

# Empirical implementation of a quantitative reverse stress test for defaultable fixed-income instruments with macroeconomic factors and principal components

Peter Grundke\*      Kamil Pliszka†

This Version: April 2013

## Abstract

This paper provides a quantitative reverse stress test approach that takes macroeconomic risk factors into account. We use principal component analysis to describe the movements of the term structure of risk-free interest rates, which, in combination with a latent systematic risk factor and an economic indicator, serve as risk factors that drive the obligors' asset returns. The sensitivities of the asset returns towards these risk factors are estimated empirically and the univariate and multivariate distribution of the risk factors is analyzed. The proposed reverse stress test evaluates the whole risk factor space and finds those scenarios that exactly lead to a presumed loss. In the last step the most plausible of these scenarios is determined and discussed. The results show that the found reverse stress test scenarios are particularly reasonable for the assumed bank portfolio. However, the results also show that reverse stress tests are exposed to considerable model and estimation risk which makes numerous robustness checks necessary.

Keywords: bottom-up approach, copula functions, extreme value theory, principal component analysis, reverse stress testing

---

\*Osnabrueck University, Katharinenstraße 7, 49069 Osnabrueck, Germany, Tel.: ++49 (0)541 969 4721, Fax: ++49 (0)541 969 6111, E-Mail: peter.grundke@uni-osnabrueck.de

†Osnabrueck University, Katharinenstraße 7, 49069 Osnabrueck, Germany, Tel.: ++49 (0)541 969 6115, Fax: ++49 (0)541 969 6111, E-Mail: kamil.pliszka@uni-osnabrueck.de

# 1 Introduction

As a reaction to the financial crisis 2007-2009, regulatory authorities have strengthened the importance of stress test methodologies. Particularly, the role of reverse stress tests was highlighted after a number of consultative papers of the Financial Services Authority (FSA (2008, 2009)) and the Committee of European Banking Supervisors (CEBS (2009, 2010)). Large banks are expected to perform reverse stress tests in a quantitative way. However, up to now, no appropriate standard for this kind of stress test has evolved and even the number of (at least published) proposals how such a test could be performed at all is very limited.

In regular stress tests, adverse scenarios are chosen upon historical observations or expert knowledge. Thus, although the choice may be reasonable, the employed scenarios remain arbitrary. In contrast, in reverse stress tests, exactly those scenarios are looked for that lead to a very unfavourable event for a bank (e. g., a very large (expected) loss, a non-fulfillment of the capital adequacy requirements or illiquidity). In the next step, the most plausible of these scenarios has to be found and evaluated by the bank's senior management (see CEBS (2009, p. 14)). Čihák (2007) calls this the "threshold approach". Reverse stress testing is mathematically and conceptually challenging, in particular, when many risk factors are relevant for the value of the bank's portfolio and when this portfolio is structured in a complex way with many different assets and types of financial instruments. For  $n$  risk factors, specific scenarios out of  $\mathbb{R}^n$  have to be found when solving the inversion problem inherent in a reverse stress test and, for each single scenario, the corresponding probability of occurrence has to be computed. Therefore, the number of used risk factors has to be kept low and a framework has to be chosen that remains numerical tractable for more sophisticated portfolios.

Most of the literature dealing with macroeconomic regular stress tests for credit risk is based on the idea of Wilson (1997a, 1997b) and extensions thereof. Within this type of models, macroeconomic variables are looked for that can explain the systematic variation of default rates across time (see, e. g., Boss (2002), Sorge and Virolainen (2006)). The current body of literature on reverse stress tests is still sparse. A discussion of a qualitative approach based on fault trees has been presented by Grundke (2012b). However, the essential conclusion of this paper is that a qualitative approach alone would not work, but, at least, would have to be supported by quantitative elements. Füsler et al. (2012a, 2012b) present a very general operating plan for (mainly qualitative) reverse

stress tests. Papers on quantitative reverse stress tests are also very rare. One approach is developed by Grundke (2011). Employing ideas from integrated risk management, he uses a bottom-up model based on CreditMetrics with correlated interest rates and rating-specific credit spreads. Later, in Grundke (2012a), this approach was expanded by more realistic assumptions, including, among others, contagion effects between single obligors and a time-varying bank rating. Drüen and Florin (2010) argue in a similar vein as Grundke (2011), but they do not use a full-fledged bottom-up approach. Instead, they rather employ two separate approaches for interest rate risk and default risk and additionally, some exogenous (not further described) functional relationship between shifts of the term structure of risk-free interest rates and the obligors' default probabilities. Furthermore, there are some case studies for simply structured portfolios with one or two risk factors (see, e. g., Liermann and Klauck (2010)). Beside this, a dimension reduction technique that yields the most relevant (based on information criteria) risk factors of a portfolio has been proposed by Skoglund and Chen (2009). The most recent contribution by McNeil and Smith (2012) introduces the concept of depth to identify the most plausible reverse stress test scenario which is called the most likely ruin event (MLRE). In a related strand of stress test literature, the worst (in the sense of expected losses for a given portfolio) scenario from a set of scenarios with a given plausibility (for example measured by the Mahalanobis-distance) is looked for. Čihák (2007) calls this the "worst case approach". An example for this approach is Breuer et al. (2008).

Our approach picks up ideas from the framework of Grundke (2011, 2012a). However, instead of performing simulation studies, we show how a quantitative reverse stress test can be implemented empirically. Furthermore, we propose to use principal component analysis for reducing the number of risk factors relevant for fixed-income portfolios. This specification keeps the dimensionality of the model low and, hence, allows us to specify and calibrate a full reverse stress test framework.

The remainder of the paper is structured as follows. In Section 2, the principal component analysis for the term structure of risk-free interest rates is carried out and the linear factor model describing the asset returns of the bank's obligors is estimated by maximum-likelihood. In Section 3, the univariate margins of the risk factors and the multivariate dependence structure are analyzed. In Section 4, it is demonstrated how a reverse stress test could be performed for a stylized fixed-income portfolio. Finally, Section 5 concludes.

## 2 Principal components and model estimation

Similar to Grundke (2011, 2012a), we assume that the credit quality of the bank's obligors is driven by their asset returns and that these asset returns are correlated with the risk-free interest rates. Furthermore, we assume that additionally, a latent systematic credit risk factor, a macroeconomic indicator and an idiosyncratic risk factor influence each obligor's asset return. The complete linear factor model for the asset return  $R_{n,t}$  of obligor  $n$ ,  $n \in \{1, \dots, N\}$ , within the time period  $[t, t + 1)$  is assumed to be given by

$$R_{n,t} = \sqrt{\rho_{n,R}} \cdot Z(t) + \alpha_n \cdot X(t) + \sum_{j=1}^p \rho'_{n,C_j,R} \cdot C_j(t) + \sqrt{1 - \rho_{n,R}} \cdot \epsilon_{n,t} \quad (2.1)$$

where  $Z(t)$  is an i.i.d. standard normally distributed random variable representing latent systematic credit risk,  $X(t)$  denotes the economic indicator (which later is assumed to be the log-return of U. S. GDP and the log-return of the S&P 500, respectively), and  $C_j(t)$ ,  $j \in \{1, \dots, p\}$  represents the principal components of the term structure of risk-free interest rates. The variable  $\epsilon_{n,t}$  denotes the idiosyncratic risk of obligor  $n$  at time  $t$  and is assumed to be an i.i.d. standard normally distributed random variable.

In order to keep the number of risk factors low, we apply principal component analysis to explain the movements of the term structure of risk-free interest rates. Principal component analysis reduces the dimensional complexity of a dataset by an orthogonal linear transformation of the original data into a new orthogonal space. The algorithm of this transformation can be described as follows:<sup>1</sup> First, a variance-maximizing linear combination of unit length representing the first principal component is obtained. Next, the remaining variance is computed and the second principal component is chosen. This is done by maximizing the remaining variance under the restriction of orthogonality to the first principal component and the standardization of unit length. This step is iterated until the whole orthogonal space is computed (this implies the same dimensionality as the one of the original dataset). The goal is to use just the first  $p$  principal components that explain a sufficient amount of the variance. Empirical studies show that in the case of the term structure of risk-free interest rates, the first two or three principal components are able to capture almost the whole variance of returns on fixed-income securities.<sup>2</sup> It can be shown that the described algorithm is equivalent to calculating the

---

<sup>1</sup>See Golub and Tilman (2000, pp. 97-98).

<sup>2</sup>See Golub and Tilman (2000, p. 94) and in detail Litterman and Scheinkman (1991); Knez et al. (1994).

variance-covariance matrix and the corresponding eigenvectors and eigenvalues whereby the eigenvectors equal the principal components. With the help of the eigenvalues, we can observe the explained fraction of variance and, therefore, determine the number of principal components to use.

Let  $r_q$ ,  $q \in \{1, 2, \dots, m\}$ , be the yield-to-maturity with time to maturity  $t_q$ . Then the  $j$ -th principal component is given by

$$C_j = \sum_{q=1}^m c_{j,q} \cdot \Delta r_q \quad (2.2)$$

where  $\Delta r_q$  denotes the change of the  $q$ -th interest rate and  $c_{j,q}$ ,  $q \in \{1, 2, \dots, m\}$ , denotes the coefficients of the  $j$ -th principal component. Due to the assumed orthogonality of the matrix of coefficients of the principal components, the interest rate changes  $\Delta r_q$ ,  $q \in \{1, 2, \dots, m\}$ , are given by linear combinations of the coefficients

$$\Delta r_q = \sum_{j=1}^m c_{q,j} \cdot C_j. \quad (2.3)$$

Further, it can be stated that the coefficients  $c_{j,1}, \dots, c_{j,m}$  are equal to the eigenvector that corresponds to the  $j$ -th eigenvalue of the variance-covariance matrix  $\Sigma$ . Thus,

$$\Sigma \cdot \begin{pmatrix} c_{j,1} \\ \dots \\ c_{j,m} \end{pmatrix} = \lambda_j \cdot \begin{pmatrix} c_{j,1} \\ \dots \\ c_{j,m} \end{pmatrix} \quad (2.4)$$

holds true. The eigenvalues are given by the solution of

$$\det(\Sigma - \Lambda \cdot \mathbf{I}) = 0 \quad (2.5)$$

where  $\mathbf{I}$  denotes a  $m$ -dimensional identity matrix and  $\Lambda$  a matrix with the eigenvalues on the main diagonal and zeros otherwise. Further, it can be shown that the eigenvalue is equal to the variance of the corresponding principal component

$$\lambda_j = \sigma^2(C_j). \quad (2.6)$$

For estimating the principal components, we use annually obtained yields of U. S. Treasury Bills (3M, 6M, 1Y) and U. S. Treasury Bonds (2Y, 3Y, 5Y, 7Y, 10Y, 30Y) ranging from

1983 to 2010. The data is provided by Datastream. To ensure stationarity, we calculate percentage changes<sup>3</sup>  $\frac{r_{q,t}-r_{q,t-1}}{r_{q,t-1}}, q \in \{1, \dots, m\}, \forall t \in \{2, \dots, T\}$ . The resulting variance-covariance matrix is shown in Table 1.

	3M	6M	1Y	2Y	3Y	5Y	7Y	10Y	30Y
3M	0.2314								
6M	0.2133	0.2183							
1Y	0.1712	0.1815	0.1690						
2Y	0.1193	0.1299	0.1305	0.1088					
3Y	0.0929	0.0995	0.1084	0.0952	0.0874				
5Y	0.0591	0.0612	0.0752	0.0712	0.0694	0.0597			
7Y	0.0478	0.0452	0.0653	0.0649	0.0670	0.0610	0.0666		
10Y	0.0354	0.0343	0.0504	0.0516	0.0535	0.0498	0.0539	0.0448	
30Y	0.0238	0.0194	0.0354	0.0379	0.0417	0.0410	0.0472	0.0390	0.0361

Table 1: Variance-covariance matrix for percentage changes of yield-to-maturities for different times to maturity.

For the variance-covariance matrix exhibited in Table 1, the eigenvectors and eigenvalues are calculated. The following matrix shows the eigenvectors column by column:

$$\begin{pmatrix} 0.4818 & 0.4104 & 0.6725 & -0.2915 & 0.1587 & 0.0024 & 0.1455 & -0.1150 & -0.0507 \\ 0.4859 & 0.3749 & -0.1861 & 0.3141 & -0.5137 & 0.0638 & -0.3888 & 0.2349 & 0.1257 \\ 0.4491 & 0.0360 & -0.3701 & 0.3877 & 0.5480 & -0.3315 & 0.3154 & 0.0065 & 0.0110 \\ 0.3514 & -0.1602 & -0.3606 & -0.2603 & -0.1269 & 0.3376 & 0.0191 & -0.4393 & -0.5741 \\ 0.2975 & -0.2796 & -0.1969 & -0.4245 & 0.1004 & 0.2905 & 0.0048 & 0.0055 & 0.7210 \\ 0.2144 & -0.3525 & 0.0255 & -0.3013 & -0.2506 & -0.2687 & 0.2835 & 0.6750 & -0.2595 \\ 0.1924 & -0.4432 & 0.2378 & 0.0644 & 0.3514 & -0.1024 & -0.7377 & 0.0683 & -0.1518 \\ 0.1508 & -0.3750 & 0.1957 & 0.1516 & -0.4426 & -0.5074 & 0.1106 & -0.5177 & 0.2044 \\ 0.1111 & -0.3528 & 0.3343 & 0.5479 & -0.0480 & 0.5895 & 0.3018 & 0.1015 & -0.0024 \end{pmatrix}.$$

The corresponding eigenvalues are:

$$\left( 0.8125 \quad 0.1761 \quad 0.0283 \quad 0.0031 \quad 0.0018 \quad 0.0002 \quad 0.0001 \quad 0.0001 \quad 0.0000 \right).$$

According to the Kaiser criterion,<sup>4</sup> which recommends to use, in case of a variance-covariance matrix, principal components with an eigenvalue exceeding the mean of the eigenvalues, we use the first two principal components as risk factors for the reverse stress test (instead of all yield-to-maturities with different times to maturity). They explain an

<sup>3</sup>Otherwise, the null hypothesis that the time series contain unit roots cannot be rejected at reasonable significance levels by the ADF-test.

<sup>4</sup>See Kaiser (1960).

amount of 96.72% of the total variance; the first three principal components would have explained 99.49%.<sup>5</sup> Figure 1 visualizes the first three principal components for times to maturity ranging from 3 months to 30 years (corresponding to the coefficients  $c_{j,q}$  for  $j \in \{1, 2, 3\}$  in Equation 2.2).

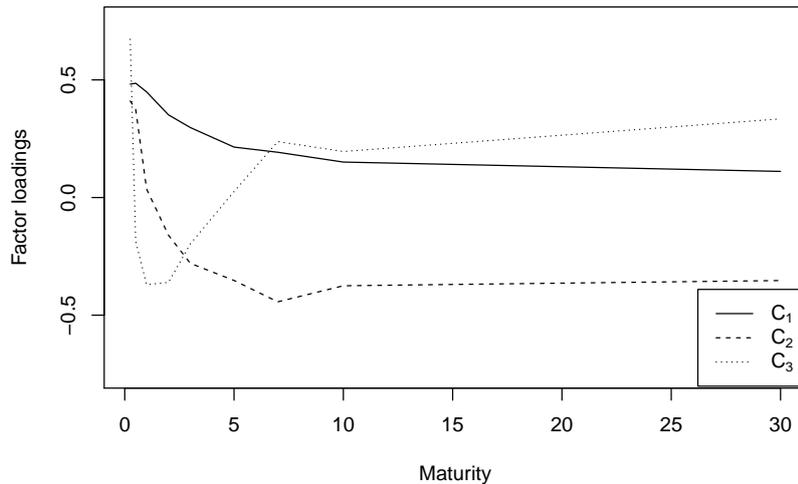


Figure 1: Principal components and their impact on interest rates for different times to maturity.

The principal components possess an economic interpretation:<sup>6</sup> The first principal component is a weighted sum of interest rate changes with the same sign for all maturities and can be interpreted as the level of the change of the term structure. The second principal component weights interest rate changes for short maturities with a positive sign and interest rate changes for long maturities with a negative sign and, thus, can be understood as the slope of the interest rate curve. The third principal component, on the one hand, associates positive signs with short-term and long-term interest rate changes and, on the other hand, negative signs with medium-term interest rate changes. Therefore, it can be interpreted as a measure of the curvature.

After determining the number of relevant principal components, which is represented by the variable  $p$ , we estimate the risk factor sensitivities in the asset return

<sup>5</sup>The third principal component is mentioned and visualized due to the fact that studies modeling stochastic movements of the term structure of risk-free interest rates by principal components use it (see, e. g., Litterman and Scheinkman (1991); Knez et al. (1994); Heidari and Wu (2003)). Nevertheless, for the later reverse stress test, we omit it for three reasons: First, the Kaiser criterion proposes to use only the first two principal components. Second, the maximum-likelihood estimation with an additional risk factor would have been more complex, and third, the evaluation of the risk factor space would have required higher computational effort.

<sup>6</sup>See Litterman and Scheinkman (1991, pp. 57-58).

Equation 2.1. The default data is taken from the annual default report of Standard & Poor's (2011a). As the historical default rates for higher (less risky) rating grades are low and partly zero, different sensitivities are estimated only for the two broad rating categories, Investment Grade and Speculative Grade. The historical default rates for these two broad rating categories are shown in Figure 2.

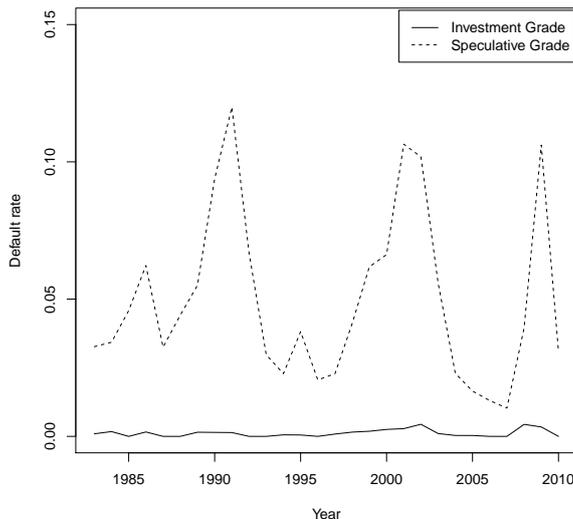


Figure 2: Historical default rates from 1983 to 2010.

For the two broad rating classes  $i \in \{1, 2\} = \{\text{Investment Grade}, \text{Speculative Grade}\}$ , the sensitivity vector

$$\left( \rho_{i,R} \quad \alpha_i \quad \rho'_{i,C_1,R} \quad \rho'_{i,C_2,R} \right)$$

and the default barrier  $R_3^i$  are estimated by maximum-likelihood. The principal components are calculated from empirical observations for interest rate percentage changes analogously to Equation 2.2

$$C_j(t) = \sum_{q=1}^m c_{j,q} \cdot \Delta r_{q,t}. \quad (2.7)$$

The log-likelihood function is given by<sup>7</sup>

$$l_i = \sum_{t=1}^T \ln \int_{-\infty}^{+\infty} \binom{N_{i,t}}{d_{i,t}} q_i(z, x(t), c_1(t), c_2(t))^{d_{i,t}} (1 - q_i(z, x(t), c_1(t), c_2(t)))^{N_{i,t} - d_{i,t}} \phi(z) dz, \quad (2.8)$$

<sup>7</sup>Estimating factor loadings in linear factor models for asset returns by maximum-likelihood (based on default data) is a frequently employed approach in the credit risk literature (see, e. g., Frey and McNeil (2003) and Hamerle and Rösch (2006)).

with the rating-specific conditional default probability<sup>8</sup>

$$\begin{aligned}
q_i(z, x(t), c_1(t), c_2(t)) &:= P(R_n \leq R_3^i | Z = z, X = x(t), C_1 = c_1(t), C_2 = c_2(t)) \\
&= \Phi\left(\frac{R_3^i - \sqrt{\rho_{i,R}}z - \alpha_i x(t) - \rho'_{i,C_1,R} \cdot c_1(t) - \rho'_{i,C_2,R} \cdot c_2(t)}{\sqrt{1 - \rho_{i,R}}}\right).
\end{aligned}
\tag{2.9}$$

The above integral is solved using methods of adaptive quadrature.<sup>9</sup>  $\phi(z)$  ( $\Phi(z)$ ) is the (cumulative) density function of a standard normally distributed random variable.  $N_{i,t}$  describes the number of obligors with rating grade  $i$  at time  $t$  and  $d_{i,t}$  is the number of defaults of obligors with rating grade  $i$  at time  $t$  within the period  $[t, t + 1)$ . The log-returns of the U. S. GDP and the S&P 500, respectively, within the period  $[t, t + 1)$  serve as the economic indicator  $X(t)$ . The data is obtained from Datastream and covers the period from 1983 to 2010.

The results of the estimation procedure for the asset return equations are summarized in Table 2.<sup>10</sup>

		$Z(t)$	$X(t)$	$C_1(t)$	$C_2(t)$
GDP	Investment Grade	0.0383 (1.61)	3.3087 (1.56)	0.1749*** (2.60)	0.2524** (1.96)
	Speculative Grade	0.0557*** (3.64)	7.8860** (3.10)	0.0963* (1.88)	0.1925* (1.73)
S&P 500	Investment Grade	0.0200 (1.16)	0.6643** (2.42)	0.1056 (1.59)	0.3217*** (3.01)
	Speculative Grade	0.0583*** (3.63)	0.0881 (0.30)	0.1116** (2.00)	0.2563** (2.28)

Table 2: Coefficients for asset returns as specified in Equation 2.1 using GDP and S&P 500 data as the macroeconomic variable  $X(t)$  for Investment Grade and Speculative Grade. The  $t$ -statistics are presented in parentheses. The symbols \*, \*\* and \*\*\* denote significance at 10%, 5% and 1% level.

For Investment Grade as well as for Speculative Grade, the sign for the macroeconomic index  $X(t)$  make economic sense. For the relationship between asset returns and interest rates, especially principal components of the interest rate, it not obvious which sign would make economic sense: On the one hand, increased interest rates lead to more expensive loans and therefore should be negatively related to asset returns. On the other hand,

<sup>8</sup>An additional constraint  $\rho_{i,R} \in (0, 1)$  ensures that we do not divide by zero or compute the square root of negative value.

<sup>9</sup>The implementation is done using the function `int` of the program R which is based on the Gauss-Kronrod quadrature (see Kronrod (1965)).

<sup>10</sup>The maximization was done using the function `constrOptim` in R and is regarded as numerically stable. Calculations were performed using the Nelder-Mead method (see Nelder and Mead (1965)) with different initial values. Numerical issues due to the improper integral were considered, too. For the integral  $\int_{-\infty}^{+\infty} \binom{N_{i,t}}{d_{i,t}} q_i(z, x(t), c_1(t), c_2(t))^{d_{i,t}} (1 - q_i(z, x(t), c_1(t), c_2(t)))^{N_{i,t} - d_{i,t}} \phi(z) dz$ , we substituted  $y = \Phi(z)$  and  $\frac{dy}{dz} = \phi(z)$ , respectively. This leads to the expression  $\int_0^1 \binom{N_{i,t}}{d_{i,t}} q_i(\Phi^{-1}(y), x(t), c_1(t), c_2(t))^{d_{i,t}} (1 - q_i(\Phi^{-1}(y), x(t), c_1(t), c_2(t)))^{N_{i,t} - d_{i,t}} dy$ . The following optimization delivered the same result as that one using the improper integral.

raising (short-term) interest rates by central banks is a tool to slow booming economies down in order to control inflation. This explanation is in line with our estimation result in which an increase of the first and the second principal component leads to increased (short-term) interest rates and is positively related to asset returns. The significance of the variables depends on the model specification. Finding significant variables for Investment Grade obligors proves as rather difficult and only two risk factors can be stated as statistically significant in each of the two specifications.<sup>11</sup> The situation is different, however, for Speculative Grade obligors. All risk factors have a significant impact when using the specification with GDP-log-return whereas three variable prove as significant in the S&P 500 specification.

### 3 Marginal distributions of the systematic risk factors and multivariate dependence

For computing the probabilities of occurrence for the reverse stress test scenarios, we need the multivariate probability distribution of the systematic risk factors of the model. These are the latent systematic credit risk factor  $Z$ , the U. S. GDP-log-return and the S&P 500-log-return  $X$ , respectively, and the first two principal components  $C_1$  and  $C_2$  of the term structure of risk-free interest rates. First, we test the null hypothesis of normality for the log-returns of the GDP, the S&P 500 and the first two principal components by means of the Kolmogorov-Smirnov test and the Jarque-Bera test.<sup>12</sup> The empirical data is visualized in Figure 3 using a QQ plot. While normality seems to be justified in the center of the distribution, the tails differ much from this assumption.

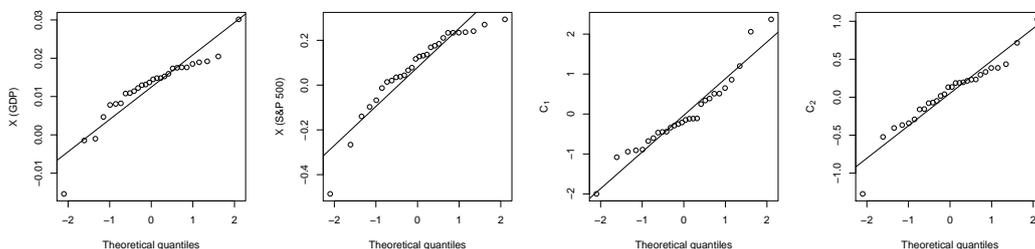


Figure 3: Quantiles of the empirical distribution function are plotted against quantiles of the normal distribution.

<sup>11</sup>As mentioned before, the coefficients are numerically stable and have the correct sign, but the default data contains some observations without defaults. For this reasons, we use this specifications, but we will draw conclusions carefully.

<sup>12</sup>The latent systematic credit risk factor  $Z$  is assumed to be standard normally distributed.

As Table 3 shows, the results of the visual inspection are only partly confirmed by the statistical tests. While the Kolmogorov-Smirnov test does not reject the normality assumption, the Jarque-Bera test rejects the null hypothesis for the GDP-log-return, the S&P 500-log-return and the second principal component at the 1% and 5% level, respectively. The Jarque-Bera test calculates skewness and kurtosis of the empirical data and carries them into the test statistic and, hence, quickly rejects normality in case of supposed fat tails.<sup>13</sup>

	X (GDP)	X (S&P 500)	$C_1$	$C_2$
D	0.1719	0.1397	0.1764	0.1132
$p$ -value(D)	0.34	0.5964	0.3108	0.8266
JB	17.8638***	15.5634***	3.6383	7.3797**
$p$ -value(JB)	0.0001321	0.0004173	0.1622	0.02498

Table 3:  $p$ -values and test statistics of the Kolmogorov-Smirnov test and the Jarque-Bera test for empirical observations of the risk factors. The symbols \*, \*\* and \*\*\* denote significance at 10%, 5% and 1% level.

In order to take into account this kind of model risk when carrying out the reverse stress test, we proceed as follows. On one hand, we assume normality of all four systematic risk factors. While the latent systematic credit risk factor  $Z$  is assumed to be standard normally distributed, the mean and the variance of the other systematic risk factors are estimated from the data by the method of moments. This yields for the mean

$$\hat{\boldsymbol{\mu}} = \left( \hat{\mu}_{X(\text{GDP})} \quad \hat{\mu}_{X(\text{S\&P 500})} \quad \hat{\mu}_{C_1} \quad \hat{\mu}_{C_2} \right) = \left( 0.0124 \quad 0.0792 \quad -0.0330 \quad 0.0512 \right) \quad (3.1)$$

and for the variance

$$\hat{\boldsymbol{\sigma}}^2 = \left( \hat{\sigma}_{X(\text{GDP})}^2 \quad \hat{\sigma}_{X(\text{S\&P 500})}^2 \quad \hat{\sigma}_{C_1}^2 \quad \hat{\sigma}_{C_2}^2 \right) = \left( 7.1 \cdot 10^{-5} \quad 0.0304 \quad 0.8125 \quad 0.1761 \right). \quad (3.2)$$

On the other hand, we employ extreme value theory to take extreme tail events into account. More precisely, we use methods based on threshold exceedances for the tails of those risk factors for which normality was rejected by the Jarque-Bera test. Tail events are especially important for us since we want to capture extreme scenarios and calculate downside risk measures. The Jarque-Bera test rejects normality for the GDP-log-return, S&P 500-log-return and for the second principal component. To take this into account, we assume the left tail of the distribution of GDP-log-returns, S&P 500-log-return, and

<sup>13</sup>Figure 3 shows that the data for GDP-log-returns includes exactly one outlier (realization in 2009), the same is true for the data of the S&P 500-log-returns (realization in 2008) and the second principal component (realization in 2009) that also include one outlier. When omitting these outliers, we could not reject normality for all risk factors at reasonable significance levels.

both tails of the distribution of the second principal component to follow the generalized Pareto distribution (GPD).<sup>14</sup>

The GPD quantifies the conditional distribution of excesses of a random variable  $X$  over a threshold  $u$  and is given by <sup>15</sup>

$$P(X - u \leq y | X > u) = G_{\xi, \beta}(y) = \begin{cases} 1 - \left(1 + \frac{\xi y}{\beta}\right)^{-\frac{1}{\xi}} & , \xi \neq 0 \\ 1 - \exp\left\{-\frac{y}{\beta}\right\} & , \xi = 0 \end{cases} \quad (3.3)$$

where  $\beta > 0$  is referred to as the shape and  $\xi$  as the scale parameter. In case of  $\xi > 0$ , fat tails are present.

The GPD tail and the normally distributed center are connected by the threshold  $u$  which is determined by mean excess plots.<sup>16</sup> The threshold  $u$  has to be chosen in such a way that the graph of the mean excess function for  $u' > u$  is (approximately) linear.<sup>17</sup> Figure 4 shows the mean excess plots for the left tail of the GDP-log-return, for the left tail of the S&P 500-log-return and for the left as well as right tail of the second principal component.

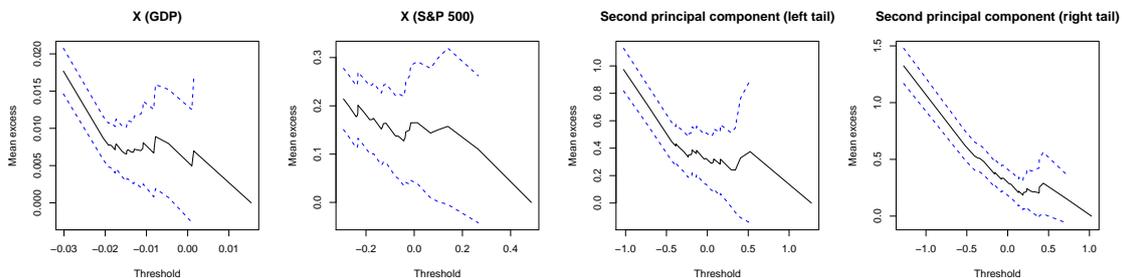


Figure 4: Mean excess plots for left tail of GDP-log-return (first), for left tail of S&P 500-log-return (second) and the left as well as right tail of the second principal component (third, fourth). The dashed line indicates the 95% confidence level. Data for the GDP, S&P 500 and the left tail of the second principal component was transformed by multiplication with  $-1$ .

<sup>14</sup>Modeling the right tail of the distribution of the GDP-log-return and the S&P 500-log-return is not necessary because we are interested in scenarios generating a sufficiently large loss. Thus, due to the positive sign of the asset return sensitivity with respect to the GDP-log-return and the S&P 500-log-return, large GDP or S&P 500-log-return increases are less relevant. The second principal component, in contrast, has an ambiguous effect on losses because it weights interest rate changes with a short time to maturity with a positive sign and interest rate changes with a long time to maturity with a negative sign. The net effect depends on the portfolio sensitivities towards interest rates for different times to maturity and, therefore, both tails should be modeled by the GPD.

<sup>15</sup>See McNeil et al. (2005, p. 275).

<sup>16</sup>These are graphs that map for every  $u$  a mean excess function  $\mathbb{E}[X - u | X > u]$  (see, e. g., Ghosha and Resnick (2010)). For an application, see, e. g., Gourié et al. (2009).

<sup>17</sup>This is required due to the linearity of the mean excess function of the GPD.

In Figure 4, it is notable that a threshold of around 0.00 for the GDP-log return and a threshold of around  $-0.2$  for the S&P 500-log return are reasonable choices.<sup>18</sup> For the tails of the second principal component, the excess return function seems to be linear when passing 0.5 ( $-0.5$ , respectively). As the dataset consists of only 28 observations, we have to choose the thresholds in such a way that on one hand, they match with the mean excess plots, and on the other hand, that estimation yields plausible results for the parameters of the GPD. For this, the estimation should be based on at least three observations. These considerations let us choose the threshold  $u = 0.00$  (GDP-log-return, left tail),  $u = -0.13$  (S&P 500-log-return, left tail),  $u^l = -0.35$  (second principal component, left tail) and  $u^r = 0.35$  (second principal component, right tail). The parameters of the GPD are shown in Table 4.<sup>19</sup>

	$\xi$	$\beta$
$X$ (GDP, left tail)	0.5703	0.0032
$X$ (S&P 500, left tail)	0.2139	0.1315
$C_2$ (left tail)	0.7779	0.1196
$C_2$ (right tail)	0.2257	0.1900

Table 4: Estimated parameters of GPD.

The resulting cumulative density function  $F_2(x)$  for the GDP-log return and the S&P 500-log-return, respectively, is given by

$$F_2(x) = \begin{cases} \Phi(u) \left(1 + \xi \frac{|x-u|}{\beta}\right)^{-\frac{1}{\xi}} & , x < u \\ \Phi(x) & , x \geq u . \end{cases} \quad (3.4)$$

The resulting cumulative density function  $F_4(c_2)$  for the second principal component is

$$F_4(c_2) = \begin{cases} \Phi(u^l) \left(1 + \xi \frac{|c_2-u^l|}{\beta}\right)^{-\frac{1}{\xi}} & , c_2 < u^l \\ \Phi(c_2) & , u^l \leq c_2 \leq u^r \\ 1 - (1 - \Phi(u^r)) \left(1 + \xi \frac{c_2-u^r}{\beta}\right)^{-\frac{1}{\xi}} & , c_2 > u^r . \end{cases} \quad (3.5)$$

QQ plots for the calibrated GPD are shown in Figure 5.

<sup>18</sup>The data was transformed by multiplication with  $-1$ .

<sup>19</sup>The parameters for the GDP-log-return were estimated by maximum likelihood and for S&P 500-log-return by probability weighted moment method.

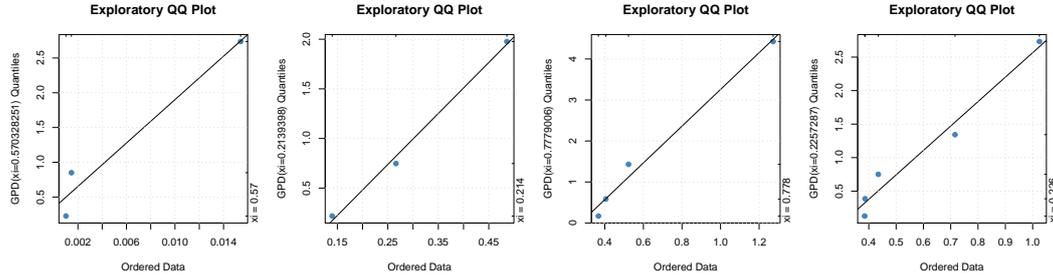


Figure 5: QQ plots for the left tail of the GDP-log-return (first), for the left tail of the S&P 500-log-return (second) and the left as well as the right tail of the second principal component (third, fourth).

In Figure 6, we visualize the left GPD tail and, in comparison, the left tail of a normal distribution, estimated with GDP-log-return and S&P 500-log-return data, respectively.

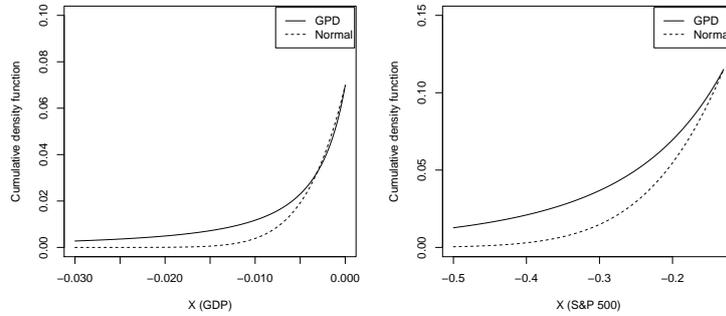


Figure 6: Left tail of the GDP-log-return and the S&P 500-log-return in comparison when modeled with GPD and normal distribution.

For extreme realizations, the GPD assigns higher probabilities than the normal distribution. At the threshold, in order to obtain a continuous cumulative density function, the GPD equals the cumulative density function of the normal distribution. A comparison of the tails of the second principal component when modeled with the GPD and the normal distribution, respectively, can be seen in Figure 7.

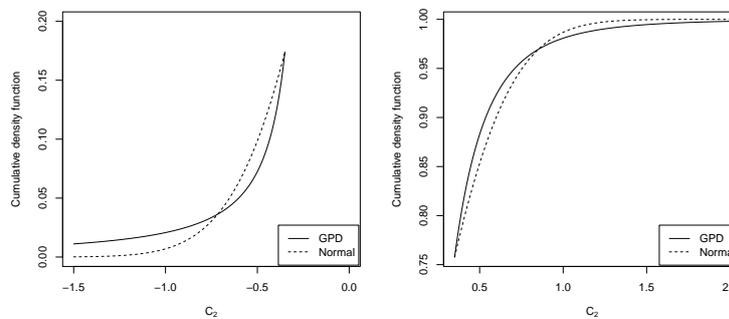


Figure 7: Left and right tail of the second principal component in comparison when modeled with the GPD and the normal distribution.

The figure shows that the GPD assigns higher probabilities for extreme realizations than the normal distribution would do on both tails.

Next, we analyze the empirical dependence structure between the systematic risk factors. For this, we do not have to take into account the latent systematic credit risk factor  $Z$  because this factor is assumed to be independent from all other variables. Multivariate dependence structures between margins can be modeled by so-called copula functions. Let  $F$  be a  $d$ -dimensional cumulative density function with margins  $F_1, F_2, \dots, F_d$ . Sklar's theorem<sup>20</sup> states that a copula function  $C : [0, 1]^d \rightarrow [0, 1]$  exists such that for all  $x_1, x_2, \dots, x_d \in \mathbb{R} \cup \{-\infty, +\infty\}$

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \quad (3.6)$$

holds true. Thus, multivariate dependence structures can be isolated from the margins and are even unique in case of continuous margins.<sup>21</sup>

Two popular types of copula functions are the elliptical and Archimedean copulas. Elliptical copulas, such as the normal copula and the  $t$ -copula, are derived from elliptical distributions. This type of copulas is characterized by a symmetry of the dependence structure and especially (in case of the  $t$ -copula) by a symmetry between the lower and upper tail dependence.<sup>22</sup> In contrast, Archimedean copulas allow for asymmetric dependence structures. Prominent representatives are the Gumbel, Clayton and Frank copulas.<sup>23</sup>

Goodness-of-fit tests describe how well empirical observations fit a supposed statistical model. We employ an approach based on the empirical copula. These approaches measure the deviation between the empirical copula and the supposed copula. The null hypothesis contains the supposed copula  $H_0 : C \in \mathcal{C}_0$  which is compared with the empirical copula

$$C_T(\mathbf{u}) = \frac{1}{T} \sum_{t=1}^T 1(\hat{U}_{t,1} \leq u_1, \dots, \hat{U}_{t,d} \leq u_d) \text{ with } \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d \quad (3.7)$$

---

<sup>20</sup>See Sklar (1959).

<sup>21</sup>See McNeil et al. (2005, p. 186).

<sup>22</sup>The normal copula does not exhibit tail dependence.

<sup>23</sup>A detailed introduction to copula functions is given, for example, in McNeil et al. (2005) or Nelson (2006).

where  $\hat{\mathbf{U}}_t = (\hat{U}_{t,1}, \dots, \hat{U}_{t,d}) = \frac{\hat{\mathbf{R}}_t}{T+1}$  are the empirical pseudo observations and  $\hat{\mathbf{R}}_t$  denotes the vector of ranks of all components at time  $t$ . The empirical copula is compared with the estimated copula  $C_{\hat{\theta}_T}$  under the null hypothesis. For estimating the parameter vector  $\hat{\theta}_T$  of the supposed copula, a variety of methods exists. We use the canonical maximum likelihood estimation (also called maximum pseudo-likelihood).<sup>24</sup> For this method, there is no need to specify the parametric form of the marginal distributions because these are replaced by the empirical marginal distributions. Thus, only the parameters of the copula function have to be estimated by maximum pseudo-likelihood (see Cherubini et al. (2004, p. 160)). The employed goodness-of-fit test based on the empirical copula uses the Cramér/von Mises<sup>25</sup> test statistic which is given by

$$S_T = T \int_{[0,1]^d} (C_T(\mathbf{u}) - C_{\hat{\theta}_T}(\mathbf{u}))^2 dC_T. \quad (3.8)$$

High values of  $S_T$  correspond with a high distance between the empirical and the supposed copula and, hence, lead to a rejection of the null hypothesis. In simulation-based power comparison studies, this method delivers more reliable results than many other goodness-of-fit test procedures (see, e. g., Berg (2009) and Genest et al. (2009)). As the probability distribution of the test statistic  $S_T$  under the null hypothesis is unknown, it has to be computed by bootstrapping.<sup>26</sup> For this, we perform 100,000 simulation runs. Table 5 shows the results of the goodness-of-fit test based on the empirical copula for different copula functions.

	GDP		S&P 500	
	Cramér/von Mises	<i>p</i> -value	Cramér/von Mises	<i>p</i> -value
Normal	0.0432	0.5435	0.0478	0.4599
$t_{2df}$	0.0659	0.1760	0.0730	0.1094
$t_{3df}$	0.0610	0.2089	0.0694	0.1195
$t_{4df}$	0.0578	0.2474	0.0662	0.1410
$t_{5df}$	0.0555	0.2455	0.0636	0.1673
Gumbel	0.0825**	0.0455	0.0667	0.1311
Clayton	0.0514	0.2908	0.0496	0.3409
Frank	0.0793*	0.0910	0.0842*	0.0810

Table 5: Cramér/von Mises test statistics and  $p$ -values for different copula functions. The symbols \*, \*\* and \*\*\* denote significance at 10%, 5% and 1% level.

As can be seen, we can only reject the Frank copula at a significance level of 10% for the GDP-log-return and the S&P 500-log-return specification and, in case of GDP-log-return, the Gumbel copula at a significance level of 5%. Further conclusions which copula describes the multivariate dependence structure best cannot be drawn. That is why we

<sup>24</sup>We apply the function `gofCopula` of the package `copula` in R in order to estimate the copula parameters as well as to perform the goodness-of-fit test.

<sup>25</sup>See Genest et al. (2009, p. 201)

<sup>26</sup>See, for a detailed description, Genest and Rémillard (2008).

use the Akaike Information Criterion (AIC) in order to find the best compromise between good approximation and compact dimensioning. It is given by

$$\text{AIC} = -2 \cdot l + 2 \cdot k \quad (3.9)$$

where  $l$  stands for the log-likelihood function and  $k$  describes the number of estimated parameters. Due to the fact that the AIC tends to overparameterize the model,<sup>27</sup> we additionally apply the Bayesian Information Criterion (BIC)

$$\text{BIC} = -2 \cdot l + k \cdot \ln \{T\} \quad (3.10)$$

where the added parameter  $T$  represents the sample size. The results of the AIC and BIC statistics are summarized in Table 6.<sup>28</sup>

		Normal	$t_{2df}$	$t_{3df}$	$t_{4df}$	$t_{5df}$	Clayton	Gumbel	Frank
GDP	ML	3.0151	5.0808	5.0902	4.8397	4.6077	1.7631	0.1018	0.1762
	AIC	-0.0302	-2.1616	-2.1804	-1.6794	-1.2154	-1.5261	1.7964	1.6476
	BIC	3.9664	3.1672	3.1484	3.6494	4.1135	-0.1939	3.1287	2.9798
S&P 500	ML	0.7611	1.9542	2.1710	2.0222	1.8572	0.3467	-	-
	AIC	4.4778	4.0916	3.6581	3.9556	4.2856	1.3066	-	-
	BIC	8.4744	9.4204	8.9869	9.2844	9.6144	2.6382	-	-

Table 6: Maximum pseudo-likelihood and information criteria for different copulas.

For the GDP-log-return specification, the  $t$ -copula with 3 degrees of freedom yields the lowest AIC value. The BIC, however, implies to choose the Clayton copula, which requires only one parameter.<sup>29</sup> In case of the S&P 500-log-return specification, the optimal choice for both, AIC and BIC, is the Clayton copula. Thus, we also face model risk on the level of the multivariate dependence between the systematic risk factors. We take this into account by carrying out the reverse stress test for both copula specifications. The estimated parameters and their significance for the chosen copulas are shown in Table 7.<sup>30</sup> As can be seen, only one parameter estimate is significant, which illustrates the considerable estimation risk (on top of the model risk) that we face when performing a reverse stress test.

<sup>27</sup>See Carter Hill et al. (2011, p. 238).

<sup>28</sup>Numerical problems prevent the calculation of parameters of the Gumbel and the Frank copula. Since the goodness-of-fit in case of the GDP-log-return specification was poor, we neglect these copula functions for S&P 500-log-return specification.

<sup>29</sup>The Clayton copula benefits of its sparse parametrization and the comparatively good fit, while, on the one hand, the elliptical copulas are punished due to their high number of parameters and, on the other hand, the other Archimedean copulas possess a much worse fit.

<sup>30</sup>The copula parameters and their significance were estimated by maximum pseudo-likelihood and by inverting Kendall's Tau (see, e. g., McNeil et al. (2005, pp. 228-237)) using the functions `gofCopula` and `fitCopula` of the package `copula` in R. Since the estimators deviated less than the standard error of each other, we only employed the maximum pseudo-likelihood estimators. The usage of different estimation techniques takes the estimation uncertainty into account and serves as an internal robustness check.

		Estimate	Standard error	<i>p</i> -value
GDP, <i>t</i> -copula	$\rho_{X,C_1}$	0.4196**	0.2046 (2.0512)	0.0402
	$\rho_{X,C_2}$	0.0603	0.2440 (0.2473)	0.8047
	$\rho_{C_1,C_2}$	-0.2509	0.1770 (-1.4182)	0.1561
GDP, Clayton copula	$\theta$	0.3783	0.2316 (1.6334)	0.1024
S&P 500, Clayton copula	$\theta$	0.1302	0.1641 (0.7934)	0.4276

Table 7: Copula parameters and their significance. The *t*-statistics are presented in parentheses. The symbols \*, \*\* and \*\*\* denote significance at 10%, 5% and 1% level.

Figure 8 illustrates random draws of the *t*-copula with 3 degrees of freedom for the GDP-log-return specification and of the Clayton copula for the GDP-log-return as well as S&P 500-log-return specification with parameters estimated by maximum pseudo-likelihood.

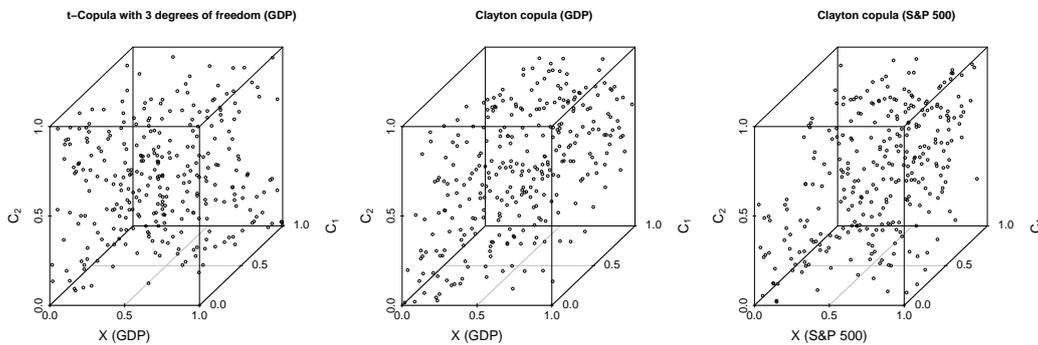


Figure 8: The first and second figure show the *t*-copula with 3 degrees of freedom and the Clayton copula for the GDP-log-return specification in comparison. The third figure shows random draws of the Clayton copula for the S&P 500-log-return specification.

The *t*-copula enables us to model lower and upper tail dependence and, hence, assumes an increased dependence in boom and bust cycles. The Clayton copula, in contrast, exhibits only lower tail dependence and is therefore well suited to model an increased dependence of joint low tail events in times of crisis.

Next, the derived information on the marginal distributions of the systematic risk factors and on their multivariate dependence structure is used to simulate the empirical distribution functions of the obligors' asset returns (as specified in Equations 2.10 and 2.11). Analogously to CreditMetrics (where standard normally distributed asset returns are assumed), the simulated empirical distribution functions for the asset returns of initially Investment Grade and Speculative Grade, respectively, rated obligors are employed for deriving asset return thresholds that correspond to specific rating grades at the end of the risk horizon of one year. For the reverse stress test, we need migration and default thresholds for an initial AA and BB rating grade because these are the assumed homogeneous credit qualities of the bank's obligors. We assume that the asset return distributions for these two rating grades are equal to those of the broader

rating categories Investment Grade and Speculative Grade, respectively. The necessary migration probabilities over a one-year risk horizon are provided by Standard & Poor's<sup>31</sup> and summarized in Table 8.

	AAA	AA	A	BBB	BB	B	C-CCC	Default
AAA	90.86%	8.35%	0.56%	0.05%	0.08%	0.03%	0.05%	0.00%
AA	0.59%	90.14%	8.52%	0.55%	0.06%	0.08%	0.02%	0.02%
A	0.04%	1.99%	91.64%	5.64%	0.40%	0.18%	0.02%	0.08%
BBB	0.01%	0.14%	3.96%	90.49%	4.26%	0.71%	0.16%	0.27%
BB	0.02%	0.04%	0.19%	5.79%	83.97%	8.09%	0.84%	1.05%
B	0.00%	0.05%	0.16%	0.26%	6.21%	82.94%	5.06%	5.32%
C-CCC	0.00%	0.00%	0.22%	0.33%	0.97%	15.20%	51.24%	32.03%

Table 8: Migration probabilities based on Standard & Poor's (2011a).

We perform 1,000,000 draws in order to determine the empirical distribution functions of the obligors' asset returns. Afterwards, the default and migration thresholds are chosen in such a way that they coincide with the appropriate (corresponding to the default and migration probabilities for initially AA- and BB-rated obligors that are presented in Table 8) quantiles of the empirical distribution functions of the obligors' asset returns. The results are summarized in Table 9.<sup>32</sup>

GDP	Thresholds for obligors with initial rating grade AA								
	Default	C-CCC	B	BB	BBB	A	AA	AAA	
<i>t</i> -copula, normal	≤ -3.56	(-3.56,-3.36]	(-3.36,-3.03]	(-3.03,-2.90]	(-2.90,-2.42]	(-2.42,-1.30]	(-1.30,2.60]	> 2.60	
<i>t</i> -copula, GPD	≤ -7.29	(-7.29,-4.53]	(-4.53,-3.26]	(-3.26,-3.07]	(-3.07,-2.48]	(-2.48,-1.31]	(-1.31,2.60]	> 2.60	
Clayton, normal	≤ -3.54	(-3.54,-3.36]	(-3.36,-3.06]	(-3.06,-2.92]	(-2.92,-2.45]	(-2.45,-1.31]	(-1.31,2.62]	> 2.62	
Clayton, GPD	≤ -8.18	(-8.18,-5.20]	(-5.20,-3.39]	(-3.39,-3.16]	(-3.16,-2.53]	(-2.53,-1.33]	(-1.33,2.62]	> 2.62	
	Thresholds for obligors with initial rating grade BB								
	Default	C-CCC	B	BB	BBB	A	AA	AAA	
<i>t</i> -copula, normal	≤ -2.22	(-2.22,-1.99]	(-1.99,-1.19]	(-1.19,1.67]	(1.67,2.93]	(2.93,3.34]	(3.34,3.67]	> 3.67	
<i>t</i> -copula, GPD	≤ -2.26	(-2.26,-2.02]	(-2.02,-1.20]	(-1.20,1.67]	(1.67,2.93]	(2.93,3.34]	(3.34,3.67]	> 3.67	
Clayton, normal	≤ -2.24	(-2.24,-2.00]	(-2.00,-1.20]	(-1.20,1.67]	(1.67,2.95]	(2.95,3.37]	(3.37,3.70]	> 3.70	
Clayton, GPD	≤ -2.29	(-2.29,-2.04]	(-2.04,-1.21]	(-1.21,1.67]	(1.67,2.95]	(2.95,3.37]	(3.37,3.70]	> 3.70	
S&P 500	Thresholds for obligors with initial rating grade AA								
	Default	C-CCC	B	BB	BBB	A	AA	AAA	
Clayton, normal	≤ -3.56	(-3.56,-3.35]	(-3.35,-3.05]	(-3.05,-2.91]	(-2.91,-2.43]	(-2.43,-1.29]	(-1.29,2.63]	> 2.63	
Clayton, GPD	≤ -9.87	(-9.87,-5.85]	(-5.85,-3.49]	(-3.49,-3.22]	(-3.22,-2.54]	(-2.54,-1.32]	(-1.32,2.64]	> 2.64	
	Thresholds for obligors with initial rating grade BB								
	Default	C-CCC	B	BB	BBB	A	AA	AAA	
Clayton, normal	≤ -2.32	(-2.32,-2.09]	(-2.09,-1.29]	(-1.29,1.59]	(1.59,2.85]	(2.85,3.27]	(3.27,3.58]	> 3.58	
Clayton, GPD	≤ -2.37	(-2.37,-2.12]	(-2.12,-1.30]	(-1.30,1.59]	(1.59,2.86]	(2.86,3.27]	(3.27,3.59]	> 3.59	

Table 9: Default and migration thresholds for initial rating grades AA and BB for normal marginal distributions with/without GPD tails and for different copulas.

<sup>31</sup>Data was adjusted for rating withdrawals.

<sup>32</sup>We use the simulated default threshold instead of the estimated one in Equation 2.8 and 2.9.

## 4 Performing the reverse stress test

Analogously to Grundke (2011, 2012a), for demonstrating the usage of the modeling framework and for performing a reverse stress test, we assume a stylized bank portfolio that exclusively consists of assets and liabilities structured as zero coupon bonds. The bank pursues a strategy of positive maturity transformation, whereby it is assumed that the term structure of the bank's assets and liabilities does not vary across time. Thus, value variations caused by a decreasing time to maturity are not considered. The assumed cash flow profile is illustrated in Figure 9.

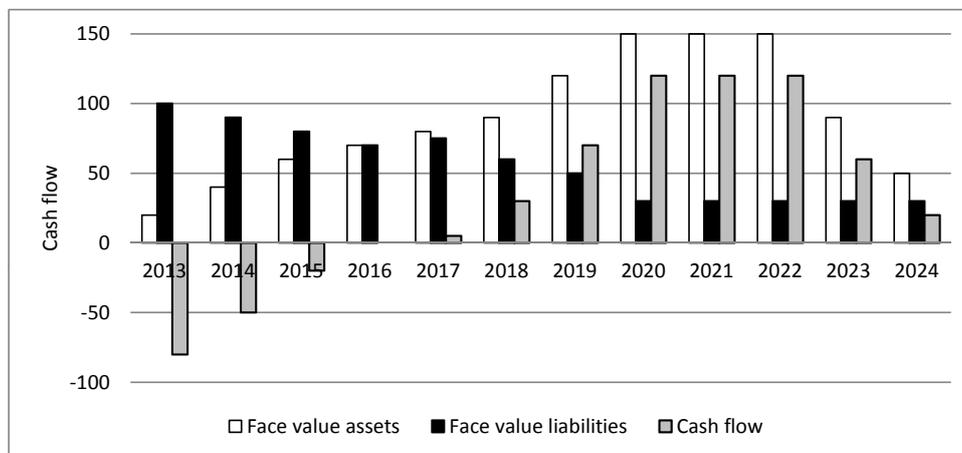


Figure 9: Cash flows of assets and liabilities of the stylized bank.

All positions  $n \in \{1, \dots, N\}$  on the asset side are assumed to be issued by different obligors with initially equal default probability. They have a standardized face value of one and a time to maturity of  $T_n \in \{1, \dots, 12\}$ .

The value of a defaultable zero-coupon bond at the risk horizon  $H$  issued by obligor  $n$  who is rated as  $\eta_H^n \in \{1, 2, 3, 4, 5, 6, 7\} = \{\text{AAA}, \text{AA}, \text{AA}, \text{BBB}, \text{BB}, \text{A}, \text{C-CCC}\}$  at the risk horizon  $H$  is given by

$$B^d(C_1(H), C_2(H), \eta_H^n, H, T_n) = \exp \left\{ -\left( R(C_1(H), C_2(H), H, T_n) + S_{\eta_H^n}(H, T_n) \right) \cdot T_n \right\} \quad (4.1)$$

where  $R(C_1(t), C_2(t), H, T_n)$  denotes the stochastic risk-free interest rate at the risk horizon  $H$  for a time to maturity of  $T_n$ , which is calculated from the last obtained empirical term structure of risk-free interest rates and the first two principal components by

$$R(C_1(H), C_2(H), H, T_n) = r_{q_n, t} \cdot (1 + \Delta r_{q_n}) = r_{q_n, t} \cdot \left( 1 + \sum_{j=1}^2 c_{q_n, j} \cdot C_j(H) \right) \quad (4.2)$$

where  $q_n$  denotes the (if necessary, linearly interpolated) time to maturity of the observed fixed-income products matching obligor's  $n$  time to maturity  $T_n$ . The expression  $S_{\eta_H^n}(H, T_n)$  denotes the non-stochastic credit spread for a time to maturity of  $T_n$  for rating grade  $\eta_H^n$  at the risk horizon  $H$ . Credit spread data is provided by Datastream and obtained from straight U. S. corporate bonds which have (as well as our assumed bank portfolio) a time to maturity up to 2024. The credit spread is calculated as the yield difference of the mid price over a similar sovereign bond.<sup>33</sup> Bonds with a negative credit spread were omitted,<sup>34</sup> half notches were upgraded (in case of -) or, respectively, downgraded (in case of +). Finally, 2,350 bonds remained. For every rating grade, the credit spread was calculated as the median to ensure an increasing credit spread with worsening rating grade. Table 10 shows the median credit spreads for all rating grades.

Rating	No. of bonds	Credit spread (in bps)
AAA	17	59
AA	57	91
A	639	132
BBB	779	208
BB	348	465
B	355	670
C-CCC	155	959

Table 10: Rating-specific median credit spreads.

To model the recovery payment to the bank in the case of a default of an obligor, we apply a modified recovery-of-treasury assumption.<sup>35</sup> In the case of a default of obligor  $n$ , the minimum of a beta-distributed fraction  $\delta_n$  with the parameters  $\mu = 0.518$ <sup>36</sup> and  $\sigma = 0.389$  of a risk-free, but otherwise identical, zero-coupon bond, and the value of the bond without any rating transition of the obligor between 0 and  $H$ , is paid. This convention ensures that the payment in case of a default is never larger than the value of the bond before. The recovery rates are assumed to be independent across issuers and independent from all other stochastic variables of the model.

<sup>33</sup>Datastream uses a linear combination of sovereign bonds in order to match the maturities of corporate bonds exactly.

<sup>34</sup>A negative spread can be explained through low liquidity shortly before the maturity date. If a bond is not traded on a day, the last observed price is taken as the current price. Therefore, the bond price does not converge against the face value, and, for bonds priced above their face value, a negative yield (and a negative credit spread) can be calculated.

<sup>35</sup>See Grundke (2011, 2012a)

<sup>36</sup>The mean and the standard deviation of the beta-distributed recovery rate equal Standard & Poor's mean and standard deviation of the recovery rate of senior unsecured bonds during 1987 to 2011 (see Standard & Poor's (2011b)).

The value of the positions  $v \in \{1, \dots, V\}$  on the liability side are given by

$$B^l(C_1(H), C_2(H), H, T_v) = \exp \{-(R(C_1(H), C_2(H), H, T_v) + S_{\eta_{AA}}(H, T_v)(H, T_v)) \cdot T_v\}. \quad (4.3)$$

This representation uses the (admittedly strong) assumption that the bank is initially rated as AA and is not exposed to migration risk until the risk horizon.<sup>37</sup>

To simplify calculations, we impose an homogeneity assumption with respect to the credit quality of the bank's asset portfolio: The obligors on the asset side are assumed to be exclusively rated as AA ( $\eta_0^n = AA \forall n \in \{1, \dots, N\}$ ) or BB ( $\eta_0^n = BB \forall n \in \{1, \dots, N\}$ ), respectively.

The market value of the bank's equity at the risk horizon  $H$  is given by the difference between the sum of the market values of the assets and the sum of the market values of the liabilities at the risk horizon  $H$ :

$$V_P(H) = \sum_{n=1}^N B^d(C_1(H), C_2(H), \eta_H^n, H, T_n) - \sum_{v=1}^V B^l(C_1(H), C_2(H), H, T_v). \quad (4.4)$$

The initial market value of the bank's equity in  $t = 0$  amounts to 236.32 (with a corresponding equity-to-asset ratio of 29.05%) in the case of initially AA-rated obligors and to 51.26 in case of initially BB-rated obligors (with a corresponding equity-to-asset ratio of 8.16%). These values can be understood as the bank's capital buffer  $B$ .

In order to perform the actual reverse stress test, a grid search in the four-dimensional space of the systematic risk factors is done. For each grid point, we calculate the conditional value-at-risk of the bank's equity value at the risk horizon  $H$  by Monte-Carlo simulation with  $S = 1,000$  draws.<sup>38</sup> For each systematic risk factor, the grid search is carried out within the interval  $[\mu - 4 \cdot \sigma, \mu + 4 \cdot \sigma]$ , which is split into equally-sized subintervals. This corresponds to an evaluation of over 99.99% of the probability space in the case of normally distributed margins and over 98.30% in case of heavier GPD tails. We choose a step size of  $0.5 \cdot \sigma$  and, therefore, get 17 equidistant grid points per risk factor. From this follows that we have to evaluate  $17^4 = 83,521$  grid points and perform

---

<sup>37</sup>For an alternative modeling with time-varying bank rating, see Grundke (2012a). With a time-varying bank rating, care has to be taken to avoid circularity problems.

<sup>38</sup>The idiosyncratic risk is the only source of uncertainty in case of the conditional distribution. Therefore, the small number of Monte-Carlo simulation runs is sufficient.

for each grid point a Monte-Carlo simulation to compute the conditional value-at-risk.<sup>39</sup>

As in Grundke (2011, 2012a), a scenario  $\omega = (z, x, c_1, c_2)$  is classified as a reverse stress test scenario when the existing capital buffer  $B$  is consumed by a decrease of the expected equity value at the risk horizon  $H$  and by the respective conditional economic capital requirement. Thus, a bank's default is understood as a non-fulfilment of the economic capital requirements according to the second pillar of Basel II. When the value-at-risk at a confidence level of  $\alpha$  is used as an economic capital measure and is defined as the difference between the conditional expected equity value at the risk horizon  $H$  and the  $(1 - \alpha)$ -quantile of the conditional probability distribution of the bank's equity value, the most likely reverse stress test scenario is given by

$$\begin{aligned} & \arg \max_{\omega \in \Omega^*} P(\omega) \\ & \text{with } \Omega^* = \left\{ \omega \in \Omega \mid \underbrace{\mathbb{E}[V_P(H)] - \mathbb{E}[V_P(H)|\omega]}_{=\text{expected loss, if } \omega \text{ occurs.}} + \underbrace{\mathbb{E}[V_P(H)|\omega] - q_{1-\alpha}(V_P(H)|\omega)}_{=VaR_{\alpha,H}(V_P(H)|\omega)} = B \right\} \\ & = \left\{ \omega \in \Omega \mid \mathbb{E}[V_P(H)] - q_{1-\alpha}(V_P(H)|\omega) = B \right\}. \end{aligned} \quad (4.5)$$

The probability that a scenario occurs is computed as follows<sup>40</sup>

$$\begin{aligned} & P(z^- < Z \leq z^+, x^- < X \leq x^+, c_1^- < C_1 \leq c_1^+, c_2^- < C_2 \leq c_2^+) \\ & = \int_{z^-}^{z^+} \int_{x^-}^{x^+} \int_{c_1^-}^{c_1^+} \int_{c_2^-}^{c_2^+} f(z, x, c_1, c_2) dz dx dc_1 dc_2, \\ & = (F_1(z^+) - F_1(z^-)) \cdot (C(F_2(x^+), F_3(c_1^+), F_4(c_2^+)) - C(F_2(x^+), F_3(c_1^+), F_4(c_2^-)) \\ & \quad - C(F_2(x^+), F_3(c_1^-), F_4(c_2^+)) - C(F_2(x^-), F_3(c_1^+), F_4(c_2^+)) \\ & \quad + C(F_2(x^+), F_3(c_1^-), F_4(c_2^-)) + C(F_2(x^-), F_3(c_1^+), F_4(c_2^-)) \\ & \quad + C(F_2(x^-), F_3(c_1^-), F_4(c_2^+)) - C(F_2(x^-), F_3(c_1^-), F_4(c_2^-))) \end{aligned} \quad (4.6)$$

where  $C$  denotes the applied copula function,  $F_1(z), F_2(x), F_3(c_1), F_4(c_2)$  denote the margins of the systematic risk factors and  $f$  denotes the joint density function of the

<sup>39</sup>A finer grid would have increased the computation time considerably.

<sup>40</sup>The expression for calculating probabilities on multi-dimensional intervals can be found in Mathar and Pfeifer (1990, p. 41). The computation of each single probability term is done using the function `pcopula` of the package `copula` in the program R. In order to calculate probabilities, `pcopula` refers to the function `pmvt` of the package `mvtnorm` which uses randomized Quasi-Monte-Carlo methods (see, e. g., Genz and Bretz (1999, 2002)). As the assigned probabilities on the edge of the considered part of the support are very low, we can get, due to numerical issues, implausible results, especially negative probabilities. To solve this problem, we calculate probabilities in case of the  $t$ -copula as the mean over several repetitions.

four systematic risk factors  $Z$ ,  $X$ ,  $C_1$  and  $C_2$ .

The border points of the intervals in the above integral are given by

$$\begin{pmatrix} z^\pm & x^\pm & c_1^\pm & c_2^\pm \end{pmatrix} = \begin{pmatrix} z & x & c_1 & c_2 \end{pmatrix} \pm 0.5 \cdot \text{factor-specific step size}. \quad (4.7)$$

Figure 10 illustrates this specification for the three dependent systematic risk factors  $X$ ,  $C_1$  and  $C_2$ .

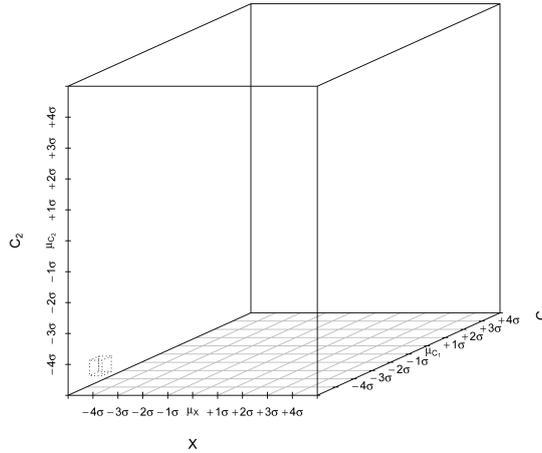


Figure 10: Specification of border points.

As described before, to consider model risk, the reverse stress test is performed, on the one hand, for the GDP-log-return specification with a  $t$ -copula and a Clayton copula dependence structure and, on the other hand, for S&P 500-log-return specification with a Clayton copula dependence structure. The risk factors are assumed to be (marginally) normally distributed and, respectively, to have heavier GPD tails. Together with the two assumed initial credit qualities (AA and BB, respectively), this yields 12 test specifications. The risk horizon is  $H = 1$  and the confidence level of the value-at-risk is set to 99%. Since, when the considered scenario set is finite, it is very likely that no scenario exhausts the capital buffer exactly, we widen our search to the interval plus/minus 5% around the capital buffer  $B$ .

For initially AA-rated obligors, none of the considered scenarios completely exhausts the capital buffer. In case of initially BB-rated obligors, however, reverse stress test scenarios exist. The most likely reverse stress test scenarios are shown in Table 11.<sup>41</sup>

<sup>41</sup>No risk factor takes its boundary value.

		$z$	$x$	$c_1$	$c_2$	Probability
GDP	$t$ -copula, normal	-0.5	-0.0086	-2.3278	-1.0172	$1.3363 \cdot 10^{-5}$
	$t$ -copula, GPD	-0.5	-0.0002	-0.4920	-1.2309	$1.5716 \cdot 10^{-5}$
	Clayton copula, normal	-1.0	-0.0044	-2.3278	-1.0172	$2.2230 \cdot 10^{-5}$
	Clayton copula, GPD	-0.5	-0.0002	-1.4099	-1.2309	$2.3887 \cdot 10^{-5}$
S&P 500	Clayton copula, normal	-0.5	-0.0952	-0.4920	-1.4446	$2.0662 \cdot 10^{-6}$
	Clayton copula, GPD	-0.5	-0.0952	-0.0330	-1.4446	$1.1456 \cdot 10^{-5}$

Table 11: Most likely reverse stress test scenarios for an initially BB-rated portfolio based on various model specifications.

A negative value of the latent systematic risk factor, a slight downturn of the economy (GDP) or, respectively, a medium downturn of the economy (S&P 500), a general decrease in the level of interest rates (first principal component), and an increased steepness of the interest rate curve through relatively decreasing interest rates for short maturities compared to increasing interest rates for long maturities (second principal component) represent the most probable scenario exhausting the capital buffer. This result is robust with respect to the employed model specification. The absolute probabilities for the occurrence of the most likely reverse stress test scenarios exhibited in Table 11 depend on the step size of the grid search. Therefore, these probabilities can only be used for finding the most likely scenario within the set of all identified reverse stress test scenarios. The stressed term structure of risk-free interest rates in the most likely reverse stress test scenarios is shown in Figure 11.

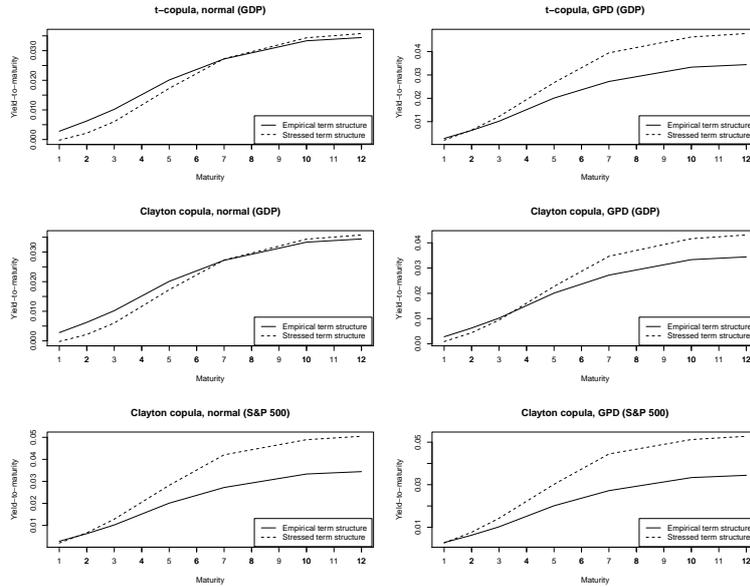


Figure 11: Impact of the most likely reverse stress test scenarios on the term structure of risk-free interest rates. The upper (left:  $t$ -copula, normal, right:  $t$ -copula, GPD) and middle (left: Clayton copula, normal, right: Clayton copula, GPD) figures show the impact in case of the GDP-log-return specification. The lower two figures illustrate the case of the S&P 500-log-return specification for a Clayton copula (left: normal, right: GPD). The empirical term structure is calculated as specified in Equation 4.2 from the last observed yield-to-maturities on January 4, 2011, and the realizations of the first and second principal components in the most likely reverse stress test scenarios.

First, changes in the first principal component that correspond to the most likely reverse stress test scenarios push the whole term structure downwards. Second, the corresponding changes in the second principal component lead also to decreased short-term interest rates and to increased long-term interest rates. The sum of these effects leads to decreased short-term interest rates and considerably increased long-term interest rates.

For the assumed bank, which is performing positive maturity transformation, this induces a double negative impact: Negative cash flows occurring at earlier points in time are discounted by an decreased short-term interest rate, while positive cash flows at later points in time are discounted by considerably increased long-term interest rates.

## 5 Conclusion

In this paper, we have presented a macroeconomic reverse stress test framework and showed how to implement it empirically. Beside this empirical implementation, the

innovation of our contribution is the usage of principal components in order to capture changes of risk-free interest rates and, hence, to keep the model tractable.

For a given stylized bank portfolio, we have determined reverse stress test scenarios for several model specifications, in particular for different marginal distributions for the systematic risk factors and for different multivariate dependence structures. In case of initially AA-rated obligors, we could not detect a reverse stress test scenario, but for initially BB-rated obligors, we found that a negative realization of the latent systematic credit risk factor, a slight, or respectively, medium downturn of the economy, and, on the one hand, decreased risk-free interest rates for short-term maturities and, on the other hand, increased risk-free interest rates for long-term maturities represent the most probable scenario exhausting the capital buffer. However, the results also show that reverse stress tests are exposed to considerable model and estimation risk which makes numerous robustness checks necessary.

Quantitative reverse stress tests confront banks with considerable challenges. Beside the problem of finding those scenarios in which the viability of the bank is threatened, probabilities of occurrence are needed to find the most likely one of these scenarios. Further research could deal for example with algorithms for finding reverse stress test scenarios that are more intelligent than the simple grid search employed in this paper. This would allow to handle extensions with more systematic risk factors and permit the usage of a smaller step size. Of course, as long as no reverse stress test standard models are approved, in addition, further research using other frameworks is needed in order to develop appropriate models meeting the regulatory requirements.

## References

- Berg, D. (2009). Copula goodness-of-fit testing: An overview and power comparison. *The European Journal of Finance*, 15, 675-701.
- Boss, M. (2002). Ein makroökonomisches Kreditrisikomodell zur Durchführung von Krisentests für das österreichische Kreditportfolio. In *Finanzmarktstabilitätsbericht*. Österreichische Nationalbank.
- Breuer, T., Jandacka, M., Rheinberger, K., & Summer, M. (2008). Hedge the stress: Using stress tests to design hedges for foreign currency loans. In D. Rösch & H. Scheule (Eds.), *Stress Testing for Financial Institutions* (p. 111-126). Risk Books, London.
- Carter Hill, R., Griffiths, W. E., & Lim, G. C. (2011). *Principles of Econometrics, Fourth Edition*. Wiley.
- CEBS. (2009). Guidelines on Stress Testing (CP32).
- CEBS. (2010). Guidelines on Stress Testing (GL32).
- Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula Methods in Finance* (1, Ed.). Wiley Finance.
- Čihák, M. (2007). Introduction to Applied Stress Testing. *IMF Working Paper 07/59*.
- Drüen, J., & Florin, S. (2010). Reverse Stresstests: Stress-Kennzahlen für die praktische Banksteuerung. *Risiko Manager [In German]*, 10, 1,6-9.
- Frey, R., & McNeil, A. J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6, 59-92.
- FSA. (2008). Stress and Scenario Testing. Consultant Paper 08/24.
- FSA. (2009). Stress and Scenario Testing, Feedback on CP08/24 and Final Rules. Policy Statement 09/20.
- Füser, K., Hein, B., & Somma, M. (2012a). Inverse Stresstests: Neue Perspektiven auf ein relevantes Thema (1). *Die Bank [In German]*, 4, 34-37.
- Füser, K., Hein, B., & Somma, M. (2012b). Inverse Stresstests: Neue Perspektiven auf ein relevantes Thema (2). *Die Bank [In German]*, 5, 45-49.
- Genest, C., & Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré: Probabilités et Statistique*, 44, 1096-1127.
- Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2), 199-213.
- Genz, A., & Bretz, F. (1999). Numerical computation of multivariate t-probabilities with application to power calculation of multiple contrasts. *Journal of Statistical Computation and Simulation*, 63, 361-378.
- Genz, A., & Bretz, F. (2002). Methods for the computation of multivariate t-probabilities. *Journal of Computational and Graphical Statistics*, 11, 950-971.
- Ghosh, S., & Resnick, S. (2010). A discussion on mean excess plots. *Stochastic Processes and their Applications*, 120(8), 1492-1517.
- Golub, B. W., & Tilman, L. M. (2000). *Approaches for fixed income markets*. New York: Wiley.
- Gourier, E., Farkas, W., & Abbate, D. (2009). Operational risk quantification using extreme value theory and copulas: from theory to practice. *The Journal of Operational Risk*, 4(3), 3-26.

- Grundke, P. (2011). Reverse stress tests with bottom-up approaches. *Journal of Risk Model Validation*, 5(1), 71-90.
- Grundke, P. (2012a). Further recipes for quantitative reverse stress testing. *The Journal of Risk Model Validation*, 6(2), 81-102.
- Grundke, P. (2012b). Qualitative inverse Stresstests mit Fehlerbäumen? *Zeitschrift für das gesamte Kreditwesen [In German]*, 65. Jg.(3), 131-135.
- Hamerle, A., & Rösch, D. (2006). Parameterizing credit risk models. *Journal of Credit Risk*, 2(4), 101-122.
- Heidari, M., & Wu, L. (2003). Are interest rate derivatives spanned by the term structure of interest rates? *The Journal of Fixed Income*, 13(1), 75-86.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Knez, P., Litterman, R., & Scheinkman, J. (1994). Explorations into factors explaining money market returns. *The Journal of Finance*, 49(5), 1861-1882.
- Kronrod, A. S. (1965). *Nodes and weights of quadrature formulas. sixteen-place tables.* Consultants Bureau New York.
- Liermann, V., & Klauck, K.-O. (2010). Banks in stress. *Die Bank [In German]*, 5, 53-55.
- Litterman, R., & Scheinkman, J. (1991). Common factors affecting bond returns. *Journal of Fixed Income (June)*, 54-61.
- Mathar, R., & Pfeifer, D. (1990). *Stochastik für Informatiker [In German]*. B.G. Teubner Stuttgart.
- McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management.* Princeton University Press, New Jersey.
- McNeil, A. J., & Smith, A. (2012). Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, 50(3), 299-308.
- Nelder, J. A., & Mead, R. (1965). A simplex algorithm for function minimization. *Computer Journal*, 7, 308-313.
- Nelson, R. B. (2006). *An introduction to copulas, Second Edition.* Springer Verlag: New York.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris [In French]*, 8, 229-231.
- Skoglund, J., & Chen, W. (2009). Risk contributions, information and reverse stress testing. *The Journal of Risk Model Validation*, 3(2), 61-77.
- Sorge, M., & Virolainen, K. (2006). A comparative analysis of macro stress-testing methodologies with application to Finland. *Journal of Financial Stability*, 2, 113-151.
- Standard & Poor's. (2011a). 2010 Annual Global Corporate Default Study And Rating Transitions.
- Standard & Poor's. (2011b). Default, Transition, and Recovery: Recovery Study (U.S.): Piecing Together The Performance Of Defaulted Instruments After The Recent Credit Cycle.
- Wilson, T. C. (1997a). Portfolio credit risk: part I. *Risk*, 10(9), 111-117.
- Wilson, T. C. (1997b). Portfolio credit risk: part II. *Risk*, 10(10), 56-61.