Risk aggregation and capital allocation using copulas

M Venter

20546564

Dissertation submitted in partial fulfilment of the requirements for the degree *Magister Scientiae* in Applied Mathematics at the Potchefstroom Campus of the North-West University

Supervisor: Prof DCJ de Jongh

May 2014

It all starts here ™

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
®

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT
POTCHEFSTROOMKAMPUS

PLEGTIGE VERKLARING

Plegtige verklaring deur student

Hiermee verklaar ek, _____,
die ondergetekende, dat die proefskrif/verhandeling/skripsie/artikel getiteld **(presies soos geregistreerde/ goedgekeurde titel)**,

_____

_____

_____

wat ek by die Noordwes-Universiteit, Potchefstroomkampus, ter voldoening / gedeeltelike voldoening aan die vereiste van die _____ -graad hiermee inhandig, my eie werk is, dit volgens die vereistes taalversorg is en dat dit nie reeds by 'n ander universiteit ingehandig is nie.

Ek verstaan en aanvaar dat die eksemplare, wat vir eksaminering ingehandig word, die eiendom van die universiteit is.

Handtekening van student_____Universiteitsnommer_____

Geteken te _____hierdie _____dag van _____20.....

Verklaar voor my hierdie _____dag van _____20.....

Kommissaris van Ede:_____

LET WEL: Indien 'n student se proefskrif/verhandeling/skripsie/artikel na die sperdatum vir indiening ingedien is, is die tydperk vir eksaminering beperk. Geen waarborg kan daarom gegee word dat (selfs sou die eksaminatorsverslae positief wees) die graad by die eerskomende gradeplegtigheid toegeken sal word nie. Dit mag ook beteken dat die student vir die volgende akademiese jaar moet herregistreer.

Plegtige verklaring deur studieleier/promotor/navorsingsdirekteur/dekaan

Hiermee verklaar ondergetekende:

		Ja	N.v.t.
1.	dat die genoemde student 'n goedgekeurde module(s) vir die betrokke kwalifikasie bygewoon het en die werk van die module uitgevoer het of werk gedoen het wat deur die Senaat goedgekeur is;		
2.	dat die student aan die minimum duur van studie voldoen soos voorgeskryf in die jaarboek;		-
3.	dat hy/sy derhalwe gemagtig word om sy/haar skripsie/verhandeling/proefskrif in te handig;		
4.	dat die registrasie/wysiging van titel goedgekeur is;		-
5.	dat die aanwysing/wysiging van eksaminatore goedgekeur is;		-
6.	dat die student se werk by TurnItIn ingedien is en 'n bevredigende verslag ontvang is; en		
7.	ander voorskrifte soos vervat in die Handleiding vir Nagraadse Studie nagekom is.		-

Handtekening Studieleier/Promotor: _____ Datum:_____

Handtekening Navorsingsdirekteur: _____ Datum:_____

Handtekening Admin Bestuurder: _____ Datum:_____

Abstract

Banking is a risk and return business; in order to obtain the desired returns, banks are required to take on risks. Following the demise of Lehman Brothers in September 2008, the Basel III Accord proposed considerable increases in capital charges for banks. Whilst this ensures greater economic stability, banks now face an increasing risk of becoming capital inefficient. Furthermore, capital analysts are not only required to estimate capital requirements for individual business lines, but also for the organization as a whole. Copulas are a popular technique to model joint multi-dimensional problems, as they can be applied as a mechanism that models relationships among multivariate distributions. Firstly, a review of the Basel Capital Accord will be provided. Secondly, well known risk measures as proposed under the Basel Accord will be investigated. The penultimate chapter is dedicated to the theory of copulas as well as other measures of dependence. The final chapter presents a practical illustration of how business line losses can be simulated by using the Gaussian, Cauchy, Student t and Clayton copulas in order to determine capital requirements using 95% VaR, 99% VaR, 95% ETL, 99% ETL and StressVaR. The resultant capital estimates will always be a function of the choice of copula, the choice of risk measure and the correlation inputs into the copula calibration algorithm. The choice of copula, the choice of risk measure and the conservativeness of correlation inputs will be determined by the organization's risk appetite.

Keywords: Copula, Gaussian, Cauchy, Student t, Clayton, dependence, correlation, capital, Basel.

Contents

1. Introduction 1

1.1. Research objectives 2

1.2. Structure of dissertation 2

2. The Basel Accord: A history of regulatory capital requirements 4

2.1. The Basel II Accord 4

2.1.1. Economic capital 4

2.1.2. Regulatory capital 7

2.2. The Basel II Accord and the financial crisis 12

2.2.1. Shortcomings of the Basel II Accord 12

2.3. The Basel III Accord: The response to the failures of Basel II 14

2.3.1. Minimum capital requirements and capital buffers 15

2.3.2. Enhanced coverage for counterparty credit risk 17

2.3.3. Leverage Ratio 18

2.3.4. Global liquidity standard 19

3. Risk based regulation and measures of risk 22

3.1. A definition of risk 22

3.2. Value at Risk 23

3.2.1. A review of Value at Risk 23

3.2.2. Risk aggregation and capital allocation 25

2.2.3. Shortcomings of VaR 26

3.3. Coherent risk measures 28

3.3.1. Worst Conditional Expectation (WCE) 29

3.3.2. Tail Conditional Expectation (TCE) 30

3.3.3. Conditional Value-at-Risk (CVaR) 30

3.3.4. α-Tail Mean (TM) and Expected Shortfall (ES) 31

3.3.5. The relationships between WCE, TCE, CVaR and ES 34

3.4. Stress Value at Risk 36

4. Copulas and dependence 36

4.1. Bivariate copulas 38

4.2. Sklar's theorem 41

4.3. Measures of dependence 44

4.3.1. Independence and dependence 44

4.3.2. Measuring the degree of association 46
4.4. Parametric classes of bivariate copulas 52
4.4.1. Elliptical copulas 53
4.4.2. Archimedean copulas 59
4.5. Multivariate copulas 62
4.5.1. Preliminary definitions 62
4.5.2. Subcopulas and copulas 63
4.5.3. Sklar's theorem 64
4.5.4. Product copula and Fréchet bounds 64
4.5.5. Parametric classes of multivariate copulas 65
5. Fitting copulas to multivariate data 67
5.1. Sample data and assumptions 67
5.2. Measuring dependence 68
5.3. Estimating business line volatilities 76
5.3.1. The GARCH(1,1) scheme 77
5.3.2. Estimating the parameters 80
5.4. Simulating business line losses using copulas 82
5.4.1. Multivariate copula calibration algorithms 82
5.4.2. Simulation of business line losses 84
6. Conclusion 91
Bibliography 93

List of figures

Figure 1: Chernobai et al. (2007): Illustration of the structure of the Basel II Capital Accord. 11
Figure 2: Capital requirements under Basel II and Basel III..... 15
Figure 3: Time lines for Basel III implementation..... 17
Figure 4: Updated Basel III Accord. 21
Figure 5: Bivariate Gaussian copula using different correlations. 54
Figure 6: Bivariate Student t copula with two degrees of freedom and different correlation inputs... 57
Figure 7: Bivariate Student t copula with five degrees of freedom and different correlation inputs. .. 58
Figure 8: Bivariate Student t copula with ten degrees of freedom and different correlation inputs.... 58
Figure 9: Bivariate Clayton copula with different values of alpha. 60
Figure 10: Bivariate Frank copula with different values of alpha. 61
Figure 11: Bivariate Gumbel copula with different values of alpha..... 62
Figure 12: Share price data from January 2000 to January 2012 for the 8 companies included in the analysis..... 68
Figure 13: Daily returns per share from January 2000 to January 2012..... 69
Figure 14: Distribution of daily returns 70
Figure 15: Comparison of AGL linear correlations over different time horizons..... 73
Figure 16: GARCH(1,1) annualized volatilities. 81
Figure 17: Capital estimates obtained by simulations using Gaussian copula, Cauchy copula, Student t copula and Clayton copula using the current linear correlation matrix as correlation input. 85
Figure 18: Comparison of capital estimates provided by different risk measures using the Gaussian, Cauchy, Clayton and Student t copulas..... 87
Figure 19: Comparison of capital estimates provided by StressVaR using the Gaussian, Cauchy, Clayton and Student t copulas. 87
Figure 20: Comparison of capital estimates obtained when using the current Kendall rank correlation matrix, current Spearman rank correlation matrix and the current linear correlation matrix..... 88
Figure 21: Comparison of capital estimates obtained using the minimum linear correlation matrix, current linear correlation matrix and maximum linear correlation matrix. 90

List of tables

Table 1: 12 year linear correlation matrix. *71*
Table 2: 12 year Spearman's Rank Correlation matrix. *72*
Table 3: 12 year Kendall's Rank Correlation matrix. *72*
Table 4: 12 year, maximum, minimum and current linear correlation matrices. *75*
Table 5: 12 year, maximum, minimum and current Spearman's Rank correlation matrices. *75*
Table 6: 12 year, maximum, minimum and current Kendall's Rank correlation matrices. *75*
Table 7: Optimized constrained values and long term variance obtained using the Maximum Likelihood Estimation (MLE) and GARCH(1,1) scheme. *81*
Table 8: Summary of the organization's value on 2 January 2012. ..... *84*

List of abbreviations

AGL	–	Anglo American PLC
AMA	–	Advanced measurement approach
AMS	–	Anglo American Platinum Corporation Ltd.
APN	–	Aspen Pharmacare Holdings
BCBS	–	Basel Committee on Banking Supervision
CEM	–	Current Exposure Method
CET1	–	Common Equity Tier I
CVA	–	Credit Value Adjustment
CVaR	–	Conditional Value at Risk
DSY	–	Discovery Holdings Ltd.
DV01	–	Dollar value of one basis point
EAD	–	Exposure At Default
EC	–	Economic Capital
EOCD	–	Organization for Economic Co-operation and Development
EPE	–	Expected Positive Exposure
ES	–	Expected Shortfall
ETL	–	Expected Tail Loss
EWMA	–	Exponentially Weighted Moving Average
FX	–	Forex
GARCH	–	Generalized AutoRegressive Conditional Heteroskedasticity
GI	–	Gross income
IMM	–	Internal Model Method
IRB	–	Internal Ratings-Based
L	–	Loss
LCR	–	Liquidity Coverage Ratio
LGD	–	Loss given a counterparty default
M	–	Maturity of exposure
MLE	–	Maximum Likelihood Estimation
MPC	–	Mr Price Group Ltd.
MPL	–	Maximum Probable Loss
MTN	–	MTN Group Ltd.
NSFR	–	Net Stable Funding Ratio

OR	–	Operational Risk
OTC	–	Over The Counter
PD	–	Probability of a counterparty defaulting
PPC	–	Pretoria Portland Cement
RC	–	Risk Capital
SBK	–	Standard Bank Group Ltd.
SIBs	–	Systemic Important Banks
SIFIs	–	Systemic Important Financial Institutions
SM	–	Standardized Method
StressVaR	–	Stress Value at Risk
TCE	–	Tail Conditional Expectation
TM	–	Tail Main
TVaR	–	Tail Value at Risk
USD	–	United States Dollar
VaR	–	Value at Risk
WCE	–	Worst Conditional Expectation
YTM	–	Yield to maturity
IID	–	Independently and identically distributed

1. Introduction

In 2007 Nassim Nicholas Taleb wrote a book called The Black Swan, where he states that: "*outlier*" events happen unexpectedly; they have an extreme impact and they cannot be predicted prior to occurring. This dilemma raises the following logical questions: (1) What causes Black Swan events? (2) Can risk measures be put in place, in order to mitigate the effect of a Black Swan? (3) Will economic capital provision be adequate in the event of a Black Swan? (4) How should policy makers address events of this magnitude?

In 2009 Carolyn Kousky and Roger M. Cooke wrote an article, referring to the unholy trinity as fat tails, tail dependence and auto correlation. These phenomena have led to question the validity of traditional risk management techniques, such as the normal distribution, linear correlation as well as Value at Risk.

Capital efficiency are two words that have greatly impacted the world of banking, following the demise of Lehman Brothers in September 2008. Capital adequacy, liquidity management as well as systematic risk have been emphasized in the lead-up to the implementation of Basel III and the resulting change in the regulatory and economic environment. Banks are now being forced to strategically review their business, or risk facing a decline in return on regulatory capital. New risk measures, such as stress VaR, have caused many financial institutions to become capital inefficient. Taleb (2001, p. 12) states: "*It does not matter how frequently something succeeds if failure is too costly to bear.*" Regulators have followed suit as first of all, new regulations have forced banks to stop activities that are no longer viable within the new capital regime. Business lines that have not produced sustainable returns on a consistent basis are being put under immense pressure and might eventually be forced to close down. Regulators have forced banks to identify high risk activities. Banks are also forced to have the capability to quantify the impact of events that could cause them to go bust.

Secondly, the new regulations have not only impacted existing activities, but it will also have an impact on the allocation of funds to new ones. Banks must not only identify key risk drivers that could have an impact on new businesses, but the degree of correlation between new business lines and existing ones must also be considered. Banks also have to be concerned with the aggregate effect that might occur over multiple business lines due to the occurrence of simultaneous extreme events. There thus exists a need to evaluate the impact of an extreme event on individual business lines as well as an entire organization. This is a primary task in establishing the degree of

diversification benefit that exists due to increasing granularity. As Hull (2007, p. 1) questions: "*When the rest of the business is experiencing difficulties, will the new venture also provide poor returns—or will it have the effect of dampening the ups and downs in the rest of the business?*"

It should however always be kept in mind that banking is first and foremost a risk and return business. In other words, in order to obtain the desired returns, banks will be required to take on risks. Risk management is thus a key function within a bank. This function is not only responsible for understanding the portfolio of all current risks that are being faced by the bank, but also all future risks that fit into the risk appetite that has been set by management.

1.1. Research objectives

As its first goal, this dissertation sets out to familiarize the reader with the pitfalls of traditional risk management techniques.

Secondly, on criticizing any methodology one should be ready to provide alternative solutions. The next goal is thus to obtain a thorough understanding of the mathematical concepts when considering copulas and to then motivate how traditional risk management techniques can be enhanced by using the copula approach.

The final aim is to then illustrate how copulas can be applied to data. Various copulas will be fitted to multivariate data in order to illustrate the functional relationship encoded within a dependence structure of the marginal distributions of several random variables.

1.2. Structure of dissertation

This dissertation starts off by considering the history of regulatory capital requirements under the Basel II Accord. Here a clear distinction will be made between regulatory capital and economic capital. This will be followed by an investigation into the failures of the Basel II Accord and its consequent role in the Financial Crises of 2008. Finally, this chapter will discuss the Basel III Accord and its response to the failures of the Basel II Accord.

Chapter 3 provides a thorough definition of risk and investigates some of the advantages and disadvantages of the best known risk measures, namely Value at Risk, coherent risk measures and Stress Value at Risk. The relationship between these risk measures will also be studied.

Chapter 4 is dedicated to the theory of copulas. This chapter provides some preliminary definitions and theorems in order to assist in defining bivariate copulas and perhaps the most important theorem in this chapter, known as Sklar's theorem. After introducing copulas, various measures of dependence will be discussed. Parametric classes of bivariate copulas will be studied next, as well as the simulation algorithms for each copula. Finally, all the proceeding theory will be extended into the multivariate case.

Having now introduced the fundamentals of the theory of copulas, chapter 5 explains how copulas can be fitted to data in order to estimate capital requirements within an organization. Here the GARCH(1,1) scheme will be used to estimate business line volatilities, in order to simulate business line losses using the Gaussian, Cauchy, Student t and Clayton copulas. These losses will then be used in determining capital requirements using 95% VaR, 99% VaR, 95% ETL, 99% ETL and StressVaR. Finally, a comparison of the capital requirements will be provided under the various copulas and risk measures.

2. The Basel Accord: A history of regulatory capital requirements

The Basel system originated from the Herstatt bank failure in 1974 (Dowd, Hutchinson & Ashby 2011). The Herstatt failure highlighted that central banks and bank managers required a greater sense of cooperation. Although Basel originally focused on creating a set of guidelines for bank closures, Basel became more concerned with the capital ratios within major banks in the 1980s. The Basel Accord was established to ensure stability within the banking system.

The Basel I Accord was published in 1988 and had to be implemented by 1992. Basel I mainly focused on weighting all risk assets on a bank's balance sheet, in order to calculate a bank's "Risk-weighted assets". Basel I stipulated a bank's minimum capital prerequisites in terms of core capital and supplementary capital (Tier I and Tier II capital, both equal to 4%).

Several revisions were published in recent years; this section will provide an outline of minimum capital requirements under the Basel II Accord, its shortcomings as well as the new definition of capital under the Basel III Accord.

In section 2.1 a clear distinction between economic and regulatory capital will be made as under the Basel II Accord. Section 2.2 investigates the role of the Basel II Accord in the Financial Crises as well as some of its shortcomings. Finally, in section 2.3 the Basel III Accord's response to these shortcomings in the Basel II Accord will be studied as well as Basel III's main focuses, namely: minimum capital requirements and capital buffers, enhanced coverage for counterparty credit risk, leverage ratio and global liquidity standard.

2.1. The Basel II Accord

Regulators' main goal when imposing a capital charge within the banking industry, is to ensure that banks will have a sufficient buffer against losses arising from both expected and unexpected losses. This section aims to provide a distinction between economic capital and regulatory capital.

2.1.1. Economic capital

The main role of economic capital is to absorb the risk faced by an institution due to market, credit, operational as well as business risks. In other words, economic capital can be seen as an estimate of the level of capital required by an organization to operate at a desired target solvency level. It is the amount of capital to be kept save and be immediately cashable, should the need arise to cover for losses.

Economic capital originated from the notion of margins used on futures exchanges. Brokers were expected to post a guarantee deposit, called a margin, at inception of a long/short position. Brokers were also required to replenish it whenever this margin fell short of a lower bound, referred to as a margin call. In the 1990s, banks incorporated the same rule into their proprietary deals. This concept which was borrowed from market risk was applied to all sources of risk in financial institutions, including credit and operational risk (Hull 2007).

An institution will never set economic capital at a confidence level of 100%; since it would be too expensive. The confidence level would rather be set at less than 100%. The confidence level must be chosen in a way that would provide a high return on capital to shareholders, protection to debt holders and confidence to depositors.

Marrison (2002) shows that if A_t and D_t denote the market values (at time t) of the assets and liabilities of an organization, the economic capital (EC_t) can be expressed as follows:
The economic capital available at the start of a year is given by

$$A_0 = D_0 + EC_0.$$

If r_D is the rate of interest payable on all debt, then the total debt to be paid at year end equals

$$D_1 = (1 + r_D) \times D_0.$$

If r_A is the interest rate receivable on all assets and λ is the rate of depreciation, then the total asset value at year end equals

$$A_1 = (1 + r_A) \times (1 - \lambda) \times A_0.$$

The economic capital at year end equals

$$EC_1 = A_1 - D_1$$
$$= (1 + r_A) \times (1 - \lambda) \times A_0 - (1 + r_D) \times D_0.$$

However, when the value of the firm's assets is equal to the value of its debt, the firm will be on the verge of bankruptcy

$$(1 + r_A) \times (1 - \lambda) \times A_0 - (1 + r_D) \times D_0 = 0.$$

From the above, the highest value of debt that can be supported by the economic capital can be denoted by

$$D_0 = \frac{(1 + r_A) \times (1 - \lambda)}{(1 + r_D)} A_0.$$

By substituting D_0 into $A_0 = D_0 + EC_0$, the economic capital required at the start of a year equals

$$EC_0 = A_0 - D_0$$
$$= A_0 - \frac{(1 - r_A) \times (1 - \lambda)}{(1 + r_D)} A_0$$
$$= \left(1 - \frac{(1 + r_A)(1 - \lambda)}{(1 + r_D)}\right) \times A_0.$$

If it is assumed that an organization only faces credit risk exposure, represented by a spread (μ) over the interest rate payable on all debt, i.e. $(1 + r_A) = (1 + r_D) \times (1 + \mu)$, then

$$EC_0 = \left(1 - \frac{(1 + r_A)(1 - \lambda)}{(1 + r_D)}\right) \times A_0$$
$$= \left(1 - \frac{(1 + r_D)(1 + \mu) \times (1 - \lambda)}{(1 + r_D)}\right) \times A_0$$
$$= (\lambda - \mu + \mu\lambda) \times A_0$$
$$\approx (\lambda_p - \mu) \times A_0$$
$$= Unexpected\ Loss - Expected\ Loss.$$

Usually, the sum of the stand-alone economic capital across all business lines would be higher than the economic capital required for a business as a whole, due to the benefits of diversification. Capital allocation methodologies that formed part of the Basel II Accord were divided into three main categories (Aziz & Rosen 2004), namely:

Stand-alone capital contribution

In this Bottom-Up approach, each business line was assigned the amount of capital that it would consume on a stand-alone basis. A disadvantage of this methodology is that it does not reflect any benefits of diversification (as mentioned above).

Incremental capital contribution (or discrete marginal capital contribution)

The total economic capital required for a single business line equals the economic capital requirement for the entire organization minus the economic capital requirement for the entire organization without this single business line. This method provides a good indication of the level of diversification benefit that each business line adds to the organization.

A disadvantage of this method is that it does not yield additive risk decomposition.

Marginal capital contribution (or diversified capital contribution)

This method portrays the measure of additivity that exists between the risk contributions of diverse business lines. In other words, this Top-Down approach allocates economic capital to a single business line, when viewed as part of a multi-business organization. Marginal contributions specifically allocate the diversification benefit among the various business lines. Under this approach, the total amount of economic capital that is allocated to an entire organization will equal the sum of the diversified economic capital for individual business lines.

Several alternative methods from game theory have been suggested for additive risk contributions (see (Denault 2001) and (Koyluoglu & Stoker 2002)). However, most of these methods have not yet been applied in practice.

Furthermore, economic capital can be estimated using a Top-Down approach or a Bottom-Up approach. The Bottom-Up approach compared to the Top-Down approach offers greater transparency when separating credit risk, market risk and operational risk.

2.1.2. Regulatory capital

Regulatory capital refers to the minimum capital requirements which banks are required to hold based on regulations established by the banking supervisory authorities. The Basel Committee on Banking Supervision (BCBS) plays an important role in creating a financial risk regulation network. Through Basel II, the BCBS attempted to create a capital requirement framework that would protect the banking industry from over exposing itself during its lending and investment practices.

Where the Basel I Accord only officially targeted minimal capital standards designed to protect the banking industry against credit risk, the Basel II Accord was aimed at credit, market and operational risk. After having undergone numerous amendments since 2001, the finalized Accord was presented in June 2006. The Basel II Accord used a three pillar approach, namely (Chernobai, Rachev & Fabozzi 2007):

- ***Pillar 1:*** Minimum risk-based capital requirements.
- ***Pillar 2:*** Supervisory review of an institution's capital adequacy and internal assessment process.
- ***Pillar 3:*** Market discipline through public disclosure of various financial and risk indicators.

The first pillar in the Basel II Accord deals with the minimum risk-based capital requirements calculated for the three main components of risk faced by a bank. Under the Basel II Accord, different approaches for estimating capital had to be followed for different components of risk.

Minimum risk based capital requirements for credit risk

Under the Basel II Accord, credit risk capital could be calculated using three different approaches, namely:

1. *Standardized approach*

This approach was first prescribed by the Basel I Accord, under which exposures were grouped into separate risk categories, each category with a fixed risk weighting. Under Basel II, however, loans to sovereigns, loans to corporates and loans to banks had risk weightings determined by external ratings.

2. *Foundation internal ratings based (IRB) approach*

This approach allowed lenders to use their own internal models in determining the regulatory capital requirement. This approach required lenders to estimate the probability of a counterparty defaulting (PD). Regulators provided set values for the loss given a counterparty default (LGD), exposure at default (EAD) as well as the maturity of exposure (M). When incorporated into the lender's appropriate risk weight function, a risk weighting for each exposure, or type of exposure could be provided.

3. *Advanced IRB approach*

Under this approach, lenders that were capable of the most advanced risk management and risk modelling techniques could themselves estimate PD, LGD, EAD and M. As the Basel II Accord promoted an improved risk management culture, lenders received a greater capital release under this approach than under the standardized approach.

Minimum risk based capital requirements for market risk

Under Basel II banks were required to develop a strategy that suited its market risk appetite. The standardized approach for calculating market risk capital varied per asset class (Maher & Khalil 2009).

1. *Interest rate and equity positions*

Capital for these instruments were calculated using two separate charges, namely a general market risk charge and a specific market risk charge. Firstly, the general market risk capital requirement was designed to offset losses that occurred due to movements in these underlying risk factors. Secondly, the specific risk capital requirement aimed at mitigating concentration risk with regards to an individual underlying risk factor.

2. *Foreign exchange positions*

Firstly, all FX exposures had to be expressed within a single currency (most commonly in USD). Secondly, banks were required to calculate capital for its net open positions when all currencies were taken into account.

3. *Commodity positions*

Capital charges for all commodity positions had to include three sources of risk, namely directional risk, interest rate risk and basis risk. Directional risk referred to the delta one exposure due to changes in spot prices. Interest rate risk aimed to capture the exposure due to movements in forward prices, as well as maturity mismatches. Basis risk was intended to capture the risk due to the association between two related commodities.

The preferred approach for estimating market risk capital under Basel II was Value at Risk. Banks however had freedom to decide on the exact nature of their models as long as the following minimum standards were adhered to:

a) VaR had to be reported on a daily basis.
b) The 99th percentile had to be used as the confidence interval.
c) Price stresses corresponding to 10-day movements had to be used.
d) Historical VaR had to use observation periods of at least one year.
e) Banks had to update their historical data sets at least once every three months.

Minimum risk based capital requirements for operational risk

Basel II recommended three methods to determine operational risk regulatory capital. Each approach required an underlying risk measure and management system, with increasing complexity and more refined capital calculations as one moved from the most basic to the most advanced approach.

1. Basic indicator approach

Under the basic indicator approach, operational risk capital is determined at $\alpha = 15\%$ of the annual gross income over the previous three years

$$RC_{BI}^{t}(OR) = \frac{1}{Z_t}\sum_{j=1}^{3} \alpha \, max\left(GI^{t-j}, 0\right)$$

where GI^{t-j} is the gross income for the year $t-j$, α is the fixed percentage of positive GI and Z_t is the number of the previous three years for which GI is positive.

2. Standardized approach

Under the standardized approach the Basel II Accord divides all activities into eight separate business lines, namely:

a) Corporate finance
b) Trading and sales
c) Retail Banking
d) Commercial banking
e) Payment and settlement
f) Agency services
g) Asset management
h) Retail brokerage

The average income over the last three years for each business line was multiplied by the "beta factor" for that business line and then these results were added. The operational risk capital under this approach in year t was given by

$$RC_{s}^{t}(OR) = \frac{1}{3}\sum_{i=1}^{3} \max\left(\sum_{j=1}^{8} \beta_j GI_j^{t-i}, 0\right)$$

where the factors β_j were between 12% and 18% depending on the risk activity.

The Basel Committee furthermore specified the following conditions when using the standardized approach:

a) The bank had to have an operational risk management function that was responsible for identifying, assessing, monitoring and controlling operational risk.
b) The bank had to keep track of relevant losses by business lines and create incentives for the improvement of operational risk.
c) There had to be regular reporting of operational risk losses throughout the bank.
d) The bank's operational risk management system had to be well documented.

e) The bank's operational risk management processes and assessment system had to be subject to regular independent reviews by internal auditors, external auditors or supervisors.

3. Advanced measurement approach (AMA)

Under the advanced measurement approach, the bank internally estimated the operational risk regulatory capital that was required, by means of quantitative and qualitative criteria, based on internal risk variables and profiles. This was the only risk sensitive approach for operational risk that was allowed and described in Basel II. The yearly operational risk exposure had to be set at a confidence level of 99.9%.

The Basel Committee also specified conditions for using the AMA approach:

a) The bank had to satisfy additional requirements.
b) The bank had to be able to specify additional requirements based on an analysis of relevant internal and external data and scenario analysis.
c) Systems had to be capable of allocating economic capital for operational risk across business lines in a way that created incentives for the business to improve operational risk management.

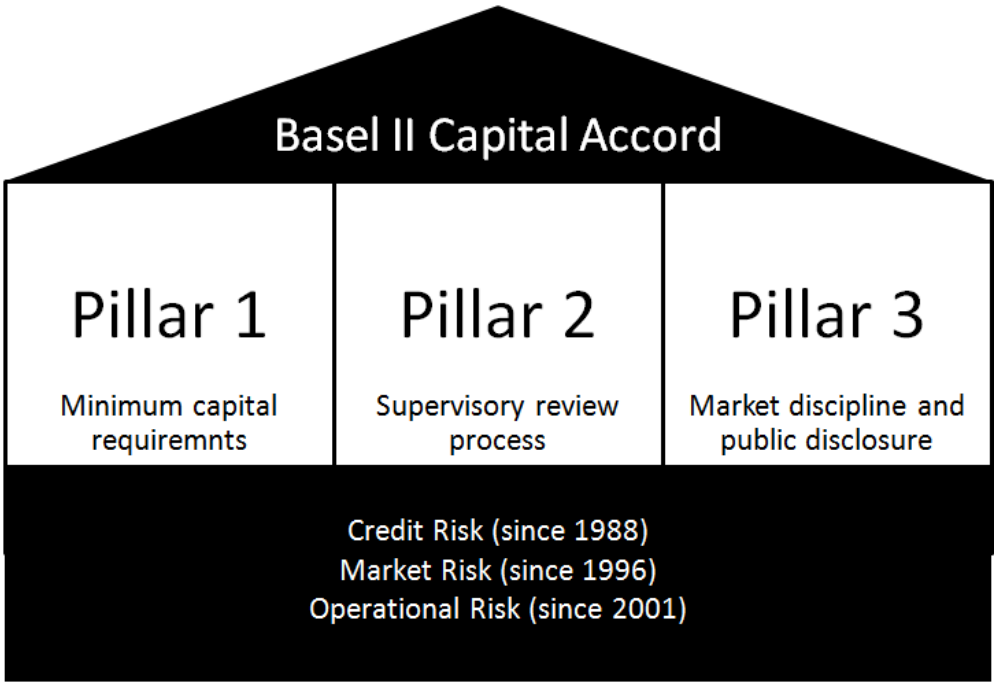

***Figure 1**: Chernobai et al. (2007): Illustration of the structure of the Basel II Capital Accord.*

Decomposition of minimum risk-based capital requirements

Under the Basel II Accord, banks were required to hold capital above the minimum required amount. According to Chernobai, Rachev and Fabozzi (2007) a definition of capital consisted of three types of capital, namely:

1. *Tier I capital*

a) Common stock (paid-up share capital)

b) Disclosed reserves

2. *Tier II capital (limited to a maximum of 100% of the total of Tier I capital)*

a) Undisclosed reserves

b) Asset revaluation reserves

c) General provisions

d) Hybrid capital instruments (debt/equaty)

e) Long-term subordinated debt

3. *Tier III capital (only eligible for market risk capitalization purposes)*

a) Short-term subordinated debt

2.2. The Basel II Accord and the financial crisis

Basel II's main goal was to prescribe banks with risk-based capital requirements that would protect the bank from going bust. At the dawn of the Credit Crises all international banks were Basel compliant, with reported capital ratios of approximately one or two times the required minimum amounts. According to Dowd et al. (2011) just five days before Lehman Brothers collapsed it possessed a Tier I capital ratio of 11%, which was close to three times the prescribed minimum regulatory requirement.

2.2.1. Shortcomings of the Basel II Accord

Dowd, Hutchinson and Ashby (2011) suggest that the Basel system suffered from three fundamental weaknesses. Firstly, financial risk models possessed numerous weaknesses and treated finance as a pure physical science. Secondly, it encouraged regulatory arbitrage. Finally, the banking industry was more concerned with short term profits than maintaining sufficient levels of capital. This section will investigate other possible shortcomings of Basel II.

Basel II failed to distinguish between normal and stress periods

Since historical VaR only required the use of one year's data, many banks excluded crisis periods that did not form part of that year's data in their models in order to produce lower VaR numbers. Consequently, if the year in question only reflected stable market conditions, the VaR numbers would not provide an accurate representation of the true risks faced by the bank.

Banks thus had pro-cyclical estimates of capital. This meant that whilst the economy was booming, no adjustment was being made to the capital estimates. In other words, when the economy reached its peak and was at its most dangerous, capital estimates were at its lowest. From Basel's point of view this defeated its main purpose, which was to stabilize the economy.

Basel II promoted frequent calibration of risk parameters

Basel II required historical data to be updated at least once every three months in order to calibrate to the current market conditions. Wilmott (2006) warns that calibration hides risk that one should be aware of. In summary, through calibration banks were effectively ignoring the fact that volatilities could rise, relationships could break down and bid-offer spreads could widen. Again, this lead to deflated capital estimates.

From a risk modelling perspective, the more conservative approach would have been to view risks over longer periods, consider the historical downside scenarios and make worst-case assumptions.

Basel II endorsed the use of VaR as primary risk measure

VaR simply reflects the highest probable loss, where the phrase probable must be understood in terms of probability. Nonetheless, VaR does not provide any indication of the size of losses that might occur given that this probability is violated. Tail events like the 2008 Credit Crises could thus not be captured by only using VaR.

Additionally, historical VaR is only a backward-looking risk measure and therefore assumes that the current distributions are a good representation for future events. Risk management therefore did not include any forward-looking or stressed scenarios that would have established how bad things could get.

Finally, VaR provided a far less intuitive expression of risk when compared to traditional trading risk measures, such as: option 'greeks', dollar value of one basis point (DV01), yield to maturity (YTM),

Macaulay duration and convexity. VaR is a much more complicated concept to understand. See Whaley (2006) for an in-depth explanation of traditional trading risk measures.

Basel II sanctioned the use of arbitrary risk weightings for credit risk

Under Basel II, the standardized approach grouped credit risk exposures into separate risk categories, each category with a fixed risk rating. This uniformed approach to credit risk was based on some terrible assumptions. Firstly, debt from the Organization for Economic Co-operation and Development (EOCD) governments were all given the same risk weighting. This thus assumed that the Greek and German governments had the same risk of defaulting. Secondly, this approach also implied that all corporate debt had equivalent credit risk. Effectively, this encouraged banks to invest in junk rated assets, as they required the same level of capital requirements as AAA-rated assets. These anomalies resulted in banks taking on excessive credit risk, as well as a deterioration of lending standards (which were both undercapitalized).

Basel II fueled the systematic instability within the financial system

Wilmott (2006) warns that the banking industry is dangerously correlated. He emphasizes this point by claiming that banks not only use the same risk models but also do the same trades. Any inherent weaknesses within the Basel regulations will thus have been forced upon all banks.

In addition, when prices started falling, this uniform approach to risk management led all banks to sell their risky positions. This caused prices to fall even further, which creates a "*vicious spiral*" as securities were being dumped (Dowd, Hutchinson & Ashby 2011).

Basel II allowed excessive levels of leverage within the banking industry

Under Basel II, banks were permitted to leverage up to 10 times in equities and up to 50 times in AAA-rated bonds. According to Sornette and Woodhard (2010) some banks held core capital of which only 3% consisted of their own assets. Even an uncomplicated scenario analysis would have indicated that banks were severely at risk.

2.3. The Basel III Accord: The response to the failures of Basel II

The recent financial crises have confirmed several weaknesses within the global regulatory framework, as well as risk management practices within the banking industry. Regulators have responded by proposing numerous measures that will provide increasing solidity in financial markets and that will assist in mitigating negative effects on the global economic environment.

In December 2010 the BCBS issued the first amendment "*Basel III: A global regulatory framework for more resilient banks and banking systems*". This was followed in June 2011 by the second amendment "*Basel III: International framework for liquidity risk measurements, standards and monitoring*". This section aims to provide insight into the newly proposed Basel III Accord and its main focuses, namely: minimum capital requirements and capital buffers, enhanced coverage for counterparty credit risk, leverage ratio and global liquidity standard.

2.3.1. Minimum capital requirements and capital buffers

This new definition of capital attempts to remove the incoherencies that existed under the previous definition of minimum capital requirements under the Basel II Accord. This aims to improve not only the estimates for minimum capital requirements, but also the quality of capital held.

The Basel III Accord aims to achieve these goals by increasing both the amount and class of Tier I capital, simplifying and decreasing Tier II capital, purging Tier III capital and bringing in new limits for elements of capital. The new definition of capital included:

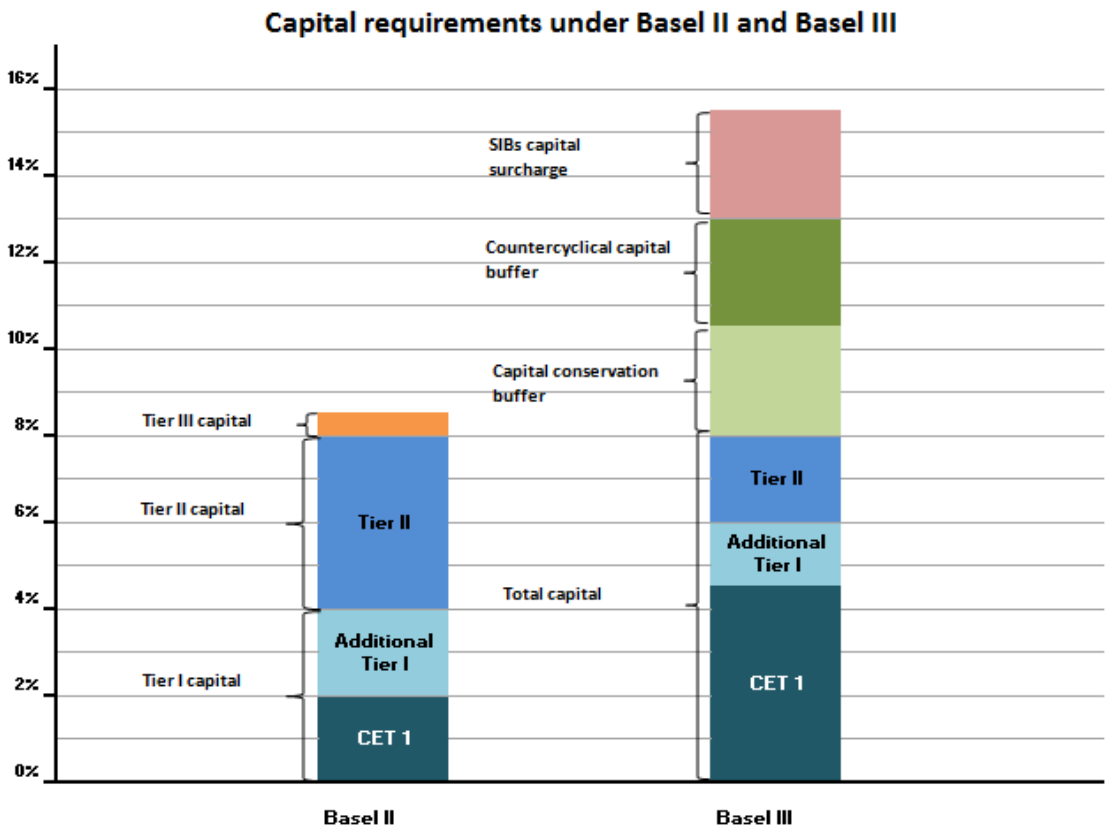

Figure 2: *Capital requirements under Basel II and Basel III.*

Total capital

Total capital consists out of Tier I and Tier II capital and will eventually be charged at 8%. In other words, *total capital* will equal the entire Basel II capital charge by 1 January 2015.

1. Tier I capital

Tier I capital should provide a bank with sufficient capital requirements to ensure solvency. This common equity Tier I capital (CET1) charge must primarily consist out of common equity and retained earnings. This capital charge will be supplemented by additional capital charges. This will result in Tier I capital being at 4.5% from 1 January 2013, 5.5% from 1 January 2014 and 6% from 1 January 2015.

2. Tier II capital

Tier II capital is aimed at guaranteeing that depositors and senior creditors get paid back in the case that a bank goes bust. However, the significance of Tier II capital lessens by decreasing the capital charge from 4% until 2012, to 3.5% in 2013, to 2.5% in 2014 and 2% from 2015 onwards.

Capital buffers

These new capital buffers are aimed at mitigating the effect of losses during future periods of financial as well as economic crises. The Basel III Accord proposes two new capital buffers namely, a capital conservation buffer and a countercyclical buffer. Furthermore, discussions are currently underway, surrounding additional capital surcharges. This surcharge involves systemic important financial institutions (SIFIs) or systemic important banks (SIBs).

1. Capital conservation buffers

Banks will be permitted to hold a 2.5% capital conservation buffer. This buffer serves as a forward-looking risk capital and aims to reduce the impact of future periods of financial turmoil. This capital conservation buffer has to be met with common equity only, increasing the total common equity prerequisite to 7%. Banks that fail to retain the capital conservation buffer risk facing restrictions on share buybacks, bonuses and even dividend payments. This capital buffer will be gradually introduced from 2016 onwards. In 2016 this capital charge will amount to 0.625% after which it will increase by the same amount every year, until reaching 2.5% in 2019.

2. Countercyclical buffers

The countercyclical buffer will be charged between 0% and 2.5%, depending on the national macroeconomic environment. This capital charge has to be exclusively met with common equity or other high quality capital (fully loss absorbing). This capital will be introduced in exactly the same manner as the capital conservation buffer (subject to the national macroeconomic conditions).

3. Additional surcharge

Additional capital surcharges for SIFIs and SIBs are still being debated. These charges will supposedly range between 1% and 2.5%, depending on the systemic importance that the institution presents. Furthermore, instruments that were part of the Basel II Accord and were issued before 12 September 2010, that do not comply with the Basel III Accord will be phased out over a ten-year period commencing in 2013.

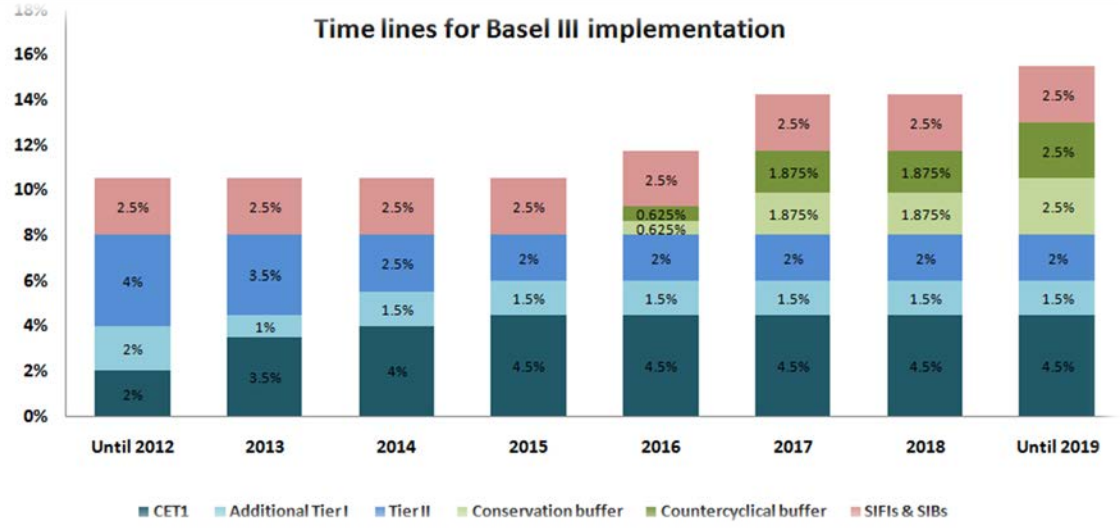

Figure 3: *Time lines for Basel III implementation.*

2.3.2. Enhanced coverage for counterparty credit risk

Under Basel III additional capital charges are added in order to mitigate the effect associated with possible losses due to a deterioration of counterparty credit quality. The updated credit risk framework provides incentives for clearing OTC derivative transactions through a central clearer. In addition, client trades as well as OTC derivative transactions that are not centrally cleared will be subject to a credit value adjustment (CVA).

Under the Basel III Accord, banks will be required to hold two forms of credit risk capital. Banks are firstly required to hold default risk capital. This capital charge is calculated using both stressed and calibrated parameters on a total portfolio level, in order to estimate the Expected Positive Exposure

(EPE) that a bank might face due to its activities. Secondly, banks are required to hold CVA capital. The CVA capital charge applies to non-centrally cleared transactions and is split up into general credit spread risk capital and specific credit spread risk capital.

The overall counterparty credit risk capital that Basel III will ultimately impose on a bank will be determined by the quality of a bank's credit risk modelling capabilities. The Basel III Accord classifies banks into three risk categories, namely:

Banks with approval for Internal Model Method and Specific-Risk VaR approaches

The default risk capital for these banks will be estimated by its EPE. The general CVA capital charge will be equal to the higher of its Internal Model Method (IMM) capital, using current market parameters or stressed parameters for exposure at default calculations. Specific risk CVA capital may be calculated using the in-house models. IMM banks are allowed to manage CVA together with pure market risk.

Banks with approval for the Internal Model Method approach

The default risk capital for these banks will also be estimated by its EPE. The general CVA capital charge will be calculated in the same manner as mentioned above. A standardized CVA capital charge will be applied for specific risk CVA capital requirements.

Other banks

These banks' default risk capital charge will be determined by summing across all counterparties using the Current Exposure Method (CEM) or the Standardized Method (SM). Non-IMM banks must estimate CVA general capital using statistical estimates of counterparty credit losses. CVA must also be treated as credit risk in these banks and will have to be managed separately from market risk. Regarding specific credit spread risk capital, a standardized CVA capital charge will be applied to such banks. Counterparty credit risk capital within these banks will generally tend to be much higher.

2.3.3. Leverage Ratio

In order to avoid the disproportionate levels of leverage, as previously seen prior to the financial crisis, the Basel III Accord established an additional non-risk based capital framework as an enhancement to the risk-based capital requirements previously mentioned.

This Leverage Ratio will be equal to the bank's total Tier I capital, expressed as a fraction of the bank's total exposure. Total exposure equals the sum of all assets and off-balance-sheet items not subtracted from the calculation of Tier I capital.

The Leverage Ratio is currently proposed at 3%. A parallel run will be introduced on 1 January 2013 that will continue until 1 January 2017. During this time regulators will track the Leverage Ratio and evaluate its performance in relation to the risk based requirements. Current proposals are to migrate to the Leverage Ratio to Pillar I treatment on 1 January 2018.

$$Leverage\ Ratio\ = \frac{Tier\ I\ capital}{Total\ exposure} \geq\ 3\%.$$

The final breakdown of total exposure and the credit risk adjustment to off-balance sheet items are still to be finalized.

2.3.4. Global liquidity standard

Finally, the Basel III Accord initiates a new liquidity standard by introducing two liquidity ratios, namely the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR). In short, the new liquidity standard aims to examine a bank's maturity mismatches, funding concentration and available unencumbered assets. Both proposals are yet to be finalized; this section presents the liquidity standard proposals as they stand in December 2012.

Liquidity Coverage Ratio

The Liquidity Coverage Ratio is aimed at improving banks' short-term liquidity risk profile. The LCR necessitates banks to hold high quality liquid assets as well as reduce asset and liability mismatches in near dated tenors.

$$LCR = \frac{High\ quality\ liquid\ assets}{Total\ net\ liquidity\ outflows\ over\ a\ 30\ day\ time\ period} \geq 100\%.$$

1. High quality liquid assets

Here high quality liquid assets must consist of assets of high liquidity and credit quality in order to ensure that an institution sustains a sufficient liquidity buffer. High quality liquid assets can be of two types:

a) Level 1 assets - These assets must consist of cash, deposits held with central banks or transferable assets of extremely high credit and liquidity quality assets. Banks will be required to hold a minimum of 60% of an organization's liquid assets. The value of the liquid assets will be

set equal to market value, subject to haircuts, ranging between 0% and 20%. Due to the superior credit and liquidity quality of Level 1 assets, it will not be subject to any haircuts.

b) Level 2 assets - These are transferable assets of high credit and liquidity quality. Level 2 assets will be subject to a minimum haircut of 15%.

2. Total net liquidity outflow over a 30-day period

The net total liquidity outflow over a 30-day period represents an organization-specific outflow as well as systematic shocks. This measure aims to protect banks from imbalances that might exist due to mismatches arising from liquidity inflows and outflows under extreme conditions over short periods of time. The net total liquidity outflow over a 30-day period of stress equals liquidity inflows minus liquidity outflows, where liquidity inflows are capped at 75% of the liquidity outflows. From 2013, banks will be required to report their LCR on a monthly basis.

$$Net\ liquidity\ outflow = Liquidity_{outflow} - \min\left(Liquidity_{inflow}; 75\%\ of\ Liquidity_{outflow}\right).$$

Net Stable Funding Ratio

The Net Stable Funding Ratio is intended to improve long-term stability through forcing banks to fund its business through more constant sources of funding. The Basel III Accord will require banks to maintain a sound funding structure over a calendar year subject to a firm-specific stress set-up. Banks will be required to report on its NSFR on a quarterly basis.

$$NSFR = \frac{Available\ stable\ funding}{Required\ stable\ funding} \geq 100\%.$$

1. Available stable funding

Banks must obtain stable funding within the 3 months, 3 – 6 months, 6 – 9 months, 9 – 12 months and after 12 months maturity buckets. Stable funding consists of:

a) Own funds
b) Retail deposits
c) Other deposits (fulfilling certain conditions)
d) Funding obtained from customers
e) Funding through secured lending
f) Liabilities resulting from covered bonds or other issued securities
g) Other liabilities

2. Required stable funding

The Basel III Accord also requires banks to determine its need for stable funding. These items must also be reported in the five maturity buckets mentioned above.

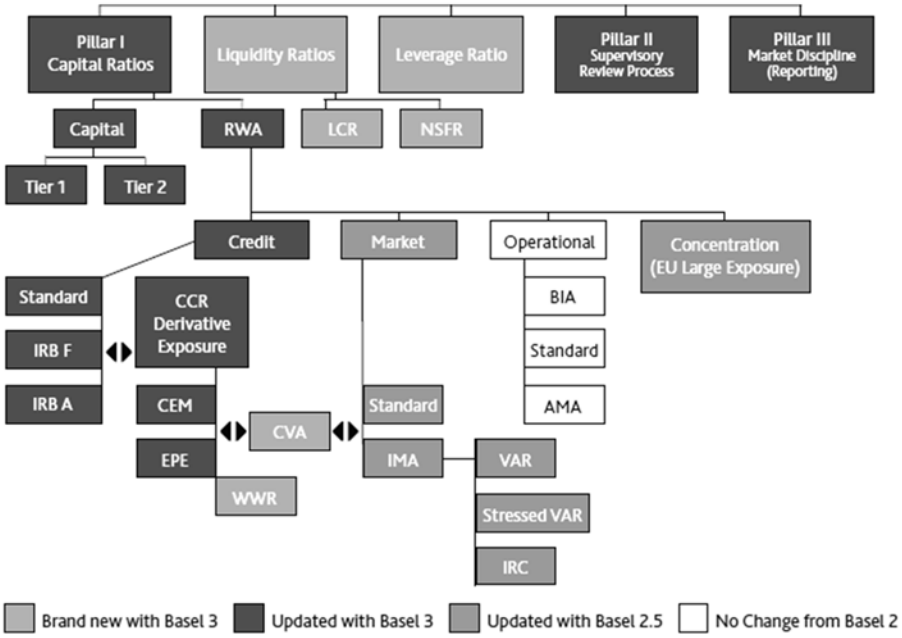

Figure 4: *Updated Basel III Accord.*

3. Risk based regulation and measures of risk

A fundamental attribute of the Basel Accord is the principle of risk based regulation. This principle aims to facilitate a capital adequacy framework where banks make use of financial modelling in order to determine its capital requirements. This principle has received much criticism, and so have the measures that it uses. This section evaluates conventional risk measures as proposed by the Basel Accord when estimating capital requirements.

This chapter starts off by looking at a thorough definition of risk; this will be followed by a review of Value at Risk, risk aggregation and capital allocation as well as a review of some of the shortcomings of Value at Risk. In section 3.3 coherent risk measures will be studied as well as the relationship between these risk measures. The coherent risk measures that will be studied include: Worst Conditional Expectation, Tail Conditional Expectation, Conditional Value at Risk, Tail mean and Expected Shortfall. In the final section, section 3.4, Stress Value at Risk will be studied.

3.1. A definition of risk

The Oxford English dictionary describes risk as "*a hazard, a chance of bad consequences, loss or exposure to mischance*". Embrechts et al. (2005, p. 1) describe risk as "*any event or action that may adversely affect an organization's ability to achieve its objectives and execute its strategies*" or "*the quantifiable likelihood of loss or less-than-expected returns*".

In any context risk can be related to uncertainty. Wilmott (2000) makes a distinction between randomness and uncertainty. Randomness not only assumes the existence of a set of different events, but also that each event has a certain probability of taking place. Whilst uncertainty similarly acknowledges the existence of a set of different events it does not make any assumptions regarding the probability of their occurrences.

In order to understand the nature of risk, certain risk concepts exist that form the foundation when assessing risk within an organization. These concepts include: exposure, probability, severity, volatility, time horizon and correlation. Exposure offers an estimate of what a company could potentially lose, whilst probability indicates how likely these losses are. Severity specifies the magnitude of possible losses and volatility unveils how uncertain the future might be. Since the duration of exposure to risks also concerns us, the concept of time horizon plays an essential role in understanding risks. In order to determine how much capital should be set aside to cover for

unexpected losses, one also needs to understand how risks in the business are related to each other and this is known as correlation.

3.2. Value at Risk

Financial institutions require a measure of risk that exemplifies the amount of money at stake in their investments; one such measure is Value at Risk (VaR). Like all other risk measures, VaR aims to quantify the riskiness of a portfolio as an executive summary. This concept was first introduced to risk management in the 1990s and was then known as Maximum Probable Loss (MPL). Prior to the 1990s, risk management mainly focused on the concept of asset and liability management (Hull 2007).

3.2.1. A review of Value at Risk

Wilmott (2006, p. 460) defines VaR as "*an estimate, with a given degree of confidence, of how much one can lose from one's portfolio over a given time horizon*". Thus, the VaR calculation is dependent on two parameters, namely the time horizon and the confidence level. Assigning the best possible values to these parameters is a non-trivial task and requires some reflection.

The time horizon will typically depend on contractual and legal constraints, liquidity considerations as well as the type of risk that is being measured (Embrechts, Frey & McNeil 2005). For instance, in operational risk the time horizon would equal to the time required to restore operations after a break in business continuity. In contrast the time horizon for market risk would equal the period related to orderly liquidation of a position. Embrechts et al. (2005) also explain that it might be optimal to use a shorter time horizon, since this leads to more historical data of risk factor changes.

In order to have a sufficient safety margin for capital adequacy purposes, a high confidence level is preferred (typically 95%, 99%, 99.9% and so on). Since quantiles play an important role in risk management, once a loss distribution has been computed, the choice of confidence level will be central in determining capital estimates. For example, under the Basel III Accord banks are required to use a time horizon of 10 days under a confidence level of 99% for Market risk.

Cherubini et al. (2011) explains that VaR mainly consists of the probability distribution of losses over a given period of time. VaR can be seen as the quantile of this measure,

$$q_\alpha = F_x^{-1}(\alpha) = inf\{x: \mathbb{P}(X \leq x) \geq \alpha\}$$

where X is a random variable representing a portfolio of exposures to risk and F_x^{-1} is the generalized inverse of $F^{-1}: (0,1) \rightarrow \mathbb{R}$. i.e. it is simply an alternative notation for the quantile function of F_x evaluated at α. Thus,

$$VaR_\alpha(X) = q_\alpha(-X) = F_{-X}^{-1}(\alpha)$$

where $VaR_\alpha(X)$ is the Value at Risk of an exposure X at confidence level α, and α close to 1. Furthermore,

$$\begin{aligned} VaR_\alpha(X) = q_\alpha(-X) &= inf\ \{x: \mathbb{P}(-X \leq x) \geq \alpha\} \\ &= inf\ \{x: \mathbb{P}(x + X \geq 0) \geq \alpha\} \\ &= inf\{x: \mathbb{P}(x + X < 0) \leq 1 - \alpha\}. \end{aligned}$$

In other words, $VaR_\alpha(X)$ is the smallest amount of money which, if added to X, keeps the probability of a negative outcome below the level $1 - \alpha$. Furthermore, if F_x is invertible

$$\begin{aligned} F_x\big(-VaR_\alpha(X)\big) &= \mathbb{P}\big(X \leq -VaR_\alpha(X)\big) \\ &= \mathbb{P}\big(-X \geq VaR_\alpha(X)\big) \\ &= \mathbb{P}\left(-X \geq F_{-X}^{-1}(\alpha)\right) \\ &= \mathbb{P}(F_{-X}(-X) \geq \alpha) \\ &= 1 - \alpha \end{aligned}$$

such that

$$VaR_\alpha(X) = -F_X^{-1}(1 - \alpha).$$

A further property of VaR is that it is homogeneous of degree one. This property will be central when using VaR estimates in determining capital allocation as illustrated in the Euler theorem. However, before this property can be proved, one first has to define when a point $x_0 \in \mathbb{R}$ is the α-quantile.

Lemma 3.1 (Embrechts, Frey & McNeil 2005):

A point $x_0 \in \mathbb{R}$ is the α-quantile of some distribution function F if and only if

$$F(x_0) \geq \alpha$$
$$F(x) < \alpha$$

for all $x < x_0$.

Following Lemma 3.1 it can now be proved that VaR is homogeneous of degree one as provided by Cherubini et al. (2011).

Theorem 3.2

VaR is homogeneous of degree one such that

$$VaR_\alpha(\lambda X) = \lambda VaR_\alpha(X).$$

Proof:

If $\lambda > 0$,

$$F_{\lambda X}^{-1} = inf\ \{t: F_{\lambda X}(t) \geq \alpha\}$$
$$= inf\ \{t: \mathbb{P}(\lambda X \leq t) \geq \alpha\}$$
$$= inf\left\{t: \mathbb{P}\left(X \leq \frac{t}{\lambda}\right) \geq \alpha\right\}$$
$$= inf\left\{t: F_X\left(\frac{t}{\lambda}\right) \geq \alpha\right\}$$
$$= \lambda\ inf\left\{\frac{t}{\lambda}: F_X\left(\frac{t}{\lambda}\right) \geq \alpha\right\}$$
$$= \lambda\ inf\ \{z: F_X(z) \geq \alpha\}\ = \lambda F_X^{-1}(\alpha)$$

with $VaR_\alpha(X) = F_{-X}^{-1}(\alpha)$.

3.2.2. Risk aggregation and capital allocation

Cherubini et al. (2007) state that risk management is an intrinsically multivariate concept and that there are numerous exposures to risk in the market as a result of the high level of interdependence in markets and risk factors.

In theory there exists a central limit theorem which allows us to accumulate all minor and independent shocks into a single variable called noise, and this would be normally distributed. Unfortunately it is not so simple in reality since markets are not normally distributed due to the fact that shocks are not independent. Thus, the amount of capital to be allocated will depend on the likelihood that losses will occur simultaneously. Association risk is the most important risk when considering capital allocation, i.e. the risk of simultaneous losses across different business lines. Diversification plays a very important role in this case; since it can decrease the amount of capital that should be assigned to each business line, as well as the organization as a whole. When determining capital allocation, three logical questions need to be answered (Cherubini et al. 2011):

1. How much capital should be devoted to the entire business?
2. How much capital should be devoted to each business line?

The order in which questions 1 and 2 are answered will determine whether a top-down or bottom-up approach will be followed. For instance, assume that there exists a financial institution that

consists of two business lines, business line A and business line B, with exposures X_i, $i = A, B$. The multivariate risk can then be expressed as

$$VaR_\alpha^f(X_A, X_B) = \{(x_A, x_B): \mathbb{P}(f(X_A + x_A, X_B + x_b) < 0) = 1 - \alpha\}$$

where $f: \mathbb{R}^2 \to \mathbb{R}$ is an aggregation function. If a risk measure is homogeneous of degree one, the Euler principle can be used to provide an answer of how much of the capital allocated to a set of risk sources is to be accounted for by each of them.

Theorem 3.3 (Euler Theorem on homogeneous functions):

Let f be a n-variate function. The function is homogeneous of degree κ, that is, for all $\lambda > 0$

$$f(\lambda u) = \lambda^\kappa f(u),$$

$$\Leftrightarrow \ \kappa f(u) = \sum_i^n \frac{\partial f(u)}{\partial u_i}$$

with $u = [u_1, u_2, \dots, u_n]$.

Since VaR is homogeneous of degree one (see theorem 3.2), this property can be applied to VaR:

If $\lambda > 0$, then

$$VaR_\alpha^f(\lambda X_1, \lambda X_2, \dots, \lambda X_n) = \{(y_1, y_2 \dots, y_n) : \mathbb{P}(f(\lambda X_1 + y_1, \lambda X_2 + y_2, \dots, \lambda X_n + y_n) < 0) = 1 - \alpha\}$$

$$= \left\{(y_1, \dots, y_n) : \mathbb{P}\left(\lambda f\left(X_1 + \frac{y_1}{\lambda}, X_2 + \frac{y_2}{\lambda}, \dots, X_n + \frac{y_n}{\lambda}\right) < 0\right) = 1 - \alpha\right\}$$

$$= \left\{(y_1, \dots, y_n) : \mathbb{P}\left(f\left(X_1 + \frac{y_1}{\lambda}, X_2 + \frac{y_2}{\lambda}, \dots, X_n + \frac{y_n}{\lambda}\right) < 0\right) = 1 - \alpha\right\}.$$

Thus,

$$VaR_\alpha^f(X) = \sum_{i=1}^m X_i \frac{\partial VaR_\alpha^f(X)}{\partial X_i},$$

in other words, the total VaR of a financial institution can be represented as a linear combination of all the business lines' VaR sensitivities.

2.2.3. Shortcomings of VaR

In his article, *Against Value–at–Risk: Nassim Taleb Replies to Philippe Jorion*, Taleb (1997) states: "*I maintain that the due-diligence VaR tool encourages untrained people to take misdirected risk with the shareholder's, and ultimately the taxpayer's, money*".

The previous sections provided a formal definition of VaR, properties of VaR as well as an evaluation of the role that VaR plays in risk aggregation and capital allocation. This section will provide insight into some of the shortcomings of this measure of risk.

The misinterpretation of the definition of VaR

A literal interpretation of the definition of VaR can be quite misleading. Acerbi et al. (2001, p. 4) state that a 95%, 7 day VaR in an organization is often expressed as "*the maximum potential loss that a portfolio can suffer in the* 5% *worst cases in* 7 *days*". They also point out that the correct version of the definition should rather be: "*VaR is the minimum potential loss that a portfolio can suffer in the* 5% *worst cases in* 7 *days*".

By definition, VaR at a confidence level α does not provide any insight regarding the severity of losses that might occur once the confidence level $1-\alpha$ has been breached (Embrechts, Frey & McNeil 2005).

Failure to use stress periods in historical VaR estimates

Prior to Basel III, VaR was calculated assuming normal market circumstances. This meant that extreme market conditions such as crashes were not considered, or were examined separately. Effectively, capital estimates only represented the risks expected during normal "day-to-day" operations of an institution. In other words, this ignored the fact that most financial time series data shows fatter tails and higher peaks. It can thus be concluded that under normal market conditions, VaR would have provided sufficient capital estimates. However, under extreme market conditions one would rather make use of measures such as stress testing[1] and crash metrics[2].

VaR neglects the effect of market liquidity

Historical VaR provides risk estimates based on historical market moves, or historical moves in the underlying risk factors. However, many financial institutions only calibrate to "mid" prices when considering historical price moves. Thus, VaR ignores the effect of bid-offer spreads that would apply when disposing of a long position and closing out a short position. A poor understanding of liquidity constraints has led to many famous financial disasters, most notably LTCM in 1998.

[1] Stress testing is a methodology for estimating a portfolio's performance during financial crises.
[2] CrashMetrics is a methodology for approximating the exposure of a portfolio to extreme market movements or crashes. For more information on this topic see Wilmott (2006).

Essentially VaR has to capture a wide range of factors, such as the complexity of financial instruments, dimensions of the portfolio and the assessment of the market. This can result in complicated computation and leads to approximations to ease the computation which ultimately leads to statistical errors in the estimation of VaR.

VaR is a non-subadditive measure of risk

A key strength of VaR lies in the fact that it can be applied to any financial instrument and that the risk associated with a portfolio of instruments can be expressed as a single number. This was one of the main reasons why the Basel II Accord chose VaR as the primary measure of risk based regulation.

Ironically, even though VaR is mainly used as an executive summary on a portfolio basis, VaR in itself has poor aggregation properties as was shown by Artzner et al. (1999) and Embrechts et al. (2005). This implies that the VaR of a portfolio is not made up of the sum of the sub-portfolios, thus, when adding a new sub-portfolio, the risk of the entire portfolio needs to be re-estimated.

3.3. Coherent risk measures

As mentioned in the previous sections, one of the main criticisms of VaR is that it is non-sub-additive. Thus, the notion of measures of coherent risk was introduced. Measures that form part of this group are: Expected Tail Loss (ETL), Conditional VaR (CVaR), Worst Conditional Expectation (WCE), Tail Conditional Expectation (TCE) and Tail Value-at-Risk (TVaR).

Artzner et al. (1999) present four axioms that must be satisfied by a risk measure in order to be classified as coherent. Let Ω be the finite set of states of nature and let ζ be the set of all real valued functions on Ω. In other words, ζ defines the set of all risks.

Definition 3.4 (Coherent risk measures)

Let X_1 and X_2 be two random variables. A risk measure (a mapping from ζ into $\mathbb{R}$) satisfying the following conditions is a coherent risk measure:

1. *Translation invariance:* for $X \in \zeta$ and all real numbers α, we have $\rho(X + \alpha \cdot r) = \rho(X) - \alpha$.
2. *Subadditivity:* for all X_1 and $X_2 \in \zeta$, $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.
3. *Positive homogeneity:* for all $\lambda \geq 0$ and all $X \in \zeta$ with $\rho(\lambda X) = \lambda\rho(X)$.
4. *Monotonicity:* for all $X \; and \; Y \in \zeta$ with $X \leq Y$, we have $\rho(Y) \leq \rho(X)$.

The first condition, translation invariance, implies that adding (subtracting) α from a current position, decreases (increases) the risk by α. The second condition indicates that the sum of the risk measures for two stand-alone portfolios is always bigger than or equal to the combined risk measure for the two merged portfolios. The third condition implies that when the size of the portfolio increases by an absolute factor λ, the risk measure associated with the portfolio will also increase by a factor λ. Finally, the fourth condition implies that a portfolio with lower returns than another portfolio (in every state of) should have a higher risk measure.

This section will provide definitions and properties of coherent risk measures as well as relationships between these risk measures as presented by Acerbi and Tasche (2002) unless otherwise cited.

For the remainder of this section, let X be a random variable on the probability space $(\Omega, \mathcal{A}, P)$ and let $\alpha \in (0,1)$. We will also make use of the indicator function

$$1_A(a) = 1_A = \begin{cases} 1, & a \in A \\ 0, & a \notin A. \end{cases}$$

Furthermore, let $x_{(\alpha)} = q_\alpha(X) = inf\{x \in \mathbb{R} : P[X \leq x] \geq \alpha\}$ be the lower α-quantile of X and let $x^{(\alpha)} = q^\alpha(X) = inf\{x \in \mathbb{R} : P[X \leq x] > \alpha\}$ be the upper α-quantile of X.

The positive part of a number x will be denoted by

$$x^+ = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

and the negative part of a number x will be denoted by

$$x^- = (-x)^+.$$

3.3.1. Worst Conditional Expectation (WCE)

The first coherent measure of risk that will be considered is Worst Conditional Expectation (WCE).

Definition 3.5 (Worst conditional expectation):

Assume $E[X^-] < \infty$. Then

$$WCE = WCE(X) = -inf\{E[X|A] : A \in \mathcal{A}, P[A] > \alpha\}$$

is the worst conditional expectation at level α of X.

Although WCE is classified as a coherent risk measure, it is not useful in practice since it could hide the fact that it does not only depend on the distribution of X but also on the structure of the underlying probability space. In order to see this, note that the value of WCE_α is finite under $E[X^-] < \infty$, since

$$\lim_{t\to\infty} P\left[X \le x^{(\alpha)} + t\right] = 1$$

implies that there exists some event,

$$A = \left\{X \le x^{(\alpha)} + t\right\}$$

where $P[A] > \alpha$ and $E[|X|1_A] < \infty$. Also, for any random variables X and Y on this probability space, WCE is subadditive:

$$WCE_\alpha(X+Y) \le WCE_\alpha(X) + WCE_\alpha(Y).$$

This measure was introduced since TCE in general does not define a sub-additive risk measure (Delbaen 1998).

3.3.2. Tail Conditional Expectation (TCE)

Unlike WCE, TCE is not only useful in a theoretical setting, but also proves to be useful in practical applications. Unfortunately TCE is not subadditive in general.

As for VaR, when referring to the quantile functions and not the proportion of the quantile functions, there also exists a choice for an upper and lower TCE.

Definition 3.6 (Tail conditional expectations):

Assume $E[X^-] < \infty$. Then

$$TCE_\alpha = TCE_\alpha(X) = -E[X|X \le x_{(\alpha)}]$$

is the lower tail conditional expectation at level α of X and

$$TCE^\alpha = TCE^\alpha(X) = -E[X|X \le x^{(\alpha)}]$$

is the upper tail conditional expectation at level α of X.

It is also obvious that $TCE_\alpha \ge TCE^\alpha$.

3.3.3. Conditional Value-at-Risk (CVaR)

Acerbi and Tasche (2002, p. 1490) state that CVaR can be "*used as a base for very efficient optimization procedures*".

Definition 3.7 (Conditional Value-at-Risk):

Assume $E[X^-] < \infty$. Then

$$CVaR = CVaR(X) = inf\left\{\frac{E[(X-s)^-]}{\alpha} - s : s \in \mathbb{R}\right\}$$

is the Conditional Value-at-Risk at level α of X.

3.3.4. α-Tail Mean (TM) and Expected Shortfall (ES)

An alternative measure of risk is Expected Shortfall (ES). According to Dowd et al. (2011), actuaries have been using this measure for many years. In contrast to VaR this measure indicates what to expect once the confidence level has been breached. In other words, ES measures what the expected loss could be in the $x\%$ worst cases in y days. Although ES can be classified as a better risk measure when compared to VaR, it is not as simplistic as VaR because it is slightly more difficult to understand and to back test (Wilmott 2006). Nonetheless ES allows us to "*look further into the tail*" (Embrechts, Frey & McNeil 2005) and $ES \geq VaR$.

Acerbi and Tasche (2002) choose to define the α-tail mean in two variants, namely the tail mean and Expected Shortfall, since the former is negative but appears to be better as defined in a statistical context and the latter is positive and represents potential loss best. Also, since the tail mean is independent on the distributions of the underlying random variables, it allows for a straightforward proof of super-additivity (negative sub-additivity). On the other hand, as will be seen, ES is coherent, continuous and monotonic in the confidence level α.

Definition 3.8 (a) (Tail mean)

Let $E[X^-] < \infty$, then

$$\bar{x}_{(\alpha)} = TM_\alpha = \alpha^{-1}\left(E\left[X\, 1_{\{X \leq x_{(\alpha)}\}}\right] + x_{(\alpha)}(\alpha - P[X \leq x_{(\alpha)}])\right)$$

is the α-tail mean at the level α of X.

Definition 3.8 (b) (Expected Shortfall)

Expected Shortfall at a confidence level α of X can be defined as

$$ES_\alpha = ES_\alpha(X) = -\bar{x}_{(\alpha)}.$$

Acerbi et al (2001) also expresses ES in terms of VaR as can be seen in definition 3.9

Definition 3.9 (Expected Shortfall and VaR):

For a loss L with $E(|L|) < \infty$, the expected shortfall at confidence level α is defined as

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)du\,.$$

Unlike VaR, ES does not depend on a particular definition of the quantile, but only depends on the distribution of X and the level of α.

Expected Shortfall as a preferable risk measure

In section 3.2.3, a discussion regarding some of the shortcomings of VaR was provided. This section will draw attention to the enhanced properties of ES as an alternative measure of risk.

1. Expected Shortfall provides insight into the severity of tail events

Embrechts et al. (2005) show how an expression can be derived for a continuous distribution, which illustrates that ES can be interpreted as the expected loss that is incurred in the event that VaR is surpassed:

Lemma 3.10:

For an integral loss L with continuous distribution function F_L and any $\alpha \in (0,1)$ it follows that

$$ES = \frac{E(L; L \geq q_\alpha(L))}{1-\alpha} = E(L|L \geq VaR).$$

Proof: See (Embrechts, Frey & McNeil 2005, p. 45)

Embrechts et al. (2005) also prove that for a discontinuous loss distribution, F_L, the above expression does not hold for all α and that the following expression holds:

$$ES = \frac{1}{1-\alpha}\Big(E(L; L \geq q_\alpha) + q_\alpha\big(1-\alpha-P(L \geq q_\alpha)\big)\Big).$$

2. Expected Shortfall is a coherent risk measure

Acerbi and Tasche (2002) also state that the most important property of expected shortfall might be its coherence:

Proposition 3.11 (Coherence of ES) (Delbaen 1998):

Let $\alpha \in (0,1)$ be fixed. Consider a set V of real-valued random variables on some probability space $(\Omega, \mathcal{A}, P)$ such that $E(|X^-|) < \infty$ for all $X \in V$. Then $\rho: V \to \mathbb{R}$ with $\rho(X) = ES_\alpha(X)$ for $X \in V$ is a coherent risk measure in the sense that it is:

Monotonic: $X \in V,\ X \geq 0 \Rightarrow \rho(X) \leq 0$,

Sub-additive: $X, Y, (X+Y) \in V \Rightarrow \rho(X+Y) \leq \rho(X) + \rho(Y)$,

Positively homogeneous: $X \in V, h > 0, (hX) \in V \Rightarrow \rho(hX) = h\rho(X)$, and

Translation invariant: $X \in V, a \in \mathbb{R} \Rightarrow \rho(X+a) = \rho(X) - a$.

Proof: see (Acerbi & Tasche 2002, p. 5)

3. Expected Shortfall is less sensitive to changes in the confidence level

Risk measures like VaR, WCE, TCE and CVaR are very sensitive to changes in the confidence level $\boldsymbol{\alpha}$ when applied to discontinuous distributions, whereas Expected Shortfall is less sensitive when there is a switch in the confidence level. This could be explained due to the fact that ES is continuous with respect to $\boldsymbol{\alpha}$.

Another property of ES can be seen in the next proposition; the smaller the level of α the greater the risk:

Proposition 3.12:

If X is a real-valued random variable on a probability space $(\Omega, \mathcal{A}, P)$ with $E[X^-] < \infty$ and $\alpha \in (0,1)$ is fixed, then

$$\bar{x}_{(\alpha)} = \alpha^{-1} \int_0^{\alpha} x_{(u)} du,$$

where $\bar{x}_{(\alpha)}$ is the tail mean given by

$$\bar{x}_{(\alpha)} = \alpha^{-1} \left(E\left[X_{1_{\{X \leq x_{(\alpha)}\}}} \right] \right) + x_{(\alpha)} \left(\alpha - P\left[X \leq x_{(\alpha)} \right] \right)$$

and $x_{(u)}$ is the lower α-quantile of X, given by

$$x_{(\alpha)} = q_{\alpha}(X) = inf\{x \in \mathbb{R}: P[X \leq x] \geq \alpha\}.$$

Proof: see (Acerbi, Nordio & Sitori 2001, p. 8)

From the definition of ES and the above proposition follows that

$$ES_{\alpha} = -\alpha^{-1} \int_0^{\alpha} q_u(X) du \, .$$

The next corollary shows that ES is continuous with respect to α:

Corollary 3.13:

If X is a real-valued random variable with $E[X^-] < \infty$, then the mappings $\alpha \to \bar{x}_{\alpha}$ and $\alpha \to ES_{\alpha}$ are continuous on $(0,1)$.

Proof: Follows directly from the above proposition and equation.

Through their names CVaR, TCE and WCE all illustrate that they are *conditional* expected values of a random variable X. On the other hand, one should note that the α-tail mean is not submitted under a representation in terms of a conditional expectation X.

3.3.5. The relationships between WCE, TCE, CVaR and ES

This paragraph is devoted to clarifying the relations between WCE, TCE, CVaR and ES. The next proposition, as presented by Acerbi and Tasche (2002) will assist in effortlessly deriving the relations that exist between these measures.

Proposition 3.14

Let $\alpha \in (0,1)$ be fixed and let X be a real-valued random variable on a probability space $(\Omega, \mathcal{A}, P)$. Furthermore, assume there exists a function $f: \mathbb{R} \to \mathbb{R}$ such that $E[(f \circ X)^-] < \infty$, $f(x) \leq f(x_{(\alpha)})$ for $x < x_{(\alpha)}$, and $f(x) \geq f(x_{(\alpha)})$ for $x > x_{(\alpha)}$. Finally, let $A \in \mathcal{A}$ be an event with $P[A] \geq \alpha$ and $E[|f \circ X|1_A] < \infty$, then

1. $TM_\alpha(f \circ X) \leq E[f \circ X|A]$,

2. $TM_\alpha(f \circ X) = E[f \circ X|A]$ if $P[A \cap \{X > x_{(\alpha)}\}] = 0$ and
   a. $P[X < x_{(\alpha)}] = 0$ or
   b. $P[X < x_{(\alpha)}] > 0, P[\Omega \backslash A \cap \{X < x_{(\alpha)}\}] = 0$, and $P[A] = \alpha$,

3. if $f(x) < f(x_{(\alpha)})$ for $x < x_{(\alpha)}$ and $f(x) > f(x_{(\alpha)})$ for $x > x_{(\alpha)}$, then $TM_\alpha(f \circ X) = E[f \circ X|A]$ implies $P[A \cap \{X > x_{(\alpha)}\}] = 0$ and either (1) or (2).

Proof: see (Acerbi & Tasche 2002, p. 11).

From this proposition the next two corollaries follow.

Corollary 3.15

Let $\alpha \in (0,1)$ be fixed and let X be a real-valued random variable on a probability space $(\Omega, \mathcal{A}, P)$ and $E[X^-] < \infty$, then

1. $TCE^\alpha(X) \leq TCE_\alpha(X) \leq ES_\alpha(X)$ *, and*

2. $TCE^\alpha(X) \leq WCE_\alpha(X) \leq ES_\alpha(X)$**.**

Proof: Follows directly from lemma 3.10 and the above proposition.

Corollary 3.16

Again let $\alpha \in (0,1)$ be fixed and let X be a real-valued random variable on a probability space $(\Omega, \mathcal{A}, P)$ and $E[X^-] < \infty$, then

1. $P\left[X \leq x^{(\alpha)}\right] = \alpha, P\left[X < x_{(\alpha)}\right] > 0$ or $P\left[X \leq x^{(\alpha)}, X \neq x_{(\alpha)}\right] = 0$ if and only if
   a. $ES_\alpha(X) = WCE_\alpha(X) = TCE_\alpha(X) = TCE^\alpha(X)$.

   Moreover, (a) holds if the distribution of X is continuous, i.e. $P[X = x] = 0$ for all $x \in \mathbb{R}$.
2. $P\left[X \leq x^{(\alpha)}\right] = \alpha, P\left[X < x_{(\alpha)}\right] = 0$ if and only if $ES_\alpha(X) = TCE_\alpha(X)$.

Proof: Follows directly from the above corollary and proposition.

Firstly, the relation between WCE and TCE will be investigated. From corollary 3.15 follows that $P\left[X \leq x_{(\alpha)}\right] > \alpha$ implies that $WCE_\alpha(X) \geq TCE^\alpha(X)$.

Furthermore, from corollary 3.16 follows that
$P\left[X \leq x_{(\alpha)}\right] = \alpha$ implies that $WCE_\alpha(X) \leq TCE_\alpha(X)$.

Also, since $TCE^\alpha \geq VaR^\alpha$, WCE_α is the "*smallest coherent risk measure dominating* VaR^α*"* (Acerbi & Tasche 2002).

Secondly, when examining the relation between WCE and ES, WCE and ES are related when the probability space is "small", i.e., when the random variables under consideration are finite and always positive it will allow us to switch to a "larger" probability space. More generally,

Proposition 3.17

Let $X, Y \in \mathbb{R}$ be random variables on a probability space $(\Omega, \mathcal{A}, P)$ and $E[X^-] < \infty$ and $\alpha \in (0,1)$ fixed. Furthermore, assume that $Y = f \circ X$ where f satisfies $f(x) \leq f\left(x_{(\alpha)}\right)$ for $x < x_{(\alpha)}$, and $f(x) \geq f(x_{(\alpha)})$ for $x > x_{(\alpha)}$.

1. If $P\left[X \leq x_{(\alpha)}\right] = \alpha$ then $ES_\alpha(Y) = -\inf_{A \in \mathcal{A}, P[A] \geq \alpha} E[Y|A]$.

2. If the X is continuous then $ES_\alpha(Y) = WCE_\alpha(Y)$.

Proof: see (Acerbi & Tasche 2002, p. 1500)

Lastly, $CVaR^\alpha(X) = TCE^\alpha(X)$ if the distribution of X is continuous, without any additional assumptions.

3.4. Stress Value at Risk

Coste et al. (2009) demonstrate that Stress Value at Risk (StressVaR) has enhanced and similar properties as a number of VaR measures. They also illustrate that a portfolio constructed using StressVaR, on average, outperforms both the market and the portfolios constructed using common VaR measures.

StressVaR uses factor models, as does the factor-based VaR estimate, but its strength lies in the modelling of nonlinearities and its ability to analyze numerous potential risk factors.

Three steps exist in modelling StressVaR which starts with the selection of a large sample of factors (Coste, Douady & Zovko 2009):

1. *Factor scoring*

A dynamic nonlinear model of the fund is estimated for each of the factors in the sample against the factor obtaining a goodness of fit t-measure. The factor p-values are then used to score and rank the factors.

2. *Estimating factor risks*

By using a calibrated model for each factor, the fund returns are predicted for all possible factor returns from the first to 99th quantile of the long-term factor return distribution.

3. *Estimating the StressVaR*

StressVaR is estimated as the maximum predicted loss across all top selected factors. This estimate for the VaR might overestimate the real VaR. Coste et al. (2009) however prove that this approach ultimately leads to better investment decisions.

The area of StressVaR is open to future research and improvements. "*With a methodology of this kind, the formerly rigid boundary between risk-management and asset allocation is arguably fading*" (Coste, Douady & Zovko 2009, p. 17).

4. Copulas and dependence

In Paul Wilmott's article, Name and *Shame in Our New Blame Game! Results Part 1*, he describes copulas as an "*abomination*". He further states that copulas are "*such abstract models that only a few people, mostly with severe emotional intelligence problems, really understand*". This section will

attempt to provide the reader with a technical background on the subject of dependence functions, better known as copula functions, to assist in making these functions less abstract.

As a mathematical tool that encodes a dependence structure, copulas are a borrowed concept from statistics. According to Embrechts et al. (2005, p. 197), "*every joint distribution function for a random vector of risk factors implicitly contains both a description of the marginal behavior of individual risk factors and a description of their dependence structure; the copula approach provides a way of isolating the description of the dependence structure".* Copulas are therefore a popular technique to model joint multi-dimensional problems, as they can be applied as a mechanism that models relationships among multivariate distributions.

Sklar introduced copulas in 1959 in the context of a probabilistic metric space. His theorem showed that any joint distribution can be written as a function of marginal distributions

$$F(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

and that the class of functions $C(.)$, denoted copula functions, may be used to extend the class of multivariate distributions well beyond those most familiar with and commonly used (Cherubini, Luciano & Vecchiato 2004). According to Nelsen (2006, p. 1), copulae can be seen from two viewpoints: "*From one point of view, copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions. Alternatively, copulas are multivariate distribution functions whose one-dimensional margins are uniform on the interval (0, 1)*".

Cherubini et al. (2011) question what copula functions can do above other techniques? They answer this question by stating that copula functions are the main tool for a bottom-up approach and that the essence of the answer lies in this fact.

Section 4.1 starts by exploring the basic definitions of bivariate copulas that will lead to perhaps one of the main theorems in this dissertation, known as Sklar's theorem. In section 4.2 the concept of survival copulas will briefly be explored. Section 4.3 is aimed at providing the reader with an overview of dependence structures, as well as some well-known measures of association. Section 4.4 will define some parametric classes of bivariate copulas. Finally, section 4.5 is aimed at translating all the previous sections to the multivariate case.

4.1. Bivariate copulas

This section is aimed at providing an overview of the technical background of copulas. This section will start by introducing the bivariate case. Once one has built an understanding of the bivariate case, this knowledge can easily be extended to the multivariate case.

Definition 4.1: Grounded function (Cherubini, Luciano & Vecchiato 2004)

Let $A, B \subset \mathbb{R}$ be two non empty subsets and G be a function such that $G: A \times B \to \mathbb{R}$. Let a be the least element of A and b be the least element of B. A function is said to be grounded if $\forall (v, z) \in A \times B$,

$$G(a, z) = 0 = G(v, b).$$

Definition 4.2: 2-increasing function (Cherubini, Luciano & Vecchiato 2004)

A function $G: A \times B \to \mathbb{R}$ is called 2-increasing if for every "rectangle" $[v_1, v_2] \times [z_1, z_2]$ such that $v_1, v_2 \in A$ with $v_1 \leq v_2$ and $z_1, z_2 \in B$ with $z_1 \leq z_2$,

$$G(v_2, z_2) - G(v_2, z_1) - \big(G(v_1, z_2) - G(v_1, z_1)\big) \geq 0 \quad (4.2.1).$$

According to the function G, the left hand side of equation 4.2.1 represents the mass or area of the rectangle $[v_1, v_2] \times [z_1, z_2]$. As a consequence, a 2-increasing function allocates mass to every rectangle within its domain.

Definition 4.3: 2-dimensional subcopula (Cherubini, Luciano & Vecchiato 2004)

A 2-dimensional subcopula is a real-valued function C' with the following properties:

C' is defined on $A \times B$ where $A, B \subset I = [0,1]$, A, B are nonempty subsets and $(0,1) \in A, B$ if

$$C': A \times B \to \mathbb{R};$$

C' is grounded, 2 –increasing and

$$C'(v, 1) = v,\ C'(1, z) = z$$

for all $(v, z) \in A \times B$.

Let $A = B = I$. According to the above definition, the functions

$$G(v, z) = max(v + z - 1, 0);$$

$$G(v, z) = min(v, z);$$

$$G(v, z) = vz$$

are subcopulas.

Definition 4.4: 2-dimensional subcopula (Cherubini, Luciano & Vecchiato 2004):

A 2-dimensional copula C is a 2-dimensional subcopula with $A = B = I$.

Since we set $A = B = I$ in our previous example, the functions

$$G(v, z) = max(v + z - 1, 0);$$
$$G(v, z) = min(v, z);$$
$$G(v, z) = vz$$

are also copulas.

Many important properties of copulas are also properties of subcopulas. Next, some important properties of subcopulas will be considered as presented in Cherubini et al. (2004).

Property 4.5:

A function $G(v, z)$ that is both grounded and 2-increasing is non-decreasing in both the variables v and z.

Property 4.6:

As a result from property 4.5, it can be noted that a subcopula is bounded by 0 and 1

$$0 \leq C'(v, z) \leq 1 \ \forall (v, z) \in A \times B.$$

Property 4.7:

A subcopula C' is uniformly continuous on $A \times B$.

Property 4.8:

In the interior of $A \times B$, both partial derivatives of C', $\frac{\partial C'}{\partial v}$ and $\frac{\partial C'}{\partial z}$, exist almost everywhere and take values on I.

The next theorem illustrates that subcopulas are bounded functions. According to Nelson (2006), these results were first published by Hoeffding in 1940 at the outbreak of the Second World War. Uninformed of Hoeffding's work, in 1951 Fréchet obtained many of these results in his own work. In acknowledgment of their mutual contribution to this result, we refer to Fréchet-Hoeffding bounds.

Theorem 4.9: Fréchet-Hoeffding bounds

Bivariate subcopulas are bounded for all $(v, z) \in A \times B$,

$$max(v + z - 1, 0) \leq C'(v, z) \leq min(v, z).$$

Proof: See (Nelson 2006, p. 47)

As seen in the previous examples, every copula is in fact a subcopula. Consequently, the above theorem also holds for copulas. As also seen, when $A = B = I$, then $C'(v, z)$, $min(v, z)$ and $max(v + z - 1, 0)$ are also copulas. It thus follows that the minimum copula, C^- is given by

$$C^-(v, z) = max(v + z - 1, 0)$$

and the maximum copula, C^+ is given by

$$C^+(v, z) = min(v, z).$$

In financial applications, these bounds for copulas are of importance since it presents precise minimum and maximum prices or risks. This result also allows us to define an association order, that is, we can compare copulas and determine which one is bigger than the other.

Definition 4.10: Comparison result (Cherubini, Luciano & Vecchiato 2004)
Let C_1 and C_2 be two copulas. We say that C_1 is smaller than C_2, denoted by $C_1 \prec C_2$, if

$$C_1(v, z) \leq C_2(v, z)$$

for every $(v, z) \in I^2$.

This result has some limitations, as not all copulas can be strictly ordered and evaluated against each other. In other words, this definition should rather be seen as a guideline and not as a strong result.
In order to obtain a better understanding of how copulas relate to probabilities, we first have to investigate how copulas can be expressed as joint density functions of standard uniform random variables

$$C(v, z) = P[U_1 \leq v, U_2 \leq z].$$

If one now assumes that the random variables X and Y have marginal probability density functions $F_1(x)$ and $F_2(y)$ then

$$\begin{aligned} C\big(F_1(x), F_2(y)\big) &= P[U_1 \leq F_1(x), U_2 \leq F_2(y)] \\ &= P\big[{F_1}^{-1}(U_1) \leq x, {F_2}^{-1}(U_2) \leq y\big] \\ &= P[X \leq x, Y \leq y] \\ &= F(x, y) \end{aligned}$$

where F is the joint cumulative distribution function of X and Y. This relationship is of primary importance when using copulas in risk management.

Now that we have defined this primary link between copulas and probability, the next logical step will be to express probabilities in terms of copulas. One can now form joint probabilities of uniform variates using copulas

$$P[U_1 \leq v, U_2 > z] = v - C(v, z)$$
$$P[U_1 > v, U_2 \leq z] = u - C(v, z)$$
$$P[U_1 \leq v, U_2 \leq z] = \frac{C(v, z)}{z}$$
$$P[U_1 < v, U_2 > z] = \frac{v - C(v, z)}{1 - z}.$$

One can also formulate a set of "new" conditional copulas (Cherubini et al. 2011), as can be seen in the following example

$C_{1,2}(v, z) = P[U_1 \leq v, U_2 = z]$

$$= \lim_{\Delta u \to 0} \frac{C(v, z + \Delta z) - C(v, z)}{\Delta z}$$
$$= \frac{\partial C(v, z)}{\partial z}$$

and

$C_{1,2}(v, z) = P[U_1 = v, U_2 \leq z]$

$$= \lim_{\Delta z \to 0} \frac{C(v + \Delta v, z) - C(v, z)}{\Delta v}$$
$$= \frac{\partial C(v, z)}{\partial v}.$$

4.2. Sklar's theorem

Cherubini et al. (2011, p. 4) explained that copulas are "*built on purpose with the goal of pegging a multivariate structure to prescribed marginal distributions*". This problem was first tackled and explained by Abe Sklar in 1959. Sklar's theorem can be seen as an essential outcome forming the foundation in the use of copulas in probability theory (Cherubini, Luciano & Vecchiato 2004). This theorem illustrates that there is a strong correspondence between the traditional joint distribution formulation and the (sub) copula. This section starts with a short overview of distribution functions as presented in Nelson (2006):

Definition 4.11:

A distribution function is a function F with domain $(-\infty, \infty)$ such that F is nondecreasing, $F(-\infty) = 0$ and $F(\infty) = 1$.

Definition 4.12:

A joint distribution function is a function F with domain $[-\infty, \infty]$ such that F is 2-increasing, $F(x, -\infty) = F(-\infty, y) = 0$ and $F(\infty, \infty) = 1$.

Let X and Y be two real-valued measurable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let F_1 and F_2 denote the two marginal distributions and F the joint distribution.

Theorem 4.13: Sklar's theorem (Nelson 2006):

If $F(x, y)$ is a joint distribution function with marginal distributions F_1 and F_2, then there exists a unique subcopula C' whose domain is equal to $Range\ F_1 \times Range\ F_2$, such that

$$F(x, y) = C'(F_1(x), F_2(y))$$

for every (v, z) in $Range\ F_1 \times Range\ F_2$. Conversely, let F_1 and F_2 be two marginal distribution functions and C' be any subcopula whose domain equals to $Range\ F_1 \times Range\ F_2$, then for all x, y in $\mathbb{R}^2$ it follows that

$$C'(F_1(x), F_2(y))$$

is a joint density function with margins $F_1(x)$ and $F_2(y)$. If $F_1(x)$ and $F_2(y)$ are continuous functions, then the subcopula C′ is indeed a copula. Otherwise there exists a copula C such that $C(v, z) = C'(v, z)$ for (v, z) in the cross product of ranges of F_1 and F_2.

Proof: see (Nelson 2006, p. 21)

Corollary 4.14 (Nelson 2006):

Let F be a joint distribution function with marginal distribution functions F_1 and F_2. There exists a unique subcopula C' such that

$$F(x, y) = C'(F_1(x), F_2(y))$$

is

$$C'^{(v,z)} = F\left({F_1}^{-1}(x), {F_2}^{-1}(y)\right).$$

It can thus be concluded that if the ranges of F_1 and F_2 are equal to I, the subcopula is a copula (Nelson 2006).

Sklar's theorem clearly illustrates that the main purpose of a copula is to provide a dependence structure and this is indeed the true power of the copula approach. "*Decomposing the multivariate distribution into the marginal distributions and the copula allows for the construction of better models of the individual variables than would be possible if we constrained ourselves to look only at*

existing multivariate distributions" (Patton 2006, p. 529). Thus, the joint distribution can accurately be split between:

a) the marginal distributions centralizing the behavior of each random variable independently from each other, and
b) the copula function centralizing the association/dependence structure of the random variables.

Cherubini et al. (2011) also state that the core advantage of these properties is that the requirement of the joint distribution can be split from the requirement of the marginal distributions. The copula approach provides the modeler with a greater degree of freedom when fitting copulas to observed data opposed to more traditional distributional approaches. This is mainly because the modeler can now use two different dimensions of dependence and marginal behavior.

Through applying Sklar's theorem, one can rewrite the minimum and maximum copulas as follows

$$C^{-}\big(F_1(x), F_2(y)\big) = max(F_1(x) + F_2(y) - 1, 0)$$

and

$$C^{+}\big(F_1(x), F_2(y)\big) = min\big(F_1(x), F_2(y)\big).$$

Theorem 4.15: Monotone transformations and copulas (Cherubini, Luciano & Vecchiato 2004):

Let X and Y be continuous random variables with marginal distribution functions F_1 and F_2 and copula C. Let g_1 and g_2 be two increasing functions. Then the transformations $g_1(X)$ and $g_2(Y)$, with marginal probability functions $G_1 = F_1({g_1}^{-1})$, $G_2 = F_2({g_2}^{-1})$ and joint probability function G, i.e.

$$G(v, z) = P[g_1(X) \leq v, g_2(Y) \leq z]$$

have copula C, i.e.

$$G(v, z) = C\big(G_1(v), G_2(z)\big).$$

This theorem is of particular interest when one wants to transform the price distribution into the distribution of log returns. One can also conclude that copulas are invariant when it comes to monotone increasing transformations.

Another central concept in copula theory is that of survival copulas. This concept is of fundamental importance in the theory underlying the pricing of credit derivatives.

Definition 4.16: Survival copulas (Cherubini et al. 2011)

The joint survival copula, $\bar{C}$, associated with the copula C is defined as

$$\bar{C}(v, z) = v + z - 1 + C(1 - v, 1 - z).$$

Property 4.17:

The survival copula associated with the minimum copula, maximum copula and product copula is the minimum copula, maximum copula and product copula themselves, i.e.

$$\bar{C}^{-} = C^{-}$$
$$\bar{C}^{+} = C^{+}$$
$$\bar{C}^{\perp} = C^{\perp}.$$

4.3. Measures of dependence

Two random variates X and Y are said to be associated if they are dependent. This section is aimed at providing a theoretical background into dependence structures as well as measures of association. Measures of dependence are of particular importance since the correlation (or rank correlation) matrix plays a crucial role when fitting copulas to multivariate data.

Since copulas will be used to evaluate a dependence structure between random variables, it makes sense to define some of the basic principles of dependence and association.

4.3.1. Independence and dependence

One of the most basic definitions in probability theory is that of independence.

Definition 4.18: Independence

Two random variables X and Y are said to be independent if their joint distributions are given by the product of their marginal distributions

$$F(x, y) = F_1(x)F_2(y).$$

In order to obtain a similar result in terms of copulas, one must first define the concept of a product copula.

Definition 4.19: Product copula

The product copula can be defined as $C^{\perp}(v, z) = vz$. The product copula has no dependence structure which implies that the dependence is equal to zero.

As an outcome of definition 4.19 and Sklar's theorem, one can now express independence in terms of copulas. In this case two random variables X and Y are independent if they have the product copula, $C^{\perp}$ defined as (Cherubini, Luciano & Vecchiato 2004)

$$C^{\perp}\big(F_1(x), F_2(y)\big) = F_1(x)F_2(y).$$

Having now defined independence, the next logical step would be to define dependence. Two events are perfectly positive dependent if one event happens whenever the other event takes place. In contrast, two events are perfectly negative dependent if one event takes place, only if the other does not take place. Perfect positive and negative dependence can formally be defined as follows:

Definition 4.20: Perfect positive dependence (Cherubini, Luciano & Vecchiato 2004)
The two random variables X and Y are then said to be perfectly positively dependent if they have the minimum copula

$$C^{+}\big(F_1(x), F_2(y)\big) = min\big(F_1(x), F_2(y)\big).$$

Definition 4.21: Perfect negative dependence (Cherubini, Luciano & Vecchiato 2004)
The two random variables X and Y are said to be perfectly negative dependent if they have the maximum copula

$$C^{-}\big(F_1(x), F_2(y)\big) = max(F_1(x) + F_2(y) - 1{,}0).$$

Cherubini et al. (2011) provide a handy example of how one can set up a copula family based on these three copulas. They show that the weighted averages of the maximum copula, the minimum copula and the product copula form the Fréchet family (Fréchet 1951) of copulas, which can be defined as follows

$$C\big(F_1(x), F_2(y)\big) \equiv \alpha\, min(F_1(x), F_2(y)) + (1-\alpha-\beta)(F_1(x)F_2(y)) + \beta\, max(F_1(x) + F_2(y) - 1, 0)$$

with $\alpha, \beta \geq 0$ and $\alpha + \beta \leq 1$.

A special case that is often used in financial applications (Zi-sheng, Hui & Xiang-qun 2009) is that with $\beta = 0$

$$C\big(F_1(x), F_2(y)\big) \equiv \alpha\, min\big(F_1(x), F_2(y)\big) + (1-\alpha)\,\big(F_1(x)\,F_2(y)\big),$$

which is known as a mixture copula.

4.3.2. Measuring the degree of association

Until now, only the cases of independence, perfect positive dependence and perfect negative dependence have been considered. However, copulas can be used to model any level of association that exists between random variables. In order to achieve this goal, one must first define measures of association.

Definition 4.22: Measure of association (Cherubini, Luciano & Vecchiato 2004)

A measure $M_{X,Y}^{C}$ between two random variables X and Y with copula C is a measure of association if it satisfies the following properties:

1. *Completeness*: the measure is defined for every pair of random variables

2. *Normalized measure:* $-1 \leq M_{X,Y}^{C} \leq 1$

3. *Symmetric:* $M_{X,Y}^{C} = M_{Y,X}^{C}$

4. If X and Y are *independent* then $M_{Y,X}^{C} = 0$

5. $M_{-X,Y}^{C} = M_{X,-Y}^{C} = M_{Y,X}^{C}$

6. *Respects association order*: if $C_1 \prec C_2$, then $M_{X,Y}^{C_1} \leq M_{X,Y}^{C_2}$

7. $M_{X,Y}^{C}$ converges (point wise) when the copula does, which implies that if $\{X_n, Y_n\}$ is a sequence of continuous random variables with copula C_n, and $lim_{n\to\infty} C_n(v,z) = C(v,z)\ \forall\ (v,z) \in\ l^2$ then $lim_{n\to\infty} M_{X_n,Y_n}^{C} = M_{X,Y}^{C}$.

Since copulas are tied to dependence structures, they must be related to dependence measures (Cherubini et al. 2011). Also, different types of copulas capture different types of dependence between variables. The most common dependence measures will be investigated, namely, linear correlation and rank correlation.

Linear correlation

Correlation plays a central role in capital allocation. Campbell et al. (1997) discuss how two models for an optimal portfolio, the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT), employ correlation as measures of dependence between different financial instruments which

is built on the assumption of multivariate normally distributed returns. Embrechts et al. (2002) describe correlation not only as a "*source of confusion*", but also a concept that is frequently misunderstood. Correlation gives an indication of what happens in the smallest, infinitesimal timescale. However, it does not provide an understanding of the bigger picture. Correlation is also one of the most unstable statistical parameters, even more unstable than volatility (Wilmott 2006).

Correlation, ρ, answers the question of how two objects, e.g. two assets, are related to each. For example, two assets can be perfectly positive correlated, $\rho = +1$, but still move in opposite directions or perfectly negative correlated, $\rho = -1$, but move in the same direction. Correlation can thus be seen as only one particular measure of stochastic dependence among many. It is thus important to note that dependence cannot be distinguished on the grounds of correlation alone. Firstly, consider a formal definition of correlation as presented by Embrechts et al. (2002):

Definition 4.23: Linear correlation
Linear correlation is defined as

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

where the variance of X is $var(X) = E[(X - E[X])^2]$, the variance of Y is $var(Y) = E[(Y - E[Y])^2]$ and $cov[X,Y] = E[XY] - E[X]E[Y]$ is the covariance between X and Y.

Embrechts et al. (2002) provides three explanations for the popularity of linear correlation in finance. The first reason is that correlation is easy to compute, unlike alternative measures of dependence, one only needs to compute the variance and covariance to derive correlation for bivariate distributions. The second reason is that the variance of any linear combination can be fully determined by the covariance between the components, a fact that is commonly exploited in portfolio theory. The last reason is that correlation is unaffected as a measure of dependence in multivariate normal, spherical and elliptical distributions.

Correlation is thus only useful to measure linear relationships/dependence between variables, however for strongly non-normal variables, linear correlation can be confusing. Also, linear correlation cannot be preserved under non-linear transformations. Cherubini et al. (2004) list the following consequences of these properties:

Property 4.24:

$\rho_{X,Y}$ is invariant under linear increasing transformations, not under non-linear increasing transformations.

Property 4.25:

$\rho_{X,Y}$ is bounded and the bounds are reached in the perfectly negative dependent case and the perfectly positive dependent case.

Property 4.26:

$\rho_{X,Y}$ may not be 1 (-1) for perfectly positive (negative) dependent variables.

These properties as well as the next two resulting theorems will assist us in solving the restriction of linear correlation, namely that linear correlation does not satisfy property 7 of definition 4.22.

Theorem 4.27: Invariance (Cherubini, Luciano & Vecchiato 2004):

If g_1 and g_2 are two increasing functions, almost everywhere respectively on the range of F_1 and F_2, then

$$M_{X,Y}^{C} = M_{g_1(X)g_2(Y)}^{C}.$$

Theorem 4.28a: Lower bound

If X and Y are perfectly negative dependent, then

$$M_{X,Y}^{C} = -1.$$

Theorem 4.28b: Upper bound

If X and Y are perfectly positive dependent, then

$$M_{X,Y}^{C} = 1.$$

Firstly, theorem 4.27 ensures that any linear or nonlinear (increasing) transformation does not affect the measure of association. Secondly, theorem 4.28a (4.28b) guarantees that the lower (upper) bounds of -1 (1) are attained in the case of perfectly negative (positive) dependence.

Rank correlation

Having now defined a revised set of desired properties with regards to measures of association, the next theorem will define a class of association measures that satisfies these revised properties.

Theorem 4.29: A class of association measures (Cherubini, Luciano & Vecchiato 2004)

Let f be a bounded, weakly monotone, odd function with $[-\frac{1}{2}, \frac{1}{2}]$ as its domain, then

$$k \iint_{I^2} f\left(v - \frac{1}{2}\right) f\left(z - \frac{1}{2}\right) dC(v, z)$$

where $k^{-1} = \int_{I^2} f\left(u - \frac{1}{2}\right) du$, is a association measure.

As an alternative measure to linear correlation, this section will investigate rank correlation. Unlike linear correlation, rank correlation measures association in terms of rank, i.e. rank correlation measures the degree to which large (small) values of one random variable associate with large (small) values of another (Nelson 2006). This is also commonly referred to as the measure of concordance. Rank correlation is invariant subject to non-linear monotonic transformations and can thus provide a finer view of the dependence structure at hand (Embrechts, McNeil & Straumann 2002). Another useful property of rank correlation in describing association between random variables is the fact that this measure is invariant to the choice of marginal distributions.

1. Spearman's rho

The first dependence measure under rank correlation that will be considered is Spearman's rho. Spearman's rho is closely related to linear correlation, with the only difference being that the calculations are done after the numbers have been ranked. Thus, Spearman's rho can be seen as the linear correlation between n associated cumulative distribution functions.

Definition 4.30: Spearman's rho (Cherubini, Luciano & Vecchiato 2004)

In order to define Spearman's rho, let $f(u) = u$ in theorem 4.29, then

$$\rho_s(v, z) = 12 \iint_{I^2} C(vz) dv dz - 3$$

$$= 12 \iint_{I^2} vz dC(vz) - 3.$$

Embrechts et al. (2005, p. 207) state that "*Spearman's rho is simply the linear correlation of the probability transformed random variables, which for continuous random variables is the linear correlation of the unique copula".* Thus, one can also express Spearman's rho as follows

$$\rho_s = 12\mathbb{E}\left[F_x(x)F_y(y)\right] - 3$$

where F_x and F_y are the marginal distributions of x and y respectively.

In order to calibrate single parameter copulas using Spearman's rho, one must calculate Spearman's rho from observed data. According to Kruskal (1958), Spearman's rho can also be defined in terms of concordance and discordance for n random independently and identically distributed (iid) pairs, which translates into the following

$$\rho_s = 1 - 6\frac{\sum_{i=1}^{n}\left(R_{x_i} - R_{y_i}\right)^2}{n(n^2-1)}$$

where $R_{x_i} = rank(x_i)$ and $R_{y_i} = rank(y_i)$.

Another fascinating property of Spearman's rho is the fact that the copula and its associated survival copula have the same Spearman's rho. This is illustrated in property 4.31.

Property 4.31:

A copula and its associated survival copula have the same Spearman's rho

$$\rho_s{}^{C} = \rho_s{}^{\bar{C}}.$$

Zi-sheng et al. (2009, p. 396) state that although linear correlation is the most commonly used measure of dependence it is not robust: "*… it can be close to 0 or close to 1, due to a single outlier*". They conclude that Spearman's rank correlation is a more robust measure of dependence.

2. Blomqvist's beta

A second, less used measure of dependence under rank correlation is Blomqvist's beta, also known as the medial correlation coefficient. This measure of dependence will briefly be considered next.

Definition 4.32: Blomqvist's beta (Nelson 2006)

Let $f(u) = sgn(u)$ in theorem 4.29, then

$$\rho_b = 4C\left(\frac{1}{2},\frac{1}{2}\right) - 1.$$

3. Kendall's tau

Kendall's tau is the third measure of dependence under rank correlation that will be considered. Theorem 4.29 cannot provide a characterization of all possible concordance measures, including Kendall's tau.

Definition 4.33: Kendall's tau (Cherubini, Luciano & Vecchiato 2004)

Kendall's tau is defined as

$$\rho_k(v,z) = 4\iint_{I^2} C(v,z)dC(v,z) - 1$$

or

$$\rho_k(v,z) = 1 - 4\iint_{I^2} \frac{\partial C(v,z)}{\partial v}\frac{\partial C(v,z)}{\partial z}dvdz.$$

According to Embrechts et al. (2005), two points (X_1, Y_1) and$(X_2, Y_2) \in \mathbb{R}^2$ are concordant if $(X_1 - X_2)(Y_1 - Y_2) > 0$ and discordant if $(X_1 - X_2)(Y_1 - Y_2) < 0$. Kendall's tau can also be seen as the difference between the probability of concordance and the probability of discordance of two iid random vectors, (X_1, Y_1) and (X_2, Y_2), with joint density function and copula C, which can be translated into

$$\rho_k = 4\mathbb{E}\big[F_{xy}(x,y)\big] - 1$$

where F_{xy} is the joint distribution of x and y.

In order to calibrate single parameter copulas, one must compute Kendall's tau from observed data. Since Kendall's tau can also be defined as the difference between the probability of concordance and the probability of discordance of n pairs (x_i, y_i), that were randomly drawn from a joint distribution, this result can be translated into the unbiased estimator

$$\rho_k = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j>i}^{n} sgn\Big(\big(x_i - x_j\big)\big(y_i - y_j\big)\Big).$$

As with Spearman's rho, a copula and its associated survival copula have the same Kendall's tau, as can be seen in property 4.34.

Property 4.34:

A copula and its associated survival copula have the same Kendall's tau

$$\rho_k{}^{C} = \rho_k{}^{\bar{C}}.$$

4. Gini's gamma

The final measure of dependence under rank correlation that will briefly be discussed is Gini's gamma. This measure can also not be expressed through theorem 4.29.

Definition 4.35: Gini's gamma (Nelson 2006)

Gini's gamma can be expressed as follows

$$\rho_g = 2 \iint_{I^2} (|v + z - 1| - |v - z|) dC(v, z) - 1.$$

As a final remark with regards to rank correlation, it is important to note that independence is sufficient but not necessary for concordance to be equal to zero. This property does not hold under linear correlation.

Property 4.36:

$\rho_{X,Y} = 0$ does not imply independence unless X and Y are Gaussian.

From now on, only Spearman's rho and Kendall's tau will be considered from the above mentioned measures of rank correlation. Embrechts et al. (2005) list the following properties that both Spearman's rho and Kendall's tau have in common:

a) Both measures of rank correlation are symmetric on the interval $[-1; 1]$.
b) Both assign a value of zero in the case of independence, however a rank correlation equal to zero does not automatically mean that the involved variables are independent (see property 4.36).
c) Both assign a value of 1 (-1) when the involved variables are comonotonic (countermonotonic).
d) For continuous marginal distributions, both are only dependent on the unique copula of the involved variables and both thus inherit their property of invariance under strictly increasing transformations.

4.4. Parametric classes of bivariate copulas

This section aims to provide a systematic development of the theory of copulas. However, this section will mainly focus on bivariate copulas. Once one has built a firm understanding of bivariate copulas, these concepts can easily be extended into the multivariate case (as discussed in section 4.5).

This section will focus on the basic definitions, properties as well as simulation algorithms for a few of the best known bivariate copulas, namely the Gaussian, Student's T, Fréchet and Archimedean copulas. Copulas that belong to the Archimedean family that will be considered in this section include the Clayton and Gumbel copulas.

4.4.1. Elliptical copulas

According to Malevergne and Sornette (2006), elliptical copulas like the Gaussian and Student t copula are derived from multivariate elliptical distributions. Elliptical copulas are simple to simulate and widely used in scenario analysis. This is due to their numerical tractability when generating a distribution of random variables.

Gaussian copula

The first copula that will be considered is the bivariate Gaussian or Normal copula. This copula forms part of the family of the implicit copulas and can be used to generate a joint normal distribution from normal marginal distributions.

Definition 4.37: Bivariate Gaussian copula (Cherubini, Luciano & Vecchiato 2004)

The bivariate Gaussian copula can be defined as

$$C^G_{\rho_{XY}} = \Phi_{\rho_{XY}}\left(\Phi^{-1}(v), \Phi^{-1}(z)\right)$$

$$= \int_{-\infty}^{\Phi^{-1}(v)} \int_{-\infty}^{\Phi^{-1}(z)} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} exp\left\{\frac{2\rho_{XY}st - s^2 - t^2}{2(1-\rho_{XY}^2)}\right\} dsdt$$

where $\Phi_{\rho_{XY}}$ is the joint density function of a standard bivariate Gaussian random variable with correlation ρ_{XY}.

According to Malevergne and Sornette (2006), the family of Gaussian copulas is fully parameterized by the degree of linear correlation. Thus, the conditional copula can be defined as

$$C^G_{1|2}(v, z) = \Phi\left(\frac{\Phi^{-1}(z) - \rho_{XY}\Phi^{-1}(v)}{\sqrt{1-\rho_{XY}^2}}\right).$$

Furthermore, if the class of Gaussian copulas is positively ordered, i.e. $\rho_1 < \rho_2$, then

$$C^G_{\rho_1} \prec C^G_{\rho_2}.$$

Also, the class of Gaussian copulas is comprehensive, i.e,

$$C^G_{\rho=-1} = C^-$$

$$C^G_{\rho=+1} = C^+.$$

Finally,

$$C^G_{\rho=0} = C^\perp.$$

According to Cherubini et al. (2011), because the Gaussian copula is completely symmetric, it will not capture any tail dependence. The Gaussian copula is most renowned for its application in pricing credit derivatives as proposed by David Li in 2000. The Gaussian copula has received wide spread criticism following the role of credit derivatives in the Financial Crisis. Amongst others, Salmon (2009) refer to the Gaussian copula as "*The formula that killed Wall Street*".

Procedure for constructing the bivariate Gaussian copula

The algorithm for generating the bivariate Gaussian copula with correlation matrix Σ proceeds as follows (Embrechts, Frey & McNeil 2005):

1. Generate a set of normally distributed variables, Z_1 and Z_2

2. Decompose Σ into A by using Cholesky decomposition, such that $\Sigma = AA'$.

3. Set $X = AZ$ to generate a correlated Gaussian vector.

4. Transform the uniform variables $U_1 = \Phi(X_1)$ and $U_2 = \Phi(X_2)$, where Φ is the standard normal distribution function, such that the pair (U_1, U_2) represents the bivariate Gaussian copula.

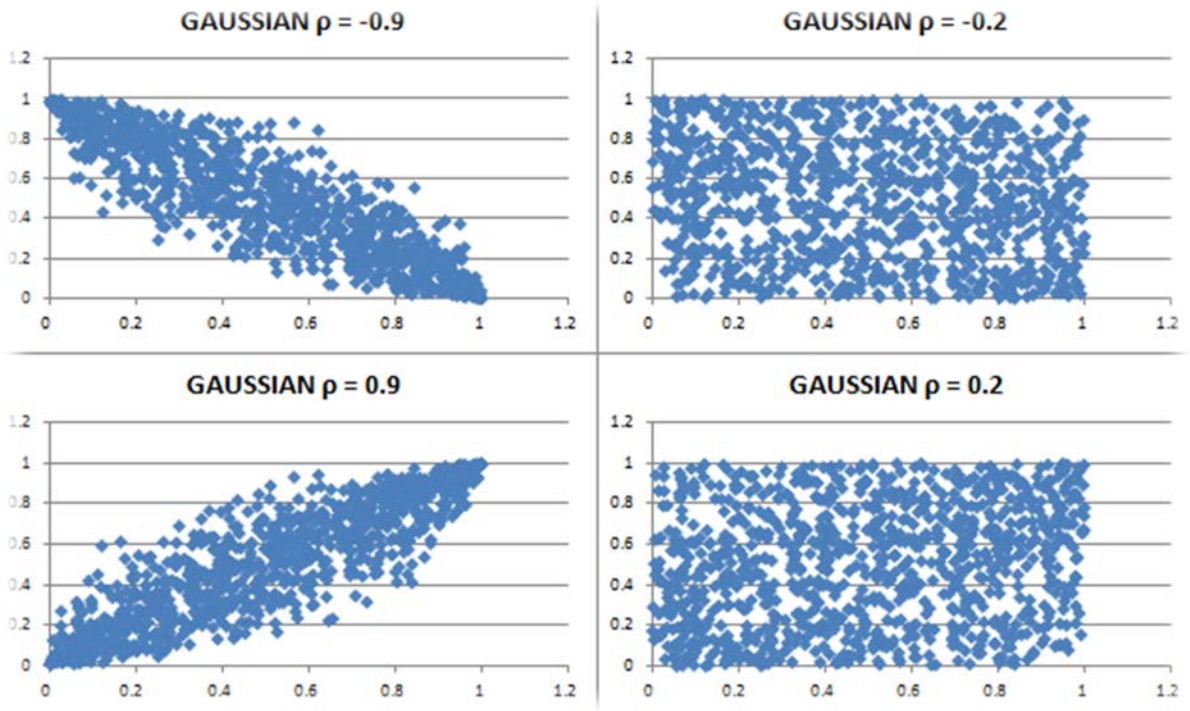

Figure 5: *Bivariate Gaussian copula using different correlations.*

Student t copula

As an alternative to the Gaussian copula, one can also use the Student t copula. The bivariate Student t copula also forms part of the family of implicit copulas, which are based on popular multivariate distributions. The Student t and Gaussian copulas are closely related in their central part. The main advantage of using the Student t copula is that, unlike the Gaussian copula, it allows for greater tail dependence. The Student t and Gaussian copulas become closer related as the degrees of freedom of the Student t copula increases (Malevergne & Sornette 2006). Even so, these two copulas could behave dissimilarly with regards to extreme dependencies.

The Student t copula is used to generate a joint t-distribution with ν degrees of freedom by using marginal t-distributions. The Student t copula with ν degrees of freedom is given by (Cherubini, Luciano & Vecchiato 2004)

$$t_\nu = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{s^2}{\nu}\right)^{-\frac{\nu+1}{2}} ds$$

where

$$\Gamma(n) = \int_0^{+\infty} t^{n-1}e^{-t}dt, \;\; n>0.$$

The bivariate t-distribution with ν degrees of freedom and correlation coefficient ρ is given by

$$T_{\rho,\nu}(v,z) = t_{\rho,\nu}\left(t_\nu^{-1}(v), t_\nu^{-1}(z)\right)$$

$$= \int_{-\infty}^{t_{\nu(v)}^{-1}} \int_{-\infty}^{t_{\nu(z)}^{-1}} \frac{1}{2\pi\sqrt{1-\rho_{XY}^2}} \left(1+\frac{s^2+t^2-2\rho st}{\nu(1-\rho^2)}\right)^{-\left(\frac{\nu+1}{2}\right)} dsdt.$$

Definition 4.38: Bivariate Student t copula (Cherubini, Luciano & Vecchiato 2004)

The bivariate t copula is defined as

$$C_{\rho_{XY}}^T(v,z) = \rho^{\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{{s_1}^2+{s_2}^2-2\rho s_1 s_2}{\nu(1-\rho^2)}\right)^{-\left(\frac{\nu+2}{2}\right)}}{\Gamma\left(\frac{\nu+1}{2}\right)^2 \prod_{j=1}^{2}\left(1+\frac{{s_j}^2}{\nu}\right)^{-\left(\frac{\nu+2}{2}\right)}}$$

where $s_1 = {t_\nu}^{-1}(v)$ and $s_2 = {t_\nu}^{-1}(z)$.

The conditional copula is given by

$$C^T_{2|1\nu,\rho}(v,z) = t_{\nu+1}\left(\sqrt{\frac{\nu+1}{\nu+t_\nu^{-1}(v)^2}}\frac{t_\nu^{-1}(z)-\rho t_\nu^{-1}(v)}{\sqrt{1-\rho^2}}\right).$$

The class of t copulas is also positively ordered and comprehensive, thus if $\rho_1 < \rho_2$, then

$$C^T_{\nu,\rho_1} \prec C^T_{\nu,\rho_2}$$

and

$$C^T_{\nu,\rho=-1} = C^-$$
$$C^T_{\nu,\rho=+1} = C^+.$$

However, note that for a finite ν

$$C^T_{\nu,\rho=0} \neq C^\perp.$$

In other words, the t copula is the copula function which joins the marginal t-distributions with the same degrees of freedom to the bivariate t-distribution. The t copula simplifies the bivariate t-distribution due to the fact that it can adopt any marginal distribution.

Unlike the Gaussian copula, the t copula does not only depend on the shape of the correlation matrix, but also on ν. This makes the t copula harder to use in applications and to fit to (Malevergne & Sornette 2006).

Procedure for constructing the bivariate Student t copula

The algorithm followed for generating the bivariate Student t copula with correlation matrix Σ proceeds as follows (Embrechts, Frey & McNeil 2005):

1. Generate a set of normally distributed variables, Z_1 and Z_2.

2. Decompose Σ into A by using Cholesky decomposition, such that $\Sigma = AA'$.

3. Draw an independent Chi-square random variable, χ^2_ν.

4. Compute a correlated standard normal vector, such that $= AZ$.

5. Compute correlated n dimensional Student's t, such that

$$X = \frac{Y}{\sqrt{\frac{\chi_\nu^2}{\nu}}}.$$

6. Map X back to the uniform vector, such that $U = t_\nu(X)$.

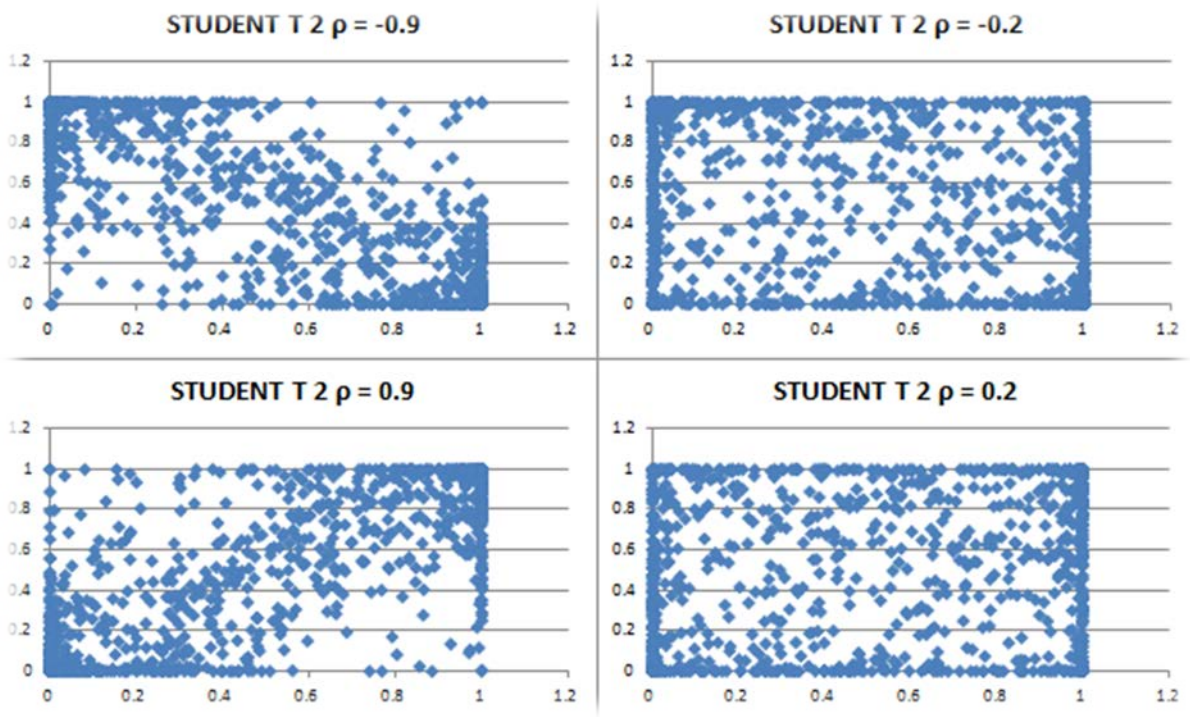

Figure 6: *Bivariate Student t copula with two degrees of freedom and different correlation inputs.*

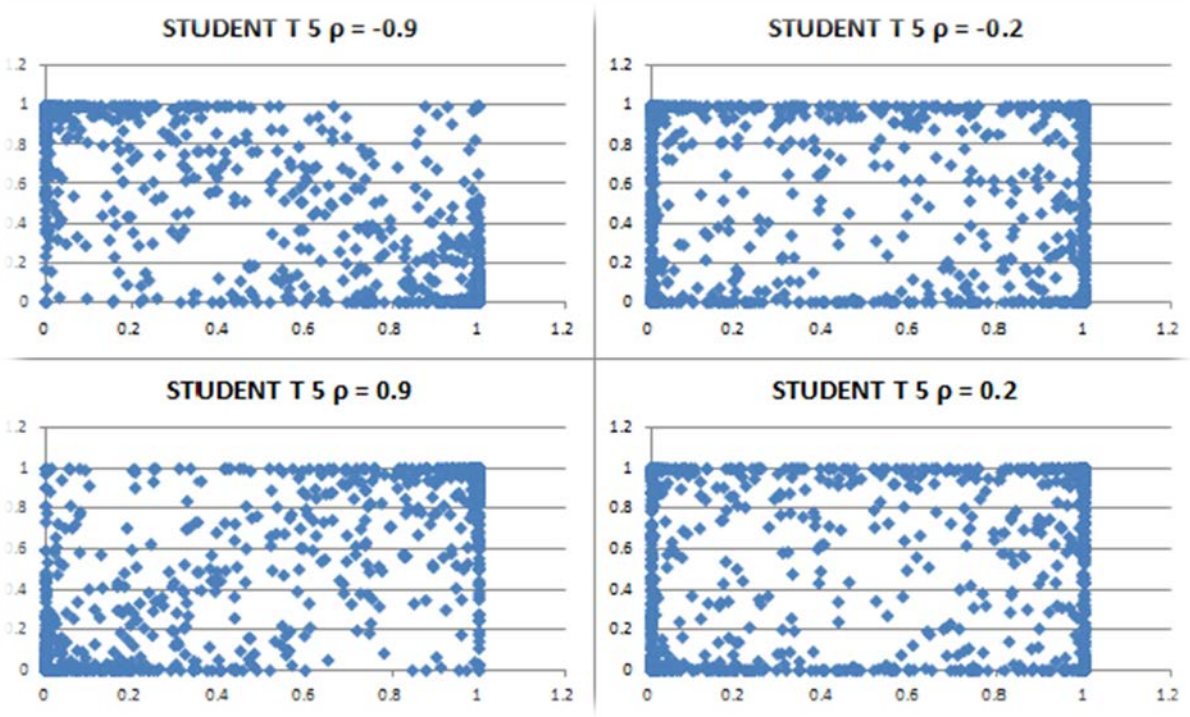

***Figure 7**: Bivariate Student t copula with five degrees of freedom and different correlation inputs.*

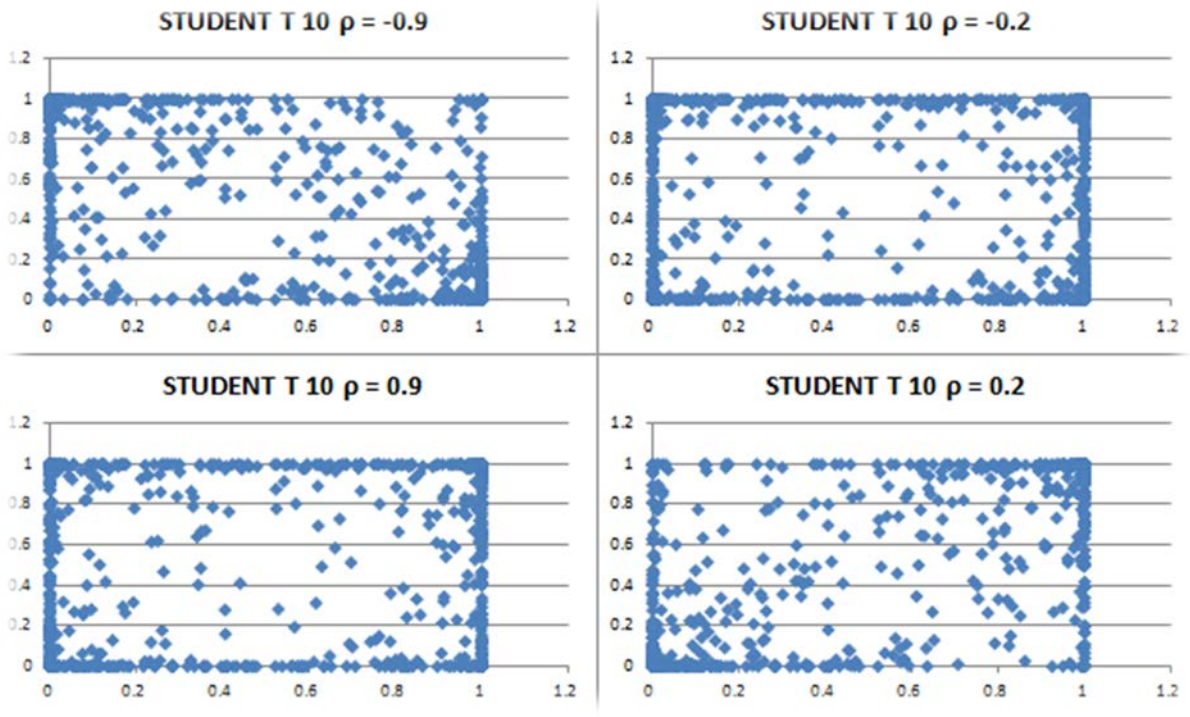

***Figure 8:** Bivariate Student t copula with ten degrees of freedom and different correlation inputs.*

Fréchet copula

The third bivariate copula that will be considered is the Fréchet copula. The Fréchet copula has two parameters, p and q such that $p, q \in [0,1]$ and $p + q \leq 1$.

Definition 4.39: Bivariate Fréchet copula (Cherubini, Luciano & Vecchiato 2004)

The bivariate Fréchet copula is given by

$$C^F(v,z) = p\ max(v+z-1,0) - (1-p-q)vz + q\ min(v,z)$$
$$= pC^- + (1-p-q)C^{\perp} + qC^+.$$

Bivariate Fréchet copulas model two risks' dependencies by means of weighting the comonotonicity, countermonotonicity and independency respectively, where p, $(1-p-q)$ and q assign weights to each dependence (Nelson 2006).

4.4.2. Archimedean copulas

The fourth class of bivariate copulas that will be considered is that of Archimedean copulas. Archimedean copulas are extensively used in actuarial science and portfolio credit risk modelling due to their analytical tractability. According to Cherubini et al. (2004), Archimedean copulas are defined based on a function ψ called a generator. This function ψ can be classified as a generator if it is from $[0,1] \to \mathbb{R}$, continuous, decreasing, convex and such that $\psi(1) = 0$. This function can further be classified as a strict generator whenever $\psi(0) = +\infty$. Furthermore, the pseudo-inverse of ψ is defined by

$$\psi^{[-1]}(u) = f(u) = \begin{cases} \psi^{-1}(u), & 0 \le u \le \psi(0) \\ 0, & \psi(0) \le u \le +\infty \end{cases}.$$

Definition 4.40: Bivariate Archimedean copula (Cherubini, Luciano & Vecchiato 2004)

Given a generator and its pseudo-inverse, a bivariate Archimedean copula is generated by the following

$$C^A(v,z) = \psi^{[-1]}(\psi(v) + \psi(z)).$$

Note that when the generator is strict, the copula can be classified as a strict Archimedean copula.

The simplest way to obtain a generator is to investigate the class of inverse Laplace transforms, as Laplace transforms always give generators. Different choices of generators will produce different types of copulas. For example, with a functional form $\psi_\alpha(u)$ (Cherubini, Luciano & Vecchiato 2004):

Definition	$\boldsymbol{\psi_\alpha(u)}$	**Range of α**	**Copula**
Gumbel	$(-u)^\alpha$	$[1,+\infty)$	$C^{A_G}(v,z) = exp\left\{-[(-\ln v)^\alpha + (-\ln z)^\alpha]^{\frac{1}{\alpha}}\right\}$
Clayton	$\frac{1}{\alpha}(u^{-\alpha} - 1)$	$[-1,0) \cup (0,+\infty)$	$C^{A_C}(v,z) = max\left[(v^{-\alpha} + z^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0\right]$

Frank	$-ln\frac{e^{-\alpha u}-1}{e^{-\alpha}-1}$	$(-\infty,0)\cup(0,+\infty)$	$C^{A_F}(v,z)=-\frac{1}{\alpha}\ln\left[1+\frac{(e^{-\alpha v}-1)(e^{-\alpha z}-1)}{e^{-\alpha}-1}\right]$

According to Malevergne and Sornette (2006), the Clayton copula behaves as a limit copula, whilst the Gumbel copula uses extreme value theory when encoding the dependence structure.

Procedure for constructing the bivariate Clayton copula

The algorithm followed for generating the Clayton copula proceeds as follows (Embrechts, Frey & McNeil 2005):

1. Generate two independent uniform random variables (Z_1, Z_2) as discussed above.
2. Set $Z_1 = U_1$ and $U_2 = \left[U_1^{-\alpha}\left(Z_2^{-\alpha/(1+\alpha)}-1\right)+1\right]^{-\frac{1}{\alpha}}$.

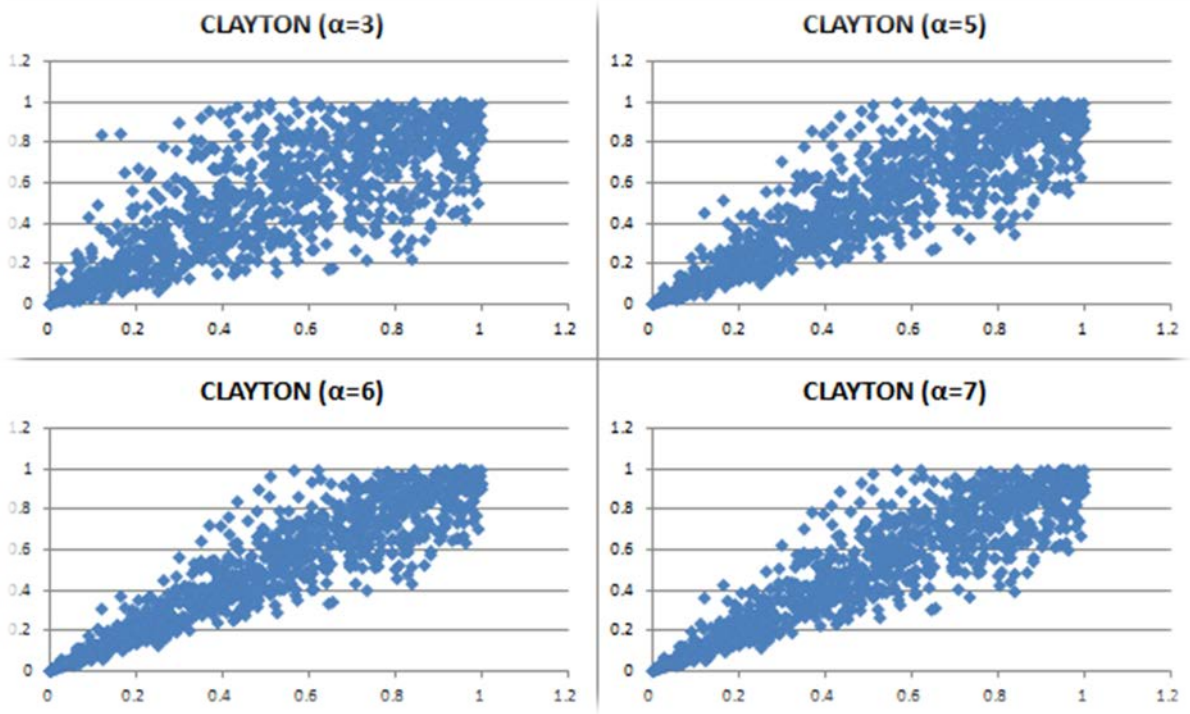

Figure 9: *Bivariate Clayton copula with different values of alpha.*

Procedure for constructing the bivariate Frank copula

The algorithm followed for generating the Frank copula proceeds as follows (Embrechts, Frey & McNeil 2005):

1. Generate two independent uniform random variables (Z_1, Z_2).

2. Set $Z_1 = U_1$ and

$$U_2 = -\frac{1}{\alpha} ln\left[1 + \frac{Z_2(1 - e^{-\alpha})}{Z_2(e^{-\alpha U_1} - 1) - e^{-\alpha U_1}}\right].$$

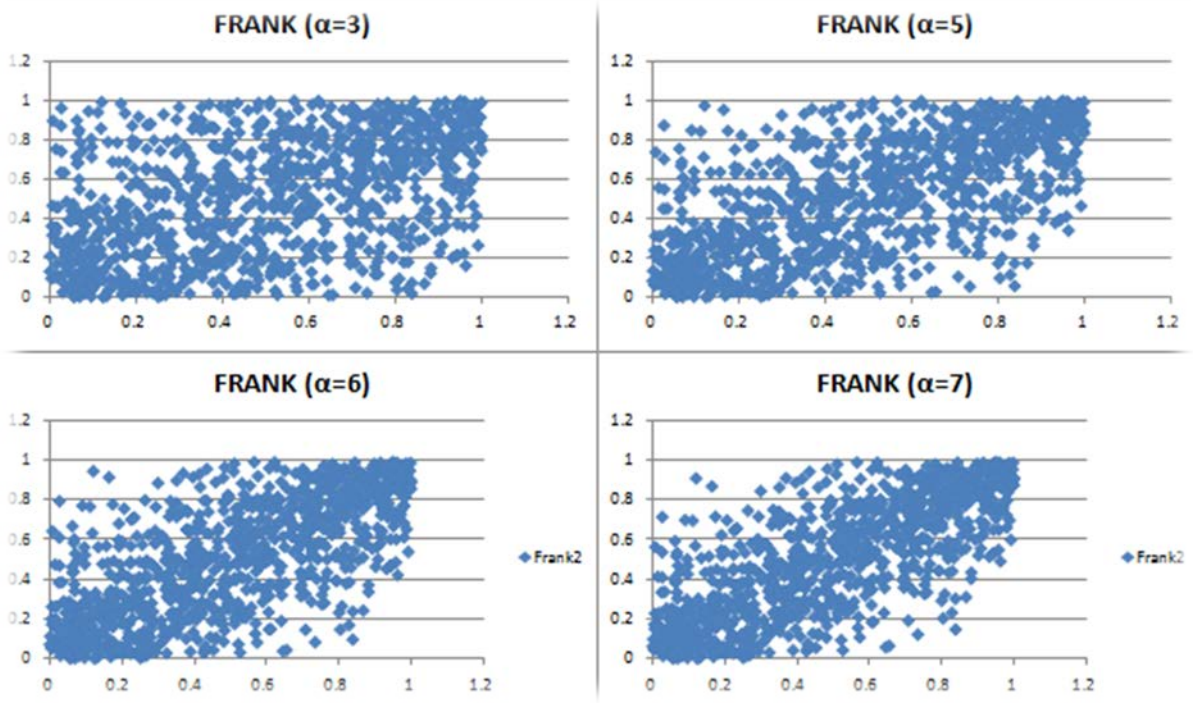

Figure 10: *Bivariate Frank copula with different values of alpha.*

Procedure for constructing the bivariate Gumbel copula

The algorithm followed for generating the Frank copula proceeds as follows ((Nelson 2006) and (Genest & Rivest 1993)):

1. Generate two independent uniform random variables (Z_1, Z_2).

2. Set $w\left(1 - \frac{\ln(w)}{\alpha}\right) = Z_2$, and solve $0 < w < 1$ numerically.

3. Set

$$U_1 = e^{\left[Z_1^{\frac{1}{\alpha}} \ln(w)\right]}$$

and

$$U_2 = e^{\left[(1-Z_1)^{\frac{1}{\alpha}} \ln(w)\right]}.$$

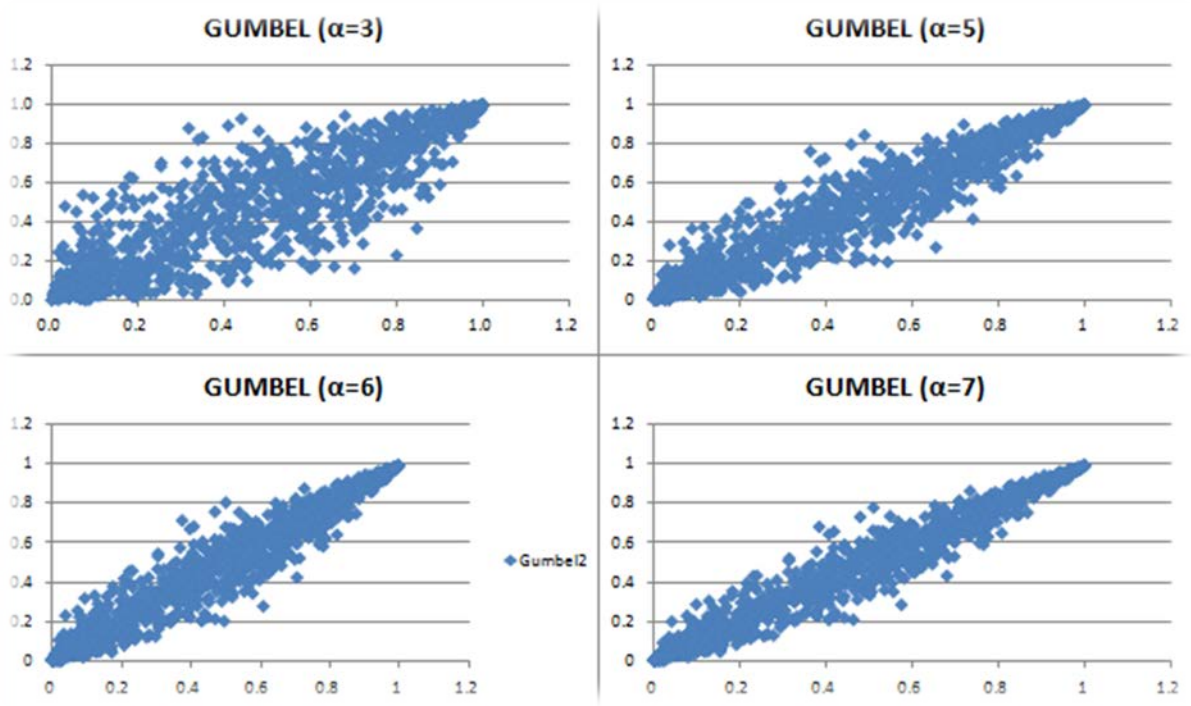

Figure 11: *Bivariate Gumbel copula with different values of alpha.*

4.5. Multivariate copulas

Having now built the foundation in terms of bivariate copulas, these ideas can easily be extended to the multivariate case. This section will provide a short overview of basic definitions and theorems in the multivariate case.

4.5.1. Preliminary definitions

The following definitions follow directly from the bivariate case as presented in section 4.4. Let $\boldsymbol{u} = (u_1, u_2, \dots, u_n)$, $\boldsymbol{v} = (v_1, v_2, \dots, v_n)$ and the *n*-dimensional box be defined as

$$A = [u_{11}, u_{12}] \times [u_{21}, u_{22}] \times \dots \times [u_{n1}, u_{n2}]$$

where $u_{i1} \leq u_{i2}$ for $i = 1,2, \dots, n$.

Definition 4.41: Grounded function (Cherubini, Luciano & Vecchiato 2004)

Let a function $G: A_1 \times A_2 \times \dots \times A_n \to \mathbb{R}$ where $A_i \subset \mathbb{R}$, for all i and where the non empty sets A_i have a least element a_i. G is said to be grounded if it is null for every vector $\boldsymbol{v}$ in its domain such that at least one of the elements $v_k = a_k$, i.e.

$$G(\boldsymbol{v}) = G(v_1, v_2, \dots, v_{k-1}, a_k, v_{k+1}, \dots, v_n) = 0\,.$$

Furthermore, a *n*-dimensional box represents the Cartesian product of n closed intervals. If any the verticals of A, denoted by μ, are in the domain of G, we can define the G-volume of A as

$$\sum_{\mu} G(\boldsymbol{\omega}) \prod_{i=1}^{n} sgn(2\omega_i - u_{i1} - u_{i2})$$

where $\boldsymbol{\omega}$ is any vertex of A.

Theorem 4.42: Grounded and n-increasing function (Cherubini, Luciano & Vecchiato 2004)

A grounded and n-increasing function $G: A_1 \times A_2 \times \ldots \times A_n \to \mathbb{R}$ is non-decreasing with respect to all its entries.

Proof: The proof of theorem 4.42 can be found in the proof of the n-dimensional version of Sklar's theorem below.

Definition 4.43: i-th one-dimensional margin (Cherubini, Luciano & Vecchiato 2004)

The i-th one-dimensional margin of the function $G: A_1 \times A_2 \times \ldots \times A_n \to \mathbb{R}$ if each $A_i \neq \emptyset$ is the function $G_i(u): A_i \to \mathbb{R}$ defined as

$$G_i(u) = G(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_{i-1}, u, \bar{a}_{i+1}, \ldots \bar{a}_n)$$

where $\bar{a}_i$ is the maximal element in A_i.

4.5.2. Subcopulas and copulas

In this section the definition of a copula and its subcopula will be extended to the multidimensional case.

Definition 4.44: Multidimensional subcopula (Cherubini, Luciano & Vecchiato 2004)

An n-dimensional subcopula is a real-valued function C defined on $A_1 \times A_2 \times \ldots \times A_n$ where $A_i \subset I$ for all i, A_i nonempty and $[0,1] \in A_i$ for all i such that:

1. $C(u_1, \ldots, u_{i-1}, 1, u_{i+1}, \ldots, u_n)$ is grounded for all i and all $u_1, u_2, \ldots, u_n$.

2. The copula's one-dimensional marginal is the identity function on I: $C_i(u) = u$ for all i.

3. The copula is n-increasing.

Definition 4.45: Multidimensional copula (Cherubini, Luciano & Vecchiato 2004)

A n-dimensional copula is a n-dimensional subcopula with $A_i = I$ for every i.

4.5.3. Sklar's theorem

Sklar's theorem in the n dimensional case will now be considered. For a distribution, F, with marginal distribution functions F_1, F_2,…, F_n, there exists a subcopula C' that couples these marginals to their joint distribution as $F(x_1, x_2, \dots, x_n) = C'\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big)$.

Theorem 4.47: Sklar's theorem in n dimensions (Cherubini, Luciano & Vecchiato 2004)

Let $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ be n marginal density functions, then for every vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$:

1. If C' is any subcopula whose domain includes the cross product of the ranges of $F_1, F_2, \dots, F_n$ it follows that
$$C'(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$$
is a joint density function with margins $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$.
2. Conversely, if $F(x)$ is a joint density function with margins $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$, then there exists a unique subcopula C' whose domain is equal to the cross product of the ranges of $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ such that
$$F(x) = C'\big(F_1(x_1), F_2(x_2), \dots, F_n(x_n)\big).$$

Moreover, if $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ are continuous, the subcopula is a copula. Otherwise, there exists a copula C such that

$$C'(u_1, u_2, \dots, u_n) = C(u_1, u_2, \dots, u_n)$$

for every $(u_1, u_2, \dots, u_n)$ in the cross product of $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$.

Proof: see (Moore & Spruill 1975, pp. 599–616).

4.5.4. Product copula and Fréchet bounds

The following theorems extend the product copula to the n-dimensional case.

Theorem 4.47: n-dimensional product copula (Cherubini, Luciano & Vecchiato 2004)

The random variables in the vector $\boldsymbol{u}$ are independent if they have the product copula defined as

$$C^{\perp}(\boldsymbol{u}) = (u_1, u_2, \dots, u_n)$$

on the cross product of ranges of F.

Proof: the proof of theorem 4.47 follows directly from Sklar's theorem.

Theorem 4.48: Bounds (Nelson 2006)

Every copula satisfies the inequality

$$max(u_1 + u_2 + \cdots + u_n - 1,0) \leq C(u) \leq min(u_1, u_2, \dots, u_n).$$

Since, the upper bound is still a copula; the definition can be extended to the maximum copula C^+ in n-dimensions. However, the lower bound is not a copula, as can be seen in the next theorem.

Theorem 4.49: Lower bound (Nelson 2006)

When $n > 2$, for every $u \in I^n$, there exists a copula C_n such that

$$C_u = max(u_1 + u_2 + \cdots + u_n - 1,0)$$

Proof: see Nelson (2006).

4.5.5. Parametric classes of multivariate copulas

In this section some of the families of copulas that were considered in the bivariate case will be extended to the multivariate case.

Gaussian copula

The Gaussian copula is used to generate a joint normal distribution from Gaussian marginal distributions.

Definition 4.50: Multivariate Gaussian copula (Cherubini, Luciano & Vecchiato 2004)

The multivariate Gaussian copula is defined as

$$C_R^G = \Phi_{\rho_{XY}}\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n)\right)$$
$$= \frac{1}{|R|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}\zeta^T(R^{-1} - I)\zeta\right\}$$

where R is the correlation matrix and $\zeta = \left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n)\right)^T$.

The family of Gaussian copulas is therefore fully parameterized by the degree of linear correlation.

Student t copula

Next the copula generated by the multivariate Student t distribution will be considered.

Definition 4.51: Multivariate Student t copula (Cherubini, Luciano & Vecchiato 2004)

The Student t copula is defined as

$$C_R^T(u_1, u_2, \dots, u_n) = |R|^{-\frac{1}{2}} \frac{\Gamma(\frac{\nu+n}{2})}{\Gamma(\frac{\nu}{2})} \left(\frac{\Gamma(\frac{\nu}{2})}{\Gamma(\frac{\nu+1}{2})} \right)^n \frac{\left(1 + \frac{1}{\nu}\zeta^T R^{-1}\zeta\right)^{-\frac{\nu+n}{2}}}{\prod_{j=1}^n \left(1 + \frac{{\zeta_j}^2}{\nu}\right)^{\frac{-\nu+1}{2}}}$$

where R is a positive definite matrix, $\zeta = \left({t_{nu}}^{-1}(u_1), {t_{nu}}^{-1}(u_2), \dots, {t_{nu}}^{-1}(u_n)\right)^T$ and ζ_j is the j-th element of the vector ζ.

Archimedean copulas

The use of Laplace transforms can help us to construct the Archimedean copula. Again these classes of copulas will be extended to the multidimensional case.

Theorem 4.52: (Cherubini, Luciano & Vecchiato 2004)

Let ψ be a strict generator. The function $C: [0,1]^n \to [0,1]$ defined by

$$C^A(\boldsymbol{u}) = \psi^{[-1]}(\psi(u_1) + \psi(u_2) + \cdots + \psi(u_n))$$

is a copula if ψ^{-1} is completely monotonic on $[0, +\infty]$.

Proof: See (Kimberling 1974, pp. 152-164)

Again, the Archimedean copulas are defined as follows in the multidimensional case (Cherubini, Luciano & Vecchiato 2004):

Definition	$\boldsymbol{\psi_\alpha(u)}$	**Range of α**	**Copula**
Gumbel	$(-\ln(u))^\alpha$	$\alpha > 1$	$C^G(\boldsymbol{u}) = exp\left\{-\left[\sum_{i=1}^n (-\ln u_i)^\alpha\right]^{-\frac{1}{\alpha}}\right\}$
Clayton	$u^{-\alpha} - 1$	$\alpha > 0$	$C^C(\boldsymbol{u}) = \left[\sum_{i=1}^n {u_i}^{-\alpha} - n + 1\right]^{-\frac{1}{\alpha}}$
Frank	$ln\left(\frac{exp(-\alpha u) - 1}{exp(-\alpha) - 1}\right)$	$\alpha > 0$ *when* $n \geq 3$	$C^{Fr}(\boldsymbol{u}) = -\frac{1}{\alpha} ln\left\{1 + \frac{\prod_{i=1}^n (e^{-\alpha u_i} - 1)}{(e^{-\alpha} - 1)^{n-1}}\right\}$

For a comprehensive list of Archimedean copulas and their attributes refer to Nelson (2006).

5. Fitting copulas to multivariate data

Until now the foundations have been laid in terms of regulatory capital requirements, risk measures, measures of dependence as well as copulas. This section aims to implement these ideas within a multivariate framework. Various copulas will be fitted to multivariate data in order to illustrate the functional relationship encoded within a dependence structure of the marginal distributions of several random variables. In other words, if the marginal distributions are known and the measure of dependence has been chosen, one can go beyond correlation when measuring the risk of co-movement that exists within an organization with multiple business lines.

Section 5.1 provides a brief discussion regarding the sample data used during the analysis. Section 5.2 considers various measures of dependence both from a theoretical and practical point of view. Section 5.3 aims to illustrate how business line volatilities can be estimated using the GARCH(1,1) scheme. Finally, after deciding on a desired correlation structure as well as estimating business line volatilities, a comparison will be drawn between capital estimates obtained using the Gaussian, Student t (with various degrees of freedom), Clayton and Cauchy copulas under VaR, ETL and StressVaR.

5.1. Sample data and assumptions

Due to a lack of loss data, share price data was used to illustrate the co-movements that could exist between various business lines. It should be noted that in using the AMA approach, internal loss data is of fundamental importance as this will be a direct input into the capital model.

Share price data for eight shares listed on the Johannesburg Stock Exchange was chosen for the analysis. The historical period between January 2000 and January 2012 was considered. Thus, the chosen data included both the 2001 and 2008 stress periods. The shares that were included in the analysis are:

- Anglo American PLC (AGL)
- Anglo American Platinum Corporation Ltd. (AMS)
- Aspen Pharmacare Holdings (APN)
- Discovery Holdings Ltd. (DSY)
- Standard Bank Group Ltd. (SBK)
- Mr Price Group Ltd. (MPC)
- MTN Group Ltd. (MTN)
- Pretoria Portland Cement (PPC)

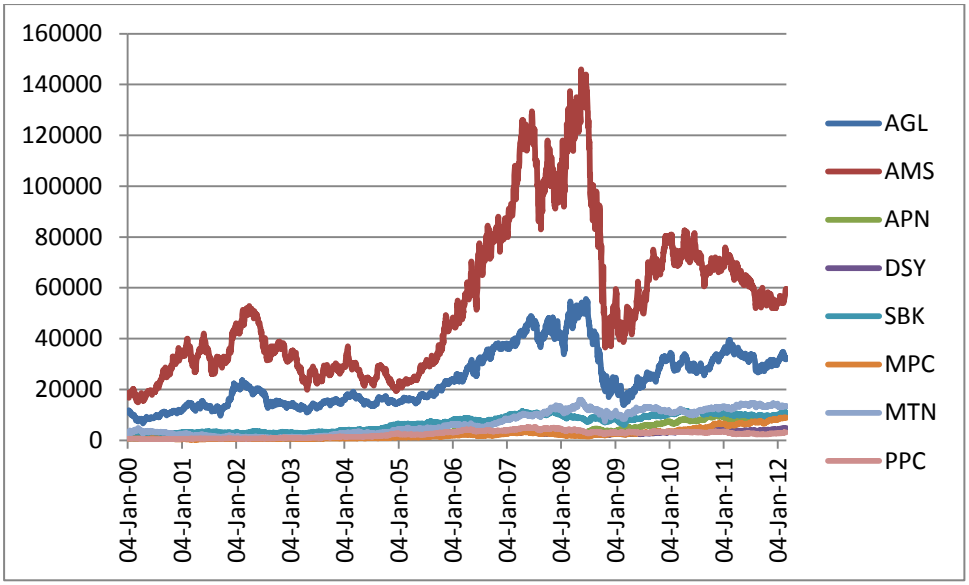

***Figure 12**: Share price data from January 2000 to January 2012 for the 8 companies included in the analysis.*

From a capital analyst's perspective the choice of data provides some interesting initial considerations. Firstly, what dependence structure existed during the South African bull market pre 2007? Secondly, what dependence structure existed during the stock market crash in 2007/08? Thirdly, were there any significant changes in these dependence structures after this stock market crash? Finally, how stable were these relationships in the first place?

Other questions that could be considered include how these relationships were impacted as a result of being exposed to the same interest rate environment, currency exposure as well as levels of inflation. Furthermore, to what extent were these companies exposed to changes in the international macroeconomic environment. Additionally, one would expect a high level of correlation between shares like DSY and APN due to their medical origin (similarly for AGL and AMS). Also, was diversification the best means of avoiding catastrophic losses and do correlations tend to one during market crashes? Finally, do some companies offer stable if not spectacular returns during any business cycle?

5.2. Measuring dependence

Bouchaud and Potters (2004, p. 91) state: "*… different stocks can have completely different prices, and therefore unrelated absolute daily price changes, but rather similar daily returns*". Thus, the raw share price data first had to be transformed into the daily log returns over the 12 year period under examination. This was done in order to strip out any drift present in the raw stock price data.

Consider a R20 drop in the share price in February 2012; this would have only represented a 3.4% drop in the AMS share price, compared to a 22.4% drop in the MPC share price. This scaling enables us to consider relative price changes instead of absolute price changes.

Let δS represent the actual price change between two intervals, separated by the time interval τ, then (Bouchaud & Potters 2004)

$$\delta S_i = S_{i+1} - S_i = S(t+\tau) - S(t),$$

where $t \equiv u\tau$. Furthermore, let u represent the relative price change (return) over the same period, such that

$$u_i = \frac{\delta S_i}{S_i} \approx \log S_{i+1} - \log S_i.$$

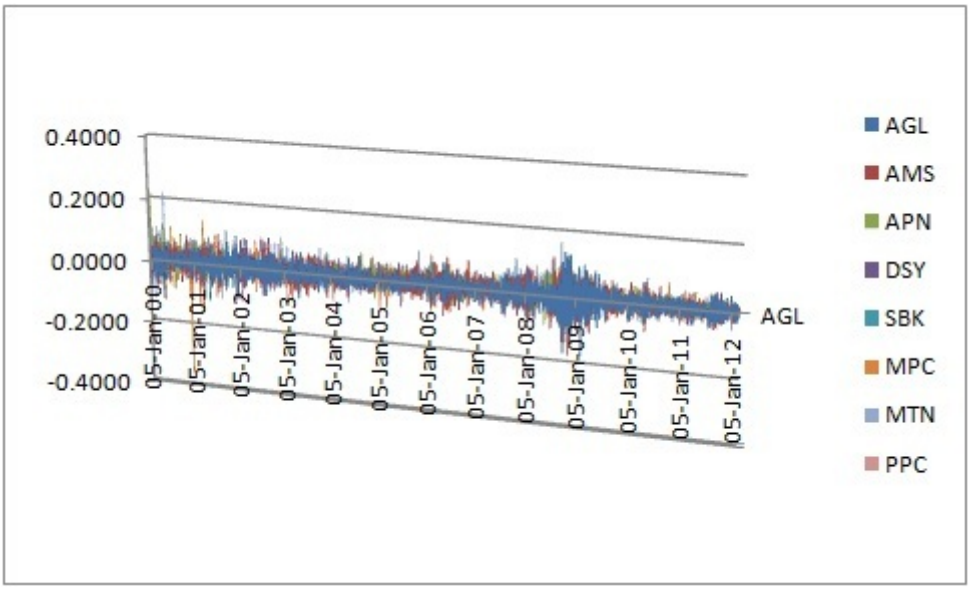

Figure 13: *Daily returns per share from January 2000 to January 2012.*

Embrechts et al. (2005) present a set of "*stylized facts*" of financial time series data. These stylized facts consist of empirical observations and inferences observed from a series of daily price changes, such as relative changes in equity, currency or commodity prices. These stylized facts are:

a) Even though return series show minimal evidence of serial correlation, return series are not iid.
b) Squared return series show significant serial correlation.
c) Conditional expected returns tend to zero.
d) Volatility changes over time (see figure 13).
e) Extreme returns appear to cluster (see figure 13). This is also referred to as volatility clustering.
f) Return series are fat tailed (see figure 14).

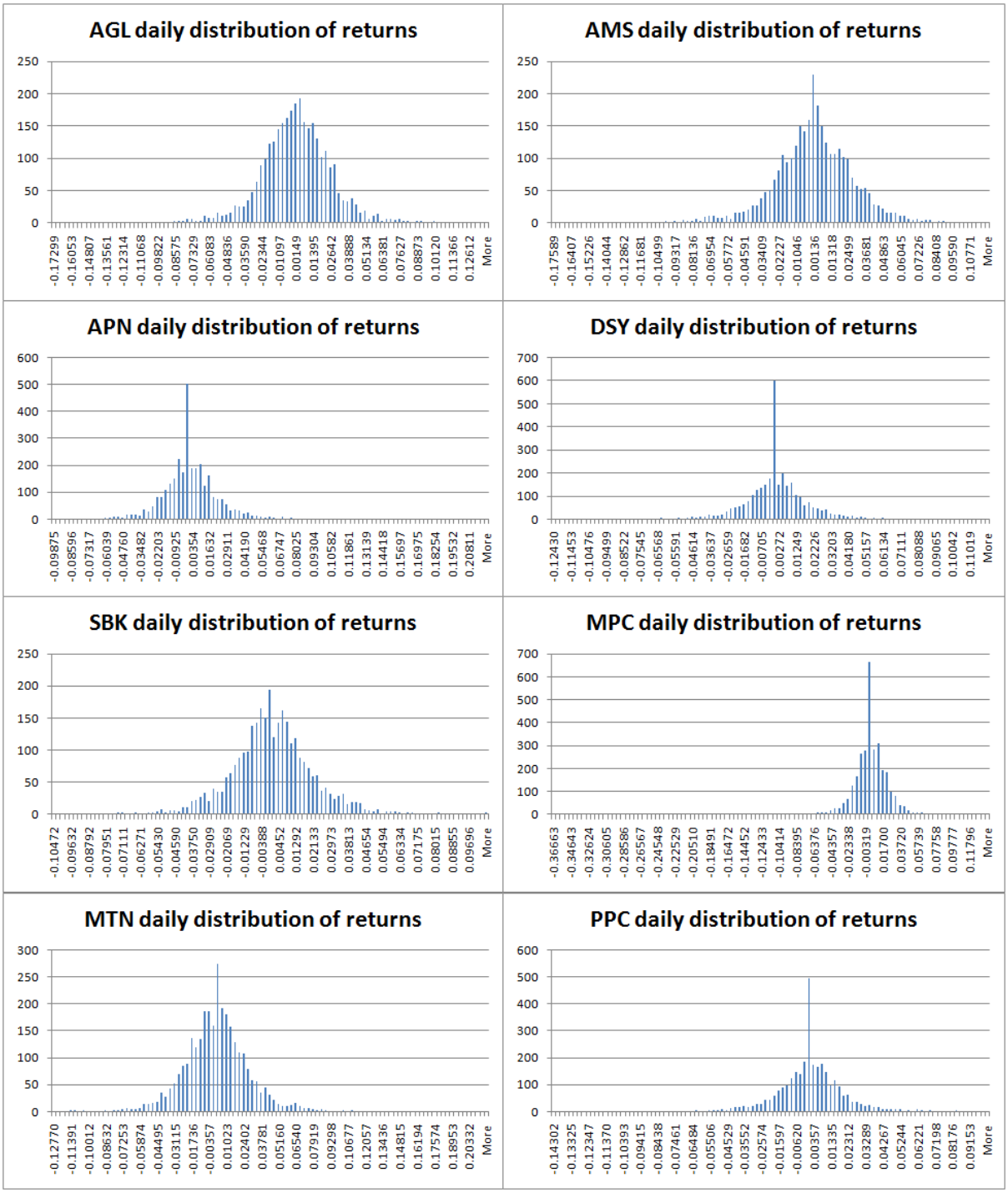

Figure 14: *Distribution of daily returns*

Figure 14 shows the shape of the daily distribution of relative returns for the eight stocks under examination. Bouchaud and Potters (2004) suggest fitting the daily distribution of price returns using a truncated Lévy distribution (TLD) or a Student t distribution. As the tails of these distributions are much broader than Gaussian, the asymptotic tails of the TLD and the power-law tails of the Student t distribution provide pleasing results.

These daily returns were then used to obtain 12-year linear correlation, Spearman's rho correlation and Kendall's tau correlation matrices. These matrices were obtained using the "corr" function in Matlab R2009b:

$$LinearCorrelation = corr(ReturnData),$$

$$SpearmanRankCorrelation = corr(ReturnData,' type',' spearman'),$$

$$KendallRankCorrelation = corr(ReturnData,' type',' kendall')$$

where

$$ReturnData = [u_{AGL}, u_{AMS}, u_{APN}, u_{DSY}, u_{SBK}, u_{MPC}, u_{MTN}, u_{PPC}]$$

and u_{AGL} represents the12-year relative price changes of AGL, u_{AMS} represents the 12-year relative price changes of AMS, and so on.

The below 12-year correlation matrices could now be seen as a reflection of the long-term relationships that exist between these companies.

Linear Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.50257	0.204833	0.181577	0.329005	0.201803	0.272443	0.243232
AMS	0.50257	1	0.166538	0.165808	0.274254	0.136741	0.275932	0.213213
APN	0.204833	0.166538	1	0.134372	0.273011	0.17918	0.220967	0.206553
DSY	0.181577	0.165808	0.134372	1	0.233402	0.181613	0.207043	0.165405
SBK	0.329005	0.274254	0.273011	0.233402	1	0.25241	0.41869	0.280476
MPC	0.201803	0.136741	0.17918	0.181613	0.25241	1	0.214498	0.210427
MTN	0.272443	0.275932	0.220967	0.207043	0.41869	0.214498	1	0.222142
PPC	0.243232	0.213213	0.206553	0.165405	0.280476	0.210427	0.222142	1

Table 1: *12-year linear correlation matrix.*

Spearman's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.483156	0.197011	0.185823	0.294862	0.218995	0.252131	0.191418
AMS	0.483156	1	0.16179	0.162267	0.25378	0.151653	0.256745	0.168972
APN	0.197011	0.16179	1	0.139936	0.243777	0.180126	0.192546	0.173547
DSY	0.185823	0.162267	0.139936	1	0.228576	0.202264	0.184321	0.161205
SBK	0.294862	0.25378	0.243777	0.228576	1	0.235503	0.389442	0.250102
MPC	0.218995	0.151653	0.180126	0.202264	0.235503	1	0.20379	0.202638
MTN	0.252131	0.256745	0.192546	0.184321	0.389442	0.20379	1	0.215808
PPC	0.191418	0.168972	0.173547	0.161205	0.250102	0.202638	0.215808	1

Table 2: *12-year Spearman's Rank Correlation matrix.*

Kendall's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.342077	0.13503	0.12824	0.204565	0.150347	0.172536	0.131282
AMS	0.342077	1	0.110922	0.111087	0.174888	0.104419	0.176878	0.115583
APN	0.13503	0.110922	1	0.0963	0.167788	0.124	0.131283	0.119304
DSY	0.12824	0.111087	0.0963	1	0.158355	0.13998	0.126395	0.111386
SBK	0.204565	0.174888	0.167788	0.158355	1	0.161988	0.270057	0.172044
MPC	0.150347	0.104419	0.124	0.13998	0.161988	1	0.139136	0.140305
MTN	0.172536	0.176878	0.131283	0.126395	0.270057	0.139136	1	0.147613
PPC	0.131282	0.115583	0.119304	0.111386	0.172044	0.140305	0.147613	1

Table 3: *12-year Kendall's Rank Correlation matrix.*

These results raise a few questions. Firstly, is a 12-year correlation matrix applicable when capital is allocated over a shorter period of time? Secondly, how stable are these correlations? Thirdly, do these correlations provide an accurate reflection of the risks that an institution might face during a black swan event? Finally, how applicable is this long-term relationship within the current macroeconomic environment, especially in the aftermath of the 2008 Credit Crunch?

In order to answer the first two questions, rolling period correlations were considered. This was done in order to establish how stable these correlations were over time. Figure 15 illustrates the results of this analysis for rolling period linear correlations with AGL. The time horizons used were nine years, six years, three years and one year.

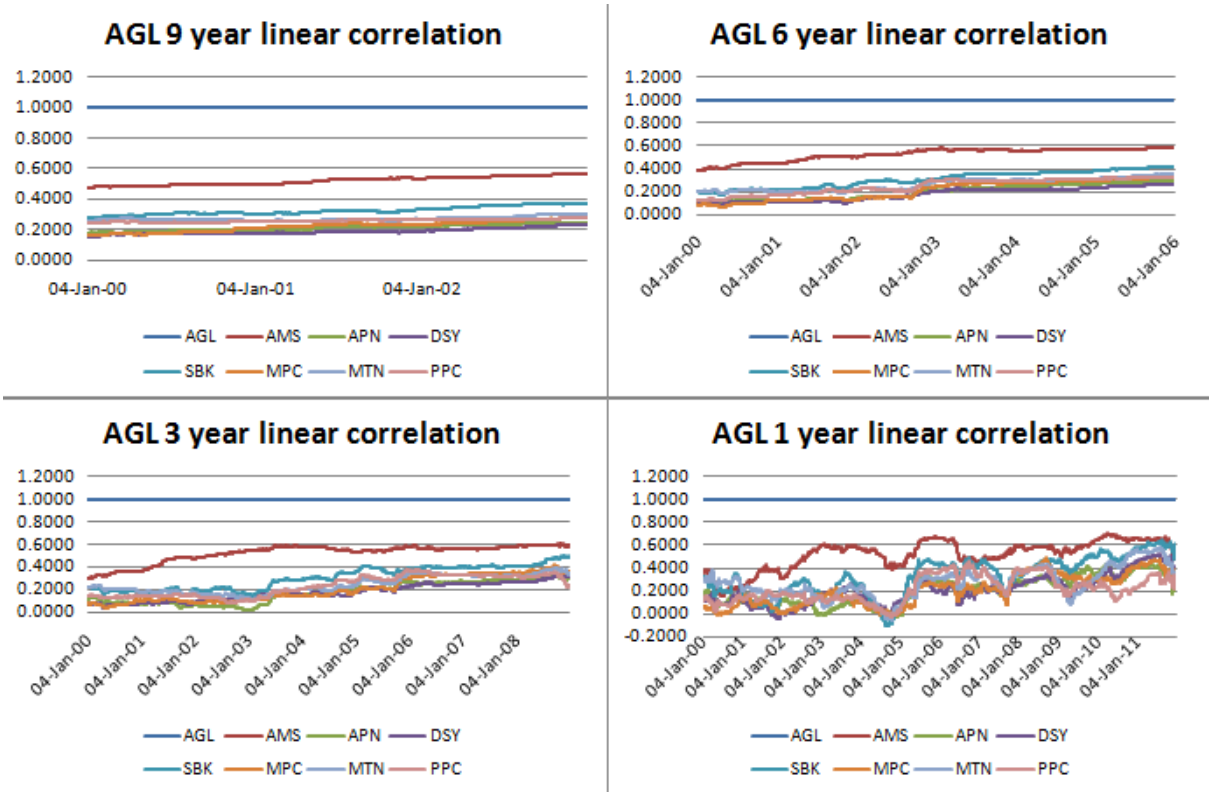

Figure 15: *Comparison of AGL linear correlations over different time horizons.*

Firstly, the results clearly indicate that correlations change over time. Secondly, correlations are most unstable when shorter time periods are measured. It would thus make sense to apply a shorter time horizon when estimating correlations, as this would provide a more accurate reflection of the current co-movement between business lines.

As mentioned above, a third question has to be answered before correlations can be used in determining capital requirements, namely how applicable are current correlation estimates with regards to extreme events? In order to provide an answer to this question, it is of great importance to have a deep understanding of what is truly meant by association risk.

The simplest rationalization of association risk can be seen at the two extremities, namely positive and negative correlation. In order to develop an understanding of the risks that an organization faces at positive correlation, consider the following scenario: Our current organization consists of eight business lines; if these business lines' returns are positively correlated the organization would expect to realize massive profits during good years, as a profit in one business line would be accompanied by profits in others. However, the converse would also be true during bad years, as losses in one business line would likely be followed by losses within other business lines. Positive

correlation thus greatly enhances the likelihood of major losses during difficult economic circumstances and during an extreme event it almost guarantees that an organization will go bust.

As mentioned above, an organization also faces risks at negative correlation. It should however be noted that exposure to negative correlation is more subtle. In order to illustrate this principle, consider the following scenario: Our current organization has eight business lines, if these business units are negatively correlated with one another we would expect to obtain some profits during good years, but certainly also some losses. This would also be true during years that are less profitable.

Furthermore, to illustrate the intricacy of risk at a negative correlation, consider a second scenario: Until now it has been assumed that our organization has an equal exposure in every business unit, however in reality this might not be the case. If our current organization for instance generates 70% of its returns within one business unit, it would not only require all the other business units to be negatively correlated with this main business unit, but it would also require an extremely high level of positive correlation between these other business units in order to sufficiently offset losses within the main business unit.

In other words, diversification benefit is not only dependent on the degree of correlation between business lines, but it is also a function of the capital that was allocated to a business line in the first place. For the rest of this dissertation it will however be assumed that each business line has been allocated an equal amount of capital.

The final question that was raised above was how relevant are long term relationships within the current economic circumstances. In order to answer this question, three additional scenarios were considered, namely:

- The one year period where correlation has been at its maximum for the given eight companies (19 March 2010 until 11 March 2011).
- The one year period where correlation has been at its minimum for the eight companies (20 August 2004 until 15 August 2005).
- The current 12 months correlation for the eight companies (28 February 2011 until 29 February 2012).

For a comprehensive review on scenario analysis and generation see Ziemba and Ziemba (2008).

Linear Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.50257	0.204833	0.181577	0.329005	0.201803	0.272443	0.243232
AMS	0.50257	1	0.166538	0.165808	0.274254	0.136741	0.275932	0.213213
APN	0.204833	0.166538	1	0.134372	0.273011	0.17918	0.220967	0.206553
DSY	0.181577	0.165808	0.134372	1	0.233402	0.181613	0.207043	0.165405
SBK	0.329005	0.274254	0.273011	0.233402	1	0.25241	0.41869	0.280476
MPC	0.201803	0.136741	0.17918	0.181613	0.25241	1	0.214498	0.210427
MTN	0.272443	0.275932	0.220967	0.207043	0.41869	0.214498	1	0.222142
PPC	0.243232	0.213213	0.206553	0.165405	0.280476	0.210427	0.222142	1

Max Linear Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.694043	0.430669	0.366793	0.547694	0.311189	0.417558	0.23258
AMS	0.694043	1	0.443302	0.403724	0.586954	0.383859	0.555259	0.265249
APN	0.430669	0.443302	1	0.395988	0.529199	0.428706	0.442907	0.441285
DSY	0.366793	0.403724	0.395988	1	0.332248	0.336791	0.356302	0.348829
SBK	0.547694	0.586954	0.529199	0.332248	1	0.390716	0.505292	0.315412
MPC	0.311189	0.383859	0.428706	0.336791	0.390716	1	0.380474	0.383056
MTN	0.417558	0.555259	0.442907	0.356302	0.505292	0.380474	1	0.343636
PPC	0.23258	0.265249	0.441285	0.348829	0.315412	0.383056	0.343636	1

Min Linear Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.410199	-0.00718	-0.01314	-0.09287	0.011223	-0.03414	0.003414
AMS	0.410199	1	0.025195	0.00491	-0.06039	-0.06406	-0.05376	-0.06805
APN	-0.00718	0.025195	1	0.000133	0.099699	0.003163	0.168347	0.032534
DSY	-0.01314	0.00491	0.000133	1	0.178264	0.120385	0.189611	-0.0745
SBK	-0.09287	-0.06039	0.099699	0.178264	1	0.166063	0.288304	0.172854
MPC	0.011223	-0.06406	0.003163	0.120385	0.166063	1	0.075102	0.149852
MTN	-0.03414	-0.05376	0.168347	0.189611	0.288304	0.075102	1	0.061341
PPC	0.003414	-0.06805	0.032534	-0.0745	0.172854	0.149852	0.061341	1

Current Linear Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.645746	0.418311	0.481531	0.587983	0.446165	0.53304	0.258992
AMS	0.645746	1	0.372133	0.362806	0.508281	0.361006	0.495547	0.213645
APN	0.418311	0.372133	1	0.371652	0.499474	0.309198	0.470872	0.114163
DSY	0.481531	0.362806	0.371652	1	0.476042	0.454868	0.454496	0.266804
SBK	0.587983	0.508281	0.499474	0.476042	1	0.334452	0.630643	0.166619
MPC	0.446165	0.361006	0.309198	0.454868	0.334452	1	0.40661	0.214726
MTN	0.53304	0.495547	0.470872	0.454496	0.630643	0.40661	1	0.114154
PPC	0.258992	0.213645	0.114163	0.266804	0.166619	0.214726	0.114154	1

Table 4: *12-year, maximum, minimum and current linear correlation matrices.*

Spearman's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.483156	0.197011	0.185823	0.294862	0.218995	0.252131	0.191418
AMS	0.483156	1	0.16179	0.162267	0.25378	0.151653	0.256745	0.168972
APN	0.197011	0.16179	1	0.139936	0.243777	0.180126	0.192546	0.173547
DSY	0.185823	0.162267	0.139936	1	0.228576	0.202264	0.184321	0.161205
SBK	0.294862	0.25378	0.243777	0.228576	1	0.235503	0.389442	0.250102
MPC	0.218995	0.151653	0.180126	0.202264	0.235503	1	0.20379	0.202638
MTN	0.252131	0.256745	0.192546	0.184321	0.389442	0.20379	1	0.215808
PPC	0.191418	0.168972	0.173547	0.161205	0.250102	0.202638	0.215808	1

Max Spearman's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.681257	0.379009	0.293474	0.52567	0.276257	0.394543	0.236747
AMS	0.681257	1	0.346688	0.340513	0.533985	0.339141	0.493878	0.237265
APN	0.379009	0.346688	1	0.330278	0.451756	0.388943	0.401058	0.422339
DSY	0.293474	0.340513	0.330278	1	0.308547	0.331789	0.3267	0.307241
SBK	0.52567	0.533985	0.451756	0.308547	1	0.371123	0.490528	0.306255
MPC	0.276257	0.339141	0.388943	0.331789	0.371123	1	0.367395	0.392436
MTN	0.394543	0.493878	0.401058	0.3267	0.490528	0.367395	1	0.314072
PPC	0.236747	0.237265	0.422339	0.307241	0.306255	0.392436	0.314072	1

Min Spearman's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.355903	-0.04108	0.009374	-0.111	0.044756	-0.04772	0.024379
AMS	0.355903	1	0.043776	-0.01591	-0.07447	-0.03706	-0.04087	-0.06685
APN	-0.04108	0.043776	1	-0.00207	0.069434	0.006066	0.119522	0.043657
DSY	0.009374	-0.01591	-0.00207	1	0.185051	0.109664	0.206646	-0.04014
SBK	-0.111	-0.07447	0.069434	0.185051	1	0.13613	0.268142	0.160474
MPC	0.044756	-0.03706	0.006066	0.109664	0.13613	1	0.051171	0.189472
MTN	-0.04772	-0.04087	0.119522	0.206646	0.268142	0.051171	1	0.065454
PPC	0.024379	-0.06685	0.043657	-0.04014	0.160474	0.189472	0.065454	1

Current Spearman's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.627992	0.352428	0.460601	0.563938	0.429189	0.490723	0.256876
AMS	0.627992	1	0.367491	0.353883	0.526628	0.324229	0.463571	0.216844
APN	0.352428	0.367491	1	0.350694	0.426112	0.279782	0.445631	0.120421
DSY	0.460601	0.353883	0.350694	1	0.470596	0.40497	0.423959	0.256277
SBK	0.563938	0.526628	0.426112	0.470596	1	0.288181	0.594014	0.193011
MPC	0.429189	0.324229	0.279782	0.40497	0.288181	1	0.378656	0.200251
MTN	0.490723	0.463571	0.445631	0.423959	0.594014	0.378656	1	0.125473
PPC	0.256876	0.216844	0.120421	0.256277	0.193011	0.200251	0.125473	1

Table 5: *12-year, maximum, minimum and current Spearman's Rank correlation matrices.*

Kendall's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.342077	0.13503	0.12824	0.204565	0.150347	0.172536	0.131282
AMS	0.342077	1	0.110922	0.111087	0.174888	0.104419	0.176878	0.115583
APN	0.13503	0.110922	1	0.0963	0.167788	0.124	0.131283	0.119304
DSY	0.12824	0.111087	0.0963	1	0.158355	0.13998	0.126395	0.111386
SBK	0.204565	0.174888	0.167788	0.158355	1	0.161988	0.270057	0.172044
MPC	0.150347	0.104419	0.124	0.13998	0.161988	1	0.139136	0.140305
MTN	0.172536	0.176878	0.131283	0.126395	0.270057	0.139136	1	0.147613
PPC	0.131282	0.115583	0.119304	0.111386	0.172044	0.140305	0.147613	1

Max Kendall's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.496007	0.26362	0.203785	0.372464	0.189198	0.271213	0.158421
AMS	0.496007	1	0.239461	0.235314	0.381457	0.234457	0.34065	0.163097
APN	0.26362	0.239461	1	0.227479	0.312946	0.266939	0.270047	0.287141
DSY	0.203785	0.235314	0.227479	1	0.212093	0.228403	0.224072	0.215885
SBK	0.372464	0.381457	0.312946	0.212093	1	0.255788	0.346548	0.207483
MPC	0.189198	0.234457	0.266939	0.228403	0.255788	1	0.25247	0.271926
MTN	0.271213	0.34065	0.270047	0.224072	0.346548	0.25247	1	0.217182
PPC	0.158421	0.163097	0.287141	0.215885	0.207483	0.271926	0.217182	1

Min Kendall's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.245366	-0.02612	0.007403	-0.0793	0.028727	-0.02972	0.017146
AMS	0.245366	1	0.029273	-0.01152	-0.04969	-0.02464	-0.02256	-0.04583
APN	-0.02612	0.029273	1	-0.00047	0.045971	0.005632	0.08263	0.031313
DSY	0.007403	-0.01152	-0.00047	1	0.126153	0.076962	0.145899	-0.02546
SBK	-0.0793	-0.04969	0.045971	0.126153	1	0.093669	0.183661	0.109072
MPC	0.028727	-0.02464	0.005632	0.076962	0.093669	1	0.035511	0.131936
MTN	-0.02972	-0.02256	0.08263	0.145899	0.183661	0.035511	1	0.044681
PPC	0.017146	-0.04583	0.031313	-0.02546	0.109072	0.131936	0.044681	1

Current Kendall's Rank Correlation								
	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
AGL	1	0.446453	0.245166	0.31891	0.402455	0.294274	0.343237	0.174314
AMS	0.446453	1	0.255472	0.237514	0.365088	0.227215	0.320144	0.147108
APN	0.245166	0.255472	1	0.241076	0.298726	0.186669	0.30779	0.081634
DSY	0.31891	0.237514	0.241076	1	0.320888	0.285236	0.293274	0.171264
SBK	0.402455	0.365088	0.298726	0.320888	1	0.198415	0.420256	0.130632
MPC	0.294274	0.227215	0.186669	0.285236	0.198415	1	0.259715	0.13658
MTN	0.343237	0.320144	0.30779	0.293274	0.420256	0.259715	1	0.085486
PPC	0.174314	0.147108	0.081634	0.171264	0.130632	0.13658	0.085486	1

Table 6: *12-year, maximum, minimum and current Kendall's Rank correlation matrices.*

From this analysis a couple of observations can be made. Firstly, there is a significant difference between the maximum and minimum correlation matrices. These differences are even more evident when using the Spearman's and Kendall's rank correlation matrices. Furthermore, for all three cases, the 12-year correlation matrix appears to be a fair representation of the long term relationships that exists between these eight business lines. This is illustrated by the fact that the 12-year correlation falls between the maximum and minimum correlations in all three cases. When considering the current correlation matrix, it is interesting to note that the current correlations are close to the highest levels of correlations that have been realized over the last 12 years.

From a capital analyst's perspective, the following logical conclusions can be drawn from this analysis. Firstly, when allocating capital, it will be done under the maximum correlation as this provides a fair representation of our current macroeconomic environment. This will be compared to the 12 year correlation matrices and the minimum correlation matrices, in order to establish what the effect of a decrease in correlations would have on the current capital estimates.

5.3. Estimating business line volatilities

After determining the dependence structures that exist between business lines, the next logical step was to estimate the volatility of business line returns. The first assumption that was made was that business line returns follow the Markov property. According to Wilmott (2006, p. 73) the Markov property holds if "*the distribution of the value of the random variable* S_i *conditional upon all of the past events only depends on the previous value* S_{i-1}*"*. In other words, business line returns have no memory beyond where it is now. Business line returns thus satisfy the following stochastic differential equation, known as a geometric Brownian motion

$$dS = \mu S dt + \sigma S \, dX \, .$$

According to Bouchaud and Potters (2004) one can summarize a geometric Brownian motion as follows:

a) Relative returns are assumed to be iid random variables.
b) The price process is a continuous time process, as it is assumed that the time scale tends to zero.
c) The process is scale invariant, in other words the process' statistical properties do not depend on the chosen time scale.

However, according to Alexander (2008), the iid assumption is not realistic in practice as the volatility of the returns of financial time series data changes over time (see figure 13). Because there

will be periods where volatility is extremely high, as well as periods where volatility is atypically low, the geometric Brownian motion must capture the effects of volatility clustering of returns.

5.3.1. The GARCH(1,1) scheme

The exponentially weighted moving average (EWMA) model is a commonly used method for estimating volatilities (Alexander 2008)

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

where $\lambda\sigma_{n-1}^2$ represents the persistence in volatility, $(1-\lambda)u_{n-1}^2$ represents the intensity of reaction of volatility to market events and $\lambda \in (0,1)$.

The EWMA model thus models the volatility as a weighted average between the previous estimate of volatility and the most recent return. Sinclair (2008, p. 33) makes the following observation when referring to the EWMA model: "*This method has the virtues of being simple to use and understand. It has the drawback of being a stupid solution*". This comment was aimed at this model's smoothing effect on jumps in volatility when large absolute returns occur.

Due to financial returns not being independently, identically or normally distributed, more practitioners use GARCH models to estimate volatilities of financial returns. According to Alexander (2008) the GARCH volatility forecasts capture volatility clustering, unlike the forecasts from moving average models that only represent current estimates. In other words GARCH volatility forecasts provide volatility estimates that can be greater or smaller than the average over the short term. Hull (2008) also states that GARCH(1,1) model is "*theoretically more appealing*" than the EWMA since it includes mean reversion. According to Alexander (2008, p. 131) it should however be noted that "*as the forecast horizon increases the GARCH volatility forecasts converge to the long term volatility*".

According to Alexander (2001) it is sufficient to use the GARCH(1,1) scheme when estimating the steady state long-term volatility. Thus, the technique that was applied when estimating business line volatilities was the GARCH$(1,1)$ scheme, that has just one lagged error square and one autoregressive term. The derivation of the GARCH(1,1) scheme is as follows (Hull 2008):

Consider a simple m period moving average, where σ_n is the volatility of returns on day n. The volatility of returns on day n can initially be expressed as

$$\sigma_n^2 = \frac{1}{m-1}\sum_{i=1}^{m}(u_{n-i} - \bar{u})^2 .$$

Now let $m-1 \approx m$ and $\bar{u} \approx 0$ since a one day mean return is negligible compared to the standard deviation of changes, such that

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}^2.$$

However, the above expression assigns equal weights to every u_i. In order to obtain a more realistic estimate for business line volatilities, one should rather assign greater weightings to the more recent u_is. In other words, let $\frac{1}{m} = \alpha_i$, where α_i is known as the GARCH error coefficient. Now, let $\sum_{i=1}^{m} \alpha_i = 1$ and choose $\alpha_i < \alpha_j$ when $i > j$.

Since volatility is a mean reverting process (Javaheri 2005), it tends to vary about a long term mean, $\bar{\sigma}$. In order to now incorporate this mean reversion around the long term mean into our model, the above expression can be rewritten as

$$\sigma_n^2 = \gamma\bar{\sigma}^2 + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \,,$$

where $\bar{\sigma}$ is the long term volatility rate, γ is the weight assigned to the long term volatility rate and all the weights must sum to one

$$\gamma + \sum_{i=1}^{m} \alpha_i = 1 \,.$$

In order to simplify the above expression, let $\omega = \gamma\bar{\sigma}^2$, where ω is referred to as the GARCH constant, such that

$$\sigma_n^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{n-i}^2.$$

Since the GARCH(1,1) is equivalent to the infinite ARCH model with exponentially decaying weights, as one moves back in time, α_i will decrease exponentially , thus

$$\alpha_{i+1} = \lambda\alpha_i$$

where $0 < \lambda < 1$.

Now

$$\sigma_n^2 = \sum_{i=1}^{\infty} \alpha_i u_{n-i}^2$$

where

$$\alpha_2 = \lambda\alpha_1$$
$$\alpha_3 = \lambda\alpha_2 = \lambda^2\alpha_1$$
$$\vdots \qquad \vdots \qquad \vdots.$$

Since

$$\sigma_{n-1}^2 = \sum_{i=1}^{\infty} \alpha_i u_{n-1-i}^2$$

one can rewrite the above equation as

$$\lambda\sigma_{n-1}^2 = \lambda\alpha_1 u_{n-2}^2 + \lambda^2\alpha_1 u_{n-3}^2 + \lambda^3\alpha_1 u_{n-4}^2 + \cdots$$

or

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + \alpha_1 u_{n-1}^2.$$

Since all $\sum_{i=1}^{m} \alpha_i = 1$, it is trivial that

$$\alpha_1(1 + \lambda + \lambda^2 + \lambda^3 + \cdots) = 1$$

and for an infinite series

$$(1 + \lambda + \lambda^2 + \lambda^3 + \cdots) = (1 - \lambda)^{-1}$$

such that

$$\alpha_1 = 1 - \lambda.$$

This translates into

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1 - \lambda)u_{n-1}^2.$$

In order to obtain the GARCH(1,1) scheme, we can generalize the above equation and add a long term volatility, $\gamma\bar{\sigma}^2$, such that

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2.$$

The above equation is subject to the following constraints

$$\gamma + \alpha + \beta = 1$$
$$\bar{\sigma}^2 = \frac{\omega}{1 - \alpha - \beta}$$
$$\alpha + \beta < 1.$$

These constraints must hold in order to ensure that the long term steady state variance, $\bar{\sigma}^2$, remains non-negative.

5.3.2. Estimating the parameters

In order to estimate business line volatilities using the GARCH(1,1) scheme, one first has to find the values of ω, β and α. The Maximum Likelihood Estimation (MLE) was used to estimate these parameters. It is important to distinguish between probability and likelihood. According to Sinclair (2008) probability refers to the chance of a future event whilst likelihood references past events. Given a set of data, this approach backs out values for the parameters in order to maximize the likelihood of the observed data occurring (Hull 2008). The MLE function can now be defined as

$$l(\theta; x_1, x_2, x_3, \dots, x_n) = f(x_1, x_2, x_3, \dots, x_n; \theta)$$

where $x_1, x_2, x_3, \dots, x_n$ are n iid pieces of data with probability density function $f(x_1, x_2, x_3, \dots, x_n; \theta)$ and θ unknown parameter(s). Furthermore, the log-likelihood function can be defined as

$$L(\theta; x_1, x_2, x_3, \dots, x_n) = \log l\,(\theta; x_1, x_2, x_3, \dots, x_n)$$

where the maximum likelihood estimate of the parameter(s) θ can be obtained by maximizing $L(\theta, x_1, x_2, x_3, \dots, x_n)$.

Assume that $X = x_1, x_2, x_3, \dots, x_n$ is a normally distributed random sample of iid observations, where $X \sim N(\mu, \sigma^2)$. In order to find the maximum likelihood estimators μ and σ^2 the log-likelihood function must be maximized

$$f(x_1, x_2, x_3, \dots, x_n; \mu, \sigma) = f(x_1; \mu, \sigma).f(x_2; \mu, \sigma) \dots .f(x_n; \mu, \sigma)$$

$$l(\mu, \sigma; x_1, x_2, x_3, \dots, x_n) = f(x_1; \mu, \sigma).f(x_2; \mu, \sigma) \dots f(x_n; \mu, \sigma)$$

$$\therefore L(\mu, \sigma; x_1, x_2, x_3, \dots, x_n) = \log l\,(\mu, \sigma; x_1, x_2, x_3, \dots, x_n)$$

$$= logf(x_1; \mu, \sigma) + logf(x_2; \mu, \sigma) + \cdots + logf(x_n; \mu, \sigma)$$

$$= \sum_{i=1}^{n} logf(x_i; \mu, \sigma)\,.$$

For the normal distribution

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

such that

$$L(\mu, \sigma; x_1, x_2, x_3, \dots, x_n) = \sum_{i=1}^{n} log\left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right].$$

Table 7 shows the results that were obtained using the GARCH(1, 1) scheme, as explained above, through the Maximum Likelihood Estimator. The size of α and β control the short-run dynamics of the resulting volatility time series. According to Alexander (2001), it is a common practice to

estimate the lag and error coefficients around 0.8 and 0.2 respectively when using financial data. Larger values of β indicate that shocks will tend to last longer, i.e. volatility is continual. On the other hand, larger values of α indicate that the volatility is influenced by market movements. In other words, α indicates how quickly volatility will react to news in the market, while β reflects how long the reaction is likely to last. Thus, when α is higher than 0.2 and β lower than 0.8, spikes in the volatility are more likely to occur (Alexander 2001). In contrast, when β is higher than 0.8 and α is lower than 0.2, less spikes will occur in the volatility, but the levels of volatility will be sustained over longer periods of time.

	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC
Longrun Var ($\bar{\sigma}^2$)	0.000671389	0.000744	0.000422263	0.000320982	0.000399599	0.00045313	0.000700657	0.000363309
Omega (ω)	7.06447E-06	1.53E-05	1.21654E-05	2.36517E-06	1.05019E-05	1.61573E-05	1.3678E-05	1.90659E-05
Alpha (α)	0.060487038	0.088267	0.106158193	0.052231075	0.089837117	0.090677907	0.089046985	0.08138537
Beta (β)	0.92899078	0.891185	0.86503188	0.940400386	0.883881887	0.873664929	0.89143131	0.866136236
Likelihood (L)	7144.698885	6932.396	7787.366796	8163.744535	7799.814629	7662.925846	7050.768602	7936.581256
Contraint	0.989477818	0.979452	0.971190072	0.992631461	0.973719004	0.964342835	0.980478295	0.947521606

Table 7: *Optimized constrained values and long-term variance obtained using the Maximum Likelihood Estimation (MLE) and GARCH(1,1) scheme.*

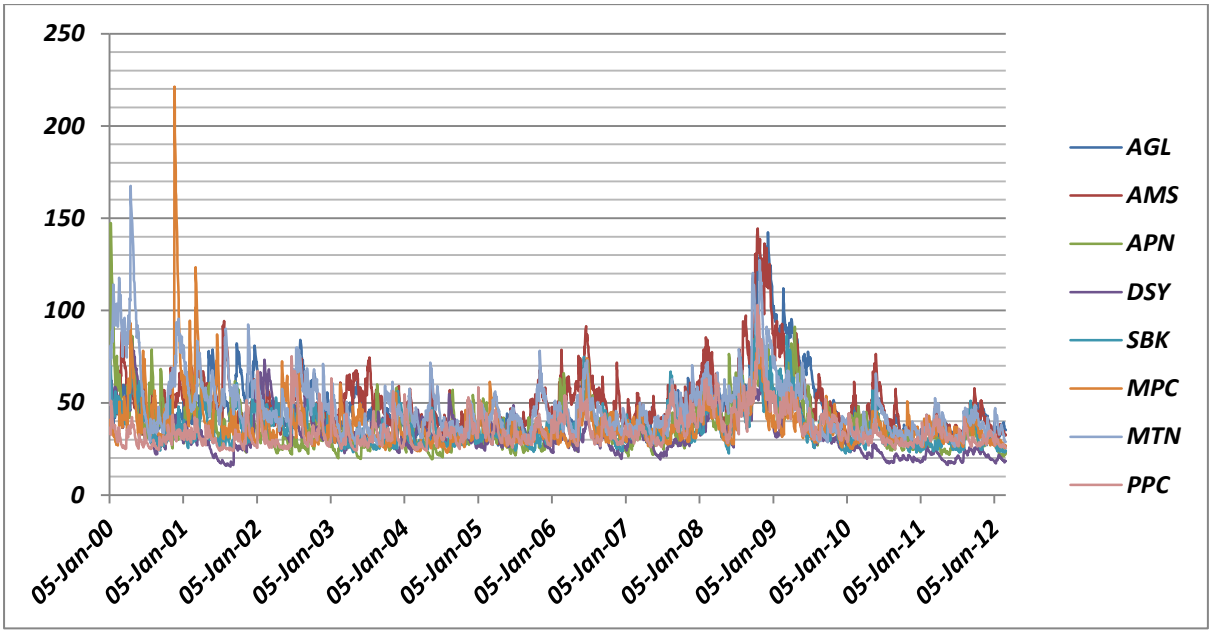

Figure 16: *GARCH(1,1) annualized volatilities.*

From figure 16 it is clear that volatility was characterized by short-term spikes from 2000 up until 2008; in other words volatility had a high error coefficient. However, since the 2008 Credit Crises the volatility levels have been much more sustainable irrespective of volatility being at high or low

levels. This would indicate a higher lag coefficient; this is also supported by the volatility estimates provided in table 7.

GARCH volatility models are simple to estimate and have robust coefficients that can be logically interpreted in terms of long-term volatilities as well as short-run dynamics. Limitations of these models include that of all three the parameters, especially ω are sensitive to the data used. Long-term volatility forecast will especially be influenced if the historical data that is used includes extreme events.

Malevergne and Sornette (2006, p. 108) state: "*When the volatility follows ARCH and GARCH processes, then the asset returns are also elliptically distributed with fat-tailed marginal distributions*". These volatilities can now be used to transform marginal distributions into joint distributions in order to estimate capital requirements.

5.4. Simulating business line losses using copulas

When dealing with heterogeneous risk factors, there seldom exists a good multivariate model. According to McNeil et al. (2005), such a model must be able to effectively describe both the marginal behavior and the existing dependence structure. Having now estimated the business line volatilities, this section aims to estimate capital requirements through simulating multivariate financial losses by Monte-Carlo simulation, using both elliptical and Archimedean copulas.

5.4.1. Multivariate copula calibration algorithms

Malevergne and Sornette (2006, p. 120) state, "*An important practical application of copulas consists in the simulation of random variables with prescribed margins and various dependence structures in order to perform Monte Carlo studies, to generate scenarios for stress-testing investigations or to analyze the sensitivity of portfolio allocations to various parameters*". This section provides simulation algorithms for the Gaussian, Student t and Clayton copulas that will be used in the next section.

According to Cherubini et al. (2004) the general method for simulating multivariate copulas is as follows:

a) Let $C_i = C(F_1, F_2, \dots, F_i, 1,1, \dots,1)$ for $i = 2, \dots, n$.

b) Draw F_1 from the uniform distribution $U(0,1)$.

c) Draw F_2 from $C_2(F_2|F_1)$.

d) Thus, in general, draw F_n from $C_2(F_n|F_1, F_2, \dots, F_{n-1})$.

Gaussian multivariate

The algorithm followed for generating the Gaussian copula with correlation matrix Σ proceeds as follows (Embrechts, Frey & McNeil 2005):

a) Find the Cholesky decomposition A from Σ, such that $\Sigma = AA'$, where A is a lower-triangular matrix.

b) Draw a n-dimensional independent standard normal vector $= (Z_1, Z_2, \dots, Z_n)'$.

c) Let $\boldsymbol{X} = \boldsymbol{AZ}$ to obtain correlated normal vector.

Student t multivariate

The algorithm followed for generating the Student t copula with correlation matrix Σ proceeds as follows (Embrechts, Frey & McNeil 2005):

a) Find the Cholesky decomposition A from Σ, such that $\Sigma = AA'$.

b) Draw a n-dimensional independent standard normal vector $= (Z_1, Z_2, \dots, Z_n)'$.

c) Draw an independent Chi-square random variable χ^2_ν.

d) Compute correlated standard normal vector $= \boldsymbol{AZ}$ **.**

e) Compute correlated n dimensional Student's

$$t = \frac{\boldsymbol{Y}}{\sqrt{\frac{\chi^2_\nu}{\nu}}}.$$

f) Map X back to uniform vector by $= t_{\boldsymbol{\nu}}(\boldsymbol{X})$.

All copulas simulated up to now, belongs to the family of elliptical copulas. Malevergne and Sornette (2006) state that the simplicity of simulating this family of copulas is one of the many appeals of using these copulas.

Clayton multivariate

The algorithm followed for generating the Clayton copula with correlation matrix Σ proceeds as follows (Cherubini, Luciano & Vecchiato 2004):

a) Draw a n-dimensional independent random vector $\boldsymbol{Z} = (Z_1, Z_2, \dots, Z_n)'$.

b) Set $U_1 = Z_1$.

c) For $n = i + 1, i = 1, \dots n$ let $U_n = \left\{(U_1^{-\alpha} + U_2^{-\alpha} + \cdots + U_3^{-\alpha} - n + 2) \cdot \left(Z_n^{\frac{\alpha}{\alpha(1-n)-1}}\right) + 1\right\}^{-\frac{1}{\alpha}}$.

5.4.2. Simulation of business line losses

This section will now illustrate how business line losses can be simulated through Monte Carlo simulation using the volatilities that were estimated in section 5.3 and the multivariate copula algorithms as presented in section 5.4.1.

As mentioned before, it was assumed that business line returns follow a geometric Brownian motion. One could thus simulate a single business line's returns by using a lognormal random walk

$$\delta S = rS\delta t + \sigma S\sqrt{\delta t}\emptyset\,,$$

where $\emptyset$ is drawn from a standard normal distribution. However, when simulating business line returns for multiple business lines one has to make use of correlated random walks. One can easily extend the lognormal random walk to the multidimensional case (this method is commonly used in practice in order to price basket options by Monte Carlo simulation)

$$\delta S_i = rS_i\delta t + \sigma_i S_i\sqrt{\delta t}\emptyset_i\,,$$

where S_i is the price of the i-th asset and σ_i volatility of the i-th asset. However, it is important to note that all the $\emptyset_i$s (known as random shocks) are now correlated, thus

$$E[\emptyset_i\emptyset_j] = \rho_{ij}.$$

It was now assumed that the organization's current value equaled the sum of the current spot prices of the eight shares that were considered in the analysis. It was also assumed that the organization owned one share of each of the abovementioned shares.

Date	AGL	AMS	APN	DSY	SBK	MPC	MTN	PPC	Total Value
02-Jan-12	31413	54000	9888	4371	10100	8250	13961	2780	134763

Table 8: *Summary of the organization's value on 2 January 2012.*

In order to simulate what the organization's value would be at the end of the year, the above copula calibration algorithms were performed in order to compute the correlated random shocks. These correlated random shocks were then used in simulating the correlated random walk one year forward. The profit or loss that was realized within a business line was determined by subtracting the original business line value from the newly simulated business line value. The change in the organization's value was determined by adding these profits and losses.

In order to calculate the organization's capital requirements, the above steps were repeated multiple times. All losses and profits obtained were then ranked from smallest to largest. In order to compute the 95% VaR and 99% VaR, one had to consider the 95th and 99th percentile respectively.

In order to compute the 95% ETL and 99% ETL, one had to consider the average of the sum of losses greater than the 95th and 99th percentile respectively. The StressVaR was computed by multiplying the 95% VaR estimate by five.

Figure 17 shows a comparison of the results obtained when using the current linear correlation matrix as correlation input for the Gaussian, Cauchy, Student t (with various degrees of freedom) and the Clayton copulas.

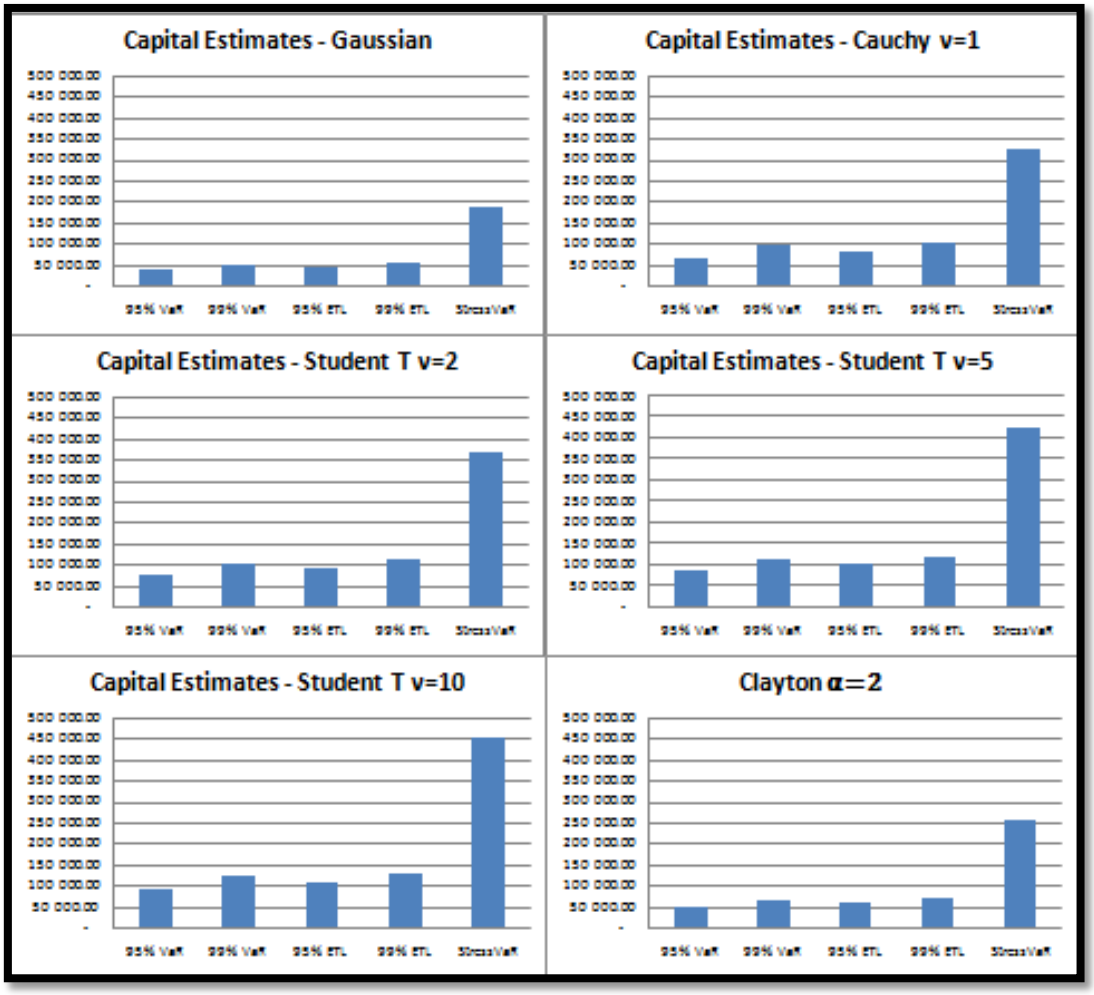

Figure 17: *Capital estimates obtained by simulations using Gaussian copula, Cauchy copula, Student t copula and Clayton copula using the current linear correlation matrix as correlation input.*

In every case, the Gaussian copula provides the smallest capital estimate regardless of the risk measure used. This reflects the fact that the Gaussian copula does not account for fat tails. Thus, when using the Gaussian copula in allocating capital, the organization could be exposed given the occurrence of an extreme event.

The Clayton copula provides the next smallest capital estimates. The Clayton copula has lower tail dependence but no upper tail dependence like the Gaussian copula. However, increasing the value of α would result in an increase in the value of the capital estimates.

The Cauchy and Student t copulas have lower and upper tail dependence. The Cauchy copula, which is just the Student t copula with one degree of freedom, has lower capital estimates than the Student t copulas with higher degrees of freedom, that is, as the degrees of freedom increases, so does the tail dependence. Even with one degree of freedom, the Cauchy/Student t still has higher capital estimates than the Gaussian and Clayton copulas.

In general, copulas provide a better alternative to linear correlation, as it extends the dependence to nonlinear cases (Chernobai, Rachev & Fabozzi 2007). In operational risk management, upper tail dependence is of utter importance. We can thus conclude that the Student t and Cauchy copulas will provide an organization with a better buffer against catastrophic events when compared to the Gaussian and Clayton copulas. When using the Student t copula, the degrees of freedom will reflect an organization's risk appetite.

In figure 18 it can be seen that the difference between 95% VaR and 99% VaR is much higher when compared to the difference between the 95% ETL and the 99% ETL. This indicates that ETL is a more suitable measure for tail risk and low probability events.

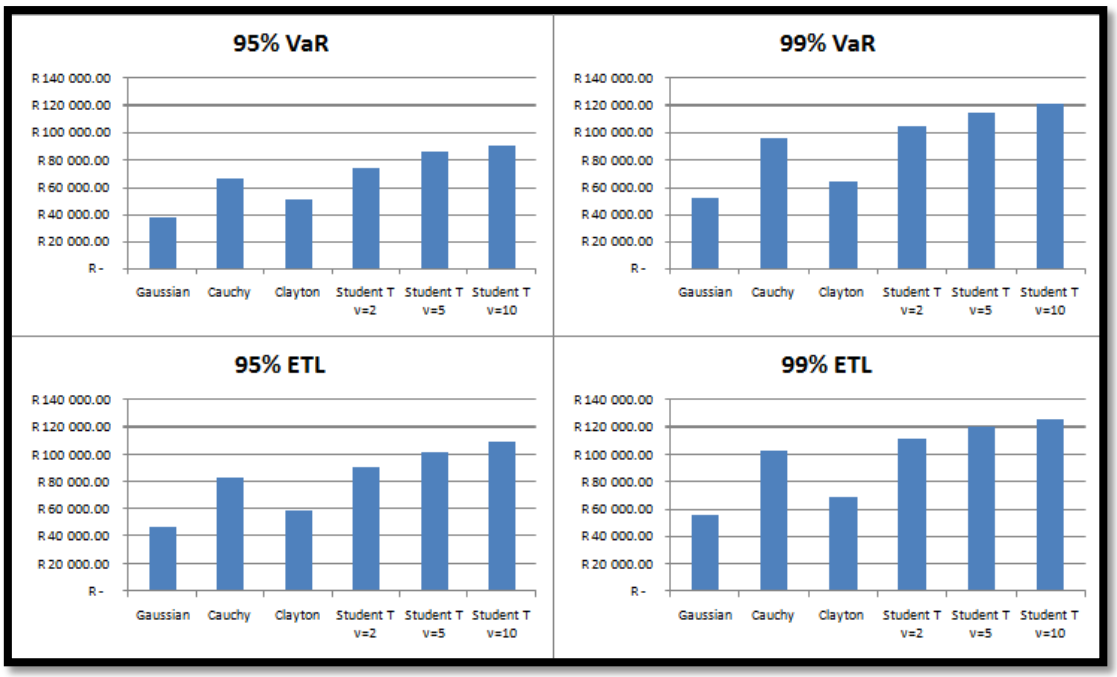

Figure 18: *Comparison of capital estimates provided by different risk measures using the Gaussian, Cauchy, Clayton and Student t copulas.*

In figure 19 it can be seen that the StressVaR estimates are much higher than all other risk measures provided above. It is always important to understand that although banks are the protectors of deposits, they are still in a risk and return business. A fundamental question that banks have answer is how much capital to hold. Too little could lead to bankruptcy, while too much would lead to inefficiencies and opportunity costs.

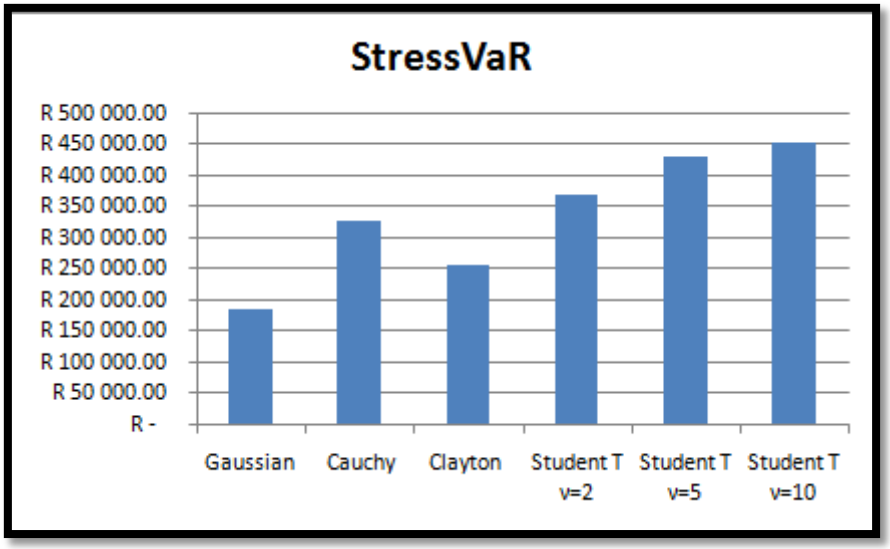

Figure 19: *Comparison of capital estimates provided by StressVaR using the Gaussian, Cauchy, Clayton and Student t copulas.*

Having now obtained a good understanding of the effect on capital estimates when using various copulas as well as different risk measures, the next step was to investigate what the effect on capital estimates would be when using different correlation inputs. This was done by computing the current linear correlation, Spearman's correlation and Kendall's correlation matrices as inputs to the above copula calibration algorithms.

The results obtained during this analysis can be seen in figure 20. This shows that the effect of using different correlation measures as correlation inputs into the copula calibration algorithms, have a smaller impact on the capital estimates than the choice of copula or risk measure.

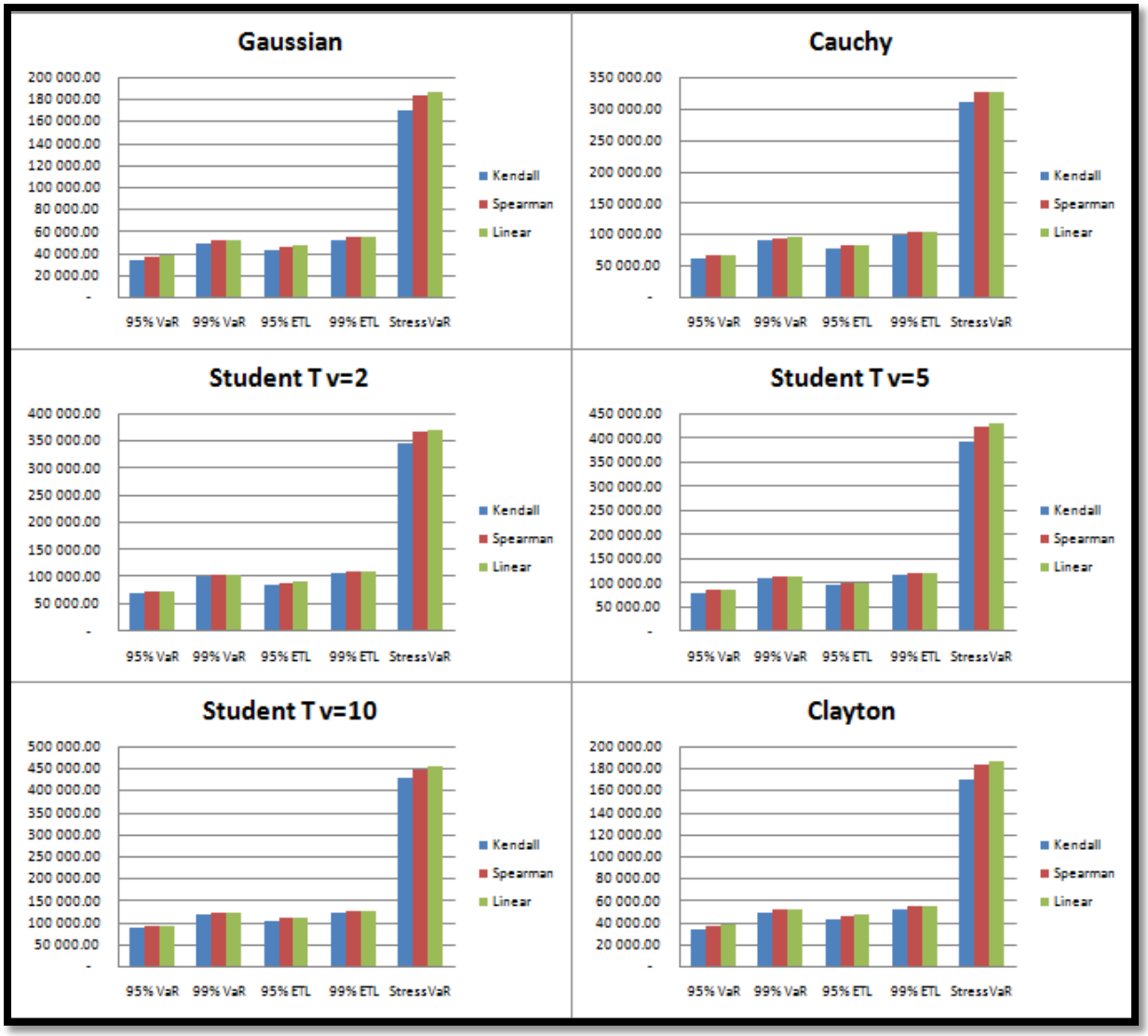

Figure 20: *Comparison of capital estimates obtained when using the current Kendall rank correlation matrix, current Spearman rank correlation matrix and the current linear correlation matrix.*

The next step was to investigate what the effect would be on capital estimates when using different inputs for correlation. This was done by using the minimum, current and maximum linear correlations as illustrated in table 4 in section 5.2.

The results obtained for this analysis can be seen in figure 21. Firstly, this illustrates that lower correlation indicates more diversification benefit among business lines and a consequent saving in capital. Secondly, the converse also holds as higher correlation indicates a greater risk of collective losses and a consequent higher capital charge. Thirdly, since correlations increase during extreme events, this also indicates the need for choosing conservative correlation inputs when determining capital requirements. Finally, the capital estimates obtained by using the current linear correlation matrix as input into the copula calibration algorithms seem to provide a fair capital charge for normal everyday business.

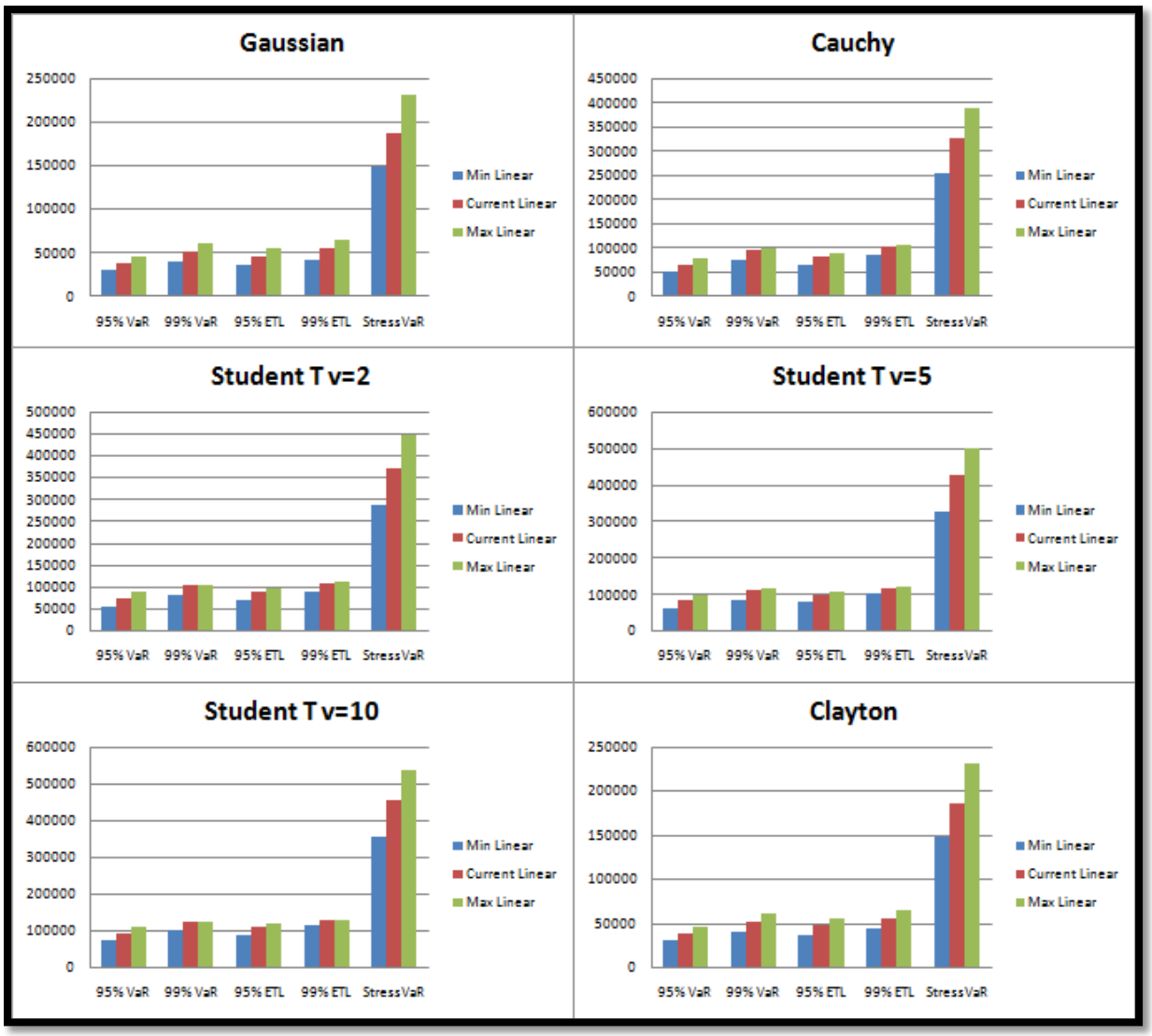

Figure 21: *Comparison of capital estimates obtained using the minimum linear correlation matrix, current linear correlation matrix and maximum linear correlation matrix.*

It can thus be concluded that the capital estimates provided are a function of the correlation input into the copula calibration algorithm, the selected risk measure and the choice of copula. In order to safeguard a bank, StressVaR would provide the most comfort to depositors, although it is a very capital inefficient risk measure. It could thus make more sense to use a conservative coherent risk measure like 99% ETL along with a stressed correlation input as well as a copula that has upper tail dependence like the Cauchy copula or Student t copula. Furthermore, by stressing the business line volatility, one could increase the capital charges even further.

6. Conclusion

In this dissertation, risk management techniques under the Basel II Accord were considered. The main finding was that financial risk models possessed numerous weaknesses. As a response to these weaknesses, the Basel III Accord proposed numerous additional regulations that will provide increasing solidity in financial markets.

The principle of risk based regulation under the Basel Accord has received much criticism and so have the measures that it uses. VaR was critiqued for its misinterpretation, its failure to use stress periods in historical VaR estimates, its inability to incorporate the effects of market liquidity and the fact that it is non-sub additive. As a result, coherent risk measures were introduced. ES possesses some enhanced properties, namely its ability to provide insight into the severity of tail events, its coherence property and the fact that it is less sensitive to changes in the confidence level. Another enhanced risk measure that was studied was StressVaR.

Three fundamental measures of dependence were considered, namely linear correlation, rank correlation and copulas. Even though easy to manipulate, dependence cannot be distinguished on the grounds of linear correlation alone. Moreover, failing to aggregate losses within an organization will lead to an overestimation of capital requirements.

Rank correlation proved to be invariant subject to non-linear monotonic transformations and invariant to the choice of marginal distributions. Copulas on the other hand extend the nature of dependence to the nonlinear case. Copulas are a popular technique to model joint multi-dimensional problems and the wide choice of dependence structures makes copula functions more attractive than the other measures of dependence.

In this dissertation, comparisons of capital estimates using different correlation inputs, risk measures and copulas were provided. The copulas used in this analysis were the Gaussian copula, the Cauchy copula, the Student t copula and the Clayton copula. Risk measures that were evaluated were VaR, ETL and StressVaR. The different correlation inputs that were considered included linear correlation, Spearman's rank correlation and Kendall's rank correlation. Finally, capital estimates were compared under stressed correlations, current correlations and relaxed correlations.

The first key observation of this dissertation was that the choice of copula has a dramatic effect on the capital estimates for a multi-business line organization. In particular, the more upper tail

dependence a copula allows, the higher the required capital estimate. It is thus imperative for capital analysts to select a copula that is most reflective of their own unique situation and risk appetite, in order to avoid the risk of miscalculating their capital requirement.

The second key observation of this dissertation was that the selection of risk measure also has a severe impact on the resultant capital estimates. When considering tail events, ETL provides a better alternative to VaR. Even though StressVaR consistently provided the highest capital estimates, it could be considered as a very capital inefficient risk measure.

The third key observation of this dissertation was that stressing the correlation inputs into the copula calibration algorithm also had an effect on the capital estimates. This effect however was less significant than that of the choice of copula and risk measure.

In conclusion, when aggregating risk and allocating capital using copulas, the resultant capital estimates will always be a function of the choice of copula, the choice of risk measure and the correlation inputs into the copula calibration algorithm. The choice of copula, the choice of risk measure and the conservativeness of correlation inputs will be determined by the organization's risk appetite. A conservative and capital efficient choice could be that of using a 99% ETL, a Cauchy/Student t copula as well as a stressed correlation input.

Further research with regards to capital allocation using copulas could be considering the effects of using other copula such as the Gumbel copula or Frank copula. Other further interesting research could be how copulas could be used to supplement traditional portfolio management, selection and optimization techniques.

Bibliography

Acerbi, C, Nordio, C & Sitori, C 2001, *Expected Shortfall as a Tool for Financial Risk Management*, Cornell University Library, Italy, viewed 15 October 2012, <http://arxiv.org/pdf/cond-mat/0102304>.

Acerbi, C & Tasche, D 2002, 'On the coherence of expected shortfall', *Journal of Banking and Finance*, vol 26, no. 7, pp. 1478-1503.

Alexander, C 2001, *Market Models: A Guide to Financial Data Analysis*, 1st edn, John Wiley and Sons Ltd, Chichester.

Alexander, C 2008, *Market Risk Analysis: Practical Financial Econometrics v. 2*, 1st edn, John Wiley and Sons Ltd, Chichester.

Artzner, P, Delbaen, F, Eber, J & Heath, D 1999, 'Coherent Measures of Risk', *Mathematical Finance*, vol 9, no. 3, pp. 203-228.

Aziz, A & Rosen, D 2004, 'Capital allocation and RAPM', in C Alexander, E Sheedy (eds.), *The Professional Risk Manager's Handbook*, PRIMA Publications, Wilmington.

Bouchaud, J-P & Potters, M 2004, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*, 2nd edn, Cambridge University Press, Cambridge.

Chernobai, AS, Rachev, ST & Fabozzi, FJ 2007, *Operational Risk. A Guide to Basel II Capital Requirements, Models, and Analysis*, 1st edn, John Wiley and Sons Ltd, New York.

Cherubini, U, Luciano, E & Vecchiato, W 2004, *Copula Methods in Finance*, 1st edn, John Wiley and Sons Ltd, Chichester.

Cherubini, U, Mulinacci, S, Gobbi, F & Romagnoli, S 2011, *Dynamic Copula Methods in Finance*, John Wiley and Sons Ltd, Chichester/GB.

Coste, C, Douady, R & Zovko, II 2009, *The StressVaR: A New Risk Concept for Superior Fund Allocation*, viewed 15 August 2012, <http://arxiv.org/abs/0911.4030>.

Delbaen, F 1998, 'Eigenössische Technische Hochschule Zurich', viewed 8 August 2012, <http://www.math.ethz.ch/~delbaen/>.

Denault, M 2001, 'Coherent allocation of risk capital', *Journal of Risk*, vol 4, no. 1, pp. 1-34.

Dowd, K, Hutchinson, MO & Ashby, SG 2011, 'Capital Inadequacies: The Dismal Failure of the Basel Regime of Bank Capital Regulation', *Cato Institute Policy Analysis*, no. 681, p. 40.

Embrechts, P 2000, 'Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool', *ETH Zurich*, <http://www.math.ethz.ch/~embrechts/>.

Embrechts, P, Frey, R & McNeil, AJ 2005, *Quantitative Risk Management: Concepts, Techniques and Tools*, 1st edn, Princeton University Press, New Jersey.

Embrechts, P, McNeil, A & Straumann, D 2002, 'Correlation and Dependence in Risk Management: Properties and Pitfalls', in HA Dempster M (ed.), *In Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge.

Fréchet, M 1951, 'Sur le tableaux de corrélation dont les marges sont donées', *Ann. Univ. Lyon*, vol 9, pp. Sect. A, 53-77.

Genest, C & Rivest, LP 1993, 'Statistical Inference Procedures for Bivariate Archimedean Copulas', *Journal of the American Statistical Association*, vol 88, no. 423, pp. 1034-1043.

Hoeffding, W 1940, 'Masstabinvariante Korrelationstheorie', *Schriften des Mathematischen Instituts und des Instituts fur Angewandte Mathematik der Universitat Berlin*, vol 5, pp. 179-233.

Hull, JC 2007, *Risk Management and Financial Institutions*, 1st edn, Pearson Education (US), Upper Saddle River.

Hull, J 2008, *Options, Futures, and Other Derivatives*, 7th edn, Pearson Education (US), Upper Saddle River.

Javaheri, A 2005, *Inside Volatility Arbitrage: The Secrets of Skewness*, 1st edn, John Wiley and Sons Ltd, New York.

Kimberling, CH 1974, 'A probabilistic interpretation of complete monotonicity', *Aequationes Math*, no. 10, pp. 152-164.

Kousky, C & Cooke, RM 2009, *The Unholy Trinity: Fat Tails, Tail Dependence, and Micro-Correlations*, viewed 13 February 2011, <http://ssrn.com/abstract=1505426 or http://dx.doi.org/10.2139/ssrn>.

Koyluoglu, U & Stoker, J 2002, 'Honour your contribution', *Risk (April)*, pp. 90-94.

Kruskal, WH 1958, 'Ordinal measures of association', *Journal of the American Statistical Association*, vol I, no. 53, pp. 814-861.

Maher, A & Khalil, L 2009, *Basel II Market Risk Framework*, viewed 20 November 2012, <https://www.kpmg.com/EG/en/issuesAndInsights/Documents/issues-insights%20PDFs/Basel%20II%20letter%204%20Market%20Risk.pdf>.

Malevergne, Y & Sornette, D 2006, *Extreme Financial Risks: From Dependence to Risk Management*, 1st edn, Springer-Verlag Berlin and Heidelburg GmbH & Co. KG, Berlin.

Marrison, C 2002, *The Fundamentals of Risk Measurement*, McGraw - Hill Education - Europe, New York.

Meucci, A 2005, *Risk and Asset Allocation*, 1st edn, Springer, Berlin.

Moore, DB & Spruill, MC 1975, 'Unified large-sample theory of general chi-squared statistics for tests of fit', *Ann Statist*, vol 3, pp. 599-616.

Nelson, RB 2006, *An Introduction to Copulas*, 2nd edn, Springer, New York.

Patton, AJ 2006, 'Modelling Asymmetric Exchange Rate Dependence', *International Economic Review*, vol 47, no. 2, p. 30.

Salmon, F 2009, 'Recipe for Disaster: The Formula that Killed Wall Street', *Wired Magazine*.

Sinclair, E 2008, *Volatility Trading*, 1st edn, John Wiley and Sons Ltd, New Jersey.

Sklar, A 1996, 'Random variables, distribution functions, and copulas - a personal look backward and forward', in *Distributions with Fixed Marginals and Related Topics*, Institute of Mathematical Statistics, Hayward, CA.

Sornette, D 2002, 'Critical market crashes', *science@direct*, no. 378, pp. 1-98.

Supervision, BCOB 1997, *bis.org*, viewed 13 August 2012, <www.bis.org/publ/bcbs30.pdf>.

Taleb, N 1997, *Against Value-at-Risk: Nassim Taleb Replies to Phiippe Jorion*, viewed 12 August 2012, <http://www.fooledbyrandomness.com/jorion.html>.

Whaley, RE 2006, *Derivatives: Markets, Valuation, and Risk Management*, 1st edn, John Wiley & Sons Ltd, New York.

Wilmott, P 2000, 'Uncertainty versus Randomness: Minimizing model dependence', *International Journal of Theoretical and Applied Finance*, vol 3, no. 3, pp. 493-500.

Wilmott, P 2006, *Paul Wilmott on Quantitative Finance*, 2nd edn, John Wiley & Sons Ltd, London.

Ziemba, RES & Ziemba, WT 2008, *Scenarios for Risk Management and Global Investment Strategies*, 1st edn, John Wiley and Sons Ltd, Chichester.

Zi-sheng, O, Hui, L & Xiang-qun, Y 2009, 'Modeling dependence based on mixture copulas and its application in risk management', *Appl. Math. J. Chinese Univ.*, vol 24(4), pp. 393-401.