Credit Risk and Asset Correlation Modelling for the Swedish Market

A Comparative Analysis

LUDVIG HAMILTON

CARL AXEL JÖNSSON
Credit Risk and Asset Correlation Modelling for the Swedish Market

A Comparative Analysis

LUDVIG HAMILTON
CARL AXEL JÖNSSON

Degree Projects in Financial Mathematics (30 ECTS credits)
Degree Programme in Industrial Engineering and Management
KTH Royal Institute of Technology year 2019
Supervisor at Svedbank AB: Niklas Neville, Henrik Ribom
Supervisor at KTH: Anja Janssen
Examiner at KTH: Anja Janssen
Abstract

In order to ensure solvency, financial institutions must evaluate their credit risk exposure and determine how much economic capital is required to hold as a cushion. This thesis compares three factor models, namely Asymptotic Single Risk Factor, Inter-sector and Intra-sector factor models and evaluates how their different characteristics affect the economic capital outcomes. The thesis also investigates how these outcomes are affected when assuming asset dependency through a Student’s-\(t\) copula. Focus will also be put on how different types and levels of asset correlation affect the models’ credit risk results. We use a fictitious loan portfolio consisting of 138 Swedish firms with equity data from between 2007 and 2019 in order to calculate asset correlations and economic capital. Our main findings are that the asset correlations severely impact the outcomes of the credit risk models and that practitioners must calibrate and stress test their models regularly with respect to how correlations vary between different firms. The thesis also finds that using copulas for credit portfolios provides more conservative risk outcomes but makes the models less sensitive to correlation level input.

**Keywords:** Credit Risk, Economic Capital, Value-at-Risk, Intra-sector correlation, Inter-sector correlation, Copula, Basel III.
Modellering av kreditrisk och tillgångskorrelationer på den svenska marknaden:
En komparativ analys

Sammanfattning


Nyckelord: Kreditrisk, Ekonomiskt kapital, Value-at-Risk, Intra-sektorkorrelation, Inter-sektor-korrelation, Copula, Basel III.
Acknowledgements

First and foremost we would like to thank our supervisor Anja Janssen from the Department of Mathematics at KTH. Her valuable input, guidance, and insightful comments made this thesis possible. At Swedbank, we would like to thank Niklas Neville and Henrik Ribom for providing the thesis subject. The thoughtful discussions, navigation, and advice they provided was of utmost importance.
<table>
<thead>
<tr>
<th>Term</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basel Committee on Banking Supervision</td>
<td>BCBS</td>
</tr>
<tr>
<td>Internal Ratings-Based</td>
<td>IRB</td>
</tr>
<tr>
<td>Exposure at Default</td>
<td>EAD</td>
</tr>
<tr>
<td>Probability of Default</td>
<td>PD</td>
</tr>
<tr>
<td>Loss Given Default</td>
<td>LGD</td>
</tr>
<tr>
<td>Economic Capital</td>
<td>EC</td>
</tr>
<tr>
<td>Internal Capital Adequacy Assessment Process</td>
<td>ICAAP</td>
</tr>
<tr>
<td>Expected Loss</td>
<td>EL</td>
</tr>
<tr>
<td>Unexpected Loss</td>
<td>UL</td>
</tr>
<tr>
<td>Value-at-Risk</td>
<td>VaR</td>
</tr>
<tr>
<td>Recovery Rate</td>
<td>RR</td>
</tr>
<tr>
<td>Asymptotic Single Risk Factor</td>
<td>ASRF</td>
</tr>
<tr>
<td>Monte Carlo (simulation)</td>
<td>MC</td>
</tr>
<tr>
<td>Herfindahl-Hirschman Index</td>
<td>HHI</td>
</tr>
</tbody>
</table>
## Contents

1 Introduction 1

2 Theoretical Framework 4
   2.1 Default and Credit Risk 4
   2.2 Risk Measures 5
   2.3 Correlation 6
   2.4 Credit Risk Modelling 8
      2.4.1 The Vasicek Model 8
      2.4.2 Single-Factor Model 11
      2.4.3 Multi-Factor Models 12
      2.4.4 Copulas and Copula Credit Risk Models 14
   2.5 Economic Capital and the IRB Approach 17
   2.6 Stochastic Simulation 18
   2.7 Portfolio Diversification 19

3 Methodology 20
   3.1 Data 20
   3.2 Credit Risk Models 20
   3.3 Asset Correlations 21
   3.4 Modelling and Simulations 23

4 Results 25
   4.1 Convergence of Monte Carlo Simulations 25
   4.2 Correlation Matrices 26
   4.3 Credit Risk Results 28
      4.3.1 Factor Models 28
      4.3.2 Copula Factor Models 29
   4.4 Loss Histograms 31
   4.5 Diversification Impact 33

5 Discussion 34

6 Conclusion and Further Research 38

A Appendix 43
1 Introduction

In order for banks and other financial institutions to assure solvency, excess capital must be held as a cushion. This regulatory capital needs to reflect the different risks and exposures which the financial institution carries. The three major risks for such institutions are market risk, operational risk, and credit risk. Market risk is defined as the risk of losses in balance sheet positions due to movement in market prices and operational risk is broadly defined as the risk of loss resulting from inadequate or failed internal processes. Credit risk is the risk that a borrower (or counterpart) fails to meet its payment obligation, resulting in a loss of principal and associated interest for the lender. This thesis has credit risk as its focal point and will examine and discuss how different approaches to credit risk modelling affect how banks analyze and conclude their required volume of excess capital.

The Basel III framework (Basel III (2010)) was developed by the Basel Committee on Banking Supervision (“BCBS”) and defines the regulatory capital requirements of financial institutions. The framework was partially developed to combat deficiencies revealed by the financial crisis 2007-2008. It provides a standardized regulatory capital calculation formula where institutions with Internal Ratings-Based (“IRB”) permission are allowed to use their own estimates on Exposure at default (“EAD”), Probability of default (“PD”), Loss given default (“LGD”) and maturity. However, banks also have to perform their own excess capital calculations as a complement to the regulatory capital. This amount of excess capital is called Economic Capital (“EC”), and is mostly used for the banks' Internal Capital Adequacy Assessment Process (“ICAAP”) and for steering purposes. The EC calculations are based on forecasting the average level of expected credit losses in their loan portfolio. This particular loss is called the Expected Loss (“EL”). However, sometimes losses exceeding the EL occur. The amount that the total losses exceed the EL is called the Unexpected Loss (“UL”). It’s the UL which determines how much excess capital, namely the EC, that the bank considers should be held as a financial cushion.

The most basic way of determining the EC is based on a single-factor framework which is dependent on an aggregation of all individual exposures and their default probabilities. However, one of the downfalls of the framework is that it naively assumes that the bank’s portfolio is perfectly diversified and that there exists a single systematic risk factor which reflects the state of the entire economy. Today, it is common to use a multi-factor approach based on Vasicek’s model as a supplement to the standard single-factor approach. One of the benefits of the multi-factor model is that it does not assume the portfolio to be perfectly diversified and utilizes risk-drivers such as inter-correlations (cross-sector correlation) and intra-correlations (internal sector correlations).
is also becoming increasingly popular to model credit risk using copulas which allow for customized dependency structures among the components in a credit portfolio. Depending on what method one uses for modelling correlations and risk factors, the credit risk results might vary dramatically.

The purpose of this thesis is to compare different alternatives on how to calculate a bank’s capital requirements. More specifically, Value-at-Risk (“VaR”) and EC will be calculated by using several different credit risk models, followed by comparisons of results and a discussion regarding how the different natures of the models affect their respective outcomes. The main fashion in which the models differ will be how they handle the correlations between the assets (obligors) in a credit portfolio. It will also be investigated how sensitive the models are to the level of correlation. The purpose is not to find an optimal “one rules all”-model, but rather to highlight the differences in outcomes generated by the models and seek to explain the differences using mathematically supported arguments. The thesis may be useful to a practitioner who seeks to get an overview over the different popular alternatives to the standard Basel III framework and how correlation input may affect credit risk modelling outcomes.

Previous research within the field has also focused on assessing how varied types of credit risk models provide different estimation results for credit risk. Düllmann et al. (2007) compares the asset correlation estimates of a single-factor (“market”) model and a multi-factor (“sector”) model and proceeds by using the results to compare what estimates of EC and VaR the models provide. The authors use monthly asset values, generated by Moody’s KMV Credit Monitor, of European firms ranging from January 1996 to February 2004 and calculate the squared sample correlations between returns of individual firms and the entire market (for the market model) and between individual firms and their respective sector index as well as between sectors (for the sector model). The authors continue by calculating VaR estimates for the models and use the standard IRB-generated VaR formula (presented in Section 2.5) as a benchmark. The main findings of the article include that the asset correlations are slightly higher when employing a sector model compared to a market model, which is in line with Lopez (2002) who also conducted similar research within the area. The authors also find that the correlations vary over time, but that the VaR generated by the benchmark IRB model is the most stable due to the smoothing effect of the PD.

Other research within the area of credit risk and asset correlations include Grundke (2007) which contains a summary of asset correlation estimation techniques and results generated by numerous other researchers. This overview includes de Servigny & Renault (2002) who find that the link between empirical equity correlation and default correlation is rather weak. However, Düllmann
et al. (2008) find that direct estimation of asset correlations from equity returns and indirect estimation from asset returns are superior to estimation from default rates, both in terms of bias and errors. The authors recommend using equity prices for asset correlation estimation but advocate that care must be taken if the correlation estimates are applied in credit risk modelling due to the simplifications made in the study.

This thesis is closely linked to the study conducted by Düllmann et al. (2007), as well as other studies who compare different correlation setups for credit risk models. However, this thesis seeks to be more extensive by including both single-factor and multi-factor modelling, as well as copula dependency modelling of credit risk. This thesis also handles a larger and more recent set of data than several other articles within the area, having stock return data spanning from 2007-2019, which includes several important financial market-related events such as the financial crisis 2007-2008 and the Euro-crisis 2009-2014. Another point which separates this thesis from other articles within the area is that it focuses solely on the Swedish equity and credit market instead of its European or American equivalents.

The remainder of the thesis is structured as follows. Section 2 explains the theory behind the most frequently used concepts of the thesis. Section 3 explains how the data was collected and adjusted, as well as the methodology of the portfolio simulations and the credit risk model comparisons. Section 4 presents the models’ results from the simulations. Section 5 contains both a technical and non-technical discussion. Section 6 concludes.
2 Theoretical Framework

2.1 Default and Credit Risk

The probability of default, denoted \( PD_i \), is the probability that an obligor \( i \) defaults over some specified time horizon. In the context of the Basel framework, \( PD_i \) represents the probability that borrower \( i \) defaults within one year. The (uncertain) default event of borrower \( i \) is represented by the default indicator variable

\[
D_i = \begin{cases} 
1 & \text{with prob. } PD_i \\
0 & \text{with prob. } 1 - PD_i 
\end{cases}
\]  

(2.1)

Note that \( \mathbb{E}(D_i) = PD_i \). The exposure at default, denoted \( \widehat{EAD}_i \) is the total value the lender is exposed to in the event of default of obligor \( i \). The expected value of the exposure at default of an individual obligor is expressed as

\[
\mathbb{E}(\widehat{EAD}_i) = \widehat{EAD}_i. 
\]

(2.2)

Here, a notational distinction shall be pointed out between the random variable \( \widehat{EAD}_i \) and its deterministic expected value \( EAD_i \). To determine the lender’s total notional exposure outstanding toward all its obligors, simply aggregate all individual exposures to obtain \( EAD \), i.e., \( EAD = \sum_{i=1}^{n} EAD_i \). The fraction of expected exposure in the event of a default \( w_i \) towards each individual obligor \( i \) can then be found by

\[
w_i = \frac{EAD_i}{\sum_{j=1}^{n} EAD_j} = \frac{EAD_i}{EAD}. 
\]

(2.3)

The loss given default, denoted \( \widehat{LGD}_i \) is the fraction of exposure a lender cannot recover in the event of obligor \( i \) defaulting. Conversely, the recovery rate ("RR") can be defined as \( \widehat{RR}_i = 1 - \widehat{LGD}_i \). Bank debt is generally one type of senior debt, which means that these types of loans have a higher claim on the obligors assets in the case of default as opposed to e.g., equity holders. With the previously introduced notational convention the expected loss given default of obligor \( i \) is given as

\[
\mathbb{E}(\widehat{LGD}_i) = LGD_i. 
\]

(2.4)

The systematic risk variable originates from the assumptions of how assets are correlated within both the Asymptotic Single Risk Factor model (referred to as the “single-factor model”) and the multi-factor models. The models are formally introduced in Section 2.4. In the single-factor model, the assets are correlated with the systematic risk variable \( X \), which can be viewed as a system-
atic market risk and reflects the state of the economy which affects all assets in a portfolio. In the multi-factor models, the systematic risk variable is divided into sector-specific risk factors $X_s$, $s = 1, 2, \ldots, S$ where $S$ is the number of sectors in the model. The multi-factor models’ risk factors reflect the state of each sector. Both $X$ and $X_s$ are assumed to be standard normally distributed, $N(0, 1)$, and might be denoted as $X$ in order to ease notation. In other words, the main underlying assumption of the models is that the default variable $D_i$ is dependent on the state of the economy or the sectors. Put differently, $X$ can be interpreted as either the systematic market risk variable or as the sector-specific systematic risk variables, depending on if the expected loss is calculated through the single-factor model or the multi-factor models.

The loss variable is defined as the outstanding amount that can not be recovered in the event of obligor $i$ defaulting. It can be written as

$$L_i = \widehat{EAD}_i \cdot \widehat{LGD}_i \cdot D_i.$$  \hspace{1cm} (2.5)

It follows that the total loss variable $L_n$ for the lender can be expressed as

$$L_n = \sum_{i=1}^{n} L_i = \sum_{i=1}^{n} \widehat{EAD}_i \cdot \widehat{LGD}_i \cdot D_i.$$  \hspace{1cm} (2.6)

This can be seen as the aggregation of all outstanding amounts that may not be recovered in the case of defaults of certain obligors. Under the assumption that $\widehat{EAD}_i$ and $\widehat{LGD}_i$ are mutually independent for $i = 1, 2, \ldots, n$ and being independent of $D_i$, we can then find the total expected loss, given by

$$EL = \mathbb{E}(L) = \mathbb{E} \left( \sum_{i=1}^{n} \widehat{EAD}_i \cdot \widehat{LGD}_i \cdot D_i \right) = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot PD_i$$  \hspace{1cm} (2.7)

For further details on the underlying theory of the most prominent concepts of credit risk we recommend Hibbeln (2010).

2.2 Risk Measures

According to Hibbeln (2010), the EL is the most important way of measuring the risk of a single loan. However, on an aggregate portfolio level, additional tools are required to capture the dynamics and interaction of the outstanding loans. A risk measure makes it possible for the lender to quantify a “worst-case” scenario in the portfolio. Artzner et al. (1999) proposed four axioms to identify a suitable risk measure and its mathematical properties. If a risk measure satisfies all of the axioms, it is called coherent. In order to state the axioms mathematically, let $\mathcal{G}$ be a set of real-valued random variables and let $\rho : \mathcal{G} \rightarrow \mathbb{R}$ be some risk measure function. A risk measure is then said to
be coherent if the following four properties are satisfied:

- **Monotonicity**: for all \( Z_1, Z_2 \in \mathcal{G} \), with \( Z_1 \leq Z_2 \) then \( \rho(Z_1) \geq \rho(Z_2) \).
- **Subadditivity**: for all \( Z_1, Z_2 \in \mathcal{G} \), \( \rho(Z_1 + Z_2) \leq \rho(Z_1) + \rho(Z_2) \).
- **Positive homogeneity**: for all \( Z \in \mathcal{G} \) and \( \lambda \in \mathbb{R}_+ \) then \( \rho(\lambda \cdot Z) = \lambda \cdot \rho(Z) \).
- **Translation invariance**: for all \( Z \in \mathcal{G} \) and \( \kappa \in \mathbb{R} \) then \( \rho(Z + \kappa) = \rho(Z) - \kappa \).

The intuitive portfolio interpretation of the four axioms is the following: The monotonicity condition states that if portfolio one has systematically lower returns than portfolio two, then its risk must be greater. The subadditivity condition states that merging portfolios cannot increase the total risk. The homogeneity condition states that if you multiply the portfolio by a factor \( \lambda \), the total risk is simply scaled by the same proportion \( \lambda \). Finally, the translation invariance condition states that if you add a cash amount \( \kappa \) to the portfolio, the total risk is reduced by \( \kappa \).

The **Value-at-Risk** is intuitively defined by Jorion (2000) as a risk measure which “summarizes the worst loss over a target horizon that will not be exceeded within a given level of confidence”. As previously mentioned, in the context of the Basel III framework the target horizon equals one year and the level of confidence is defined as \( q = 0.999 \). The formal definition of the VaR is

\[
\text{VaR}_q(Z) = \inf \{ z \in \mathbb{R} : 1 - F_Z(-z) \geq q \},
\]

where \( F_Z \) is the cumulative distribution function of \( Z \). Artzner et al. (1999) show that the VaR does not satisfy the subadditivity condition for a coherent risk measure. However, the methodology is still very popular among practitioners globally due to its simplicity. Engle & Manganelli (2001) describe VaR as extraordinarily user-friendly since it reduces the risk of a portfolio to just one number and requires very little computing power.

### 2.3 Correlation

In this thesis, dependence is interpreted as the statistical association between two or more random variables. This implies that there exists some sort of relationship between them. Measuring the relationship can be done using **linear correlation**.

First and foremost, a distinction between the **equity correlation**, **asset correlation**, and **default correlation** is warranted. The equity correlation represents the correlation between two series of equity price returns. The asset correlation represents the correlation between two firms’ assets. Lastly,
the default correlation represents the correlation between the events of default of two firms. As the interaction between different types of assets and liabilities held by financial institutions have come under more scrutiny post financial crisis, the asset correlation in the context of credit risk has been researched more. However, since a company’s asset returns cannot be observed, the asset value has to be decomposed into liabilities and equity, which are both continuously observable in the market. Hence, to approximate the asset correlations a natural step would be to use the default correlation inferred from the market value of the firm’s debt, or the equity correlation inferred from the market value of the firm’s equity.

The basic idea behind modelling the default correlations of two obligors is rather intuitive. Merton (1974) tackled the problem on how to price corporate debt with repayment uncertainty. One of the key insights, which acted as the model’s foundation, was that a firm’s equity holders have a residual claim to a company’s assets in the event of a default. This allowed Merton to define the pay-off function for equity holders in firm $i$ at time $t$ as

$$E_{i,t} = \max\{A_{i,t} - L_{i,t}, 0\},$$

which is recognized as the pay-off function for a call option. $A_{i,t}$, $E_{i,t}$, and $L_{i,t}$ represent firm assets, equity and liabilities, respectively. Note that if $A_{i,t} \leq L_{i,t}$ the firm is valued at zero and consequently defaults. This is referred to as the firm’s default point. It follows that the joint probability distribution of two obligors during the same time span represents the likelihood of both obligors’ asset values falling below the default point threshold. In specific, the joint probability distribution of obligor $i, j$ for some time $t \leq T$ is defined as

$$\mathbb{P}(A_{i,t} \leq L_{i,t}, A_{j,t} \leq L_{j,t}).$$

With the joint probability distribution and the firms' individual default rate there are a number of ways to extract the default dependencies between firms. Such methodologies, among other, include the Maximum Likelihood, the Method-of-Moments, and the Vasicek Model, as derived by Vasicek (1987). The Vasicek Model will be presented in Section 2.4.1.

As opposed to the number of ways of inferring asset dependencies from default dependencies, the reviewed literature has proposed two main approaches with equity as the starting point. In accordance with Düllmann et al. (2007) these approaches are referred to as “direct” and “indirect”, respectively. The direct approach utilizes equity returns and prices without converting them to observable asset returns. Hence, the asset correlation will be approximated using the pair-wise linear correlation of
equity returns for all firms. In specific, let \( \mathbf{r}_i = (r_{i,1}, r_{i,2}, \ldots, r_{i,T})^\top \) and \( \mathbf{r}_j = (r_{j,1}, r_{j,2}, \ldots, r_{j,T})^\top \) represent two vectors of firms \( i \) and \( j \)'s equity log-returns in \( \mathbb{R}^T \). Define an appropriate rolling window size \( k \) and let \( k \leq t \leq T \). Note that from the original length \( T \) we will yield \( T - k + 1 \) rolling windows. Let the rolling equity log-return window number \( t-k \) for firm \( i \) and \( j \) be defined as \( \mathbf{r}_i^{(t-k)} = (r_{i,t-k+1}, \ldots, r_{i,t})^\top \) and \( \mathbf{r}_j^{(t-k)} = (r_{j,t-k+1}, \ldots, r_{j,t})^\top \) in \( \mathbb{R}^k \). It follows that for \( t = k, k+1, \ldots, T \) the rolling minimum, mean, and maximum pair-wise linear correlations for assets \( i \) and \( j \) can be defined as

\[
\rho_{i,j}^{(\text{min})} = \min_{t=k,\ldots,T} \left\{ \text{Corr}(\mathbf{r}_i^{(t-k)}, \mathbf{r}_j^{(t-k)}) \right\},
\]

\[
\rho_{i,j}^{(\text{mean})} = \frac{1}{T-k+1} \left( \sum_{t=k}^{T} \text{Corr}(\mathbf{r}_i^{(t-k)}, \mathbf{r}_j^{(t-k)}) \right),
\]

\[
\rho_{i,j}^{(\text{max})} = \max_{t=k,\ldots,T} \left\{ \text{Corr}(\mathbf{r}_i^{(t-k)}, \mathbf{r}_j^{(t-k)}) \right\}.
\]

There are three different types of correlations which will be considered in this thesis; the market-, inter-sector, and intra-sector correlations. The market correlation is the firm’s correlation to the overall market, the inter-sector correlation is the correlation between different sectors consisting of a number of firms, and lastly, the intra-sector correlation is the correlation between a firm and the sector it operates in. Consequently, Equation (2.11), (2.12), and (2.13) can be customized for calculation of all three different correlation types through substituting e.g., \( \mathbf{r}_i^{(t-k)} \) by the market and \( \mathbf{r}_j^{(t-k)} \) by a firm which operates in that market to obtain the market correlation. Similarly, you can set \( \mathbf{r}_i^{(t-k)} \) and \( \mathbf{r}_j^{(t-k)} \) to two different sectors, or a sector and a firm which operates in said sector to obtain the inter-sector or intra-sector correlation, respectively.

In this thesis, we will employ the “direct” correlation approach and compute asset values using Vasicek’s framework. The “indirect” approach shall only be mentioned here: It consists of two broad steps where in the first you approximate asset values and returns using the firm’s equity and liabilities. The second step is to estimate asset correlation using the approximated asset returns. For more details, see Bohn & Crosbie (2003) or Düllmann et al. (2007).

### 2.4 Credit Risk Modelling

#### 2.4.1 The Vasicek Model

Vasicek (1987) extended the Merton framework presented in the previous section by realizing it could be used to model dependencies. The value of company \( i \)’s assets at time \( t \) is modelled by the
multivariate Wiener process

\[ dA_{i,t} = \mu_i A_{i,t} dt + A_{i,t} \sum_{k=1}^{m} \sigma_{i,k} dW_{k,t}^{(1)} + \eta_i A_{i,t} dW_{i,t}^{(2)}, \]

where \( \mu_i, \eta_i, \) and \( \sigma_{i,1}, \ldots, \sigma_{i,m} \) are unknown constants a priori and \( W_{1,t}^{(1)}, \ldots, W_{m,t}^{(1)}, W_{i,t}^{(2)} \) are mutually independent Wiener processes. \( W_{1,t}^{(1)}, \ldots, W_{m,t}^{(1)} \) represent the systematic risk factors which are shared between all obligors and \( W_{i,t}^{(2)} \) represents the firm-specific idiosyncratic risk. By Proposition 5.2 in Bjork (2009), the solution to Equation (2.14) can be expressed as

\[ A_{i,1} = A_{i,0} \exp \left( \mu_i + \sum_{k=1}^{m} \left( \sigma_{i,k} Y_k - \frac{1}{2} \sigma_{i,k}^2 \right) + \eta_i \epsilon_i - \frac{1}{2} \eta_i^2 \right). \]  (2.15)

Here, \( Y_k \) and \( \epsilon_i \) are independent identically distributed standard normal variables. Let \( L_{i,t} = L_i \) be known and constant over \( t \). We can thus express the probability of default as

\[ PD_i = \mathbb{P} \left( A_{i,0} \exp \left( \mu_i + \sum_{k=1}^{m} \left( \sigma_{i,k} Y_k - \frac{1}{2} \sigma_{i,k}^2 \right) + \eta_i \epsilon_i - \frac{1}{2} \eta_i^2 \right) < L_i \right) \]

\[ = \mathbb{P} \left( \sigma_{i,1} Y_1 + \cdots + \sigma_{i,m} Y_m + \eta_i \epsilon_i \leq \ln \left( \frac{L_i}{A_{i,0}} \right) + \frac{1}{2} \sum_{k=1}^{m} \sigma_{i,k}^2 + \frac{1}{2} \eta_i^2 - \mu_i \right) \]

\[ = \mathbb{P} \left( \frac{\sigma_{i,1} Y_1 + \cdots + \sigma_{i,m} Y_m + \eta_i \epsilon_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} \leq \frac{\ln \left( \frac{L_i}{A_{i,0}} \right) + \frac{1}{2} \sum_{k=1}^{m} \sigma_{i,k}^2 + \frac{1}{2} \eta_i^2 - \mu_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} \right) \]  (2.16)

Note that the company defaults if

\[ \frac{\sigma_{i,1} Y_1 + \cdots + \sigma_{i,m} Y_m + \eta_i \epsilon_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} \leq \frac{\ln \left( \frac{L_i}{A_{i,0}} \right) + \frac{1}{2} \sum_{k=1}^{m} \sigma_{i,k}^2 + \frac{1}{2} \eta_i^2 - \mu_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} \].  (2.17)

This expression can for our purposes be rewritten in a more suitable way. Define

\[ \rho_i = \frac{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} \quad \text{and} \quad \alpha_i = \frac{\sigma_{i,k}}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} \]  (2.18)
However, since we prefer to have a vector of risk factors. Moreover, we get

\[
\frac{\sigma_{i,1}Y_1 + \cdots + \sigma_{i,m}Y_m + \eta_i \epsilon_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}} = \frac{\sigma_{i,1}Y_1 + \cdots + \sigma_{i,m}Y_m}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} + \frac{\eta_i \epsilon_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}}
\]

\[
= \rho_i \sqrt{\frac{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} \frac{\sigma_{i,1}Y_1 + \cdots + \sigma_{i,m}Y_m}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} + \sqrt{1 - \rho_i^2} \epsilon_i
\]

\[
= \rho_i \left( \frac{\sigma_{i,1}}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} Y_1 + \cdots + \frac{\sigma_{m,1}}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2}} Y_m \right) + \sqrt{1 - \rho_i^2} \epsilon_i
\]

(2.19)

where the more compact notations \( \alpha_i = (\alpha_{i,1}, \ldots, \alpha_{i,m})^\top \) and \( Y = (Y_1, \ldots, Y_m)^\top \) have been used.

Since the expression on the right hand side of Equation (2.19) follows a standard normal distribution, we get

\[
\Phi^{-1}(PD_i) = \frac{\ln \left( \frac{L_i}{A_i} \right) + \frac{1}{2} \sum_{k=1}^m \sigma_{i,k}^2 + \frac{1}{2} \eta_i^2 - \mu_i}{\sqrt{\sigma_{i,1}^2 + \cdots + \sigma_{i,m}^2 + \eta_i^2}},
\]

(2.20)

followed by the default variable

\[
D_i = \begin{cases} 
1 & \text{if } \rho_i \alpha_i^\top Y + \sqrt{1 - \rho_i^2} \epsilon_i \leq \Phi^{-1}(PD_i), \\
0 & \text{if } \rho_i \alpha_i^\top Y + \sqrt{1 - \rho_i^2} \epsilon_i > \Phi^{-1}(PD_i).
\end{cases}
\]

(2.21)

Notice that \( D_i|Y = y, \ldots, D_n|Y = y \) are mutually independent, this in turn implies that \( Y \) is a vector of risk factors. Moreover, we get

\[
P(D_i = 1|Y = y) = P \left( \rho_i \alpha_i^\top Y + \sqrt{1 - \rho_i^2} \epsilon_i \leq \Phi^{-1}(PD_i) | Y = y \right)
\]

\[
= P \left( \rho_i \alpha_i^\top Y + \sqrt{1 - \rho_i^2} \epsilon_i \leq \Phi^{-1}(PD_i) \right)
\]

\[
= \Phi \left( \frac{\Phi^{-1}(PD_i) - \rho_i \epsilon_i}{\sqrt{1 - \rho_i^2}} \right)
\]

(2.22)

where the final expression in Equation (2.22) is generally referred to as the Vasicek formula. Note, that in the original paper, Vasicek used \( \rho_i \) instead of \( \rho_i^2 \) for rewriting the left hand side in Equation (2.18). However, since we prefer to have \( \rho_i \) as the correlation coefficient instead of \( \sqrt{\rho_i} \) we’ve chosen...
to use $\rho_i^2$ in the derivation.

### 2.4.2 Single-Factor Model

The single-factor model, also known as the Asymptotic Single Risk Factor model ("ASRF"), is based on a special case of Vasicek’s model (presented in the previous section) and was introduced by Gordy (2003). This framework includes the two main assumptions that the portfolio is *infinitely fine-grained* and that only a *single systematic risk factor* influences the credit risk of all loans in the portfolio. The first assumption can, according to Bluhm et al. (2003), be expressed as the following two conditions:

\[
\lim_{n \to \infty} \sum_{i=1}^{n} EAD_i = \infty, \quad (2.23)
\]

and

\[
\sum_{n=1}^{\infty} \left( \frac{EAD_n}{\sum_{j=1}^{n} EAD_j} \right)^2 < \infty. \quad (2.24)
\]

The second assumption implies that all dependencies across credit events can be expressed by a systematic risk factor $X$ so that all the credit events are mutually independent conditional on the risk factor. For more details on the assumptions behind the single-factor model, see Bluhm et al. (2003).

In order to model asset values of obligors in a credit portfolio, the single-factor model assigns the following expression for the asset value $A_i$ of obligor $i$:

\[
A_i = \rho_i \cdot X + \sqrt{1 - \rho_i^2} \cdot \zeta_i \quad (2.25)
\]

where $\rho_i$ is the correlation between the obligor’s asset value and the market, $X \sim N(0, 1)$ is the systematic risk factor (or the “market”) which is shared by all obligors in the portfolio, and $\zeta_i \sim N(0, 1)$ is the idiosyncratic risk factor, which is individual for all obligors. Using Equation (2.6), the expression for the loss variable of the single-factor model becomes

\[
L_n = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot D_i = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot 1_{\{A_i \leq \Phi^{-1}(PD_i)\}} \quad (2.26)
\]

where $\Phi^{-1}$ is the quantile function of the standard normal distribution and $PD_i$ is the probability of default for obligor $i$. Note, that the default event occurs if the asset value $A_i$ falls below the quantile generated by the probability of default $PD_i$ of obligor $i$.

The single-factor model’s assumption in Equation (2.23) of infinite granularity of the portfolio with
n obligors implies, according to Bluhm et al. (2003), that the loss variable \( L_n \) of a portfolio with \( n \) obligors converges almost surely to the expected loss conditional on the systematic risk factor, which can be expressed as

\[
P \left( \lim_{n \to \infty} [L_n - \mathbb{E}(L_n | X)] = 0 \right) = 1
\]  

(2.27)

which means that, in terms of VaR, we get

\[
\lim_{n \to \infty} \text{VaR}_q(L_n) = \text{VaR}_q(\mathbb{E}(L_n | X))
\]

(2.28)

where \( q \) symbolizes the quantile level.

Using the second assumption expressed in Equation (2.24) of the single-factor model, regarding that the systematic risk factor is the only risk factor of the portfolio, it can be shown that

\[
\text{VaR}_q(\mathbb{E}[L|X]) = \mathbb{E}(L|X = \text{VaR}_{1-q}(X)).
\]

(2.29)

This provides the following pivotal equality:

\[
\text{VaR}^{ASRF}_q = \lim_{n \to \infty} \text{VaR}_q(L_n) = \text{VaR}_q(\mathbb{E}[L|X]) = \mathbb{E}(L|X = \text{VaR}_{1-q}(X))
\]

(2.30)

which can be used to justify, using the previously defined concepts of \( EAD_i, LGD_i \) and \( D_i \), for large values of \( n \), the approximation

\[
\text{VaR}^{ASRF}_q(L) = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot \mathbb{E}(D_i|X = \text{VaR}_{1-q}(X)).
\]

(2.31)

This formula is the foundation of the VaR and EC calculations used in the Basel II generated IRB framework, which is described in detail in Section 2.5.

2.4.3 Multi-Factor Models

The single-factor model (ASRF) presented in the previous section is, as mentioned, a special case of Vasicek’s model, assuming only one risk factor affecting the loans in a portfolio. However, the general Vasicek formula allows for several risk factors. This enables a framework for incorporating sector specific correlations, namely intra-sector correlation and inter-sector correlation. Intra-sector correlation is the correlation between the obligors and the sector which they operate in. This reflects that, in reality, all firms are to some extent affected by the overall market behaviour of their respective industries. On the other hand, inter-sector correlation captures how the different industries are correlated to each other, meaning that in reality, the market behaviour of an individual sector will
to some extent affect other sectors. In order to describe this mathematically we recall Vasicek’s model from Section 2.4.1 and re-write it in order to highlight the correlations. We will focus on two different multi-factor models: an inter-sector model and an intra-sector model.

We begin with the inter-sector model. The asset value $A_{i,s}$ of obligor $i$ operating in sector $s$ in a credit portfolio in the inter-sector model is expressed as:

$$A_{i,s} = \rho_s \cdot X_s + \sqrt{1 - \rho_s^2} \cdot \zeta_{i,s}$$

(2.32)

where $X_s$ is the systematic risk factor of sector $s$ with $s = 1, \ldots, S$ and $\rho_s$ is the mean intra-correlation between an obligor in sector $s$ and the sector itself. Note that all obligors operating within the same sector share this intra-correlation in the inter-sector model. The sector risk factor $X_s$ can be presented as a combination of independently and standard normally distributed factors $Y_k$ with $k = 1, \ldots, K$. The expression for $X_s$ becomes:

$$X_s = \sum_{k=1}^{K} \alpha_{s,k} Y_k \quad \text{with} \quad \sum_{k=1}^{K} \alpha_{s,k}^2 = 1$$

(2.33)

Note that both $X_s$ and $\zeta_{i,s}$ are mutually independent normally distributed variables with mean zero and standard deviation one. Also, note that the subscript $s$ of $\alpha_{s,k}$ denotes the specific sector $s$ in which obligor $i$ operates. The factor weights $\alpha_{s,k}$ are found via a Cholesky decomposition of the inter-sector correlation matrix, which means that the inter-sector correlations are given by

$$\rho_{s,u}^\text{inter} := \text{Corr} (X_s, X_u) = \sum_{k=1}^{K} \alpha_{s,k} \cdot \alpha_{u,k} \quad \text{where} \quad \sum_{k=1}^{K} \alpha_{s,k}^2 = 1 \quad \text{for} \quad s, u = 1, \ldots, S.$$

(2.34)

Hibbeln (2010) underlines that obligors in the same sector are highly correlated with each other when their intra-sector correlation is large and that correlations of obligors in different sectors depends on the factor weights which are derived from the inter-sector correlation matrix. This means that the dependence structure in the multi-factor model is completely described by the intra- and inter-sector correlations. The author adds that the Cholesky decomposition approach is common for generating correlated random variables and that the procedure leads to an identical number of independent risk factors $Y_k$ and dependent risk factors $X_s$. In other words, $K$ equals $S$. Another common method (which is not used in this thesis) is the principal component analysis which, in contrast, determines a reduced number of independent risk factors.

For the intra-sector model we use that this model allows for individual intra-sector correlations
which are unique for each obligor, yielding the following expression for the asset value of obligor $i$ operating in sector $s$:

$$A_{i,s} = \rho_{i,s} \cdot X_s + \sqrt{1 - \rho_{i,s}^2} \cdot \zeta_{i,s}$$  \hspace{1cm} (2.35)

where $\rho_{i,s}$ is the individual correlation between each asset and its sector, while $X_s$ is constructed as in the inter-sector model.

If we incorporate the multi-factor framework into the loss distribution function given by Equation (2.6) we get the following expression

$$L = \sum_{s=1}^{S} \sum_{i=1}^{n_s} \tilde{EAD}_{i,s} \cdot \tilde{LGD}_{i,s} \cdot D_{i,s} = \sum_{s=1}^{S} \sum_{i=1}^{n_s} \tilde{EAD}_{i,s} \cdot \tilde{LGD}_{i,s} \cdot \mathbf{1}_{\{A_{i,s} \leq \Phi^{-1}(PD_{i,s})\}}, \hspace{1cm} (2.36)$$

where $n_s$ is the number of obligors within sector $s$ and $D_{i,s}$ is the default variable of obligor $i$ in sector $s$. Note, that the default event occurs if the asset value $A_{i,s}$ falls below the quantile generated by the asset’s probability of default, $PD_{i,s}$.

### 2.4.4 Copulas and Copula Credit Risk Models

In order to model dependency between multivariate random variables, one can apply a multivariate distribution function called a **copula**. The copula has marginal distributions which are standard uniform. Hibbeln (2010) uses the following notation for the distribution of a copula:

$$C(u_1, \ldots, u_n) : [0, 1]^n \to [0, 1]. \hspace{1cm} (2.37)$$

Here, $n$ is the number of obligors or assets.

To further understand the fundamentals of copulas, we present the relationship between copulas, multivariate distribution functions and marginal distributions, introduced by Sklar (1959), who states that for every multivariate distribution function $F$, there exists a copula $C$ such that

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \hspace{1cm} x_1, \ldots, x_n \in \mathbb{R}, \hspace{1cm} (2.38)$$

where $F_1, \ldots, F_n$ are the marginal distributions of $F$.

One can also show that every multivariate distribution with continuous marginals admits a unique copula representation. The copulas and distribution functions are the building blocks when deriving new multivariate distributions with prescribed correlation structure and marginal distributions.
Bluhm et al. (2003) argue that the copula representation introduced by Sklar (1959) indicates why copulas became prominent within the field of risk modelling. The authors continue by stating that the overall purpose of fitting a credit model to data remains an old fashioned problem of finding a best fitting multivariate distribution, no matter if one looks at the copula and margins in joint or separate ways. However, separating margins and copula of a multivariate distribution is especially useful in situations when the marginals are known but the multivariate dependence is still to be determined. In such cases, like the framework used in this thesis, it makes sense to focus on the copula part of a multivariate distribution.

The copula models of this thesis aim to construct loss distributions for our credit portfolio with more tail dependence than it would be for a normally distributed portfolio. There exist a variety of different copulas with different characteristics, but this thesis will solely focus on Student’s-\( t \) copulas since they are the most common within the field of credit risk.

A multivariate Student’s-\( t \) copula with \( n \) dimensions and \( m \) degrees of freedom with normal marginal distributions has the following representation of its multivariate distribution function:

\[
F(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n)
\]

\[
= \mathbb{P}(T_1 \leq F_{t,m}^{-1}\left(\Phi(x_1)\right), \ldots, T_n \leq F_{t,m}^{-1}\left(\Phi(x_n)\right))
\]

(2.39)

where \( \Phi \) is the standard normal distribution function, \( F_{t,m}^{-1} \) denotes the inverse of the \( t \)-distribution function with \( m \) degrees of freedom and \((T_1, \ldots, T_n) \sim t(m, \Gamma)\), where \( t(m, \Gamma) \) is the \( n \)-variate \( t \)-distribution with \( m \) degrees of freedom and correlation matrix \( \Gamma \). From this, we finally achieve the copula representation:

\[
F(x_1, \ldots, x_n) = C_{m, \Gamma}\left(\Phi(x_1), \ldots, \Phi(x_n)\right).
\]

(2.40)

In order to get insight from a more practical perspective on how to model credit risk using the copula shown above we introduce our three different “copula models” which are all based on Student’s-\( t \) copulas with normally distributed margins. The three models are the single-factor copula model, the inter-sector copula model and the intra-sector copula model.

For the single-factor copula model we start by modelling the asset value of obligor \( i \) as in the single-factor case given in Equation (2.25):

\[
A_i = \rho_i \cdot X + \sqrt{1 - \rho_i^2} \cdot \zeta_i,
\]

(2.41)
where $X \sim N(0, 1)$ is the systematic risk factor representing the “market”, $\rho_i$ is the correlation between each obligor and the market and $\zeta_i \sim N(0, 1)$ is the idiosyncratic risk factor. However, in order to achieve a $t$-copula dependence we multiply by a scaling factor to get

$$\tilde{A}_i = \sqrt{\frac{m}{V}} \cdot A_i = \sqrt{\frac{m}{V}} \cdot \rho_i \cdot X + \sqrt{\frac{m}{V}} (1 - \rho_i^2) \cdot \zeta_i \quad V \sim \chi^2(m)$$

(2.42)

where the chi-square distributed variable $V$ has $m$ degrees of freedom and provides a $t$-copula dependence with $m$ degrees of freedom. Here, $V$ is independent of $X$ and $\zeta_i$.

In the inter-sector copula model, we start from the asset value of obligor $i$ in sector $s$ as in the inter-sector model given in Equation (2.32)

$$A_{i,s} = \rho_s \cdot X_s + \sqrt{1 - \rho_s^2} \cdot \zeta_{i,s}.$$  

(2.43)

where $X_s \sim N(0, 1)$ is the systematic risk factor of sector $s$ with $s = 1, \ldots, S$ and $\rho_s$ is the median correlation between an obligor in sector $s$ and the sector itself. Note that all obligors operating within the same sector share this intra-correlation in the inter-sector copula model, similarly to the inter-sector model introduced in Section 2.4.3. Using a scaling factor similar to the one in Equation (2.42) we achieve the following adjusted asset value expression:

$$\tilde{A}_{i,s} = \sqrt{\frac{m}{V}} \cdot A_{i,s} = \sqrt{\frac{m}{V}} \cdot \rho_s \cdot X_s + \sqrt{\frac{m}{V}} (1 - \rho_s^2) \cdot \zeta_{i,s} \quad V \sim \chi^2(m).$$  

(2.44)

For the intra-sector copula model we use the same expression for the asset value of obligor $i$ operating in sector $s$ as in Equation (2.35):

$$A_{i,s} = \rho_{i,s} \cdot X_s + \sqrt{1 - \rho_{i,s}^2} \cdot \zeta_{i,s}.$$  

(2.45)

Where $\rho_{i,s}$ is the correlation between obligor $i$ in sector $s$ and sector $s$ itself. If we once again use the scaling factor from Equation (2.42), we get

$$\tilde{A}_{i,s} = \sqrt{\frac{m}{V}} \cdot A_{i,s} = \sqrt{\frac{m}{V}} \cdot \rho_{i,s} \cdot X_s + \sqrt{\frac{m}{V}} (1 - \rho_{i,s}^2) \cdot \zeta_{i,s} \quad V \sim \chi^2(m).$$  

(2.46)

which is the adjusted asset value of obligor $i$ operating in sector $s$ which now has a $t$-copula dependency with $m$ degrees of freedom. Note that the $V$ is the same for all asset values in the model, introducing dependency between all obligors in the credit portfolio. This applies to all three copula models.
The loss variable for each copula model is similar to the one presented in Equation (2.36) except for the major difference that the default indicator variable $D_{i,s}$ is now given as

$$D_{i,s} = \begin{cases} 
1 & \text{if } \tilde{A}_{i,s} \leq F_{t,m}^{-1}(PD_{i,s}) \\
0 & \text{else} 
\end{cases}$$

(2.47)

where $F_{t,m}^{-1}$ is the quantile function of the $t$-distribution with $m$ degrees of freedom and $PD_{i,s}$ is the probability of default for obligor $i$ in sector $s$. Taking this into account, the loss variable for the single-factor copula model becomes

$$L = \sum_{i=1}^{n} \tilde{EAD}_i \cdot \tilde{LGD}_i \cdot D_i = \sum_{i=1}^{n} \tilde{EAD}_i \cdot \tilde{LGD}_i \cdot 1\{\tilde{A}_i \leq F_{t,m}^{-1}(PD_i)\},$$

(2.48)

and respectively for the inter-sector copula model and intra-sector copula model the loss variable becomes

$$L = \sum_{s=1}^{S} \sum_{i=1}^{n_s} \tilde{EAD}_{i,s} \cdot \tilde{LGD}_{i,s} \cdot D_{i,s} = \sum_{s=1}^{S} \sum_{i=1}^{n_s} \tilde{EAD}_{i,s} \cdot \tilde{LGD}_{i,s} \cdot 1\{\tilde{A}_{i,s} \leq F_{t,m}^{-1}(PD_{i,s})\}.$$ 

(2.49)

For more details on copulas and how to use them for credit risk simulations we recommend Hibbeln (2010), Bluhm et al. (2003) and Hult et al. (2012).

2.5 Economic Capital and the IRB Approach

Hibbeln (2010) shows that under the previous definitions of the loss variable and VaR, the single-factor model’s VaR under the IRB framework can be expressed as

$$\text{VaR}_q^{\text{IRB}}(L) = \text{VaR}_q^{\text{ASRF}}(L) = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot \mathbb{E}(D_i|X = \text{VaR}_{1-q}(X)).$$

(2.50)

Since $X$ follows a standard normal distribution (in the standard model), the $\text{VaR}_{1-q}(X)$ is easy to compute using the symmetry of the inverse cumulative distribution function, $\Phi^{-1}(1-q) = -\Phi^{-1}(q)$.

As mentioned in Section 1, the Economic Capital is the capital cushion a financial institution should hold as security against unexpected losses. Bluhm et al. (2003) defines the relationship between the Loss Variable, Expected Loss and Economic Capital as

$$EC_q = Q_q - EL$$

(2.51)
where $Q_q$ is the $q$-quantile of the loss variable $L_n$, given as

$$Q_q = \inf\{Q > 0| \mathbb{P}(L_n \leq Q) \geq q\}. \quad (2.52)$$

Now, in the IRB framework, this expression for the economic capital becomes:

$$EC_q = \text{VaR}_q(L_n) - EL. \quad (2.53)$$

This expression represents the amount of capital required for a bank to hold under the Basel III framework, given no regulations on capital requirement volume. Note that the relationship between the Loss Variable, Expected Loss and Economic Capital presented in Equation (2.51) and (2.53) holds also for other types of credit risk models (e.g., multi-factor and copula models) and will be used throughout the thesis.

Using the results from Equation (2.7), (2.50) and (2.53), the economic capital in the IRB setting, with $q = 0.999$, can be expressed as

$$EC_{\text{IRB}}^{0.999} = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot \left[ \Phi \left( \frac{\Phi^{-1}(PD_i) + \rho_i \cdot \Phi^{-1}(0.999)}{\sqrt{1 - \rho_i^2}} \right) - PD_i \right], \quad (2.54)$$

where $\rho_i$ is the correlation for corporate, sovereign, and bank exposure defined by the regulatory framework BSBC International Convergence of Capital Measurement and Capital Standards (2004) §272 and is expressed as

$$\rho_i = 0.12 \cdot \frac{1 - e^{-50PD_i}}{1 - e^{-50}} + 0.24 \cdot \left( 1 - \frac{1 - e^{-50PD_i}}{1 - e^{-50}} \right). \quad (2.55)$$

Note that this correlation is only dependent on the $PD$ of the specific obligor. This is fundamentally different from the correlations used in the models of this thesis since they are dependent on the specific obligor’s asset (equity) correlation with the market or its sector. For more details on the correlation formula presented in Equation (2.55), please see Basel Committe on Banking Supervision (2004).

### 2.6 Stochastic Simulation

The idea behind stochastic simulation, commonly known as the Monte-Carlo (“MC”) method, is fairly simple and has a broad range of applications in various fields such as engineering, statistics, physics, and finance. The MC method relies on a number of (user specified) random samplings from some distribution to obtain numerical results. In finance, the MC method has been used for
roughly 40 years and has become one of the most important tools in areas such as derivative pricing and scenario analysis, as pointed out by Bolder (2018). In the context of credit risk, MC simulation can be utilized in order to simulate any number (limited only by computational power) of possible future losses in a portfolio of obligors. The simulation results can thereafter be assessed in order to determine the risk. Exactly how this thesis has set up its MC simulations will be explained in Section 3.4.

2.7 Portfolio Diversification

The diversification measure considered is the *Herfindahl-Hirschman Index* (‘HHI’), which was introduced by Herfindahl (1951). The HHI was originally developed as a tool to measure sector concentration and competitiveness for antitrust issues. Concentration and competitiveness describe the amount of companies operating within a sector and how large their respective market share is. Moreover, the HHI is sometimes used by financial institutions to measure the level of diversification within a portfolio. A low HHI indicates a highly competitive sector - and by extension, a diversified portfolio. Mathematically, the HHI is defined as

\[
HHI = \sum_{i=1}^{n} w_i^2, \tag{2.56}
\]

where \( w_i \) represents the weight of company \( i \) in the portfolio (or sector). In the context of this thesis, it follows that \( w_i = EAD_i / EAD \) which is consistent with Equation (2.3). The HHI diversification thresholds presented in Table 1 are defined in line with the U.S. Department of Justice and Federal Trade Commission (2010).

<table>
<thead>
<tr>
<th>HHI</th>
<th>Diversification</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 0.01 )</td>
<td>Very High</td>
</tr>
<tr>
<td>( \leq 0.15 )</td>
<td>High</td>
</tr>
<tr>
<td>( \leq 0.25 )</td>
<td>Moderate</td>
</tr>
<tr>
<td>( &gt; 0.25 )</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 1: HHI thresholds

For the purpose of this thesis, the HHI will *not* be used to measure concentration risk, but rather as a tool to evaluate the overall portfolio and individual sector diversification, and how they might have impacted the results.
3 Methodology

3.1 Data

The data used for this thesis consists of weekly stock prices from Swedish companies listed on the Stockholm stock exchange (Nasdaq Stockholm). The prices are used to calculate both equity log-returns and equity correlations. In order to divide the corporates into sectors, like in the sector-models presented in Section 2.4.3 and Section 2.4.4, the Nasdaq-generated indices shown in Table 2 are used. Based on these indices, all companies which have continuous price data between 2007-01-01 and 2019-01-01 are selected. This means that our sample is smaller than the combined original indices. In the end, 138 firms from 8 different sectors were selected. All firms and their tickers can be found in Section A1. Price data from the indices themselves was also fetched, as well as prices from the OMXSPI all-share index (representing the “market” in the modelling) with 364 constituents. In total, we have 92,022 data points, ranging from 2007-01-01 to 2019-01-01.

<table>
<thead>
<tr>
<th>Index Name</th>
<th>Sector</th>
<th>Short Name</th>
<th>Components</th>
<th>Selected</th>
<th>Portfolio Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX0001PI</td>
<td>Oil and Gas</td>
<td>SX0001</td>
<td>6</td>
<td>2</td>
<td>3.6%</td>
</tr>
<tr>
<td>SX1000PI</td>
<td>Basic Materials</td>
<td>SX1000</td>
<td>23</td>
<td>9</td>
<td>3.6%</td>
</tr>
<tr>
<td>SX2000PI</td>
<td>Industrials</td>
<td>SX2000</td>
<td>93</td>
<td>52</td>
<td>45.4%</td>
</tr>
<tr>
<td>SX3000PI</td>
<td>Consumer Goods</td>
<td>SX3000</td>
<td>38</td>
<td>14</td>
<td>8.5%</td>
</tr>
<tr>
<td>SX4000PI</td>
<td>Healthcare</td>
<td>SX4000</td>
<td>55</td>
<td>20</td>
<td>9.4%</td>
</tr>
<tr>
<td>SX5000PI</td>
<td>Consumer Services</td>
<td>SX5000</td>
<td>38</td>
<td>17</td>
<td>12.9%</td>
</tr>
<tr>
<td>SX6000PI</td>
<td>Telecommunication</td>
<td>SX6000</td>
<td>5</td>
<td>3</td>
<td>9.2%</td>
</tr>
<tr>
<td>SX9500PI</td>
<td>Technology</td>
<td>SX9500</td>
<td>36</td>
<td>21</td>
<td>7.5%</td>
</tr>
</tbody>
</table>

Table 2: Sector indices

3.2 Credit Risk Models

The models which will be assessed in this thesis can be divided into two main categories, namely factor models and copula models. The two main model categories will create portfolio distributions which are normal for the factor models and Student’s-t for the copula models. However, the marginal distributions of the copula models are also normal, so only the dependence structure differs. The two main model categories can also be divided into two sub-groups, namely single-factor models and multi-factor models, having one and eight factors respectively. In total, six models are used for the credit risk simulations. The models are summarized in Table 3.
### Table 3: Model categorization

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Type</th>
<th>Number of Factors</th>
<th>Portfolio Distribution</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-factor (ASRF)</td>
<td>Factor</td>
<td>1</td>
<td>Normal</td>
<td>(2.25)</td>
</tr>
<tr>
<td>Inter-sector</td>
<td>Factor</td>
<td>8</td>
<td>Normal</td>
<td>(2.32)</td>
</tr>
<tr>
<td>Intra-sector</td>
<td>Factor</td>
<td>8</td>
<td>Normal</td>
<td>(2.35)</td>
</tr>
<tr>
<td>Single-factor copula</td>
<td>Copula</td>
<td>1</td>
<td>Student’s-t</td>
<td>(2.42)</td>
</tr>
<tr>
<td>Inter-sector copula</td>
<td>Copula</td>
<td>8</td>
<td>Student’s-t</td>
<td>(2.44)</td>
</tr>
<tr>
<td>Intra-sector copula</td>
<td>Copula</td>
<td>8</td>
<td>Student’s-t</td>
<td>(2.46)</td>
</tr>
</tbody>
</table>

#### 3.3 Asset Correlations

Firstly, the weekly stock prices are converted into log-returns through the expression:

\[ r_{i,t} = \ln \left( \frac{P_{i,t}}{P_{i,t-1}} \right) \] (3.1)

where \( r_{i,t} \) is the log-return and \( P_{i,t} \) is the equity price at time \( t \) of company \( i \). Secondly, the different types of correlations are computed using the log-return time series. For the market correlations in the single-factor models, the linear correlation between the returns of each company and the returns of OMXSPI was computed in accordance with Equation (2.12) as

\[ \rho_i^{(\text{mean})} = \frac{1}{626 - 52 + 1} \left( \sum_{t=52}^{626} \text{Corr} \left( r_i^{(t-52)}, r_{\text{OMXSPI}}^{(t-52)} \right) \right) \] (3.2)

which is the mean of the time series given by computing return correlation through a 52-week rolling window over 626 weeks. This procedure was also performed with maximum and minimum operators in accordance with Equation (2.11) and (2.13), yielding:

\[ \rho_i^{(\text{min})} = \min_{t=52, \ldots, 626} \left\{ \text{Corr} \left( r_i^{(t-52)}, r_{\text{OMXSPI}}^{(t-52)} \right) \right\} \] (3.3)

and

\[ \rho_i^{(\text{max})} = \max_{t=52, \ldots, 626} \left\{ \text{Corr} \left( r_i^{(t-52)}, r_{\text{OMXSPI}}^{(t-52)} \right) \right\} \] (3.4)

This provides us with three different market correlation levels for every company and consequently allows us to test three different “scenarios” for the overall market correlation level.

For the inter-sector models, inter-sector correlation matrices were created through running pairwise linear sector index return correlation calculations from 52-week rolling windows. The mean
inter-correlations were calculated in accordance with Equation (2.12) as

\[
\rho_{\text{inter}(\text{mean})}^{s,u} = \frac{1}{626 - 52 + 1} \left( \sum_{t=52}^{626} \text{Corr} \left( r_{s}^{(t-52)}, r_{u}^{(t-52)} \right) \right)
\]  

(3.5)

where \( \rho_{\text{inter}}^{s,u} \) is the inter-sector correlation between sector \( s \) and sector \( u \) and \( r_{s}, r_{u} \) are the return vectors of the indices of sector \( s \) and \( u \) respectively. This was done between all eight sectors, yielding an \( 8 \times 8 \) inter-sector mean correlation matrix. The procedure was repeated but with minimum and maximum operators in accordance with Equation (2.11) and (2.13) which yielded:

\[
\rho_{\text{inter}(\text{min})}^{s,u} = \min_{t=52,\ldots,626} \{ \text{Corr} \left( r_{s}^{(t-52)}, r_{u}^{(t-52)} \right) \} 
\]  

(3.6)

and

\[
\rho_{\text{inter}(\text{max})}^{s,u} = \max_{t=52,\ldots,626} \{ \text{Corr} \left( r_{s}^{(t-52)}, r_{u}^{(t-52)} \right) \}.
\]  

(3.7)

This provided three separate inter-sector correlation matrices in total, namely a mean-matrix, a max-matrix and a min-matrix. Having three different correlation matrices allow for further analysis of how the overall correlation levels affect the risk and the EC.

Mean intra-sector correlation for each sector was also calculated through running 52-week rolling windows of linear correlation between all companies and their respective sectors and taking the mean for all individual company time series. This was followed by taking the mean of all companies within the same sector, thus achieving a single intra-correlation which all companies within the same sector share. Mathematically, this looks like

\[
\rho_{\text{intra}(\text{mean})}^{s} = \frac{1}{n_{s}} \frac{1}{626 - 52 + 1} \left( \sum_{i=1}^{n_{s}} \sum_{t=52}^{626} \text{Corr} \left( r_{i}^{(t-52)}, r_{s}^{(t-52)} \right) \right)
\]  

(3.8)

where \( \rho_{\text{intra}(\text{mean})}^{s} \) is the mean intra-sector correlation of sector \( s \) which all firms within the sector share, \( n_{s} \) is the number of firms operating in sector \( s \) and \( r_{s} \) is the return vector of the index of sector \( s \). Similarly to the previous correlation types, the calculation was repeated but with minimium and maximum operators in order to get three different correlation levels. These calculations have the following expressions:

\[
\rho_{\text{intra}(\text{min})}^{s} = \min_{i=1,\ldots,n_{s}} \min_{t=52,\ldots,626} \{ \text{Corr} \left( r_{i}^{(t-52)}, r_{s}^{(t-52)} \right) \} 
\]  

(3.9)

and

\[
\rho_{\text{intra}(\text{max})}^{s} = \max_{i=1,\ldots,n_{s}} \max_{t=52,\ldots,626} \{ \text{Corr} \left( r_{i}^{(t-52)}, r_{s}^{(t-52)} \right) \}.
\]  

(3.10)
The intra-sector models use the same inter-correlation matrices as the inter-sector models. However, the intra-sector models have different intra-correlations since they allow for individually different correlations for all companies. These intra-correlations were generated through running 52-week rolling windows of linear correlation between each company and its sector, followed by taking the mean of the time series to get one value of intra-correlation per company. To see this more clearly, it can be expressed in accordance with Equation (2.12) as

\[
\rho_{i,s}^{\text{intra(mean)}} = \frac{1}{626 - 52 + 1} \left( \sum_{t=52}^{626} \text{Corr} \left( r_i^{(t-52)}, r_s^{(t-52)} \right) \right).
\]  

(3.11)

Analogously with the previous correlation measures, the intra-correlation calculation is repeated but with minimum and maximum operators in accordance with Equation (2.11) and (2.13). This gives us the following expressions:

\[
\rho_{i,s}^{\text{inter(min)}} = \min_{t=52, \ldots, 626} \left\{ \text{Corr} \left( r_i^{(t-52)}, r_s^{(t-52)} \right) \right\},
\]  

(3.12)

and

\[
\rho_{i,s}^{\text{inter(max)}} = \max_{t=52, \ldots, 626} \left\{ \text{Corr} \left( r_i^{(t-52)}, r_s^{(t-52)} \right) \right\}.
\]  

(3.13)

Note, that the correlations are calculated the same way for the factor models and the copula models since they are all based on Vasicek’s asset value modelling.

### 3.4 Modelling and Simulations

After calculating all different types of correlations, the asset value modelling was conducted. This was done through sampling the random variables needed in the expressions for the asset values of all companies, followed by checking each asset value if it came out below the quantile-level corresponding to its individual PD. If the asset value was simulated below this level, the default indicator variable corresponding to the company is set to one in a binary default vector. When all companies’ asset values have been simulated, the default vector is multiplied by both a vector containing all individual exposures (EAD) and a vector containing the LGD values. This yielded the loss for this particular simulation.

The simulation procedure described above is repeated \( K \) times, in accordance to a standard Monte Carlo simulation, providing a loss vector of length \( K \). In order to get the EL and the 99.9% VaR, the mean loss and the empirical 99.9%-quantile of the loss vector are determined, followed by taking the difference between them to finally achieve the EC. In specific, for each scenario \( k = 1, \ldots, K \)
we get the loss

\[ L^{(k)} = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot D_i, \]  

(3.14)

where \( D_i \) is dependent on the threshold given by the correlations from the previous section and asset threshold model equations referenced in Table 3. This yielded the loss vector \( \mathbf{L} = (L^{(1)}, L^{(2)}, \ldots, L^{(K)})^\top \). From this, the mean loss was calculated in order to get the EL, followed by calculating the empirical VaR through sorting the losses and taking the 99.9%-quantile.

All PD values were set as normally distributed random variables with means and standard deviations set at different levels depending on the company in question’s market capitalization. The firms are sorted after market capitalization as of 2019-01-01 and are divided into the three categories, namely large, medium and small. The categorization of the firms can be seen in Table 4. The randomly generated PDs for each category are set in line with Swedbank’s Pillar 3 (2017). The PDs in the Pillar 3 report are internally estimated probabilities of default for the bank’s exposures on the Swedish market.

<table>
<thead>
<tr>
<th>Category</th>
<th>Market Cap</th>
<th>Distribution</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large</td>
<td>≥ SEK 30 bn</td>
<td>Normal</td>
<td>0.39%</td>
<td>0.17%</td>
</tr>
<tr>
<td>Medium</td>
<td>≥ SEK 3 bn</td>
<td>Normal</td>
<td>1.17%</td>
<td>0.47%</td>
</tr>
<tr>
<td>Small</td>
<td>&lt; SEK 3 bn</td>
<td>Normal</td>
<td>3.51%</td>
<td>1.42%</td>
</tr>
</tbody>
</table>

Table 4: PD Distributions

The EAD values in the simulations are set equivalently to the weight of each firm in the credit portfolio containing all 138 firms of the sample. The weights are calculated based on the companies’ market capitalizations as of 2019-01-01. It is assumed that the credit exposure of a bank towards a firm is proportional to the size of the firm, i.e. larger firms are assumed to carry larger absolute debt from banks. The portfolio weights are also used for the calculation of the portfolio concentration. The exposure weighted average probability of default, i.e. the average weighted probability of default given that \( w_i = EAD_i/EAD \) for the entire portfolio was 0.55%.

Finally, the LGD values are all set to 45%. This is corresponding to the supervisory value for a senior unsecured loan in the Foundation IRB approach of the Basel II framework which was stated by Basel Committe on Banking Supervision (2004). This preset level of LGD was also used for the credit risk modelling conducted by Düllmann & Masschelein (2006).
4 Results

In the following section the results from the MC-simulation will be presented. Moreover, it will be shown how the EC determination based on the different methodologies converges for a sufficiently large number of iterations. The different levels (minimum, mean, and maximum) of inter-sector and intra-sector correlations will be presented along with the EC, VaR, and EL corresponding to the different levels. It will also be presented how the VaR and EC results are affected when varying the degrees of freedom for the copula models. In the subsection after that, the distribution of the losses in the different models will be presented through histograms. Lastly, we'll show the sector HHIs and the portfolio HHI.

4.1 Convergence of Monte Carlo Simulations

![Figure 1: Convergence of factor model MC simulations](image)

Figure 1: Convergence of factor model MC simulations
In the convergence plots for the factor and copula models presented by Figure 1 and Figure 2 respectively, the mean correlation has been used. For the copula models, three degrees of freedom were assigned. As expected, the copula models in Figure 2 yield a significantly higher EC than their factor-counterparts in Figure 1. The Intra-sector model converges in a more stable manner and yield a lower EC than the corresponding Single-factor model. However, the Inter-sector model has the lowest EC and lowest variance in its convergence. All models seem relatively stable with an acceptable variance around 100,000 simulations. Hence, this amount of simulations will be the basis for the results of the different models.

4.2 Correlation Matrices

The inter-sector correlation matrices needed for the Inter-sector model, Intra-sector model, Inter-sector copula model and Intra-sector copula model are presented below. The elements in Table 5 are the average correlations between the sectors over time and are calculated with rolling windows in accordance with Equation (3.5). Similarly, the elements in Table 6 and Table 7 are the minimum and maximum correlations between the sectors over time, calculated with rolling windows in accordance with Equation (3.6) and (3.7) respectively.
Table 5: Mean inter-sector correlation matrix

<table>
<thead>
<tr>
<th>Sector</th>
<th>SX0001</th>
<th>SX1000</th>
<th>SX2000</th>
<th>SX30000</th>
<th>SX4000</th>
<th>SX50000</th>
<th>SX6000</th>
<th>SX95000</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX0001</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX1000</td>
<td>0.55</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX2000</td>
<td>0.47</td>
<td>0.80</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX3000</td>
<td>0.40</td>
<td>0.67</td>
<td>0.73</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX4000</td>
<td>0.32</td>
<td>0.50</td>
<td>0.55</td>
<td>0.58</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX5000</td>
<td>0.32</td>
<td>0.50</td>
<td>0.57</td>
<td>0.59</td>
<td>0.46</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX6000</td>
<td>0.40</td>
<td>0.57</td>
<td>0.59</td>
<td>0.61</td>
<td>0.48</td>
<td>0.51</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SX9500</td>
<td>0.23</td>
<td>0.50</td>
<td>0.57</td>
<td>0.52</td>
<td>0.45</td>
<td>0.41</td>
<td>0.39</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6: Minimum inter-sector correlation matrix

<table>
<thead>
<tr>
<th>Sector</th>
<th>SX0001</th>
<th>SX1000</th>
<th>SX2000</th>
<th>SX30000</th>
<th>SX4000</th>
<th>SX50000</th>
<th>SX6000</th>
<th>SX95000</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX0001</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX1000</td>
<td>0.16</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX2000</td>
<td>-0.06</td>
<td>0.32</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX3000</td>
<td>-0.22</td>
<td>-0.01</td>
<td>0.29</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX4000</td>
<td>-0.13</td>
<td>0.16</td>
<td>0.03</td>
<td>0.22</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX5000</td>
<td>-0.13</td>
<td>-0.24</td>
<td>-0.07</td>
<td>0.16</td>
<td>0.12</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX6000</td>
<td>-0.19</td>
<td>-0.03</td>
<td>0.09</td>
<td>0.08</td>
<td>-0.02</td>
<td>0.07</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SX9500</td>
<td>-0.10</td>
<td>0.14</td>
<td>0.34</td>
<td>0.26</td>
<td>0.03</td>
<td>-0.13</td>
<td>-0.03</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 7: Maximum inter-sector correlation matrix

<table>
<thead>
<tr>
<th>Sector</th>
<th>SX0001</th>
<th>SX1000</th>
<th>SX2000</th>
<th>SX30000</th>
<th>SX4000</th>
<th>SX50000</th>
<th>SX6000</th>
<th>SX95000</th>
</tr>
</thead>
<tbody>
<tr>
<td>SX0001</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX1000</td>
<td>0.83</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX2000</td>
<td>0.80</td>
<td>0.97</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX3000</td>
<td>0.81</td>
<td>0.91</td>
<td>0.94</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX4000</td>
<td>0.72</td>
<td>0.82</td>
<td>0.80</td>
<td>0.78</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX5000</td>
<td>0.68</td>
<td>0.82</td>
<td>0.86</td>
<td>0.87</td>
<td>0.85</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SX6000</td>
<td>0.77</td>
<td>0.81</td>
<td>0.83</td>
<td>0.85</td>
<td>0.75</td>
<td>0.77</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SX9500</td>
<td>0.59</td>
<td>0.81</td>
<td>0.82</td>
<td>0.77</td>
<td>0.77</td>
<td>0.71</td>
<td>0.65</td>
<td>1</td>
</tr>
</tbody>
</table>

27
When inspecting the correlation matrices, the first thing one might notice is that, quite intuitively, the correlation is the largest between sectors which are closely connected business-wise, such as SX1000 (Basic Materials) and SX2000 (Industrials) or SX3000 (Consumer Goods) and SX2000 (Industrials). These pairs have quite high correlations even in the minimum case, never going lower than 32% and 29% respectively. The opposite holds for sectors which are quite far apart business-wise, such as SX9500 (Technology) and SX0001 (Oil & Gas) which have a correlation of only 23% on average and a maximum of only 59%.

Inspecting the minimum correlation matrix (Table 6), one can see that many of the sectors have reached negative correlation levels at some point. This concerns especially the pairs containing SX0001 (Oil & Gas). When comparing the three different correlation matrices it becomes very evident that the overall correlation levels between sectors on the Swedish market vary dramatically and can both reach very high maximum levels and drop to below zero. These differences in correlation levels hugely impact the results when modelling credit risk using sector-sensitive models, as seen in Section 4.3 where the outcomes of the models are compared for minimum, mean and maximum levels of correlation.

Minimum, mean and maximum levels are also generated for the correlation types that are individual for each asset, as described in Section 3.3. However, these correlations are not presented here because of the large number of firms in the portfolio.

### 4.3 Credit Risk Results

In this section, the results from the simulations of the credit risk models are presented. The VaR, the EL and the EC are all presented as percentages of the total exposure. Firstly, the results of the factor models with different levels of correlation are presented in Table 8. Secondly, the same type of results are presented but for the copula factor models in Table 9. Finally, the results from varying the degrees of freedom of the copula factor models are presented in Table 10.

#### 4.3.1 Factor Models

The first obvious result seen in Table 8 is that the $EL$ is basically the same for all models at around 0.25%. This is due to the fact that the expected value of the loss is just dependent on the LGDs, EADs and PDs of all individual firms (see Equation (2.7)). The LGDs and EADs are all deterministic and the individual PDs are the same for all simulations of a specific model. Another clearly visible result is that the VaR (and analogously the EC) increases when the correlation level increases. This holds for all factor models. The Inter-sector model stands out from the other two
when comparing their VaR and EC generated from mean correlation since it gives significantly lower values. On the other hand, the Inter-sector model gives the significantly largest values of VaR and EC compared to the other models when maximum correlation is used. This sensitivity of the Inter-sector model might be due to the double operator methodology of generating its intra-sector correlations. More on this in Section 5.

As a benchmark to the results in Table 8, the IRB-based single-factor model provided in Basel III (shown in Equation (2.54)) gives, with the same input as our Single-factor model with mean correlation, an EC of 6.58%. This is lower than what our Single-factor model yields, but higher than the Inter-sector model.

<table>
<thead>
<tr>
<th>Model</th>
<th>VaR</th>
<th>EL</th>
<th>EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-factor (ASRF)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_i^{(\text{min})} )</td>
<td>4.49%</td>
<td>0.25%</td>
<td>4.24%</td>
</tr>
<tr>
<td>( \rho_i^{(\text{mean})} )</td>
<td>8.83%</td>
<td>0.24%</td>
<td>8.59%</td>
</tr>
<tr>
<td>( \rho_i^{(\text{max})} )</td>
<td>17.85%</td>
<td>0.24%</td>
<td>17.61%</td>
</tr>
<tr>
<td>Inter-sector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(min)})} )</td>
<td>4.57%</td>
<td>0.25%</td>
<td>4.33%</td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(mean)})} )</td>
<td>5.08%</td>
<td>0.25%</td>
<td>4.83%</td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(max)})} )</td>
<td>25.99%</td>
<td>0.24%</td>
<td>25.75%</td>
</tr>
<tr>
<td>Intra-sector</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(min)})} )</td>
<td>4.61%</td>
<td>0.25%</td>
<td>4.36%</td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(mean)})} )</td>
<td>7.88%</td>
<td>0.24%</td>
<td>7.63%</td>
</tr>
<tr>
<td>( \rho_{s,u}^{(\text{inter(max)})} )</td>
<td>19.01%</td>
<td>0.25%</td>
<td>18.76%</td>
</tr>
</tbody>
</table>

Table 8: 99.9%-VaR, EL and EC from min, mean and max correlation level(s)

### 4.3.2 Copula Factor Models

The results of the copula factor models with three degrees of freedom shown in Table 9 share some characteristics with the results of the factor models. The VaR and EC values also increase when the correlation levels increase, and the EL is on the same level as for the previously presented results at around 0.25%. However, there are a couple of important differences. Firstly, as seen already in Section 4.1, the overall VaR and EC levels are significantly higher for the copula factor models.
than for the factor models. This holds for all correlation levels. Secondly, the VaR and EC levels increase slower (relatively) when the correlation levels increase compared to the factor models. In other words, the copula factor models seem to be less sensitive to changes in correlation levels than the factor models.

When comparing the copula factor models individually we see that the Intra-sector copula model has the tightest span between minimum and maximum correlation generated VaR and EC. On the other hand, the Inter-sector copula model has the largest span between minimum and maximum, while having significantly lower values from mean correlation. This correlation-sensitive behaviour is recognized from its factor model counterpart and might also be due to its double operator methodology.

Finally, recall that the correlation levels are identical for the factor models and the copula factor models since they are all generated according to the methodology in Section 3.3. Also, note that these results are generated from the simulations of the copula factor models.

<table>
<thead>
<tr>
<th>Model</th>
<th>VaR</th>
<th>EL</th>
<th>EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-factor copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\rho_i^{(\text{min})})</td>
<td>14.73%</td>
<td>0.25%</td>
<td>14.47%</td>
</tr>
<tr>
<td>(\rho_i^{(\text{mean})})</td>
<td>22.15%</td>
<td>0.25%</td>
<td>21.90%</td>
</tr>
<tr>
<td>(\rho_i^{(\text{max})})</td>
<td>31.83%</td>
<td>0.25%</td>
<td>31.58%</td>
</tr>
<tr>
<td>Inter-sector copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\rho_{\text{inter(min)}}^{s,u}) &amp; (\rho_{\text{intra(min)}}^{s})</td>
<td>15.54%</td>
<td>0.25%</td>
<td>15.28%</td>
</tr>
<tr>
<td>(\rho_{\text{inter(mean)}}^{s,u}) &amp; (\rho_{\text{intra(mean)}}^{s})</td>
<td>16.85%</td>
<td>0.25%</td>
<td>16.60%</td>
</tr>
<tr>
<td>(\rho_{\text{inter(max)}}^{s,u}) &amp; (\rho_{\text{intra(max)}}^{s})</td>
<td>36.67%</td>
<td>0.26%</td>
<td>36.41%</td>
</tr>
<tr>
<td>Intra-sector copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\rho_{\text{inter(min)}}^{s,u}) &amp; (\rho_{\text{intra(min)}}^{i,s})</td>
<td>14.95%</td>
<td>0.24%</td>
<td>14.70%</td>
</tr>
<tr>
<td>(\rho_{\text{inter(mean)}}^{s,u}) &amp; (\rho_{\text{intra(mean)}}^{i,s})</td>
<td>20.01%</td>
<td>0.25%</td>
<td>19.76%</td>
</tr>
<tr>
<td>(\rho_{\text{inter(max)}}^{s,u}) &amp; (\rho_{\text{intra(max)}}^{i,s})</td>
<td>30.34%</td>
<td>0.25%</td>
<td>30.09%</td>
</tr>
</tbody>
</table>

Table 9: 99.9%-VaR, EL, and EC from min, mean and max correlation level(s). \(m = 3\)

In order to investigate the sensitivity of the copula factor models concerning the assigned degrees of freedom we run simulations of the models using mean correlation levels while letting \(m\) be 3, 10 and 1000. The results in Table 10 show clearly that the VaR and EC levels decrease when the degrees of freedom.
freedom increase. This is consistent for all models. One can also see how the span between lowest and highest VaR and EC levels are tighter for the Single-factor copula model and Intra-sector copula model compared to the Inter-sector model. Thus, the Inter-sector model turns out to be the most sensitive even in this aspect compared to its peers.

As expected, the copula factor models' VaR and EC levels converge towards their factor model counterparts when the degrees of freedom becomes large. This is due to the relationship between the Student's $t$-distribution and the normal distribution, shown in 2.4.4 and discussed further in Section 5.

<table>
<thead>
<tr>
<th>Model</th>
<th>VaR</th>
<th>$EL$</th>
<th>$EC$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-factor copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m = 3$</td>
<td>22.15%</td>
<td>0.25%</td>
<td>21.90%</td>
</tr>
<tr>
<td>$m = 10$</td>
<td>13.60%</td>
<td>0.25%</td>
<td>13.35%</td>
</tr>
<tr>
<td>$m = 1000$</td>
<td>9.18%</td>
<td>0.25%</td>
<td>8.93%</td>
</tr>
<tr>
<td>Inter-sector copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m = 3$</td>
<td>16.85%</td>
<td>0.25%</td>
<td>16.60%</td>
</tr>
<tr>
<td>$m = 10$</td>
<td>8.86%</td>
<td>0.25%</td>
<td>8.62%</td>
</tr>
<tr>
<td>$m = 1000$</td>
<td>5.17%</td>
<td>0.25%</td>
<td>4.92%</td>
</tr>
<tr>
<td>Intra-sector copula</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m = 3$</td>
<td>20.01%</td>
<td>0.25%</td>
<td>19.76%</td>
</tr>
<tr>
<td>$m = 10$</td>
<td>12.48%</td>
<td>0.25%</td>
<td>12.24%</td>
</tr>
<tr>
<td>$m = 1000$</td>
<td>8.04%</td>
<td>0.25%</td>
<td>7.79%</td>
</tr>
</tbody>
</table>

Table 10: 99.9%-VaR, $EL$ and $EC$ from mean correlation level(s)

4.4 Loss Histograms

The losses generated from the MC simulation of each model are presented as histograms in Figure 3 and Figure 4. Note that only the losses exceeding the EL of each model are included.
When inspecting the histograms we see that the tails are significantly fatter for the copula models compared to the factor models. This is in accordance with the VaR and EC results presented in the previous sections. We can also see a smoother distributional decay for the copula models compared to the factor models, highlighting the tail dependency of the assets in the copula models’ portfolios. One histogram which stands out from the rest is the one for the Inter-sector factor model. Not a single loss exceeds 13%, yielding a very short tail for the loss distribution of the model. Finally,
we can see that the loss distributions of the factor models are somewhat irregular in their shapes. The majority of the losses are around the minimum levels just above the EL, but there’s a bulk of slightly larger losses at around 3% for all three factor models. This irregular behaviour of the loss distributions of the factor models are in strong contrast to the smoothness of the copula models’ distributions.

### 4.5 Diversification Impact

The portfolio diversification $HHI$ and the different sector spreads are presented in Table 11. Note that the simulated portfolios and sectors might not truly reflect a real world outstanding loan portfolio for a major bank. It follows that the approximated intra-sector correlation (used in the inter-sector model) will be heavily affected by a few outliers. Similarly, a default event for a large firm will have higher impact on the EC and thus drive the loss up due to its larger weight.

<table>
<thead>
<tr>
<th></th>
<th>$HHI$</th>
<th>Diversification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Portfolio</td>
<td>0.04</td>
<td>High</td>
</tr>
<tr>
<td>SX0001</td>
<td>0.95</td>
<td>Low</td>
</tr>
<tr>
<td>SX1000</td>
<td>0.24</td>
<td>Moderate</td>
</tr>
<tr>
<td>SX2000</td>
<td>0.10</td>
<td>High</td>
</tr>
<tr>
<td>SX3000</td>
<td>0.23</td>
<td>Moderate</td>
</tr>
<tr>
<td>SX4000</td>
<td>0.17</td>
<td>Moderate</td>
</tr>
<tr>
<td>SX5000</td>
<td>0.31</td>
<td>Low</td>
</tr>
<tr>
<td>SX6000</td>
<td>0.54</td>
<td>Low</td>
</tr>
<tr>
<td>SX9500</td>
<td>0.56</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 11: Portfolio and Sector Diversification
5 Discussion

We would like to commence the discussion by comparing the single-factor models with the inter- and intra-factor models. First of all, the single-factor models are more sensitive than their peers to systematic risk. This comes naturally from the fact that the assets in the single-factor models are all assumed to be affected by the same "market" condition. The other models are more flexible in this sense. However, the credit risk results, presented in Section 4.3, show that the differences in VaR and EC estimates between the single-factor models and the intra-sector model are rather small. This holds for both the factor and the copula models. This was a rather counter-intuitive result, but might indicate that the assumption of all firms being affected by the “market” is rather reasonable, or at least not less reasonable than assuming that firms are affected by their sector.

The model type which stands out from its peers (in both the factor models and the copula models) is the inter-sector model. Both the convergence plots in Section 4.1 and the credit risk result tables in Section 4.3 show clearly that the inter-sector models behave differently compared to their peers. The VaR and EC outcomes from mean correlation levels are significantly smaller compared to the other models, but the span between minimum correlation and maximum correlation results are remarkably larger. This behaviour of the inter-sector model is evident both when looking at the factor models, the copula models with $m = 3$, and when alternating the degrees of freedom, thus indicating a consistent model specific characteristic. The reason for this model characteristic is most likely found in the methodology of how to generate the intra-correlation for the model. Recalling Equation (3.8), (3.9) and (3.10) we see that the intra-correlations on all levels are generated through double operators, either min, mean or max. This methodology might be considered dangerous and non-realistic since it generalizes the calculations severely. More specifically, in a sector where there are numerous small firms with low intra-correlation, the mean will be misguidedly low. Since all firms within the same sector share the same intra-correlation in the inter-sector model, large firms with naturally high correlations will have too low intra-correlation. The results from this are the low values of VaR and EC from the simulation results based on mean correlation levels. On the other hand, the double maximum operators give the small firms with a naturally low correlation a high intra-correlation which, in reality, suits the bigger firms better. This leads to remarkably large estimates of VaR and EC. From a practitioner’s perspective, the use of double maximum or minimum operators might be considered inappropriate. Instead, one might consider to combine the minimum or maximum over time with using average over the firms. In this way, the results might be more stable. This mixed approach is not investigated in this thesis. However, in Appendix A2 we have plotted the minimum, average and maximum intra-correlation for all sectors over our time.
interval. These plots show clearly how much it differs between the company with the highest and lowest intra-correlation.

When comparing the factor models with their copula counterparts one can clearly see that the copula models (in Table 9) generate larger values for the VaR and EC. This holds for all three main types of models. This result was rather expected since the Student’s-t copula introduces a tail dependence among the assets, leading to larger risk estimates. However, the copula models provide smaller spans between minimum correlation level VaR (and EC) and maximum correlation level VaR (and EC), compared to the factor models. In other words, the copula models seem to be less sensitive to the correlation level input. This might be explained by the dependency introduced by the Student’s-t copula where all assets are affected by the same factor which might remove some of the model’s reliance on the correlation input.

An interesting aspect of the copula models is that one can adjust the degrees of freedom for achieving different risk results, as shown in Table 10. The rather obvious conclusion one can make from the table is that the VaR and EC values decrease significantly when increasing the degrees of freedom. When reaching degrees of freedom as high as 1000 one can see that the copula models’ results clearly approach the VaR and EC values of the factor models, which have normally distributed portfolios. This result was rather expected since it was pointed out by Bluhm et al. (2003) that the Student’s-t distribution approaches the normal distribution when it reaches “infinite” degrees of freedom and that this phenomenon can be shown through simulating portfolios having the two different distributions. This thesis does not investigate what the “optimal” level of degrees of freedom is for copula credit risk modelling. It is all up to the practitioner to decide a suitable degrees of freedom for his or her copula models. The research performed here aims to create awareness of how the chosen degrees of freedom will impact the VaR and EC estimates of such models.

The differences between the factor models’ and copula models’ loss distributions, as seen in Figure 3 and Figure 4, might be explained by the distribution characteristics of the portfolios. The copula models’ smoother decay of the loss distribution’s right tail is most likely explained by the fact that the assets are tail dependent through the Student’s-t copula. However, the irregular behaviour of the decay of the right tail for the factor models is rather less intuitive. In fact, it is difficult to explain exactly why there are losses of around 3% occurring at a higher frequency than e.g. 2%. Perhaps, this is sample related. A rather large company which is highly correlated with the market or its sector might alternate between defaulting and not defaulting in the simulations, giving rise to irregular sizes of the portfolio losses. Other than that, we find no explanation for this phenomenon.
This thesis revolves quite heavily around correlation, namely different levels of correlation and different types of correlation and how they affect the outcomes when modelling credit risk. It is important to recall that we have used linear equity correlation as a proxy for asset correlations and that we are modelling asset values and defaults of companies - not equity prices. The choice of using equity correlation for credit risk modelling is discussed in Section 2.3 and supported by Düllmann et al. (2007). An alternative to the linear correlation approach would be to use e.g. Kendall’s Tau or Spearman’s rank correlation. These correlation measures might have provided different correlation estimates, leading to contrasting credit risk outcomes. Especially Kendall’s Tau could have been relevant for the copula models since it is used in copula simulations by e.g. Hult et al. (2012). However, we decided to stick with linear correlation since it is used in numerous other studies within the field of credit risk.

The correlations used for all models are static since they are the minimum, mean or maximum of rolling correlation time series. One could argue that it would have been interesting to look at how the risk calculations differ over time, namely if the VaR or EC would have been calculated for all time steps using time varying correlations. This would have been rather interesting to see since this type of approach would be able to show how the risk exposure of a portfolio evolves over time. Unfortunately, this would require extremely heavy computing power since the VaR and EC would have to be simulated for every time step, meaning that in our case, we would have to perform $100,000 \times 626$ simulations multiplied by our 626 weeks of data - plus all the rolling correlation windows. We do unfortunately not have access to computing power of this scale, but with the resources of a practitioner it might be possible.

It is important to highlight that the results regarding the VaR and EC estimates might differ significantly in scale from a “standard” corporate credit portfolio of a Swedish bank. This has several explanations. One of them is that all our values on EAD, LGD and PD are based on our own assumptions and estimations. We have not had access to any “real” credit portfolio with actual recovery rates or exposures towards the companies in our indices. Even for practitioners it is impossible to know the “true” probability of default of a company, but having access to historical default data and analyst recommendations will at least provide reasonable PD estimates which can be used in practice by credit risk teams. In our case, we have had to use public data for the PDs which has just been average values for analyst constructed sectors which have not matched our own sectors perfectly. In order to compensate for this, the PDs in our simulations have been given as random variables with means based on the averages of the sectors in Swedbank’s Pillar 3 (2017).
is also worth repeating that the EADs in this thesis are simply each company’s market capitalization weight in the total portfolio. A practitioner with access to real exposure weights will most likely get more balanced results since it is not reasonable that a company’s size is perfectly proportional to its loans. This will also mean that a practitioner’s model will be less dependent on large firms and will probably have more accurately adjusted PDs.

Another reason for why our results might differ significantly in size from typical credit risk values of practitioners’ portfolios is that we have customized the constituents of our indices in order to match our data requirements. We stated in Section 3.1 that all firms must have continuous equity price data during the entire time period in order to be included. In other words, our sample became a trade-off between component amount and time period length. Since we wanted to include at least one financial crisis, we decided to set 2007-01-01 as our starting date. This lead to our portfolio being constituted of only 138 firms. Also, this made a couple of our our indices very small component wise, with e.g. the Oil & Gas having only two firms and Telecommunication having only three. Looking at the HHI values (Table 11) of the individual sectors and the portfolio in its entirety we conclude that most indices have rather low diversification. This might have inflated the intra-sector correlation for these indices, leading to larger values of VaR and EC. However, the portfolio in its entirety has a rather high diversification according to the HHI measure. This means that the single-factor models (which are not dependent on the sectors, but rather on the “market” as a whole) might not have been affected by the small indices in our sample.

One aspect to have in mind regarding the data used in this thesis is that we are in one of the longest periods of bull market ever. More specifically, the equity prices have on average been rising since 2009 - a record long time, according to Egan (2018). This abnormally long bull market might have affected the equity correlations generated from our data. Some researchers, including Longin & Solnik (2001), argue that equity correlation increases during bear markets (price decline trends) but not bull markets. Thus, the data set might have been more balanced if it contained e.g. the stock market crash in 2000. Hopefully, practitioners are aware of the special environment on the current stock market and are continuously adjusting their credit risk models accordingly.
6 Conclusion and Further Research

In order to draw conclusions from the research performed, we recall that the purpose of this thesis is to give the reader some insight into how to calculate a bank’s Economic Capital using different credit risk models, based on a fictitious corporate loan portfolio. The purpose is also to show how the models differ in nature and how correlation levels affect the credit risk outcomes. As stated in the introduction, the aim is not to find an optimal model, but rather to make a practitioner aware of what to expect when modelling credit risk using e.g. multi-factor settings and/or copulas.

We reiterate the importance of correlations and their impact on the outcome when modelling credit risk. The results show how the models’ results vary dramatically with different correlation structures (market vs. inter and intra) and correlation levels (min, mean and max). Practitioners are recommended to stress test their models with this in mind. Since correlations might differ over time, and thus impact the EC, it’s of paramount importance to regularly analyze the credit portfolio’s risk exposure for different levels of asset correlation.

We conclude that the single-factor (ASRF) and the intra-sector model do not differ significantly regarding VaR and EC estimates - at least not for our portfolio. However, the inter-sector model stands out by being a lot more sensitive to correlation levels due to its double-operator methodology. The generalizations that this model involves might harm the reliability of its results. The other two model types are more stable, although the intra-sector model demands more correlation data and is slightly more complex than the single-factor model. A practitioner should contemplate whether to go for the straightforward approach of the single-factor model or if the detail of the intra-sector model methodology is more attractive.

Another conclusion we draw from the modelling is that using a copula-based framework will provide more conservative risk outcomes. The tail dependency introduced by the Student’s-\( t \) copula brings larger simulated portfolio losses and smoother decay of the loss distribution’s right tail. A practitioner who seeks to try modelling for more extensive “worst-outcome-scenarios” might consider using the copula models (or at least use them for comparison with the factor models). The copula models’ increasing popularity among professional credit risk analysts might be explained by the combination of the models’ conservativeness and their flexibility regarding the manual calibration of the degrees of freedom. The copula models are also less sensitive to correlation level input than factor models, which should be an attractive aspect for practitioners who don’t want to be too affected by small changes in correlation between equity prices.
We recommend further research on several topics related to this thesis. One example is that researchers could analyze credit risk in a broad international credit portfolio, including both European, American and Asian companies. This type of research might also involve comparing how e.g. inter-sector correlation differs when having a broader scope of international companies, compared to having domestic companies and sectors only. One might also look at how inter- or intra-sector correlation differs between different continents. Another area for further research could be to apply different types of copulas on credit risk modelling. The copula family includes e.g. Gaussian, Student’s-\(t\), Clayton and Gumbel types, which all have different characteristics. Comparing these could help practitioners find a framework which might suit their risk profile the best. We also suggest an extensive study of credit risk spanning an even longer time period than ours. This would result in data sets including more economic downturns and crashes, providing a more reliable empirical modelling environment. Finally, we suggest that future research could focus on modelling credit risk using other types of asset correlation measures, such as Kendall’s Tau and Spearman’s Rho. This could provide interesting contrasts to standard linear correlation results. One could put this in a historical context and determine which one tends to yield the most “true” proxy for asset correlations which could help financial institutions to optimally secure their future solvency.
References


40


## A Appendix

### A1

<table>
<thead>
<tr>
<th>Company</th>
<th>Ticker</th>
<th>Company</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lundin Petroleum</td>
<td>LUPE</td>
<td>Mekonomen</td>
<td>MEKO</td>
</tr>
<tr>
<td>Tethys Oil</td>
<td>TETY</td>
<td>Midsona</td>
<td>MSON-B</td>
</tr>
<tr>
<td>BE Group</td>
<td>BEGR</td>
<td>New Wave</td>
<td>NEWA-B</td>
</tr>
<tr>
<td>BillerudKorsnäs</td>
<td>BILL</td>
<td>Nobia</td>
<td>NOBI</td>
</tr>
<tr>
<td>Bergs Timber</td>
<td>BRG-B</td>
<td>Strax</td>
<td>STRAX</td>
</tr>
<tr>
<td>Holmen</td>
<td>HOLM-B</td>
<td>Swedish Match</td>
<td>SWMA</td>
</tr>
<tr>
<td>Lundin Mining Corp.</td>
<td>LUMI</td>
<td>Tretion</td>
<td>TRENT</td>
</tr>
<tr>
<td>Profilgruppen</td>
<td>PROF-B</td>
<td>VBG Group</td>
<td>VBG-B</td>
</tr>
<tr>
<td>Rottneros</td>
<td>RROS</td>
<td>Active Biotech</td>
<td>ACTI</td>
</tr>
<tr>
<td>SSAB</td>
<td>SSAB-B</td>
<td>AstraZeneca</td>
<td>AZN</td>
</tr>
<tr>
<td>Stora Enso</td>
<td>STE-R</td>
<td>BioInvent International</td>
<td>BINV</td>
</tr>
<tr>
<td>ABB Ltd.</td>
<td>ABB</td>
<td>BioGaia</td>
<td>BIOG-B</td>
</tr>
<tr>
<td>Addtech</td>
<td>ADDT-B</td>
<td>Biotage</td>
<td>BIOT</td>
</tr>
<tr>
<td>ÅF Pöyry</td>
<td>AF-B</td>
<td>Elekta</td>
<td>EKTA-B</td>
</tr>
<tr>
<td>Alfa Laval</td>
<td>ALFA</td>
<td>Elos Medtech</td>
<td>ELOS-B</td>
</tr>
<tr>
<td>Assa Abloy</td>
<td>ASSA-B</td>
<td>Feelgood Svenska</td>
<td>FEEL</td>
</tr>
<tr>
<td>Beijer Alma</td>
<td>BEIA-B</td>
<td>Getinge</td>
<td>GETI-B</td>
</tr>
<tr>
<td>Beijer Ref</td>
<td>BEIJ-B</td>
<td>Karo Pharma</td>
<td>KARO</td>
</tr>
<tr>
<td>Beijer Electronics Group</td>
<td>BELE</td>
<td>MedCap</td>
<td>MCAP</td>
</tr>
<tr>
<td>Bergman &amp; Beving</td>
<td>BERG-B</td>
<td>Medivir</td>
<td>MVIR-B</td>
</tr>
<tr>
<td>Bong</td>
<td>BONG</td>
<td>Oasmia Pharmaceutical</td>
<td>OASM</td>
</tr>
<tr>
<td>BTS Group</td>
<td>BTS-B</td>
<td>Otrivus</td>
<td>ORTI-B</td>
</tr>
<tr>
<td>Concordia Maritime B</td>
<td>CCOR-B</td>
<td>Orexo</td>
<td>ORX</td>
</tr>
<tr>
<td>Consilium</td>
<td>CONS-B</td>
<td>Probi</td>
<td>PROB</td>
</tr>
<tr>
<td>CTT Systems</td>
<td>CTT</td>
<td>RaySearch Laboratories</td>
<td>RAY-B</td>
</tr>
<tr>
<td>Duroc</td>
<td>DURC-B</td>
<td>SECTRA</td>
<td>SECT-B</td>
</tr>
<tr>
<td>Elanders</td>
<td>ELAN-B</td>
<td>Swedish Orphan Biovitrum</td>
<td>SOBI</td>
</tr>
<tr>
<td>Fagerhult</td>
<td>FAG</td>
<td>Vitrolife</td>
<td>VITR</td>
</tr>
<tr>
<td>Fingerprint Cards</td>
<td>FING-B</td>
<td>Axfood</td>
<td>AXFO</td>
</tr>
</tbody>
</table>

*cont. on next page*
<table>
<thead>
<tr>
<th>Company</th>
<th>Ticker</th>
<th>Company</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gunnebo</td>
<td>GUNN</td>
<td>Betsson</td>
<td>BETS-B</td>
</tr>
<tr>
<td>ICTA</td>
<td>ICTA</td>
<td>Bilia</td>
<td>BILI-A</td>
</tr>
<tr>
<td>Indutrade</td>
<td>INDT</td>
<td>Clas Ohlson</td>
<td>CLAS-B</td>
</tr>
<tr>
<td>Image Systems</td>
<td>IS</td>
<td>Eniro</td>
<td>ENRO</td>
</tr>
<tr>
<td>Largercrantz Group</td>
<td>LAGR-B</td>
<td>Hennes &amp; Mauritz</td>
<td>HM-B</td>
</tr>
<tr>
<td>Lindab International</td>
<td>LIAB</td>
<td>ICA Gruppen</td>
<td>ICA</td>
</tr>
<tr>
<td>Malmbergs Elektriska</td>
<td>MEAB-B</td>
<td>KappAhl</td>
<td>KAHL</td>
</tr>
<tr>
<td>Mycronic</td>
<td>MYCR</td>
<td>Kindred Group</td>
<td>KIND-SDB</td>
</tr>
<tr>
<td>NCC</td>
<td>NCC-B</td>
<td>Moment Group</td>
<td>MOMENT</td>
</tr>
<tr>
<td>Nibe Industrier</td>
<td>NIBE-B</td>
<td>Modern Times Group</td>
<td>MTG-B</td>
</tr>
<tr>
<td>Nolato</td>
<td>NOLA-B</td>
<td>RNB Retail and Brands</td>
<td>RNBS</td>
</tr>
<tr>
<td>NOTE</td>
<td>NOTE</td>
<td>SAS</td>
<td>SAS</td>
</tr>
<tr>
<td>OEM International</td>
<td>OEM-B</td>
<td>SkiStar</td>
<td>SKIS-B</td>
</tr>
<tr>
<td>Opus Group</td>
<td>OPUS</td>
<td>Swedol</td>
<td>SWOL-B</td>
</tr>
<tr>
<td>Peab</td>
<td>PEAB-B</td>
<td>TradeDoubler</td>
<td>TRAD</td>
</tr>
<tr>
<td>Poolia</td>
<td>POOL-B</td>
<td>Venue Retail Group</td>
<td>VRG-B</td>
</tr>
<tr>
<td>Precise Biometrics</td>
<td>PREC</td>
<td>Tele2</td>
<td>TEL2-B</td>
</tr>
<tr>
<td>Pricer</td>
<td>PRIC-B</td>
<td>Telia</td>
<td>TELIA</td>
</tr>
<tr>
<td>Rejlers</td>
<td>REJL-B</td>
<td>Millicom Int. Cellular</td>
<td>TIGO-SDB</td>
</tr>
<tr>
<td>SAAB</td>
<td>SAAB-B</td>
<td>Acando</td>
<td>ACAN-B</td>
</tr>
<tr>
<td>Sandvik</td>
<td>SAND</td>
<td>Addnode Group</td>
<td>ANOD-B</td>
</tr>
<tr>
<td>Securitas</td>
<td>SECU-B</td>
<td>Anoto Group</td>
<td>ANOT</td>
</tr>
<tr>
<td>Semcon</td>
<td>SEMC</td>
<td>DORO</td>
<td>DORO</td>
</tr>
<tr>
<td>Sensys Gatso Group</td>
<td>SENS</td>
<td>Empir Group</td>
<td>EMPIR-B</td>
</tr>
<tr>
<td>SinterCast</td>
<td>SINT</td>
<td>Enea</td>
<td>ENEA</td>
</tr>
<tr>
<td>Skanska</td>
<td>SKA-B</td>
<td>Ericsson</td>
<td>ERIC-B</td>
</tr>
<tr>
<td>SKF</td>
<td>SKF-B</td>
<td>Formpipe Software</td>
<td>FPIP</td>
</tr>
<tr>
<td>Svedbergs</td>
<td>SVED-B</td>
<td>Hexagon</td>
<td>HEXA-B</td>
</tr>
<tr>
<td>Studsvik</td>
<td>SVIK</td>
<td>HIQ International</td>
<td>HIQ</td>
</tr>
<tr>
<td>SWECO</td>
<td>SWEC-B</td>
<td>Invisio Communications</td>
<td>IVSO</td>
</tr>
<tr>
<td>Trelleborg</td>
<td>TREL-B</td>
<td>Knowit</td>
<td>KNOW</td>
</tr>
<tr>
<td>Volvo</td>
<td>VOLV-B</td>
<td>Micro Systemation</td>
<td>MSAB-B</td>
</tr>
<tr>
<td>Viking Supply Ships</td>
<td>VSSAB-B</td>
<td>MultiQ International</td>
<td>MULQ</td>
</tr>
</tbody>
</table>

cont. on next page
<table>
<thead>
<tr>
<th>Company</th>
<th>Ticker</th>
<th>Company</th>
<th>Ticker</th>
</tr>
</thead>
<tbody>
<tr>
<td>XANO Industri</td>
<td>XANO-B</td>
<td>Net Insight</td>
<td>NETI-B</td>
</tr>
<tr>
<td>AAK</td>
<td>AAK</td>
<td>NOVOTEK</td>
<td>NTEK-B</td>
</tr>
<tr>
<td>Björn Borg</td>
<td>BORG</td>
<td>Proact IT Group</td>
<td>PACT</td>
</tr>
<tr>
<td>Electrolux</td>
<td>ELUX-B</td>
<td>Prevas B</td>
<td>PREV-B</td>
</tr>
<tr>
<td>Husqvarna</td>
<td>HUSQ-B</td>
<td>Softronic B</td>
<td>SOF-B</td>
</tr>
<tr>
<td>KABE Group</td>
<td>KABE-B</td>
<td>Stockwik Förvaltning</td>
<td>STWK</td>
</tr>
<tr>
<td>Lammhults Design Group</td>
<td>LAMM-B</td>
<td>Tieto Oyj</td>
<td>TIETOS</td>
</tr>
</tbody>
</table>
All graphs constitute of Top: the min, max, and mean intra-correlations and Bottom: Weekly return (left $y$-axis) and min-max intra-correlation spread (right $y$-axis).

Figure 5: SX0001PI

Figure 6: SX1000PI

Figure 7: SX2000PI

Figure 8: SX3000PI
Figure 9: SX4000PI
Figure 10: SX5000PI
Figure 11: SX6000PI
Figure 12: SX9500PI