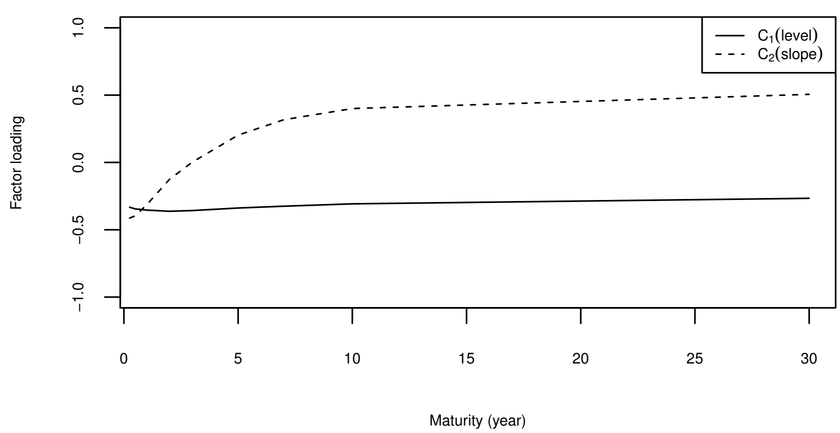

Figure 3: Factor loadings for each maturity of the interest rate for the two principal components. The first principal component can be interpreted as the level of the interest rate term structure, while the second principal component can be interpreted as the slope of the term structure.

Figure 3 clearly shows that both the level and slope component of the term structure are inversed, compared to economic theory. This is due to the fact that when both the loadings and principal components, resulting from a principal component analysis, are multiplied by -1, the same results are accomplished. Because of this, the interpretation of the components is also inversed. In figure 3 it can be seen that the level of the interest rate is negative and increases slightly with the maturity. The slope component first increases rapidly and then slowly increases further with maturity. These two principal components form the starting point for the selection of risk factors to include in the model. However, for the first principal component first differences are used to remove the trend from the data and create a stationary variable. To select the measure for equity risk and the macroeconomic variables for the model, the specific-to-general approach is used. Thus, the optimal combination of a measure for equity risk with one or more macroeconomic variables has to be found. When adding only a measure for equity risk to the model already including the principal components, the Dow Jones index and the S&P 500 index give similar results. To make a choice between these two risk factors, the combination of these factors with the macroeconomic variables is investigated. Doing so, the unemployment rate seems to be the optimal choice, looking only at the statistical measures. However, adding the unemployment rate to the model results in coefficients in the asset return model contradicting the economic theory. As a result, the second best choice regarding the statistical properties, the GDP, will be evaluated based on economic interpretability. When adding the GDP to the model as risk factor, the coefficients of the model are in line with the existing economic theory, so the GDP is added to the model. As a next step, incorporating an additional macroeconomic variable in the model is investigated, but none of the remaining variables can add sufficient additional information. So, following the specific-to-general approach, the S&P 500 index and the GDP are selected as additional risk factors for the model.

Given these four risk factors, the model for asset returns, as given by equation (5), can be estimated by maximum likelihood. The coefficients of this model are depicted in table 4, with the corresponding t-statistic between brackets.

Table 4: Maximum likelihood parameters

	$\mathbf{C_1}$	$\mathbf{C_2}$	**S&P 500 Index**	**GDP**
Investment Grade	0.5931	-1.3445	2.6070	15.881
	(0.320)	(-0.290)	(2.294)	(18.438)
Speculative Grade	-3.2794	-26.0915	1.1187	21.972
	(-2.570)	(-9.669)	(2.264)	(6.916)

Parameters estimated using Maximum Likelihood for the Investment and Speculative grade, where the first two columns give the asset return sensitivity of the principal components, where for the first principal component the first differences are used in the model. The last two columns give the asset return sensitivity of the macroeconomic variables included as risk factor. The values between brackets give the t-statistics.

Table 4 shows that the sensitivity of the asset returns is more significant in case of an initial speculative grade rating. For the investment grade only the S&P 500 index and the GDP have a significant influence on the asset return, while for the speculative grade all risk factors have a significant effect. This difference is caused by the fact that companies in the speculative grade are less stable, so changes in the economic environment have a larger influence on the return. This effect is also visible in the value of the coefficients, which are, in absolute value, higher for the speculative grade, except for the S&P 500 index. The S&P 500 index consists of more stable companies, so the influence in this case is larger for the companies in the investment grade. Both the S&P 500 index and the GDP have a significant positive effect on the asset returns, for both rating classes, which was expected from an economic perspective. The sign of the first principal component is negative for the speculative grade and positive for the investment grade, but in the last case the coefficient is not significant. Due to the negative principal component loadings, a negative sign indicates that when the level of the interest rate term structure increases, the asset returns will also increase. This aligns with economic theory, because when asset returns do not follow the development of the interest rate, investors will only invest in risk-free bonds and no longer in assets. The model coefficients suggest that this effect is stronger for assets in the speculative grade than for assets in the investment grade. The second principal component has a negative influence on the asset returns for both investment grade and speculative grade. This implies that if the interest rate curve flattens, so the long term interest rate decreases compared to short term interest rates, the asset returns will increase. This seems plausible, as it becomes less interesting to investing in the riskfree rate on the long term in this case.

To determine the marginal distribution of each of the included risk factors, the assumption of a normal distribution is investigated. To do so first the QQ-plot of each risk factor is evaluated. These QQ-plots are shown in figure 4.

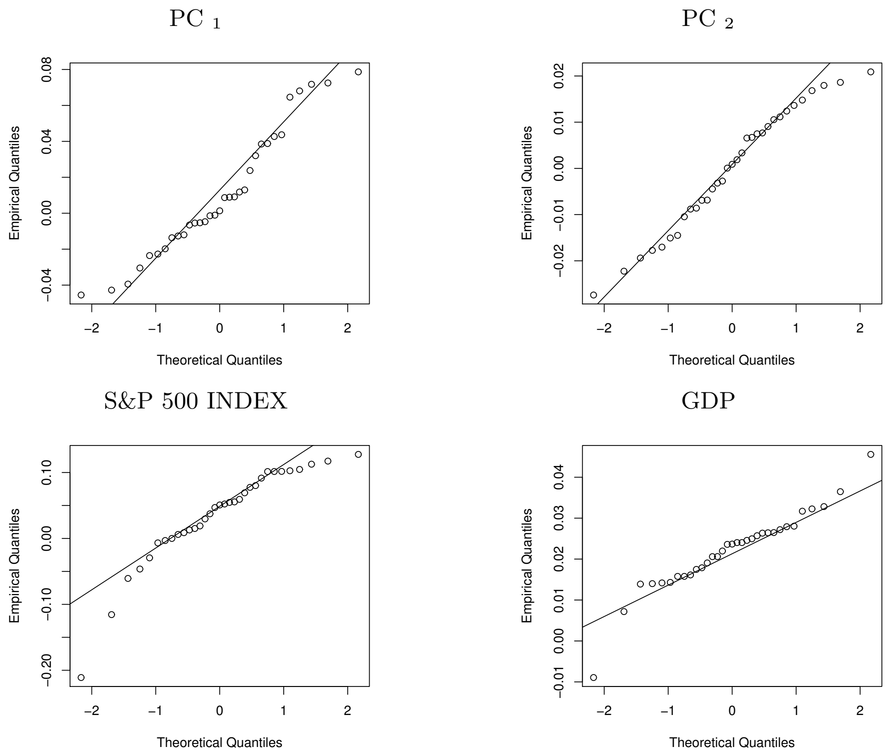

Figure 4: Quantiles of the empirical distributions of the risk factors plotted against the quantiles of a normal distribution. The risk factors evaluated are the two principal components, the S&P 500 index and the GDP.

Figure 4 shows that for all risk factors the normality assumption seems to be appropriate in the middle part of the distribution, but less in the tail part. Especially for the two macroeconomic variables a few observations in the tail part provide strong evidence against the assumptions of a normal distribution. For the principal components the normal distribution seems to be a better fit, although also here a few tail observations differ from the normal distribution. To formally test the results of the QQ-plot also a Jarque-Bera test is performed. Table 5 shows the test statistics resulting from this test, together with the corresponding p-values.

Table 5: Jarque-Bera test

	$\mathbf{C_1}$	$\mathbf{C_2}$	**S&P 500 Index**	**GDP**
JB	1.811	1.630	20.675	11.204
p-value	0.404	0.443	0.000	0.004

Test statistics and p-values for the Jarque-Bera test, to test the normality assumptions for the marginal distributions of the risk factors. The first two columns give the two principal components, while the last two columns give the macroeconomic variables.

The results of the Jarque-Bera test supplement the conclusion of the QQ-plots. For the two S&P 500 index and the GDP the normality assumption is rejected, while for the principal components the empirical distribution does not significantly differ from the normal distribution. As already described in section 4, the two selected macroeconomic variables possess excess kurtosis and skewness. To include these characteristics in the marginal distribution a skewed student-t distribution will be estimated for these risk factors. The corresponding parameters for the marginal distribution are estimated using maximum likelihood and can be found in table 6.

Table 6: Estimated parameters for marginal distributions

	μ	σ	λ	η	**Kolmogorov-Smirnov**	**p-value**
$\mathbf{C_1}$	0.0103	0.0345			0.1339	0.595
$\mathbf{C_2}$	-0.0002	0.0130			0.1188	0.741
S&P 500 Index	0.1220	0.1004	-12.4201	10.4612	0.0910	0.948
GDP	0.0249	0.0068	-0.4200	3.6741	0.0983	0.908

The estimates for the parameters of the marginal distributions for the two principal components and the macroeconomic variables, which are added to the model as risk factors. Here μ is the mean of the distribution and σ is the standard deviation. The degrees of freedom is denoted by η and λ gives the additional skewness parameter of the skewed student-t distribution. These last two parameters are only estimated when the normality tests are rejected and the skewed student-t distribution is necessary for the particular risk factor. The last two columns present the results, the test statistic and the p-value, of the Kolmogorov-Smirnov test, which measures the goodness-of-fit of the marginal distribution.

Table 6 shows that for most risk factors the mean is quite close to zero. Another important observation is the fact that both skewed-t distributed risk factors have a skewness parameter $\lambda < 0$, which indicates that they are skewed to the left. This means that the probability of a value lower than the mean is higher than the probability of a value higher than the mean. Table 6 also depicts the results of the Kolmogorov-Smirnov test, which show that for the GDP and the S&P 500 index the skewed student-t distribution is an appropriate distribution.

Using the marginal distributions for the risk factors the joint distribution can be determined. When defining the joint distribution, the dependency between the different risk factors has to be taken into account. This can be done using a copula, which is chosen using a goodness-of-fit test. Table 7 presents the test statistic of this goodness-of-fit test, the Cramér-von Mises test statistic, and the corresponding p-value.

Table 7: Selection criteria copulas

	Cramér-von Mises test statistic	**p-value**	**AIC**	**BIC**
Normal copula	0.0340	0.8367	4.3095	13.2885
t_{2df} copula	0.0232	0.9396	10.7137	19.6927
t_{3df} copula	0.0256	0.9236	5.1227	14.1017
t_{4df} copula	0.0272	0.9086	3.6388	12.6178
t_{5df} copula	0.0284	0.9176	3.1600	12.1391
Gumbel copula	0.0282	0.8167	1.8494	3.3459
Clayton copula	28.4830	0.1653	2.0000	3.4965

The selection criteria used to choose the most appropriate copula to model the dependence between the risk factors. The first two columns give the Cramér-von Mises test statistic, a goodness-of-fit measure for the copulas, with the corresponding p-value. This statistic tests the null hypothesis that the respective copula is appropriate for the used data. When for more than one copula the null hypothesis is not rejected, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are used. For these information criteria the copula with the lowest value should be chosen.

The results in table 7 show that none of the copulas can be rejected using the goodness-of-fit test. This is not surprising, as very few observations are available to estimate the dependence. To still be able to choose the most appropriate copula, information criteria are used. The last two columns in table 7 display the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), where for both the value should be as low as possible. For both the AIC and the BIC the Gumbel and Clayton copula both have the lowest value. However, when looking at the estimated coefficients for these copulas, $\hat{\theta}_{Gu} = 1.0197$ and $\hat{\theta}_{Cl} = 0.0000$, it turns out that they both represent the independence copula. Economic theory indicates at least a clear dependence between the S&P 500 index and the GDP, so applying an independence copula would contradict economic theory. The fact that both these copulas are estimated as independence copulas can be explained by the implicit assumption of equal dependence between the different risk factors, which is made for the Clayton and Gumbel copula. The theoretical motivation to choose for an elliptical copula is thus emphasized by the estimated parameters. As a result, these copulas are excluded from choice. For both information criteria the t-copula with five degrees of freedom has the lowest value, besides the Gumbel and Clayton copula and thus is chosen in this framework. The estimated coefficient ρ_c for this copula is given by

$$\rho_c = \begin{pmatrix} 1 & 0.2410 & -0.1285 & -0.4510 \\ 0.2410 & 1 & -0.1141 & -0.2293 \\ -0.1285 & -0.1141 & 1 & 0.1797 \\ -0.4510 & -0.2293 & 0.1797 & 1 \end{pmatrix}.$$

Using the joint distribution of the risk factors and the coefficients in the asset return model, estimated using maximum likelihood, a Monte Carlo simulation can be performed. In this simulation 100,000 random samples are drawn from the joint distribution. Using these simulations of the risk factors the asset return for the investment grade and the speculative grade are calculated. Using the transition probabilities as depicted in table 3 in section 4, the default and migration thresholds can be calculated. Here it is assumed that the investment grade follows the transition probabilities of an A-rated company, while the speculative grade follows these probabilities of a B-rated company. The migration thresholds can be found in table 8.

Table 8: Default and migration thresholds

Initial rating	**Default**	**Speculative grade**	**Investment grade**
Investment Grade	≤ -9.443	$(-9.443, -7.478]$	> -7.478
Speculative Grade	≤ -4.892	$(-4.892, 8.902]$	> 8.902

Default and migration thresholds for both rating classes, based on a Monte Carlo simulation with 100,000 draws. A distinction is made between the two rating classes, where for the investment grade the transition probabilities of an A-rated company are used. For the speculative grade the probabilities of a B-rated company are applied.

Table 8 shows that, as expected, the default threshold for the speculative grade is higher than for the investment grade. Companies in the investment grade need a lower return to go into default than the speculative grade. Another observation is that companies in the speculative grade need a quite high asset return to move to the investment grade, while companies in the investment grade need a quite low asset return to migrate to the speculative grade. This is not surprising, as the transition probabilities already showed a low possibility of migration between the two rating classes.

5.2 Reverse stress test on stylized bank portfolios

This section presents the results of the reverse stress test, using the parameters calibrated in section 5.1. First the scenario set is reduced using a reduction algorithm, which is chosen after comparison. Using this reduced set a set of stress scenarios is identified, where the most plausible stress scenario is selected.

The first step in the reverse stress test is to set up a stylized portfolio. Here a portfolio is created consisting of 200 asset positions. The proportion of investment grade rated positions in this portfolio will vary in this report, to investigate the effect of the risk factors on a few different portfolios. The number of liability positions depends on the initial value of these asset positions, as the value of these liability positions should add up to 95% of the total asset value. The remaining 5% of the total asset value forms the initial capital buffer of the institution. The maturities of these asset and liability positions are drawn uniformly between 1 and 10. This randomization of the maturities for the asset and liability positions creates a mismatch between both sides of the balance sheet. The reverse stress test is applied on three different portfolios, each with another composition regarding the number of obligors with initial investment grade and initial speculative grade rating. First of all an investment portfolio is evaluated, where 75% of the obligors is rated as investment grade and the remaining obligors are speculative grade rated. Furthermore, a speculative portfolio is analyzed, where the portfolio composition is the exact opposite. Finally, a portfolio with equal proportions is assessed. The composition of the portfolio also influences the initial characteristics of the portfolio. Table 9 depicts the characteristics of the three evaluated portfolios.

Table 9: Characteristics evaluated portfolios

	Number of asset positions	**Number of liability positions**	**Duration assets**	**Duration liabilities**
Investment portfolio	200	184	5.53	5.56
Equally proportioned portfolio	200	180	5.37	5.67
Speculative portfolio	200	171	5.58	5.46

Characteristics of the three evaluated portfolios. The investment portfolio contains 75% investment grade rated obligors and 25% speculative grade rated obligors. For the speculative portfolio, the compositions is the exact opposite. The equally proportioned portfolio contains the same amount of speculative grade and investment grade rated obligors. For all portfolios, the number of asset and liability positions is shown, followed by the durations of these positions.

Table 9 shows that the number of liability positions decreases when the number of speculative grade rated obligors increases. Furthermore, the ratio between the duration of the assets and liabilities differs over the portfolios.

In this report four risk factors are chosen in the model for the asset returns. To create scenarios, for each of these risk factors a grid is constructed, consisting of 17 grid points. Combining these risk factors into scenarios, 83,521 scenarios are created. In an attempt to reduce the computation time for identifying the stress scenarios, two reduction algorithms are applied, as discussed in section 2.3.1. These algorithms are the backward reduction and forward selection algorithm, where also the output of each of the scenarios, in this case the asset return, is used. However, for these algorithms the distance matrix of the complete scenario set is necessary, which cannot be calculated due to computational limits. Computing the distances step by step does not pose an alternative, as in this case the computation time exceeds the runtime of the full scenario set. Thus, reduction algorithms do not offer a solution to the computational requirements at the moment, because they experience these limitations themselves. Using more advanced software can provide a solution in the future, where these algorithms might once again be an option to reduce the scenario set.

For now, however, the reverse stress test needs to be performed on the full scenario set. In the reverse stress test a risk horizon of one year is used. As mentioned before, the reverse stress test is applied on three different portfolios. For each of these portfolios a most plausible stress scenario is found, which can be found in table 10.

Table 10: Most plausible stress scenario

Fraction initial rating			**Risk factor value**			
Inv. Grade	**Spec. Grade**	**Probability**	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{X_1}$	$\mathbf{X_2}$
75 %	25 %	-	-	-	-	-
50 %	50 %	0.000281	0.1503	0.0459	0.2145	-0.0157
25 %	75 %	0.000352	0.1328	0.0525	0.1069	0.0127

The most plausible stress scenario, using the standard model, for three different proportions of investment grade and speculative grade positions in the portfolio, displayed in the first two columns. For each combination the most plausible stress scenario is depicted, with in the third column the probability of all risk factors being in the neighborhood of the scenario. This is only the probability within the evaluated set of scenarios and does not contain information about the probability of occurrence in the real world. Furthermore, in the last columns the yearly value of each of the risk factors is depicted for this most plausible scenario. The risk factors used here are the two principal components (C_1,C_2), the S&P 500 index (X_1) and the GDP (X_2).

The first result that can be noted in table 10 is the absence of stress scenarios for the investment portfolio, based on the evaluated scenario set. This is not extremely surprising, as this portfolio consists for the largest part of assets issued by quite stable obligors, resulting in a stable portfolio. However, even for such a stable portfolio, it might be expected that at least a few stress scenarios are found. A potential reason for the fact that no stress scenarios are identified is the width of the grid for each of the risk factors. In this framework, the grid width for all risk factors equals four standard deviations to each side of the mean, based on the leftover tail probability for the normal distribution outside these boundaries. As for the skewed student-t distribution the leftover tail probability is higher when using these boundaries, using this grid for risk factors following this distribution might be unfair. For this reason, it might be more appropriate to find the quantiles of the skewed student-t distribution resulting in the same leftover tail probability. This can possibly lead to identification of stress scenarios for the investment portfolio, using the more extreme values for these risk factors. When this wider grid is applied for the risk factors following a skewed student-t distribution, the step size in this grid should be increased, to prevent the dimensionality from increasing sharply.

Also when looking at the other two portfolios, it can be observed that the composition of the portfolio influences the most plausible stress scenario. The most plausible scenario for the speculative portfolio and the equally proportioned portfolio describe a similar economic environment. The probability of having all risk factors in the neighborhood of the most plausible scenario is for both portfolios quite low, which indicates that a default of the portfolio in one year is not very likely. The differences between the two portfolios are small, but for most risk factors the value has to be a bit more extreme in case of the equally proportioned portfolio to cause a default. Only for the S&P 500 index the value needs to be less hurtful in case of the equally proportioned portfolio. This can be explained by the larger influence of this risk factor on investment grate rated obligors, as is already seen in table 4.

Looking at the values of the risk factors in the most plausible scenario for the three different portfolios shows the changes in the economic environment that are most likely to cause a default. For both compositions of the portfolio, the principal components have a lot higher value than in the economic environment at the starting point, 31 December 2015. These values indicate a steeper interest rate curve than at the starting point. This is the results of a decrease of the short term interest rate, while the long term interest rate increases. Due to the fact that the short term interest rate is already quite low at the time the reverse stress test is performed, this results in negative short term rates. Because no restrictions are applied on the interest rate regarding negativity, the model allows for these negative interest rates. The term structure of the interest rate in the most plausible scenarios is depicted in figure 5, compared to the interest rate term structure in the initial situation, measured at 31 December 2015.

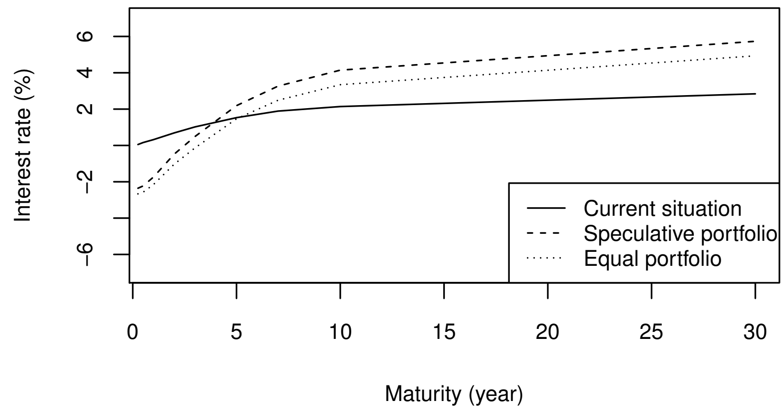

Figure 5: The term structure of the interest rate. The solid line represents the current term structure, at 31 December 2015. The dashed line gives the term structure in the most plausible stress scenario for the speculative portfolio, while the dotted line depicts the term structure in the same scenario for the equally proportioned portfolio.

As can be seen in figure 5, for the speculative portfolio the interest rate curve is slightly higher than for the equally proportioned portfolio, which suggests that this portfolio goes into default sooner. Additional to the interest rate, also the other risk factors can provide some information regarding the economic environment in the most plausible stress scenario. For the speculative portfolio the GDP increases very slowly, while for the equally proportioned portfolio the GDP even has to decrease slightly. The development of the S&P 500 index is quite surprising, as for both portfolios the index increases quite strong in the most plausible stress scenario. Even though this scenario is identified as the most plausible in the stress scenario set, it has to be assessed whether a strong increase for S&P 500 index can occur at the same time as a decrease of the short term interest rates and the GDP.

Although the most plausible stress scenario already provides some information, the stress scenarios that are a little less plausible can still contain a lot of information. As mentioned earlier, for the investment portfolio no stress scenarios are identified. The number of stress scenarios for each of the other portfolio compositions show that the requirements for the speculative portfolio to exhaust the capital buffer are slightly less strict than for the equally proportioned portfolio. A first step in investigating the stress scenarios is analyzing the risk factors individually, which displays the stress thresholds for each of the risk factors and can also give some insight on the impact of the risk factor on the default of a bank. This evaluation is similar to a sensitivity analysis, but in this case it is also possible to see what value the remaining risk factors in the model have to reach at the threshold for a specific risk factor, to cause a stress situation for the portfolio. In this investigation the equally proportioned portfolio is compared to the speculative portfolio, so also the influence of rating of the obligors can be analyzed. Figure 6 shows histograms for each of the risk factors, based on the subset of stress scenarios for the equally proportioned portfolio.

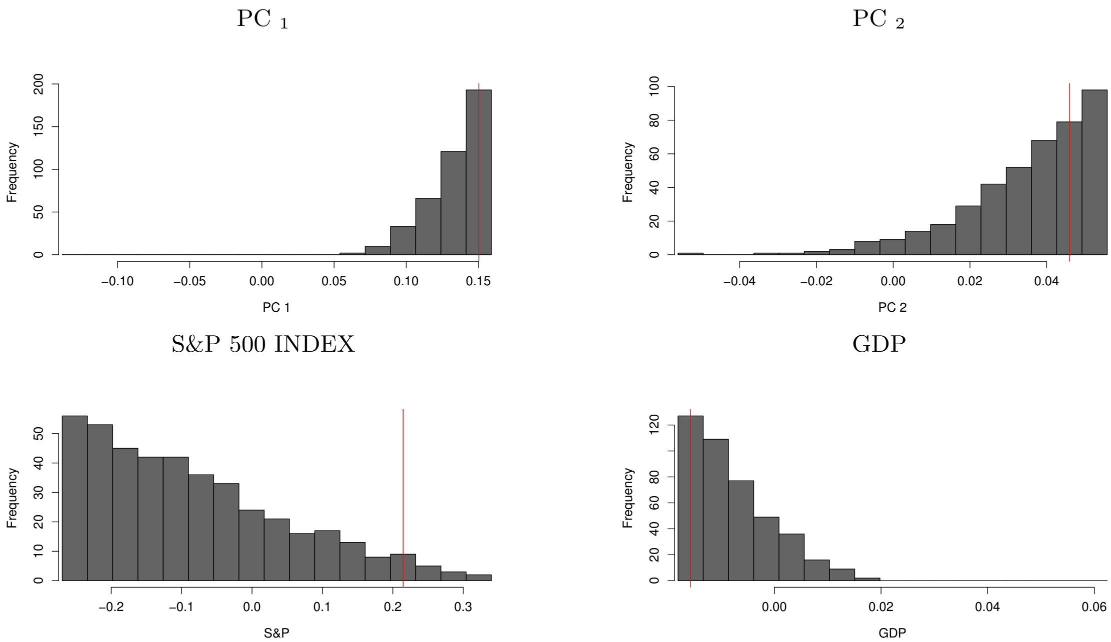

Figure 6: Histograms for each risk factor to display the empirical distribution of the risk factors in the set of stress scenarios. Each of the bars represents one grid point of the respective risk factor. The red line is located at the value of the specific risk factor in the most plausible stress scenario. These histograms are based on the equally proportioned portfolio, where half of the obligors is initially rated as investment grade and half as speculative grade. The risk factors evaluated are the two principal components, the S&P 500 index and the GDP.

The histograms in figure 6 display a clear range for the risk factors in the set of stress scenarios, except the S&P 500 index. This suggests once again that the influence of the S&P 500 index on the default of a portfolio is smaller than the influence of the other risk factors. The histogram for the first principal component shows that stress scenarios only occur with a high level of the first principal component. Because the principal components are inversed, this implicates a decrease of the level of the interest rate term structure. However, this does not mean that for all maturities the interest rate decreases, because the long term interest rate can still increase due to a positive value for the second principal component. A positive second principal component indicates decreasing short term interest rates, along with increasing long term interest rates. Furthermore, the histogram for the GDP shows a clear threshold, above which value no stress scenarios exist for the investment portfolio. Thus, for the investment portfolio to go into default, the GDP can only slightly increase or even has to decrease. For the S&P 500 index such a clear threshold is not present, although the histogram shows that in most stress scenarios a decrease of the index is experienced. Looking at the value of the risk factors in the most plausible stress scenario in these histograms, shows that for most risk factors this value is located in the tail of the distribution. Only for the S&P 500 index this is not the case, which is in line with earlier findings regarding this risk factor.

To be able to assess the influence of the composition of the portfolio on the set of stress scenarios, the individual risk factors are also evaluated for the speculative portfolio. Figure 7 shows once again the histograms for each of the risk factors, yet now based on the subset of stress scenarios for the speculative portfolio.

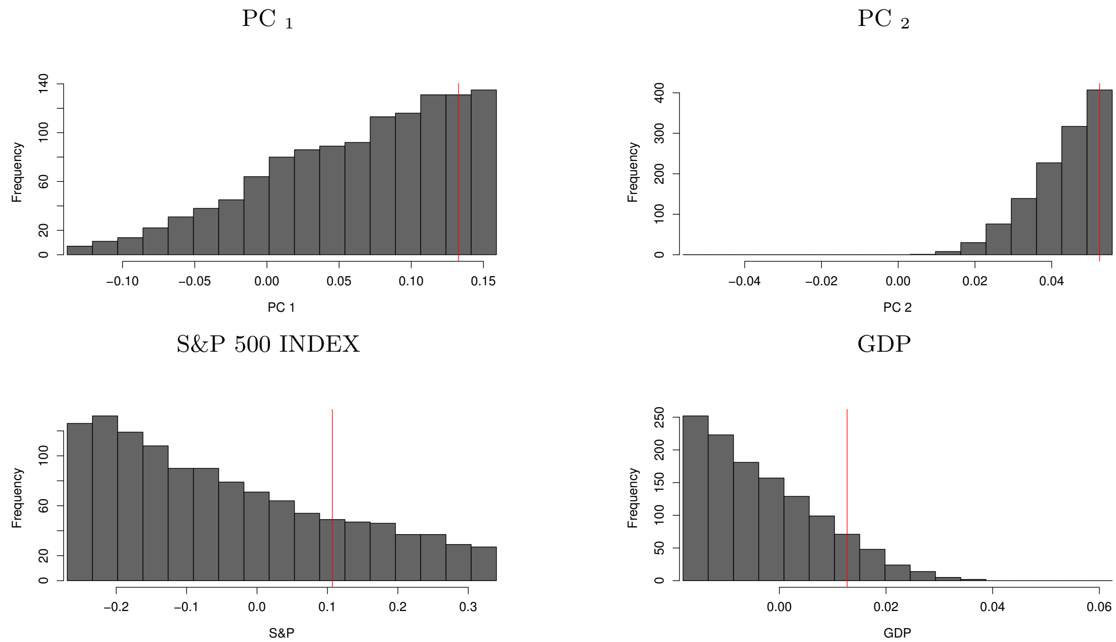

Figure 7: Histograms for each risk factor to display the empirical distribution of the risk factors in the set of stress scenarios. Each of the bars represents one grid point of the respective risk factor. The red line is located at the value of the specific risk factor in the most plausible stress scenario. These histograms are based on the portfolio that contains 25% initially investment grade rated assets. The risk factors evaluated are the two principal components, the S&P 500 index and the GDP.

Figure 7 shows that for the S&P 500 index and the GDP the empirical distribution in the set of stress scenarios does not change much. However, for the interest rate the changes are more significantly, especially regarding the level component of the curve. For the speculative portfolio the requirements regarding the level component of the curve are less strict, which could be expected, because the portfolio is less stable. Even in a scenario where the complete interest rate structure increased, the bank could still go into default. It has to be noted that in this case both the S&P 500 index and GDP have to decrease, so the plausibility of such a scenario is questionable. Another striking difference with the equally proportioned portfolio is that the value of the GDP in the most plausible stress scenario is no longer in the tail of the distribution. This indicates that the value of the GDP needs to be less extreme to cause a default for the speculative portfolio.

To get a better insight in the economic developments that cause a bank to go into default, the patterns in the set of stress scenarios need to be identified. Patterns in the stress scenario set rely on dependencies between the risk factors. Figure 8 shows the value for the second principal component, S&P 500 index and GDP for the set of stress scenarios in case of an equally proportioned portfolio. The first principal component is not shown, as for all stress scenarios, its value is in the same region.

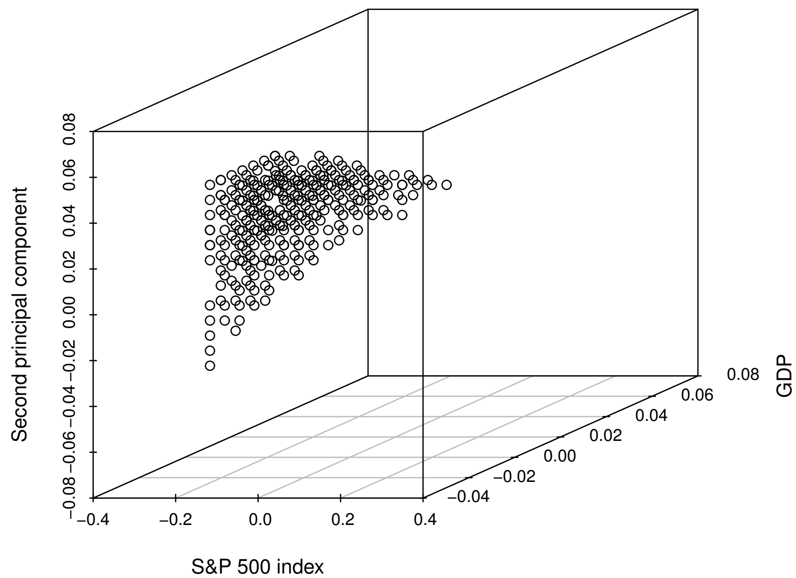

Figure 8: Set of stress scenarios for a portfolio with an equal proportion of initially investment grade rated obligors and initially speculative grade rated obligors. The stress scenarios are visualized based on the value of the second principal component, the S&P 500 index and the GDP.

Figure 8 shows that the stress scenarios are located in a corner of the scenario space. This stress scenario space can be summarized in a few scenarios using a scenario clustering algorithm, as discussed in section 2.3.1. The number of clusters can be chosen by calculating the within group sum of squares for various values for the number of clusters, which should be as low as possible. However, this within group sum of squares will keep declining with the number of clusters. This means that a trade-off exists between the number of clusters and the within group sum of squares. Evaluating this sum of squares reveals that when creating more than three clusters, the sum of squares does not decrease very strong. Thus, three clusters are formed containing the stress scenarios for the equally proportioned portfolio. These clusters provide information on the various types of scenarios that can cause a bank to exhaust the capital buffer. Thus, the set of stress scenarios can be summarized by the cluster scenarios depicted in table 11, which can also be used as input for a regular stress test if deemed plausible.

Table 11: Cluster scenarios for equally proportioned portfolio

	Number of stress scenarios	**PC_1**	**PC_2**	**S&P 500 index**	**GDP**
Cluster 1	156	0.135	0.038	-0.049	-0.0081
Cluster 2	196	0.130	0.029	-0.202	-0.0064
Cluster 3	73	0.142	0.045	0.150	-0.0115

The three cluster scenarios summarizing the set of stress scenarios for the equally proportioned portfolio. The first column presents the number of scenarios contained in the specific cluster. The remaining four columns give the risk factor values in these cluster scenarios.

Table 11 shows cluster 1 and 2 seem to be the most dominant in the stress scenario set, as they contain the largest part of the set. This also indicates that the cluster scenarios for these clusters have a higher probability to go into default, as more scenarios close to this cluster scenario exhaust the capital buffer of the bank. This suggests that these scenarios should be analyzed more thoroughly to further assess the impact of these scenarios on the organization. The number of scenarios in each cluster and the probability of the scenarios in the specific cluster already provide an insight in the plausibility of these cluster scenarios. However, it should be noted that the probability only assesses the probability of having all risk factors in the neighborhood of the specific scenario, using the current scenario grid.

The main difference of these clusters with the third cluster is the development of the S&P 500 index. A stress scenario is clearly more likely when the S&P 500 index decreases over the year, which could be expected. Furthermore, table 11 reveals that in all cluster scenarios a strong decrease of the level component of the interest rate curve is present, combined with a steepened curve over maturity. These developments result in decreasing short term interest rates and increasing long term interest rates for all cluster scenarios. The degree of steepness, and thus the degree of increase of the long term interest rates, differs between the cluster scenarios and appears to be higher in case of an increase in the S&P 500 index. Moreover, in all cluster scenarios the GDP decreases over the year, which is an expected development in a stress scenario. The extent of the decrease differs between the cluster scenarios and seems a bit stronger when the decrease of the level of the interest rate curve is also stronger. In the third scenario, the development of the S&P 500 index and the GDP are contradictory. For this scenarios, the plausibility should be assessed using expert opinion to verify whether the scenario poses a realistic threat to the bank. In the remaining cluster scenarios the movement of the risk factors seems to be more in agreement, which indicates a more plausible economic environment.

Also in case of the speculative portfolio a cluster algorithm is applied to summarize the set of stress scenarios. The number of clusters is chosen based on the sum of squares within each cluster, which decreases with the number of clusters. Reducing the sum of squares presents a trade-off with limiting the number of clusters. Once again, evaluating this sum of squares reveals that when creating more than four clusters, the sum of squares does not decrease very strong. Figure 9 shows a scatter plot for the four clusters created. Due to the small differences between the clusters regarding the second principal component, the figure is based on the first principal component, the S&P 500 index and the GDP.

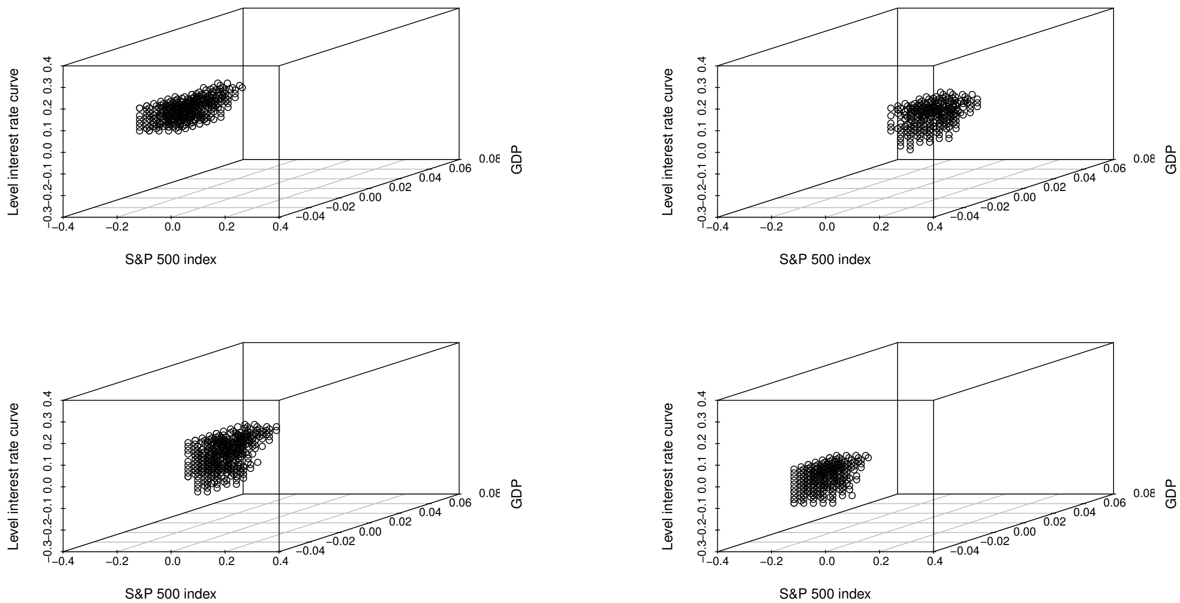

Figure 9: The set of stress scenarios for the speculative portfolio divided in four clusters, used to summarize the scenarios and add interpretation to the set of stress scenarios. The stress scenarios are visualized based on their value for the first principal component, the S&P 500 index and the GDP.

As can be seen in figure 9, each of the clusters contains scenarios in a specific region in the scenario grid. The clusters can be represented by center scenario of the cluster, which can also be used as input for a regular stress test. Thus, the set of stress scenarios can be summarized by the cluster scenarios depicted in table 12.

Table 12: Cluster scenarios for speculative portfolio

	Number of stress scenarios	$\mathbf{PC_1}$	$\mathbf{PC_2}$	**S&P 500 index**	**GDP**
Cluster 1	334	0.101	0.038	-0.185	0.0038
Cluster 2	258	0.096	0.046	0.205	-0.0063
Cluster 3	358	0.069	0.044	-0.002	-0.0039
Cluster 4	255	-0.028	0.047	-0.181	-0.0076

The four cluster scenarios summarizing the set of stress scenarios for the speculative portfolio. The first column presents the number of scenarios contained in the specific cluster. The remaining four columns give the risk factor values in these cluster scenarios.

Table 12 shows that also for the speculative portfolio, two clusters are more dominant in proportion of the stress scenario set, in this case cluster 1 and 3. So, for this portfolio the two cluster scenarios representing these clusters should be evaluated more closely to monitor the total impact of these clusters. However, in this case the differences of the more dominant clusters with the other scenarios are less clear. The most notable difference is the fact that in the more dominant scenarios the value of the GDP development is less negative from an economic perspective. In one of these cluster scenarios, the GDP even increases over the year, but still the capital buffer of the bank is exhausted. However, in this scenario the development of the S&P 500 index is contradictory, so once again, the plausibility should be assessed using expert opinion to verify whether the scenarios in this cluster pose a realistic threat to the bank. In cluster scenario 3 the economic developments seem more aligned, as the GDP decreases slightly over the year and the S&P 500 index remains approximately constant. This implies that this cluster describes a more plausible stress scenario for the bank.

Table 12 also reveals that in all cluster scenarios a steepening interest rate curve is visible, but the development of the level component does not decrease over the year in each of the cluster scenarios, as was the case for the equally proportioned portfolio. Figure 10 shows the interest rate curve in the four cluster scenarios.

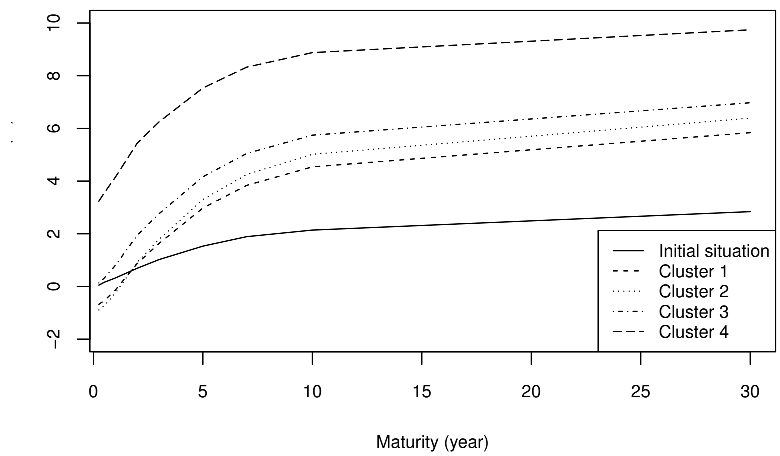

Figure 10: Term structure of the interest rate for the four cluster scenarios that summarize the set of stress scenarios for the speculative portfolio, containing 75% initially speculative grade rated obligors and 25% initially investment grade rated obligors. The straight line presents the term structure in the initial situation. The other lines depict the term structure in case of one of the cluster scenarios mentioned above.

Figure 10 shows that in cluster scenario 3 and 4 even the short term interest rate lies above zero. This indicates that also in less extreme interest rate scenarios, without negative short term rates, the speculative portfolio can still go into default. The long term interest rates are higher than in the initial situation for all cluster scenarios. Furthermore, just as for the first cluster scenario, the second scenario shows contradictory developments of the S&P 500 index and the GDP. So also here the plausibility should be assessed using expert opinion.

Looking more closely at the set of stress scenarios for both portfolios shows that some stress scenarios contain quite positive values for the S&P 500 index and the GDP. A positive value for both these risk factors indicates a positive economic environment, where normally a bank would not be exposed to the risk of default. However, due to the term structure of the interest rate in these scenarios, the scenario still leads to a default. This demonstrates that only looking at the market risk and economic indicators, like the S&P 500 index and the GDP, can results in missing potential risks. The risk of missing stress scenarios also exists when only looking at the term structure of the interest rate. In some stress scenarios the level of the interest rate curve increases over the year, but nonetheless the capital buffer is exhausted, due to a negative development of the other risk factors. As these scenarios were not expected up front, it is useful to assess the plausibility of these scenarios using expert opinion. If these scenarios are deemed plausible, they provide a beneficial extension to the set of stress scenarios used in stress testing at the moment.

6. EMPIRICAL RESULTS MIDAS EXTENSION

This section describes the empirical results of the reverse stress testing framework, when the MIDAS extension is added to the framework. In section 6.1 the parameters in the framework are estimated and evaluated, where also a comparison with the calibration of the model without MIDAS extension is made. In section 6.2 the reverse stress testing framework is applied to a stylized bank portfolio and the results of the test are presented.

6.1 Framework calibration

Just as for the framework without MIDAS specification, the first step in calibrating the framework is to select the risk factors. To make sure that the two approaches are comparable, the same risk factors are used as in the framework without MIDAS extension. So, the framework is calibrated using the first two principal components for the interest rate term structure, together with the S&P 500 Index and the GDP. These principal components are generated by performing a principal component analysis on the term structure of the interest rates. The resulting principal components can be added to the model as risk factors. Just as for the model without MIDAS specification, the first two principal components are added to the model, which represent the level and the slope of the interest rate term structure. These principal components together explain 99.8% of the total variance. Figure 11 shows the factor loading for the two principal components for the different maturities.

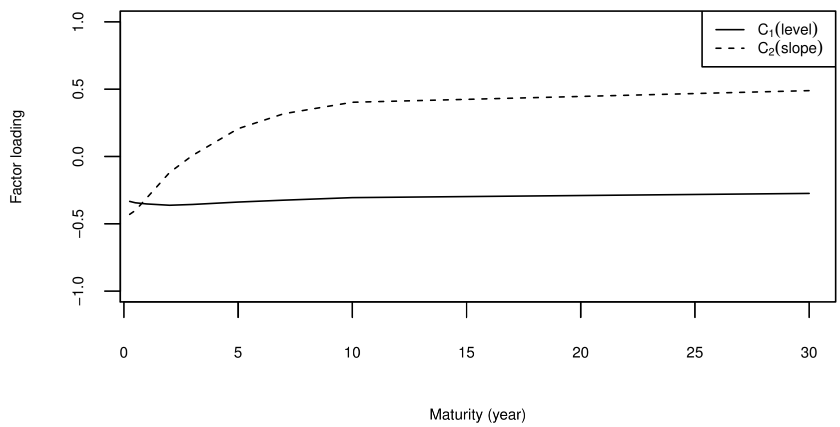

Figure 11: Factor loadings for each maturity of the interest rate for the two principal components. The first principal component can be interpreted as the level of the interest rate term structure, while the second principal component can be interpreted as the slope of the term structure.

Figure 11 shows that also in case of quarterly observations, the level and slope component of the interest rate are inversed. The level component is negative and slowly increases with maturity. The slope component increases rapidly with maturity in the short term and slowly increases with maturity for the longer maturities. The factors loadings over the maturity are quite similar for the principal components on a yearly basis and on a quarterly basis, which was to be expected. These two principal components are added to the set of risk factors. Just like for the model without MIDAS specification, the first difference is used for the first principal component.

With these risk factors, the parameters of the linear model for the asset returns need to be estimated. For the model including the MIDAS specification, this can be done using the log-likelihood given in equation 20. The estimated parameters are depicted in table 13, where the t-statistic of each parameter is shown between brackets.

Table 13: Maximum likelihood parameters

	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{X_1}$	$\mathbf{X_2}$	θ_1	θ_2
Investment Grade	-0.5348	-3.0187	5.5585	13.7037	0.0338	0.1036
	(-0.1766)	(-0.7591)	(1.2633)	(5.0835)	(0.0216)	(0.3683)
Speculative Grade	0.0644	-31.1892	8.2633	32.9104	-0.4704	0.2845
	(0.0208)	(-5.0935)	(3.2337)	(5.1417)	(-0.3456)	(1.1317)

Parameters estimated using Maximum Likelihood for the Investment and Speculative grade, where the first four columns give the asset return sensitivity of the respective risk factor. These risk factors are the principal components (C_1, C_2), the S&P 500 index (X_1) and the GDP (X_2). The last two columns give the parameters for the weighting polynomial, used in the MIDAS specification. The values between brackets give the t-statistics.

Table 13 shows that also with MIDAS specification it is difficult to find significant parameters for the asset return model. This was to be expected, as only 35 yearly observations are available for the number of defaults. For the speculative grade only the first principal component is not significant, while for the investment grade the GDP is the only significant coefficient. Once again, this difference between the rating grades can be explained by the stability of companies in the investment grade rating class. For both the investment grade as the speculative grade the sign of the parameter for the second principal components is negative. Just as in the model without MIDAS specification, this indicates that a flattening interest rate curve results in an increase in asset returns. The sign for the first principal component differs between the rating classes and both are not significant, implying that the influence of the level of the interest rate term structure is a lot weaker when using quarterly values. For the S&P 500 index and the GDP the positive sign as depicted in table 13 was to be expected, looking from an economic perspective. Also the parameters θ_1 and θ_2, used in the weighting polynomial for the MIDAS specification, are insignificant. For the investment grade most weight is placed on the first quarterly value and it decreases towards the most recent value. This suggests the existence of a delay in the effect of the risk factors. In case of the speculative grade this delay is even stronger, with a higher weight on the first quarterly value than in case of the investment grade. This indicates that even for the less stable companies it takes some time till the developments in the market have an effect on the asset return of the company.

When the model with MIDAS specification is compared with the standard model, as discussed in section 5, the significance of the model is a little lower when quarterly values are used. The decreased significance of the model suggests that the MIDAS specification is not able to include all additional information contained in the quarterly values in the model. This could be caused by the assumption that the weighting structure is the same for each of the risk factors, which is used to reduce the number of parameters that needs to be estimated. However, due to the low number of observations, estimating a separate structure for each risk factor is not possible.

When the model for the asset returns is estimated based on the risk factors, the joint distribution of these risk factors is necessary to simulate data and determine a probability for each of the scenarios. For each of these risk factors in the model first the marginal distribution should be estimated. The first option as distribution is, just as for the yearly observations, the normal distribution. So, for each risk factor the null hypothesis of a normal distribution is tested. The QQ-plots, which can be found in figure 12, provide a visual test for this null hypothesis. These QQ-plots show that the normal distribution seems to be appropriate for the middle part of the distribution, but not for the tail section. This suggests that for each of the risk factors the null hypothesis of a normal distribution should be rejected.

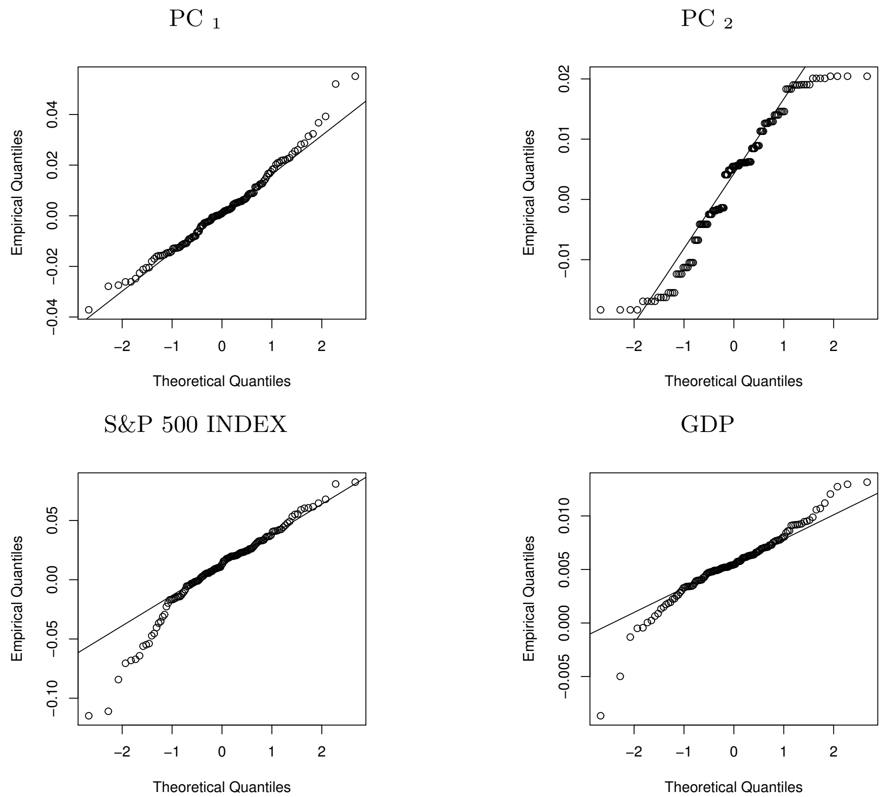

Figure 12: Quantiles of the empirical distributions of the risk factors plotted against the quantiles of a normal distribution. The risk factors evaluated are the two principal components, the S&P 500 index and the GDP.

To formally test whether the null hypothesis is rejected for the risk factors, once again a Jarque-Bera test is performed. The results of this test can be found in table 14.

Table 14: Jarque-Bera test

	$\mathbf{C_1}$	$\mathbf{C_2}$	**S&P 500 Index**	**GDP**
JB	8.354	7.123	33.769	73.039
p-value	0.0153	0.0284	0.000	0.000

Test statistics and p-values for the Jarque-Bera test, to test the normality assumptions for the marginal distributions of the risk factors. The first two columns give the two principal components, while the last two columns give the macroeconomic variables.

The results from the Jarque-Bera test confirm the conclusion of the QQ-plots. The empirical distribution of all risk factors differs significantly from a normal distribution, so a skewed student-t distribution will be estimated. However, for the second principal component estimation of this distribution suggested the absence of skewness in the data, so a skewed student-t distribution is not necessary to model this variable. Instead, a standard student-t distribution is fitted on the observations of this variable. Table 15 depicts the parameters of the marginal distribution for each risk factor, which are estimated by maximum likelihood.

Table 15: Estimated parameters for marginal distributions

	μ	σ	λ	η	**Kolmogorov-Smirnov**	**p-value**
$\mathbf{C_1}$	-0.0131	0.0210	1.7418	24.0148	0.0510	0.882
$\mathbf{C_2}$	0.0030	0.0115		25.9612	0.1134	0.067
S&P 500 Index	0.0356	0.0343	-1.4192	4.8963	0.0439	0.962
GDP	0.0060	0.0022	-0.2111	3.5098	0.0474	0.928

The estimates for the parameters of the marginal distributions for the two principal components, the S&P 500 index and the GDP. Here μ is the location parameter and σ is the scale parameter. Compared to the normal distribution, also the degrees of freedom, denoted by η, and an additional skewness parameter λ are necessary to fit a skewed student-t distribution. Including these last two parameters makes sure the heavy tails and asymmetry are included in the distribution. However, for the second principal component the standard student-t distribution is estimated, so the additional skewness parameter λ is not needed in that case. The last two columns present the results, the test statistic and the p-value, of the Kolmogorov-Smirnov test, which measures the goodness-of-fit of the marginal distribution.

Table 15 shows that both macroeconomic variables have the skewness parameter $\lambda < 0$, which indicates that they are skewed to the left. This is an important property of the distribution to take into account, because it implicates that values below the mean have a higher probability. In combination with the positive correlation of these variables with the asset returns, it results in a higher likelihood of values with a large negative impact on the asset return. Therefore, these variables are more likely to take a value that creates a stress scenario. The distribution for the first principal component is skewed to the right, indicating a higher probability for values higher than the mean. The results of the Kolmogorov-Smirnov test show that the estimated distribution are appropriate, based on the empirical observations, as none of the estimated distributions is rejected.

Using the marginal distribution for each of the risk factors also the parameters of the autoregressive model for the risk factors, which determines the structure of the quarterly values, can be estimated. This model contains one lagged term and a constant, while the error term follows the same marginal distribution of the specific risk factor, but with zero mean. This AR(1) model is used to estimate the last three quarterly values for the risk factors, based on the first quarterly value. The parameters for these models, estimated again using maximum likelihood, are shown in table 16, with the t-statistics between brackets.

Table 16: Parameters autoregressive model for risk factors

	$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{X_1}$	$\mathbf{X_2}$
α	-0.0132	0.0005	0.0361	0.0031
	(-9.9035)	(0.4614)	(13.5309)	(6.4186)
ϕ	0.0652	0.8313	-0.0667	0.5092
	(0.8158)	(9.7240)	(-0.9428)	(6.4275)

Parameters estimated using Maximum Likelihood for the AR(1) model for the quarterly values of the risk factors. The first two columns are the principal components, while the last two columns are the macroeconomic variables: The S&P 500 index (X_1) and the GDP (X_2). The values between the brackets are the t-statistics.

Table 16 shows that only for the GDP and the second principal component a significant correlation between consecutive quarterly values exists. Furthermore, the autocorrelation seems to be quite strong, indicated by the high values of the parameters itself, especially for the second principal component. This is endorsed by economical literature, which already states the persistence for these variables. For the first principal component, as well as the GDP, the correlation seems limited, which also is visible in the value of the parameter itself. This suggests that for these three risk factors the AR(1) structure does not add much information to determine the next quarterly value. The absence of autocorrelation for the first principal component can be explained by the fact that first differences are used for this variable, which removes the trend from the data. Furthermore, it can be noted that the constant for each of the autoregressive models is close to zero, but in most cases still significant. This is not surprising, as the mean of these data series is generally close to zero. This is partly due to the fact that in most cases first differences or log returns are used to transform the data.

Using the marginal distributions and the corresponding distribution parameters for each of the risk factors, also the joint probability can be estimated. Just as for the yearly observations, a copula can be used to model the dependency. For the copula a few options are available, each with their own characteristics regarding the dependency. To choose the appropriate copula in this case once again a goodness-of-fit test is applied. The results of this test, the Cramér-von Mises test statistic and the corresponding p-value, are shown in table 17.

Table 17: Selection criteria copulas

	Cramér-von Mises test statistic	**p-value**	**AIC**	**BIC**
Normal copula	0.0304	0.2892	-18.7573	-1.4605
t_{2df} copula	0.0271	0.4171	36.1698	53.4666
t_{3df} copula	0.0272	0.3951	-0.0204	17.2764
t_{4df} copula	0.0275	0.3591	-11.7483	5.5485
t_{5df} copula	0.0278	0.3541	-16.6812	0.6156
Gumbel copula	0.0308	0.4211	2.0000	4.8828
Clayton copula	119.7977	0.0115	2.0000	4.8828

The selection criteria used to choose the most appropriate copula to model the dependence between the risk factors. The first two columns give the Cramér-von Mises test statistic, a goodness-of-fit measure for the copulas, with the corresponding p-value. This statistic tests the null hypothesis that the respective copula is appropriate for the used data. When for more than one copula the null hypothesis is not rejected, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are used. For these information criteria the copula with the lowest value should be chosen.

As can be seen in table 17, only the Clayton copula can be rejected for the used dataset. Looking at the estimated coefficient for this copula shows that this copula actually represents the independence copula. This explains the rejection, as economic theory clearly states a correlation between the various risk factors. It is not surprising that the remaining copulas cannot be rejected, as the dataset still contains relatively few observations, even though quarterly observations are used. To make a choice for the most appropriate copula nonetheless, information criteria are applied. The value of the the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are presented in table 17. These information criteria both prefer the normal copula. For this copula the necessary parameter is estimated again using maximum likelihood. The estimated coefficient ρ_c is given by

$$\rho_c = \begin{pmatrix} 1 & 0.0684 & -0.0511 & -0.4347 \\ 0.0684 & 1 & -0.1997 & -0.0235 \\ -0.0511 & -0.1997 & 1 & 0.0713 \\ -0.4347 & -0.0235 & 0.0713 & 1 \end{pmatrix}.$$

Now that both the linear model and the joint distribution of the risk factors are estimated, a Monte Carlo simulation can be performed to determine the default and migration thresholds for each of the rating classes. In this simulation 100,000 draws of the risk factors and idiosyncratic risk are generated and for each of these draws the return is calculated for an investment grade rated and a speculative grade rated obligor. Based on the transition probabilities as given in table 3 in section 4, the default and migration thresholds can be determined. As discussed before, the it is assumed that the investment grade follows the transition probabilities of an A-rated company, while for the speculative grade these of a B-rated company are used. The resulting thresholds are depicted in table 18.

Table 18: Default and migration thresholds

Initial rating	**Default**	**Speculative grade**	**Investment grade**
Investment Grade	≤ -14.588	$(-14.588, -11.512]$	> -11.512
Speculative Grade	≤ -7.937	$(-7.937, 10.470]$	> 10.470

Default and migration thresholds for both rating classes, based on a Monte Carlo simulation with 100,000 draws. A distinction is made between the two rating classes, where for the investment grade the transition probabilities of an A-rated company are used. For the speculative grade the probabilities of a B-rated company are applied.

As can be seen in table 18, the default threshold is lower for obligors with an investment grade rating, which was also the case in the model without MIDAS specification. These obligors need a more extreme negative return to go into default, as they are more stable and are better able to cope with low returns. Furthermore, the transition thresholds indicate that it is very unlikely for obligors to switch to another rating class within one year. For both rating classes, the thresholds are more extreme than in the model without MIDAS specification, indicating that scenarios need to have an even more negative effect on the asset returns to be marked as stress scenario.

6.2 Reverse stress test on stylized bank portfolios

This section presents the results of the reverse stress test, using the parameters calibrated in section 6.1. First the scenario set is reduced using the same reduction algorithm as used for the yearly observations. Using this reduced set a set of stress scenarios is identified, where the most plausible stress scenario is selected. The first step in the reverse stress test is to set up a stylized portfolio. Here the same portfolio is considered as in section 5.2, which consists of 200 asset positions. The proportion of investment grade rated positions in this portfolio will vary in this report, to investigate the effect of the risk factors on a few different portfolios.

In this report four risk factors are chosen in the model for the asset returns. To create scenarios, for each of these risk factors a grid is constructed for the first quarterly value, consisting of 17 grid points. Combining these risk factors into scenarios, 83,521 scenarios are created. After this, for each scenario the other three quarterly values are generated for the risk factors, using the estimated autoregressive structure. As discussed in section 5, reduction algorithms do not provide a reduction of the computation time yet, due to computational limitations. Therefore, the grid search is applied on the full set of scenarios. To make the results of the two models, with or without MIDAS specification, comparable, the same portfolios are evaluated. Because of this, the initial portfolio value, the number of liability positions and the initial capital buffer are the same as for the portfolios in section 5.

When performing the reverse stress test using the framework with MIDAS extension, an issue regarding this extension presents itself. Due to the fact that only the first quarterly value is included in the scenario generation process and the remaining three values are estimated using the autoregressive structure, the created scenarios do not necessarily cover the complete grid of the risk factors. The consecutive quarterly values follow the respective marginal distribution, resulting in a lower probability for the extreme values. Because of this, the three estimated quarterly values can smooth the extreme value of the first quarterly value. Inspection of the estimated quarterly values shows that this effect is indeed present in the data, which strongly influences the outcomes of the reverse stress test. As a result of this smoothing effect, none of the generated scenarios causes a default of the portfolio, regardless of the portfolio compositions. This leads to the conclusion that the MIDAS extension using the current assumptions does not provide additional information to the model. This does not mean that the MIDAS extension is unable to include additional information regarding the risk factors in the model. However, to ensure the MIDAS scenarios to exhibit extreme behavior, the generation of the quarterly risk factor values for the scenarios needs to be improved. A method to ensure this extreme behavior could be a more sophisticated model to determine the consecutive quarterly values. Another possibility is estimating the weighting polynomial individually for each risk factor, as this might reduce the smoothing effect. These possibilities to improve the MIDAS extension are discussed in more detail in section 8.

7. CONCLUSION

This report presented a quantitative reverse stress testing framework that can be used to identify the set of stress scenarios for a specific portfolio. Subsequently, the framework can be used to find the most plausible stress scenario within this subset, using an estimated distribution for the selected risk factors. The resulting stress scenarios provide information on the economic environment in which a bank exhausts its capital buffer and goes into default. Looking at risk factors individually, the stress scenarios determine a default threshold. When the value of the risk factor crosses this threshold, it can lead to a default of the bank. It can also turn out that the threshold for a particular risk factor depends on the value of another risk factor, due to the dependency between the risk factors. These thresholds can be used as early warning signs for the banks, indicating an economic environment that threatens the viability of the bank. Furthermore, these scenarios can be used as input for a traditional stress testing model, when they uncover hidden risks for the portfolio.

The reverse stress testing model created using the framework is based on two principal components, representing the term structure of the interest rate, the S&P 500 index and the GDP. Using a copula function also the dependency between these variables, which is necessary to determine the plausibility of scenarios, is included in the model. To identify the stress scenarios a grid search is applied. To decrease the number of scenarios that need to be evaluated using this grid search, a few reduction algorithms were used. Due to computational limitations these algorithms cannot reduce the computation time at the moment. However, when more risk factors are added to the model, reduction algorithms might be necessary to perform a reverse stress test.

The framework is split in a few components, which gives the ability to alternate one of the components in a more bank-specific method when implementing the framework. This can for instance apply to the valuation of the portfolio, which in this framework is based on zero coupon bonds. When implementing the framework, this can be substituted by the valuation model used at the bank. However, including a more complex valuation method can increase the computation time of the reverse stress test significantly, so a trade-off exists. A less complicated alternative to create a more bank-specific framework is by selecting risk factors appropriate for the evaluated bank. Furthermore, the framework can also be applied to a sub portfolio instead of the full portfolio, analyzing only the risks associated with this sub portfolio. This is particularly interesting for larger banks, as they use multiple different systems within the organization to measure risk. In that case, a full scale reverse stress is hardly possible, so the various systems should be evaluated separately. Moreover, when applying this reverse stress test framework to a bank portfolio, this quantitative method should be combined with a qualitative reverse stress test. By combining these approaches, also the influence of risk factors that cannot be quantified can be measured.

The report also shows the empirical results of this framework on a few standardized portfolios, consisting of only zero coupon bonds. The created framework is applied on three portfolios, with different proportions of initially investment grade and initially speculative grade rated obligors. For the investment portfolio, mainly consisting of investment grade rated obligors, no stress scenarios are identified. In addition, the number of stress scenarios is, as expected, higher for the speculative portfolio than for the equally proportioned portfolio. The most plausible stress scenario for these portfolios contains decreasing short term interest rates, along with increasing long term interest rates. Furthermore, the S&P 500 index increases for both portfolios, which indicates a limited influence on the default of a portfolio. The GDP decreases slightly in case of the equally proportioned portfolio, while a slight increase is present for the speculative portfolio. These scenarios show that in particular the interest rate term structure needs to be shocked intensively to exhaust the capital buffer of the bank.

The complete set of stress scenarios leads to similar conclusions regarding the importance of the risk factors. While the components of the interest rate term structure and the GDP have clear thresholds a scenario needs to reach before it can lead to a default, such a threshold is not visible for the S&P 500 index. Furthermore, when summarizing the stress scenario set in a few cluster scenarios, the importance of the interest rate curve is also apparent. For the equally proportioned portfolio, each of these scenarios consist of decreasing short term interest rates, combined with increasing long term rates. In addition to this, the GDP decreases over the year in each of these cluster scenarios. For the S&P 500 index such a restriction is not applicable, although the decrease of the

GDP has to be stronger when the S&P 500 index increases. In case the proportion of initially speculative grade rated obligors in the portfolio increases, the influence of the interest curve is slightly less strong. In some of the cluster scenarios the level of the curve even increases. However, when this is the case, both the S&P 500 index and the GDP suggest a negative economic environment. The complete set of stress scenarios thus also shows that looking only at one type of risk, might lead to missing a scenario that can exhaust the capital buffer of the bank.

After this, the reverse stress testing framework is extended using a Mixed Data Sampling (MIDAS) specification. Using this specification, the explanatory variables can be incorporated in the model in a higher frequency than the dependent variable. Thus allows for more information to be added to the model, while limiting the additional parameter uncertainty. However, due to the smoothing effect of the autoregressive structure used to generate the quarterly scenario values, no stress scenarios are identified for each of the portfolio compositions. To be able to find stress scenarios, the structure of the quarterly values has to be improved to ensure extreme values for the risk factors.

8. DISCUSSION

This report presented a framework to apply a quantitative reverse stress test on the portfolio of a bank. Although this framework provides a general approach to define stress scenarios for this portfolio, there is still some room for improvement. This section discusses a number of suggestions to further improve the framework in future research.

Most of the problems concerning reverse stress testing are caused by the limited availability of data. Because for the number of defaults only yearly data is available, the model for the asset returns is estimated based on only a few observations. This causes quite a lot estimation risk in the sensitivity coefficients. A possible improvement can be found in the fact internal default data at financial institutions is often available at higher frequencies, but still the time span of the historical data is short. However, increasing the frequency of the default data also results in an increase of the probability of time periods without any defaults, especially regarding the investment grade. When a lot of time periods without defaults are present within the data, estimating the model could prove to be hard, due to statistical difficulties.

The limited number of observations also restricts the number of risk factors that can be included in the model for the asset returns. In this report only four risk factors are included in the model and incorporating more explanatory variables would quickly increase the estimation uncertainty. When more data becomes available regarding the number of defaults, also more risk factors can be included in the model. This way also variables that cover other types of risk, for instance exchange rate risk, credit risk or credit spread risk, could be added to the model. The types of risk factors selected can also differ in each situation, depending on the type of portfolio evaluated. A reverse stress test does not necessarily need to cover the full portfolio of an institution, but can also cover only a sub-portfolio. In this case some risk types might be irrelevant. Also the selecting procedure of the risk factors could be changed. This choice could be made exclusively based on statistical and quantitative measures, instead of also including a qualitative assessment. However, due to the limited number of observations, it is possible that the relation between the asset returns and the selected risk factors would not have been in agreement with the economic theory.

One additional risk factor that could possibly incorporate additional information in the model is a measure for market credit risk. The credit risk of the asset positions itself is already partially measured by including a credit spread based on the credit rating of the respective obligor. However, the general credit risk in the total market can also have an influence on the asset returns and thus on the default probability of the obligors that issued the asset positions. A potential measure for the credit risk are CDS spreads, which can provide an assessment of the credit conditions in the market (Angelini, 2012). Even though the financial crisis showed some shortcomings in the market practice and structure of CDS, these spreads can still provide a valuable insight.

One approach used to solve the problem concerning the data limitations is to apply the MIDAS specification to the asset return model. However, a few assumptions were made to simplify the implementation of this specification. First of all it was assumed that the weighting polynomial has the same structure for each of the risk factors. It would be more accurate to estimate this polynomial individually for each risk factor, as the influence of the quarterly values could differ between the different risk factors. With the current number of observations this leads to a huge amount of estimation uncertainty, as a lot of extra parameters need to be estimated. When more data becomes available, the precision of the weighting polynomial could be improved. However, when quarterly values are included in a model for yearly observations, it is questionable whether the MIDAS specification is more efficient than including all quarterly values into the model separately. Because for this structure for the weighting polynomial two parameters should be estimated for each risk factor, the reduction power of the MIDAS specification is limited in this case.

Using the MIDAS specification also has an influence on the scenario generation in the reverse stress testing framework. When this specification is applied the number of risk factor values that need to be included in a scenario is multiplied by four. This increase of scenario variables leads to an exponential increase of the number of scenarios, which results in additional dimensionality problems. In this report the dimensionality problem is solved by only including the first quarterly value in the scenario generation process. The remaining quarterly values are then simulated using an autoregressive structure. Due to the idiosyncratic term in the autoregressive structure, movements between the quarterly values are still possible, but the number of investigated paths is limited. The results of the reverse stress test show that this structure leads to a smoothing effect in the quarterly values of the risk factors, which causes the absence of extreme values in the scenario set. Due to the fact that the evaluated scenarios are less extreme, no stress scenarios are identified using the MIDAS specification in this format. Including all four quarterly values in the scenario generation process offers the opportunity to analyze more possible paths for the risk factors and ensures the presence of extreme values for the risk factors in the scenario set. However, the computational requirements this would invoke are not manageable at the moment. Another possibility of improvement of the MIDAS specification is changing the model used to simulate the quarterly values. A more sophisticated model might be able to better capture the autocorrelation in the variables. However, including more coefficients in the models for the risk factors also leads to more parameter uncertainty. So, once again, the optimal position in this trade-off needs to be found.

Another problem arising in the reverse stress testing framework is the computation time needed to perform the actual reverse stress test. In the current framework a grid search is applied to search for stress scenarios and to find the most plausible stress scenario. To employ this grid search, the portfolio value of each of the scenarios is necessary, which also requires a Monte Carlo simulation. This grid search is quite inefficient, as for a number of scenarios it is already known up front that they do not pose a threat to the viability of the bank. By applying a more focused grid search, the computation time for the full reverse stress test can be reduced. However, in this case it needs to be determined which parts of the grid need to be searched. One approach to identify the region of the joint distribution of the risk factors that is most interesting, is a Bayesian approach. Using Bayesian, expert opinion will be used to determine this region, based on economic theory. A potential downside of this approach is the fact that once again the hidden risks are ignored, as is done in traditional stress testing, because some stress scenarios might not be expected by experts. An alternative approach is employing a two-step grid search, where in the first step a grid search is applied on the full range of the grid for risk factors, but using a less detailed grid. This first step provides information on the regions of the grid that contain stress scenarios and which regions do not need further investigation. The second step then focuses on the most interesting regions by creating a more detailed grid for these regions. Furthermore, the regions that did not contain stress scenarios in the first step will be excluded from the grid search in the second step. This approach results in a finer grid for the most promising regions of the joint distribution and a more high-level grid for the other parts of the distribution.

However, when additional risk factors prove to be necessary to perform a reverse stress test on a particular portfolio, only reducing the grid used in the grid search might be insufficient. Due to the fact that the dimensionality increases exponentially when more risk factors are added to the model, also reduction methods will require a sharply increasing amount of computation time. In this case a more sophisticated technique to identify the set of stress scenarios is essential. In this new approach not only the most plausible stress scenario needs to determined, but also the remaining scenarios in this set need to be analyzed, along with their respective plausibility. This last constraint automatically complicates the search for a more advanced optimization method, as in addition to the optimal solution for the optimization problem in equation (7) also the rest of scenarios satisfying the stress scenario definition in equation (6) has to be identified.

References

S. Almon. The distributed lag between capital appropriations and expenditures. *Econometrica: Journal of the Econometric Society*, pages 178–196, 1965.

E. Andreou, E. Ghysels, and A. Kourtellos. Regression models with mixed sampling frequencies. *Journal of Econometrics*, 158(2):246–261, 2010.

E. Angelini. Credit default swaps (CDS) and their role in the credit risk market. *International Journal of Academic Research in Business and Social Sciences*, 2(1):584, 2012.

Basel Committee on Banking Supervision. Principles for sound stress testing practices and supervision. 2009.

D. Berg. Copula goodness-of-fit testing: an overview and power comparison. *The European Journal of Finance*, 15(7-8):675–701, 2009.

BofA Merrill Lynch. BofA Merrill Lynch US Corporate Master Option-Adjusted Spread© [BAMLC0A0CM]. `https://fred.stlouisfed.org/series/BAMLC0A0CM`, 2017a. Retrieved 29 May 2017.

BofA Merrill Lynch. BofA Merrill Lynch US High Yield Option-Adjusted Spread© [BAMLH0A0HYM2]. `https://fred.stlouisfed.org/series/BAMLH0A0HYM2`, 2017b. Retrieved 29 May 2017.

T. Breuer, M. Jandacka, K. Rheinberger, and M. Summer. How to find plausible, severe, and useful stress scenarios. *International Journal of Central Banking*, 5(3):205–224, 2009.

T. Breuer, M. Jandačka, J. Mencía, and M. Summer. A systematic approach to multi-period stress testing of portfolio credit risk. *Journal of Banking & Finance*, 36(2):332–340, 2012.

M. Čihák. Introduction to applied stress testing. *IMF Working Paper 07/59*, 2007.

Committee of European Banking Supervision. Guidelines on stress testing (CP32). European Banking Authority. 2009.

IV CRD. Directive 2013/36/EU of the European Parliament and of the Council of 26 June 2013 on access to the activity of credit institutions and the prudential supervision of credit institutions and investment firms, amending Directive 2002/87/EC and repealing Directives 2006/48/EC and 2006/49, 2013.

J. Dupačová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming. *Mathematical programming*, 95(3):493–511, 2003.

Financial Services Authority. Stress and scenario testing. Consultation paper 08/24. 2008.

Financial Services Authority. Stress and scenario testing: Feedback on CP08/24 and final rules. Policy statement 09/20. 2009.

C. Foroni and M.G. Marcellino. A survey of econometric methods for mixed-frequency data. 2013.

L. Freitag. Default probabilities, CDS premiums and downgrades: A probit-MIDAS analysis. 2014.

C. Genest and B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. In *Annales de l'IHP Probabilités et statistiques*, volume 44, pages 1096–1127, 2008.

C. Genest, J. Quessy, and B. Rémillard. Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33(2):337–366, 2006.

C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2):199–213, 2009.

E. Ghysels, P. Santa-Clara, and R. Valkanov. The MIDAS touch: Mixed data sampling regression models. *Finance*, 2004.

E. Ghysels, A. Sinko, and R. Valkanov. MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90, 2007.

N. Gröwe-Kuska, H. Heitsch, and W. Romisch. Scenario reduction and scenario tree construction for power management problems. In *Power tech conference proceedings, 2003 IEEE Bologna*, volume 3, pages 7–pp. IEEE, 2003.

P. Grundke. Top-down approaches for integrated risk management: How accurate are they? *European Journal of Operational Research*, 203(3):662–672, 2010.

P. Grundke. Reverse stress tests with bottom-up approaches. *The Journal of Risk Model Validation*, 5(1):71–90, 2011.

P. Grundke. Further recipes for quantitative reverse stress testing. *The Journal of Risk Model Validation*, 6(2): 81–102, 2012.

P. Grundke and K. Pliszka. A macroeconomic reverse stress test. 2015.

B.E. Hansen. Autoregressive conditional density estimation. *International Economic Review*, pages 705–730, 1994.

H. Heitsch and W. Römisch. Scenario reduction algorithms in stochastic programming. *Computational optimization and applications*, 24(2-3):187–206, 2003.

J. Huang and M. Huang. How much of the corporate-treasury yield spread is due to credit risk? *Review of Asset Pricing Studies*, 2(2):153–202, 2012.

Z. Li and C.A. Floudas. Optimal scenario reduction framework based on distance of uncertainty distribution and output performance: I. Single reduction via mixed integer linear optimization. *Computers & Chemical Engineering*, 70:50–66, 2014.

A.J. McNeil and A.D. Smith. Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, 50(3):299–308, 2012.

B. Mirzai and U. Müller. On reverse stress testing. *Intelligent risk*, 3(1):8–11, 2013.

Standard & Poor's. Default, transition and recovery: Recovery study U.S.: Piecing together the performance of defaulted instruments after the recent credit cycle. 2011.

Standard & Poor's. 2015 annual global corporate default study and rating transitions. 2016.

G. Studer. *Maximum loss for measurement of market risk.* PhD thesis, Swiss Federal Institute of Technology Zürich, 1997.

I.H. Witten, E. Frank, and M.A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques.* Elsevier, 2011.

APPENDIX A. SYMBOL OVERVIEW

Table 19: Symbols section 2.1

Symbol	**Description**
$n \in 1, ..., N$	Index for asset positions (obligor)
N	Number of asset positions (obligors)
$m \in 1, ..., M$	Index for liability positions
M	Number of liability positions
H	Risk horizon
t	Time index
$V_p(t), V_p(H)$	Total value of the portfolio (at t/at H)
$V_n(H)$	Value of asset position issued by obligor n at H
$V_m(H)$	Value of liability position m at H
i_n^0	Current credit rating of obligor n
i_n^H	Credit rating of obligor n at H
$R_n(t), R_n(H)$	Return of asset position issued by obligor n (at t/at H)
RD_i	Default threshold for rating class i
$S_{i_n^H}$	Credit spread, dependent on credit rating at H
q	Maturity
$r_q(H)$	End-of-year interest rate at H for maturity q
P_r	Recovery payment
δ_n	Recovery fraction
V_{rf}	Value of risk-free zero coupon bond with appropriate maturity
S_{bank}	Credit spread for credit rating of the bank
k	Index for risk factors
K	Number of risk factors
$\beta_{i_n^0,k}$	Sensitivity of asset return with respect to risk factor k for rating class i
$X_k(t)$	Risk factor k at t
$\sigma_{i_n^0,k}$	Standard deviation of asset return for rating class i
$\epsilon_n(t)$	Idiosyncratic risk for asset return of obligor n
$N_i(t)$	Total number of issuers in rating class i at t
$d_i(t)$	Number of defaults in rating class i at t
$\Phi(\cdot)$	CDF of the standard normal distribution
B_0	Initial capital buffer
B_H	Required capital buffer at H
$\boldsymbol{\Omega}^*$	Subset containing stress scenarios
$\boldsymbol{\Omega}$	Set of all scenarios
ω	Scenario
α	Confidence level
$P(\omega)$	Probability of scenario ω

Table 20: Symbols section 2.2

Symbol	**Description**
C_j	j-th principal component of interest rate term structure
$\mathbf{c_j}$	Coefficients for the j-th principal components
$\mathbf{\Sigma}$	Covariance matrix of interest rate term structure
Z	Number of principal components
$r_q(t)$	End-of-year interest rate at t for maturity q

Table 21: Symbols section 2.3

Symbol	**Description**
k	Index for risk factors
μ_k	Sample mean for risk factor k
σ_k	Sample standard deviation for risk factor k
x_k	Grid point for risk factor k
ω	Scenario
D	Kantorovich distance
s, j, l	Indexes for scenarios
J	Index set of deleted scenarios
D_{sj}	Distance between scenarios s and j
ω_s	Scenario
p_j	Probability of scenario ω_j
i	Index for steps in the reduction algorithms
j^i	Deleted scenario (Backwards reduction) or added scenario (Forward selection) in step i
p_j^*	New probability of scenario ω_j
$J(j)$	Subset of deleted scenarios that have scenario ω_j as closest scenarios still in scenario set
R_{max}^{error}	Difference in maximum output value between full scenario set and reduced set
R_{min}^{error}	Difference in minimum output value between full scenario set and reduced set
R_{exp}^{error}	Difference in expected output value between full scenario set and reduced set
R_{max}^{full}	Maximum output value in full scenario set
R_{min}^{full}	Minimum output value in full scenario set
R_{exp}^{full}	Expected output value in full scenario set
R_{max}^{new}	Maximum output value in reduced scenario set
R_{min}^{new}	Minimum output value in reduced scenario set
R_{exp}^{new}	Expected output value in reduced scenario set
R_j	Asset return for Investment grade rated issuer in scenario ω_j

Table 22: Symbols section 2.4

Symbol	**Description**
k	Index for risk factors
m	Ratio between frequency of explanatory variables and frequency of the dependent variable
$X_k^{(m)}$	Weighted value for risk factor k
$w_j(\theta)$	Polynomial function used to weight quarterly risk factor values
L	lag-operator
$X_k^{(m)(t)}$	Value for risk factor k on quarterly basis at t
θ_1	Parameters in weighting polynomial
Q	Number of parameters in weighting polynomial
$R_n(t)$	Return of asset position issued by obligor n at t
RD_i	Default threshold for rating class i
$N_i(t)$	Total number of issuers in rating class i at t
$d_i(t)$	Number of defaults in rating class i at t
α_k	Constant in AR(1) model for risk factor k
ϕ_k	Autoregressive sensitivity coefficient for risk factor k
q	Index for number of quarters into the future
u_{t+1q}	Error term

Table 23: Symbols section 3.1

Symbol	**Description**
JB	Jarque-Bera test statistic
K	Kurtosis
S	Skewness
n	Number of observations
μ	Location parameter in skewed student-t distribution
σ	Scale parameter in skewed student-t distribution
η	Degrees of freedom in skewed student-t distribution
λ	Extra skewness parameter in skewed student-t distribution
$\Gamma(\cdot)$	Gamma distribution

Table 24: Symbols section 3.2

Symbol	**Description**
K	Number of risk factors in asset return model
k	Index for risk factors
$F(x_{1t},...,x_{Kt}$	Joint distribution of risk factors
$\hat{F}_k(x_{kt})$	Estimated marginal distribution for risk factor k
T	Sample size
$\hat{T}_{CM}$	Cramér-von Mises test statistic
$\hat{C}(u)$	Empirical copula
$C_1(u)$	Estimated copula
θ	Copula parameter(s)
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
l_C	Log-likelihood of the estimated copula
k_C	Number of parameters in estimated copula