Open University Cyprus **Hellenic *Open University***

Master's join degree/post graduate Programme
Enterprise Risk Management (ERM)

MASTER THESIS

Minimum Capital Requirements for Banks in Market Risk: The Challenge of a Shift from Value at Risk to Expected Shortfall Risk Measure

Chatzivasileiadis Ioannis

Supervisor Professor
Dr. Gerontogiannis Dionysios

May 2018

Open University Cyprus **Hellenic *Open University***

Master's join degree/post graduate Programme
Enterprise Risk Management (ERM)

MASTER THESIS

Minimum Capital Requirements for Banks in Market Risk: The Challenge of a Shift from Value at Risk to Expected Shortfall Risk Measure

Chatzivasileiadis Ioannis

Supervisor Professor
Dr. Gerontogiannis Dionysios

This thesis submitted for partial fulfilment of the requirements
Master's join degree/post graduate programme
«Enterprise Risk Management (ERM)»
Faculty of Economics and Management

Open University of Cyprus **Hellenic Open University**

May 2018

BLANK PAGE

Summary

Financial Institutions today are, more than ever before, subject to various types of risks which shareholders, investors and regulators expect to be timely and accurately evaluated in order to prevent financial crises. Market Risk has become one of the most important risk sources for Financial Institutions and Basel Committee on Banking Supervision Accords have been evolved to take into account lessons from the past. This Thesis analyzes Expected Shortfall, the new risk measure introduced by BCBS in a recent paper (Fundamental Review of the Trading Book, 2012) in order to capture 'tail risk' more effectively. Various models for Value at Risk and Expected Shortfall estimation are presented and evaluated by backtesting techniques in two separate periods, representing US subprime loan crisis and after crisis market conditions respectively, for a univariate equity portfolio, using data of S&P index returns. The findings show that parametric methods like normal and t distributions as well as the non parametric model of historical simulation fail to produce reliable VAR and Expected Shortfall estimates for the crisis period, as they do not respond timely in growing market volatility. Contrary to these results, econometric models capturing volatility dynamics, represented by an EGARCH model, seem to perform best. For this kind of models, Expected Shortfall and VAR estimation and backtesting results indicate that these models could have acted proactively in the beginning of the US subprime loan crisis, determining reliable and accurate market risk limits and equivalent capital requirements. This conclusion becomes very important for future revisions of the BCBS framework, as the vast majority of banks adapt the simple but inefficient method of Historical Simulation for the calculation of their Market Risk capital requirements (EBA Report, 2017: 31, Perignon&Smith, 2009: 367). Contrary to that, the EGARCH model fails to produce reliable risk measure estimates in a low volatility, tranquil market condition.

Acknowledgments

Foremost, I would like to express my deep gratitude to my supervisor professor Dr. Gerontogiannis for his valuable guidance and suggestions. His encouragement during the Master Programme led me to the decision of undertaking a demanding but at the same time interesting and challenging field of research in my Master Thesis, the empirical part of which was conducted mainly with the use of Matlab code, a powerful tool in the analysis which I had to effectively learn how to use.

Also I want to thank all my professors and classmates for our collaboration.

Finally, I express my gratitude to my wife Alexandra for her support, especially during my studying hours after work.

Table of Contents

1	**Introduction**	1
2	**The Evolution of Market Risk Regulatory Framework**	7
2.1	The Basel Committee on Banking Supervision 1996 Amendment	7
2.1.1	Value-at-Risk	8
2.2	Basel 2.5	11
2.3	The Fundamental Review of the Trading Book	12
2.3.1	Expected Shortfall	13
3	**VAR and Expected Shortfall Estimation Methods**	17
3.1	Non Parametric Methods – The Historical Simulation Approach	18
3.2	Parametric Methods	21
3.2.1	Normal Distribution Approach	21
3.2.2	Student's t Distribution Approach	22
3.3	The Monte Carlo Simulation Approach	23
3.4	Methods Capturing Volatility Dynamics	24
3.5	The Filtered Historical Simulation Approach	29
4	**Backtesting**	33
4.1	Backtesting Value at Risk	34
4,1.1	Regulatory Framework – The Traffic Light Test	35
4.1.2	Unconditional Coverage Test – Kupiec Test	37
4.1.3	Conditional Coverage Test – Christoffersen Test	38
4.1.4	The Empirical Coverage Probability	39
4.2	Backtesting Expected Shortfall	39
5	**Empirical Study Framework**	43
5.1	Purpose of the Study	44
5.2	Study Limitations	45
5.3	Study Methodology	45
6	**Empirical Results – VAR and Expected Shortfall Estimation**	47
6.1	Data Analysis	47
6.2	Crisis Period – Models	53
6.2.1	Historical Simulation	53
6.2.2	Normal Distribution	54
6.3.2	Student's t Distribution	55
6.2.4	EGARCH Model	56
6.3	After Crisis Period – Models	59
6.3.1	Historical Simulation	59
6.3.1	Normal Distribution	60
6.3.2	Student's t Distribution	61
6.3.3	EGARCH Model	62
6.4	Comparison of VAR and Expected shortfall Estimation Results	64
7	**Empirical Results – Backtesting**	69
7.1	VAR Backtesting Results	70
7.2	Expected Shortfall Backtesting Results	73
7.3	Comparison of Backtesting Results	74
8	**Conclusions**	75
Appendices		
A	**Matlab Code, Crisis Period**	77
A.1	Matlab Variables	77
A.2	Risk Measure Estimation under Different Models	77

A.2.1	Normal Distribution Model	77
A.2.2	Historical Simulation Model	78
A.2.3	Student's t Distribution Model	78
A.2.4	EGARCH model	79
A.2.5	Comparison of Risk Measure Estimation Results under Different Models	80
A.3	Risk Measure Backtesting	81
B	**Matlab Code, After Crisis Period**	83
B.1	Matlab Variables	83
B.2	Risk Measure Estimation under Different Models	83
B.2.1	Normal Distribution Model	83
B.2.2	Historical Simulation Model	84
B.2.3	Student's t Distribution Model	84
B.2.4	EGARCH model	85
B.2.5	Comparison of Risk Measure Estimation Results under Different Models	86
B.3	Risk Measure Backtesting	87
Bibliography		89

Tables

Table 1	The Basel Committee on Banking Supervision traffic light test	36
Table 2	S&P squared return correlogram, Crisis Period	52
Table 3	S&P squared return correlogram, After Crisis Period	52
Table 4	S&P Crisis Period EGARCH model estimation results	57
Table 5	S&P After Crisis Period EGARCH model estimation results	62
Table 6	S&P VAR backtesting results, Crisis Period	70
Table 7	S&P VAR backtesting results, Crisis Peak Period	71
Table 8	S&P VAR backtesting results, After Crisis Period	72
Table 9	S&P Expected Shortfall backtesting results, Crisis Peak Period	73
Table 10	S&P Expected Shortfall backtesting results, After Crisis Period	74

Figures

Figure 1	The Evolution of Market Risk Management Framework	4
Figure 2	Calculation of VAR from the probability distribution of portfolio losses	9
Figure 3	Probability distribution of losses in portfolio value during time T	14
Figure 4	Distribution of losses (negative returns) for a S&P index tracking portfolio during year 2008	19
Figure 5	Comparison of normal distribution to a heavy-tailed distribution exhibiting leptokurtosis	23
Figure 6	S&P negative log returns, Crisis Period	48
Figure 7	S&P negative log returns, After Crisis Period	48
Figure 8	Descriptive statistics for S&P negative log returns, Crisis Period	49
Figure 9	Descriptive statistic for S&P negative log returns, After Crisis Period	49
Figure 10	Quantile-Quantile plot, S&P index negative log returns, Crisis Period	50
Figure 11	Quantile-Quantile plot, S&P index negative log returns, After Crisis Period	50
Figure 12	S&P squared returns, Crisis Period	51
Figure 13	S&P squared returns, After Crisis Period	51

Figure 14	S&P VAR estimation under Historical Simulation method, Crisis Period	53
Figure 15	S&P VAR estimation under Normal Distribution method, Crisis Period	55
Figure 16	S&P VAR estimation under t Distribution method, Crisis Period	56
Figure 17	S&P forecasted conditional variances under EGARCH model, Crisis Period	58
Figure 18	S&P VAR estimation under EGARCH model, Crisis Period	59
Figure 19	S&P VAR estimation under Historical Simulation method, After Crisis Period	60
Figure 20	S&P VAR estimation under Normal Distribution method, After Crisis Period	61
Figure 21	S&P VAR estimation under t Distribution method, After Crisis Period	62
Figure 22	S&P forecasted conditional variances under EGARCH model, After Crisis Period	63
Figure 23	S&P VAR estimation under EGARCH model, After Crisis Period	64
Figure 24	S&P VAR estimation under different models at 97,5% confidence level, Crisis Period	65
Figure 25	S&P VAR estimation under different models at 99% confidence level, Crisis Period	65
Figure 26	S&P VAR estimation under different models at 97,5% confidence level, After Crisis Period	66
Figure 27	S&P VAR estimation under different models at 99% confidence level, After Crisis Period	66
Figure 28	S&P Expected Shortfall estimation under different models at 97,5% confidence level, Crisis Period	67
Figure 29	S&P Expected Shortfall estimation under different models at 97,5% confidence level, After Crisis Period	67

BLANK PAGE

Chapter 1
Introduction

Financial risk management as a discipline has been evolving since late 1960s as the outcome of, among other factors, increased trading activity, markets interconnectedness, volatile global economic environment and the advent of financial derivatives. Banks, Mutual Funds and investment firms are now, more than ever before, vulnerable to various types of risk that threaten their mare existence. Major Financial Institutions have been hit by great losses in trading, such us Barings Bank, the UK merchant bank that was forced into bankruptcy in 2005, mainly as a result of huge losses in Japanese stock index futures, whereas in 1996 Sumitomo Corp (a Japanese bank) lost $2,6 billion in commodity futures (Saunders & Cornett, 2003: 232). The burst of Credit Crisis in USA in 2008 led several institutions to insolvency and stock market capitalization was wiped out in a short period.

The most important risks faced by Financial Institutions can be classified into four categories (RiskMetrics, 1996: 5):

- Credit risk, which is the potential loss which stems from the counterparty inability to meet its obligations.
- Operational risk, which results from errors made in payments, transactions and procedures in general.
- Liquidity risk, which arises from the inability to meet short term financial demands because of the level of illiquid assets.
- Market risk, which involves the uncertainty of a Financial Institution trading portfolio returns, caused by changes in market conditions, such as stock indexes, exchange rates, interest rates and commodity price volatilities.

Income from trading activities has increased to a considerable extent compared to income from traditional Financial Institution activities of deposit taking and lending. Market risk arises when assets and liabilities are actively traded than held on the balance sheet for longer term investment, funding or hedging purposes (Saunders & Cornett, 2003: 233).

In the late 1970s and 1980s, many Financial Institutions started developing internal models for the measurement and aggregation of market risk in their trading portfolios. This became appropriate, as the underestimation of the underlying risk could have had severe consequences. Senior management of these institutions needed a sense of the probability of losses at the firmwide level in order to set appropriate limits, to allocate resources efficiently and to evaluate performance in a risk-return basis (Dowd, 2005: 9). The RAROC formula developed by Bankers Trust (an American bank) in the late 1970s can be placed in the direction described above.

The advent of the Value at Risk measure originated, according to a banking industry legend, when the chairman of JP Morgan, Dennis Weatherstone asked his staff to submit a short scaled report indicating risk and potential losses over the next 24 hours, across the bank's trading portfolio. This led to the development of Value of Risk by RiskMetrics™ in 1994. Essentially, VAR gives a dollar denominated amount that can be potentially lost with x% probability over a given time horizon.

The widespread use of VAR as an internal measure of risk was given regulatory recognition by the BCBS under the 1996 Market Risk Amendment to the Capital Accord (BCBS, 1996). Under these guidelines, banks were expected to estimate their Market Risk regulatory capital based on a standardized approach or an internal models approach with the use of VAR. The Variance Covariance or Normal distribution approach of RiskMetrics and the non parametric and simple to apply Historical Simulation methods have dominated, since then, the bank internal models approach.

Apart from its advantages though, VAR was soon criticized of having three serious drawbacks. The first was that the estimation error in VAR produced by different VAR models could send inaccurate information to senior management of Financial Institutions, leading them to inappropriate investment decisions (Jorion, 1996, Berkowitz & O'Brien, 2001). The second was that it failed to account for losses above the X percentile of the loss distribution and thus significantly underestimates tail risk. Finally, its third shortcoming is that it lacks the property of subadditivity that coherent measures of risk have (Artzner et al., 1999: 209). This means that it does not account for diversification benefits after the merge of two or more portfolios.

The last two VAR drawbacks led to the development of a candidate risk measure that accounts for tail risk and is at the same time coherent (Acerbi & Tasche, 2001: 6), Expected Shortfall. It is also referred as Conditional VAR or Expected Tail Loss and it is the expected loss during time T conditional on the loss being greater than a certain quantile of the loss distribution (Hull, 2012: 187). Despite these superior to VAR characteristics, Expected Shortfall was proved by Gneiting (2011) to lack a mathematical property called 'elicitability'. This result sparked a confusing debate as to whether it is also backtestable, until Acerbi & Szekely (2014) proved that the lack of elicitability property is more a concern for model selection than for model testing.

Meanwhile, the US financial crisis in 2008 signaled significant weaknesses in the 1996 BCBS framework, leading to a revision of Market Risk regulatory framework in 2011 referred to as 'Basel 2.5 (BCBS, 2011). Among other changes this revision involved the calculation of an additional VAR measure by banks called 'stressed VAR' which would account for market losses in a stressed period of 250 days.

Following previous revisions, the BCBS incorporated finally the expected shortfall risk measure in its Market Risk regulatory framework in 2012 (BCBS, 2012: 20). In this consultative document the Committee recognized the fact that 'Expected Shortfall accounts for tail risk in a more comprehensive manner'. The Committee preserved the VAR though as a measure to backtest in the common way. This was attributed by some researches (eg Righi & Caretta, 2014: 15) to the on going investigation for Expected Shortfall certain estimators' superiority relative to others as well as to the ambiguity as to whether it can be applied for backtesting.

In 2016, the BCBS issued its 'Minimum Capital Requirements for Market Risk' (BSBS, 2016). Within this document the Committee sets out the transitional arrangements for the revised Market Risk framework which national supervisors are expected to put into practice and finalize their implementation by January 2019, while the banks under their supervision are expected to report under the new standards by the end of 2019.

The figure below depicts the most important changes in the Market Risk framework through its evolution.

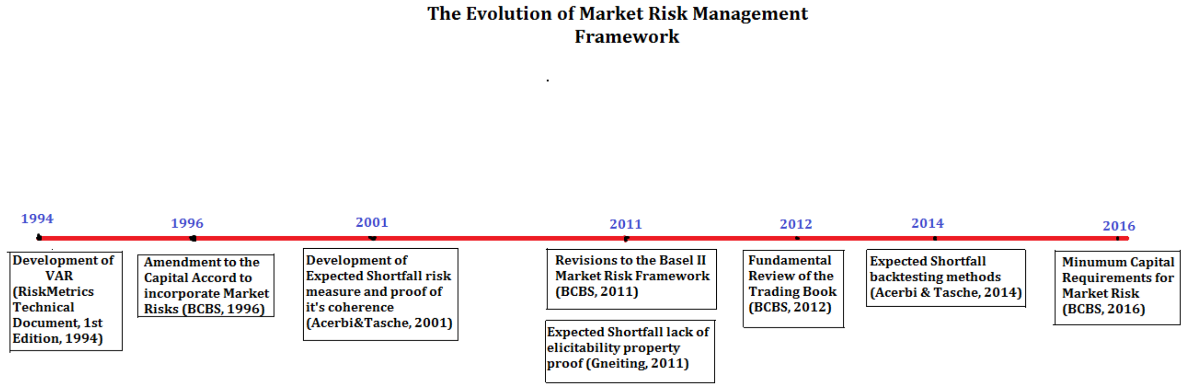

Figure 1: The Evolution of Market Risk Management Framework

During these years, much research has been done upon the accuracy and efficiency of various VAR models in correctly forecasting adverse return movements. The superiority of certain models over others was examined considering three important characteristics of financial returns described by Manganelli & Engle (2001: 8):

- Financial return distributions are leptokurtic, that is they present fat tails
- Equity returns are typically negatively skewed.
- Squared returns present serial autocorrelation, which means that return volatility tends to cluster in certain periods.

These characteristics led to the advent of econometric models for VAR estimation that account for volatility dynamics, like ARCH (Autoregressive Conditional Heteroskedasticity) and GARCH (Generalized Conditional Hetersoskedasticity), following the pioneering works of Engle (1982) and Bollerslev (1986) respectively.

Although VAR has been exhaustively studied, with respect to Expected Shortfall risk measure there is still an on going academic research on the superiority of certain estimators relative to others (Righi & Caretta, 2014: 15). This is probably due to the fact that Expected Shortfall is a new risk measure that banks are required to estimate according to the resent BCBS regulatory framework.

This Master Thesis is planned to quote a significant part of the relevant literature on certain VAR and Expected Shortfall estimation models and their accuracy. In the empirical part of the Thesis four types of models are tested during and after the US credit crisis of 2008 so as to assess their reliability through a stressed and a tranquil period respectively for an equity portfolio represented by S&P index, simulating a bank's equity trading desk. The Thesis is organized as follows. Chapter 2 describes the evolution of Market Risk regulatory framework in more detail and also presents the properties of VAR and Expected Shortfall risk measures. Chapter 3 introduces the theoretical background of various VAR and Expected Shortfall estimation methods, based on certain models. It also refers to the relevant literature on the test of these models. Chapter 4 describes the backtesting methods for VAR and Expected Shortfall. Chapter 5 analyzes the purpose of the empirical study and the methodology used as well as the study limitations. Chapter 6 presents the estimation results with the equivalent figures. Chapter 7 reviews the empirical backtesting results. Finally Chapter 8 concludes the Thesis.

BLANK PAGE

Chapter 2
The Evolution of Market Risk Regulatory Framework

In this chapter the evolution of BCBS framework for market risk is presented and the two risk measures used in the framework, VAR and Expected Shortfall, are analyzed. The major changes in the BCBS standards for market risk are found in 1996 Amendment (BCBS, 1996), Basel 2.5 (BCBS, 2011) and the Fundamental Review of the Trading Book (BCBS, 2012).

2.1 The Basel Committee on Banking Supervision 1996 Amendment

Following the Basel Accord 'International Convergence of Capital Measurement and Capital Standards' in 1988, known as Basel I, which stipulated international standards for the estimation of regulatory capital for credit risk in banks, the Committee issued an Amendment to the Basel I Accord in 1996 (BCBS, 1996), known as the '1996 Amendment' in order to set a framework for the estimation of a capital charge for market risk. In this way the Committee acknowledged the increasing importance of market risk which grew in accordance with bank trading activities.

The Amendment required banks to use 'marking to market' practice of revaluing assets and liabilities on a daily basis. This meant that banks were expected to use fair value accounting for all assets and liabilities held in their trading book. The Committee also distinguished among four types of market risk: Interest rate risk, equity risk, foreign exchange risk and commodities risk.

Two methods for measuring the capital charge for market risk were outlined by the framework: The Standardized Approach and the Internal Models Approach. The Standardized Approach assigns capital separately for each of interest rate risk, equity risk, foreign exchange risk, commodities risk and options, due to a certain methodology. More sophisticated banks with well-established risk management systems could use the Internal Models approach, upon approval from their supervisory authorities. This involves a Value at Risk measure computed on a daily basis at the 99th percentile one tailed confidence interval, for a historical observation period constrained to a minimum length of one year and a 10 day holding period (BCBS, 1996 :44). The capital requirement for the Internal Models Approach is calculated as follows:

$$\mathbf{max\ (VaR_{t-1},\ m_c \times VaR_{avg}) + SRC}$$

where m_c is a multiplicative factor and SRC is a specific risk charge set by individual supervisory authorities on the basis of their assessment of the bank's risk management systems. The variable VaR_{avg} is the average Value at Risk over the last 60 business days and VaR_{t-1} is the previous day VaR. Finally, according to the Amendment, a 10 day 99% VaR can be calculated as $\sqrt{10}$ times the one day 99% VaR.

The 1996 Amendment officially introduced Value at Risk measure in an international regulatory market risk framework.

2.1.1 Value at Risk

VAR is probably the most widely used risk measure in Financial Institutions. In the Basel regulatory framework VAR has a prominent role, as it is used for the calculation of the market risk weighted capital and also for the credit risk weighted capital, the later after the Internal Ratings Based approach of Basel II.

VAR summarizes the worst expected loss over a specified horizon within a confidence interval. It summarizes in a single number the probability of adverse moves in financial variables (Jorion, 1996: 47). When using VAR as a risk measure, we are interested in making a statement of the form: 'We are X percent certain that we will not lose more than V dollars in time T' (Hull, 2012: 183).

VAR can be calculated from either the probability distribution of gains or losses (negative gains) during time T. In the later case the calculation of VAR is depicted in the following diagram.

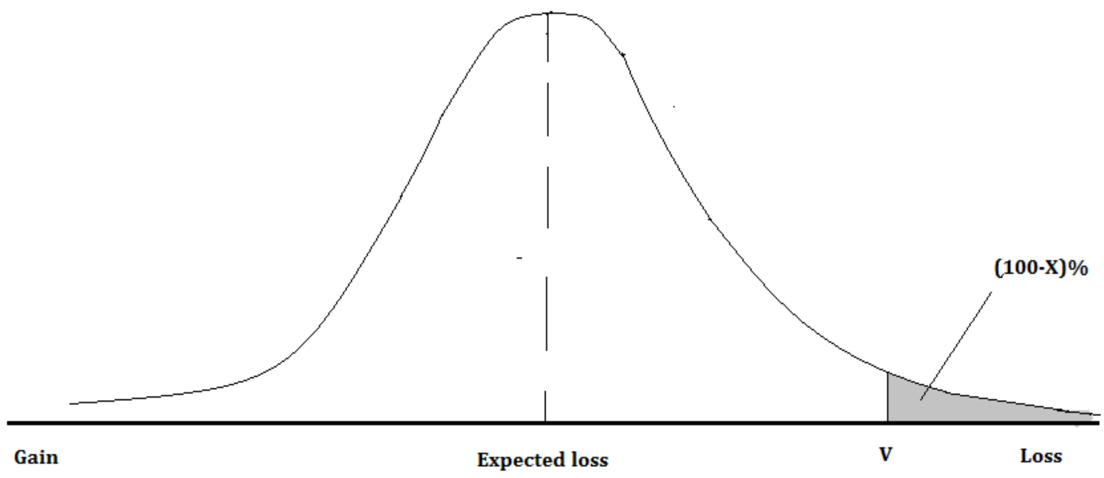

Figure 2. Calculation of VAR from the probability distribution of portfolio losses. Confidence level is X%. VAR level is V (Hull, 2012: 185).

It becomes obvious that VAR calculation requires an assumption about the distribution of portfolio returns. Additionally, the fact that return distributions might change over makes its estimation even more challenging (Manganelli & Engle, 2001: 6).

Given random numbers X with distribution function F_X, and a confidence level α%, $VAR_\alpha(X)$ can be expressed by the following mathematical formula (Kellner & Rosch, 2016: 47):

$$VAR_a(X) = \inf(x \in \Re : F_X(x) \geq a)$$

According to Dowd (2005: 12), VAR has a number of significant attractions which can be summarized as follows:

- It provides a common consistent measure of risk across different positions.
- It enables a risk manager to aggregate the risks of subportfolios into an overall measure of total portfolio risk.
- It takes full account of all driving risk factors inherent in a portfolio.
- It is probabilistic and gives useful information on the probabilities associated with specified levels of portfolio losses.

- It can simply be expressed in a money unit measure of risk, concentrating daily informative reports for market risk to risk executive officers into a single monetary value.

However VAR exhibits the serious drawback of not examining the tail risk, which is expressed by losses above the confidence level. In addition to that, VAR lacks 'coherence', which is summarized on a number of properties that best risk measures should have (Artzner et al., 1999: 203-228). In mathematical terms, these properties are presented by Acerci & Tasche (2001: 3). Considering a set V of real-valued random variables and X,Y two portfolios' return outcomes, a function ρ: V→R is called a coherent risk measure if it is

- Monotonous: $\mathrm{X}, Y \in V, Y \geq X \Rightarrow \rho(Y) \leq \rho(X)$
- Translation invariant: $\mathrm{X} \in V, \alpha \in \Re \Rightarrow \rho(\mathrm{X}+\alpha) = \rho(\mathrm{X}) - \alpha$
- Positive homogeneous: $\mathrm{X} \in V, h > 0, hX \in V \Rightarrow \rho(hX) = h\rho(\mathrm{X})$
- Subadditive: $\mathrm{X}, \mathrm{Y}, (\mathrm{X}+\mathrm{Y}) \in V \Rightarrow \rho(\mathrm{X}+\mathrm{Y}) \leq \rho(\mathrm{X}) + \rho(\mathrm{Y})$

VAR lacks the last property of subadditivity and thus it is not a coherent risk measure. This missing property of VAR constitutes it as an improper measure of risk, as it neglects the diversification benefits of aggregating two of more portfolios. The merge of these portfolios should produce a risk measure minor to the sum of the individual ones. This can be intuitively conceived by the use of the following characteristic example described by Acerbi & Tasche (2001: 3): if we consider a bank with several branches, each having its own capital requirement, the regulating authority should be confident that the overall bank risk should be more than sufficiently covered by the sum of branch capital requirements. Nevertheless, if the subadditivity property is not fulfilled, the total bank risk could turn out to be bigger than the sum of its branches and therefore VAR could imply an underestimation of the real inherent risk.

Yamai & Yoshiba (2005: 1000) also show that utility maximizing investors of a concentrated credit portfolio, choose to invest in securities with a high potential for large losses beyond the VAR level, when they are imposed VAR constraints in their investment.

We will see that Expected Shortfall does not present the above drawbacks and this is one of the main reasons Basel Committee on Banking Supervision has recently incorporated it in the market risk regulatory framework (BCBS, 2012).

2.2 Basel 2.5

Starting in 2007, the United States experienced the 'worst financial crisis since the 1930s' (Hull 2012: 121). The crisis spread rapidly throughout the world and from the financial markets to the real economy, posing a great deal of criticism to the risk management practices of Financial Institutions and the effectiveness of the risk management regulatory framework.

It was soon recognized that some changes were necessary to the calculation of capital for market risk in the BCBS framework. In 2011, the Committee published the 'Revisions to the Basel II market risk framework', a set of modifications to the 1996 Amendment, known also as 'Basel 2.5'. The Committee recognized that the financial crisis began in the mid-2007 accumulated excessive risks in the banks' trading books and that 'the capital framework for market risk, based on the *1996 Amendment to the Capital Accord to incorporate market risks*, does not capture some key risks' (BCBS, 2011:1).

The major modifications to the older framework involved:

1. The calculation of a 'stressed VAR'
2. A new incremental risk charge

As most banks, during the financial crisis, used the historical simulation method to calculate VAR, this led to a significant negligence of rising market volatility in stressed market conditions. Under the new framework, the Committee required banks to calculate a stressed VAR, taking into account a one-year observation period relating to significant portfolio losses (BCBS, 2011: 1).

Basel 2.5 actually required banks to calculate 2 VARs. One is the usual VAR (based on the previous one to four years of market return observations). The other one is a stressed

VAR , calculated from a stressed period of one year). The two measures are combined to form a total capital charge c. This is calculated according to the following formula:

$$\mathbf{c = max\ (VaR_{t-1},\ m_c \times VaR_{avg}) + max\ (s\ VaR_{t-1},\ m_s \times sVar_{avg})} \text{ (BCBS, 2011: 15)}$$

The multiplication factors m_c and m_s are set by individual supervisory authorities on the basis of their assessment of the quality of the bank's risk management system, subject to an absolute minimum of 3.

Because stressed VAR is at least as great as VAR, the formula shows that assuming $m_c=m_s$, the capital requirement under Basle 2.5 is at least doubled (Hull, 2012: 287).

Additionally, the new guidelines accounted for risk inherent in instruments in the trading book sensitive to default risk. As credit-dependent instruments were criticized for being responsible, to a large extent, for the advent of the crisis, regulators accounted for this fact by demanding the calculation of an incremental risk charge for credit sensitive products in the trading book. For the calculation of this charge banks would be expected to estimate a liquidity horizon for each product as well, representing the time required to sell the position or hedge effectively.

2.3 The Fundamental Review of the Trading Book

Although the revisions in Basel 2.5 constituted an important improvement in market risk regulatory framework that could act proactively in stressed periods of increased market volatility, the Committee soon recognized that they did not fully address the shortcomings inherent in the framework. In 2012, the BCBS published the first of a series of consultative documents named 'The Fundamental Review of the Trading Book' (BCBS, 2012). Under the new guidelines, VAR was maintained as a risk measure in market risk regulatory framework, but also Expected Shortfall was introduced in the framework for the first time.

According to this document 'a number of weaknesses have been identified with using value-at-risk (VAR) for determining regulatory capital requirements, including its inability to capture tail risk' (BCBS, 2012: 3). For this reason, the Committee introduced an alternative metric, in particular Expected Shortfall (ES). ES simply measures the expected value of losses beyond a certain confidence level.

The Committee acknowledged the effectiveness of a 10-day holding period VAR calculation for internal day to day risk management purposes but it also questioned its ability to meet the objectives of prudential regulation 'which seeks to ensure that banks have sufficient capital to survive low probability, or tail events' (BCBS, 2012: 9). One of the key weaknesses of the VAR metric includes the provision of incentives to banks to take on tail risk, which is risk above the threshold of VAR. Indeed VAR does not examine the distribution of losses above its threshold.

Under a second consultative document published in 2013 (BCBS, 2013), Expected Shortfall was set to be calculated at a 97,5% confidence level for the calibration of market risk capital requirements of the internal models approach, using current 12 month (most recent) observation period and also for a stressed period, following guidelines for VAR in Basel 2.5 framework.

Finally, the 'Minimum capital requirements for market risk' standards published in 2016, (BCBS, 2016) required banks to calculate two daily VARs for each trading desk, calibrated to one tail 99.0 and 97.5 percent confidence level and a daily ES calibrated to an equivalent 97.5 percent level (BCBS, 2016: 56). The Committee expects national supervisors to set the implementation of the new standards by January 2019.

2.3.1 Expected Shortfall

In the previous section we saw that Basel Committee on Banking Supervision adopted Expected Shortfall as a new risk measure in its regulatory market risk framework primarily because of the inefficiency of VAR to account for tail risk. Basel 2.5 framework was soon criticized for this drawback. According to Danielsson (2016: 13), this control mechanism produced moral hazard for senior management or traders who were incentivized to target low probability – high impact portfolio outcomes.

This significant inefficiency of VAR as a risk measure is plainly presented by Hull (2012, 187) through a representative example. Suppose that a bank' trader is imposed a risk limit described by a one day 99% VAR of $10 million. In order to increase its expected return, the trader might construct a portfolio where there is a 99.1% probability that daily loss is less than $10 million and a 0.9% probability that it is $500. Although the bank risk limit is satisfied by the trader, it is obvious that is taking unacceptable risks above this limit. VAR in figure 2 below is the same as VAR in figure 1, but it obvious that the riskiness of the portfolio is increased.

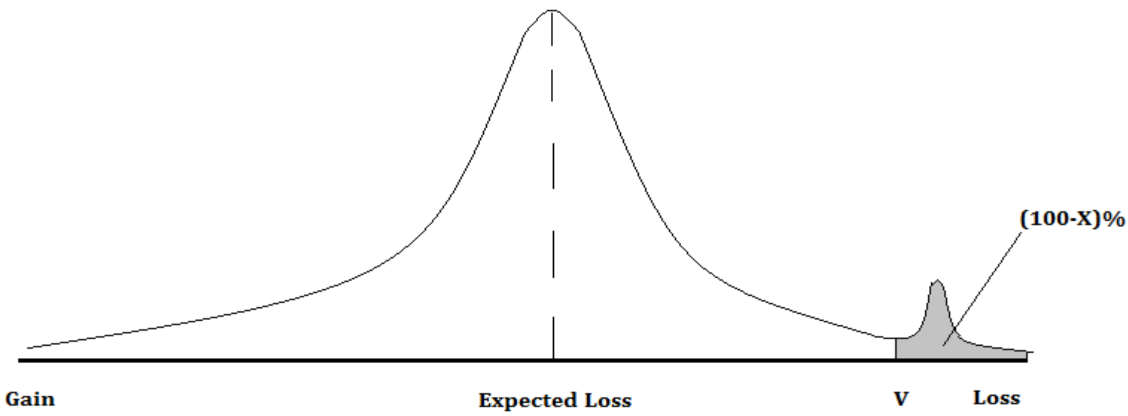

Figure 3. Probability distribution of losses in portfolio value during time T. Confidence level is X%. Portfolio has the same VAR level V, as in figure 1, but a larger loss is more likely (Hull, 2012: 187).

Expected Shortfall is the expected loss during time T conditional on the loss being greater than the X% percentile of the loss distribution. It is also known as conditional VAR.

Given random numbers X with distribution function F_x and a confidence level α%, the Expected Shortfall can be defined by the following mathematical formulas, according to Kellner & Rosch (2016: 47) and McNeal et al. (2015: 69) respectively:

$$\text{(i)} \qquad ES_a(\mathrm{X}) = \frac{1}{1-\alpha}\int_a^1 VAR_v(\mathrm{X})dv$$

$$\text{(ii)} \qquad ES_\alpha(\mathrm{X}) = \frac{1}{1-\alpha}\int_a^1 q_v F(X)dv$$

where $q_v F(X)$ is the quantile function of F(x).

Expected shortfall was proved to be a coherent risk measure in 2001 (Acerbi & Tasche, 2001). Thus it satisfies the properties of monotonicity, subadditivity, positive homogeneity and translation invariance that a coherent risk measure should exhibit. As it also accounts for tail risk in a portfolio loss distribution, the Basel Committee on Banking Supervision incorporated Expected Shortfall in its regulatory framework in 2012 (BCBS, 2012).

Despite its significant benefits, Expected Shortfall lacks a mathematical property known as elicitability. Risk professionals had ignored this property until 2011, when Gneiting (2011) proved that Expected Shortfall is not elicitable, as opposed to VAR.

Elicitability allows a measure to have a scoring function that makes comparison of different models in their forecasting ability possible. Much like the squared or absolute error, point forecasts are compared and assessed with the use of an error measure, which is averaged over forecast cases (Gneiting, 2011: 746). We say that we have forecasts that we call x and verifying observations that we call y. The scoring function now is S(x,y) and the performance criterion takes the following form:

$$\overline{S} = \frac{1}{n} \sum_{i=1}^{n} S(x_i, y_i)$$

The lack of elicitability for Expected Shortfall means that it is not possible to find a scoring function S(x,y) such that Expected Shortfall is defined as the forecast x, given a distribution of verifying observations y, that minimizes the above scoring function. According to Gneiting (2011: 756), this negative result could impact its use as a predictive measure of risk and explain the lack of academic literature, until then, on the evaluation of Expected Shortfall forecasts, as opposed to other risk measures (e.g. VAR).

Intuitively, this drawback can be understood as follows: Expected Shortfall aims at capturing the full extent of risks beyond the VAR level, in the tail of a loss distribution, including severe losses that are theoretically possible but not observed in the relevant backtesting sample period (Ming, 2014: 197).

The discovery that Expected Shortfall is not elicitable led many researchers to conclude that it could not be backtestable and sparked a debate in academic literature. Under a second consultative document published in 2013, Basel Committee on Banking Supervision added Expected Shortfall to VAR, as a risk measure that banks have to calculate in the internal models approach. At the same time, the Committee maintained VAR as the measure to backtest in the usual way.

Finally, Acerbi & Szekely (2014) proved that elicitability is irrelevant to backtesting (Acerbi & Szekely, 2014: 9) and proposed three tests for Expected Shortfall backtesting procedure. This work was pioneering, as it gave an end to a strong criticism of the inclusion of Expected Shortfall in the Basel regulatory framework.

Chapter 3
VAR and Expected Shortfall Estimation Methods

Due to Basel Committee on Banking Supervision market risk regulatory framework, no particular type of VAR or Expected Shortfall model is prescribed for banks to adopt. 'Supervisors may permit banks to use models on either historical simulation, Monte Carlo simulation, or other appropriate analytical methods' (BCBS, 2016: 54).

While VAR and Expected Shortfall are the risk measures that the Basel regulatory framework requires banks to estimate, there is little consensus on the preferred method for calculating them. Traditional methods like historical simulation and Normal distribution were criticized for producing too conservative VAR calculations during periods of unusual market movements. This gave rise to new approaches such as econometric models capturing volatility dynamics, extreme value theory (Embrechts, Resnick, & Samorodnitsky, 1999) and the filtered historical simulation method (Barone-Adesi et al., 1998).

In this chapter various models for the estimation of VAR and Expected Shortfall are presented. The fact that financial return distributions exhibit leptokurtosis and their volatilities tend to change over time, has embarked an academic research on the superiority of certain models relative to others.

3.1 Nonparametric Methods – The Historical Simulation Approach

While VAR and Expected Shortfall are mathematically defined and intuitively conceived, their calculation has been proved a challenging statistical problem. According to Manganelli & Engle (2001: 8), their calculation entails the following steps:

- Marking-to Market the portfolio
- Estimation of the portfolio returns distribution
- Estimation of VAR and Expected Shortfall

The historical simulation is a nonparametric method that makes no assumption about the distribution of portfolio returns. It rather implies that history repeats itself in a way that the calculation of a risk measure for a particular day t is based on the empirical distribution of portfolio returns throughout the previous 250 days (if a rolling window of this width is chosen). According to the Basel regulatory framework, the choice of the sample period (observation period) for calculating both risk measures is confined to a minimum length of one year. The regulatory backtesting procedure of models reliability is based on the previous 250 day period. Banks may also apply a one year rolling window of observations to calculate VAR and Expected Shortfall.

The procedure of calculating risk measures according to historical simulation starts with the sorting of negative returns in the window in an ascending order (in this way we calculate VAR and Expected Shortfall from the probability distribution of losses in portfolio value). For VAR calculation, the $\alpha = (1\text{-}p)$ quantile of interest is deduced by leaving p% observations on the right side of the distribution and (1-p)% observations on the left side. Expected Shortfall can be calculated as the average of the highest p% observations. To compute these risk measures for the following day, the whole window is moved forward by one observation and the procedure is repeated.

VAR calculation according to the historical simulation approach is plainly described by the following example: Suppose we are at time t: 02/01/2009 (first business day of 2009). A bank wants to calculate the daily VAR at time T on its equity portfolio that tracks the S&P index, based on the last 250 daily observations (observation window=250). The histogram of portfolio losses in this window is depicted below in diagram 1.

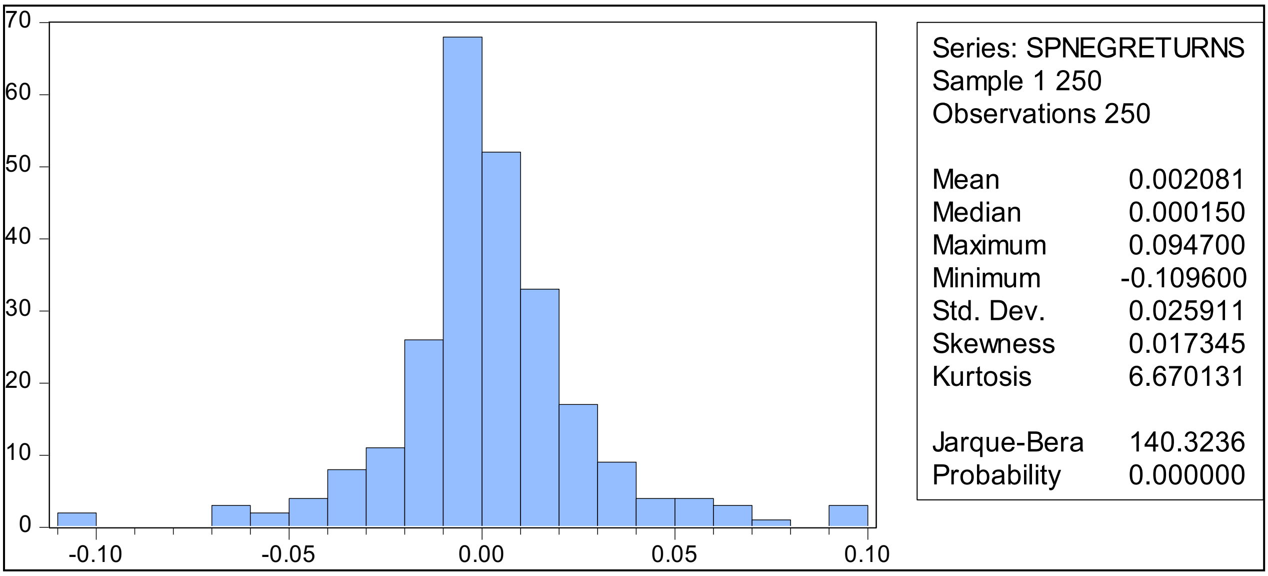

Figure 4: Distribution of losses (negative returns) for a S&P index tracking portfolio during year 2008.

We can observe that the distribution of losses presents positive skewness and excess kyrtosis, as the observation period coincides with the peak of the US financial crisis. A bank would simply calculate the 99% quantile of the distribution, that is, the negative return that cuts off the 1% of highest losses. This can be easily computed to be 0,0851.

Given a distribution of portfolio losses f, a confidence level of α = (1-p)%, and t=1,2,…..,T observations in the rolling window, the VAR of a portfolio at time t can be estimated as the α = (1-p) percentage point of the empirical distribution of losses in the window, according to the following formula:

$$VAR_t^{(1-p)} = f_{(1-p)}(\{y_{t-\tau}\}_{\tau=1}^{\mathrm{T}})$$

Historical simulation approach was criticized for being slowly reactive to fast changing market conditions. The simplifying assumption that the empirical distribution of losses in the observation window can be used to make reliable forecasts of extreme losses, has concentrated much of this criticism. This assumption means that the particular approach considers portfolio return distribution to be constant over the window.

Moreover, assuming market is moving from a period of low volatility to a period of high volatility, VAR estimates based on this approach would be biased downwards, since it would take a period of time before the observations of the low volatility window leave the window (Manganelli & Engle, 2001: 10).

For the historical simulation approach, the decision about rolling window size is crucial about the accuracy of VAR and Expected Shortfall estimation. A larger estimation window has the advantage of not being too sensitive to extreme observations, while at the same time risk measures take longer to adjust to structural changes in risk (Danielsson, 2011: 98).

Another important disadvantage of historical simulation approach is that it assigns an equal probability weight of 1/T in each observation in the window, which is equivalent to assuming that returns are independently and identically distributed (Pritsker, 2006: 563). This particular disadvantage can be omitted by placing more weight in more recent returns and represent today's portfolio risk better, following the approach that was introduced by Boudoukh, Richardson, & Whitelaw (1998: 6). For each of the most recent K returns of the portfolio, y_t, y_{t-1},....,y_{t-K+1}, an associated weight of, $\frac{1-\lambda}{1-\lambda^{K}}, (\frac{1-\lambda}{1-\lambda^{K}})\lambda, \ldots, (\frac{1-\lambda}{1-\lambda^{K}})\lambda^{K-1}$, is placed respectively. To find the (1-p)% VAR of the portfolio we just sum the corresponding weights until p% is reached, starting from the largest negative returns. This method has been proved to respond strongly, rising in magnitude, to crisis period market conditions (Pritsker, 2006: 564).

As the majority of banks use the methodological approach of historical simulation for the calculation of market risk regulatory capital (EBA Report, 2017: 31, Perignon & Smith, 2009: 367), serious considerations arise as to whether this capital is efficiently adjusted in periods of volatility clusters. Indeed, during the 1998 Russian crisis, market risk was modeled with relatively stable financial data, but when the crisis hit, volatility for some assets increased sharply, breaking many risk limits (Danielsson, 2002: 1276).

3.2 Parametric Methods

Unlike historical simulation approach which makes no assumption about the distributional properties of portfolio returns in the rolling window, parametric methods consider that returns follow a certain distribution and its parameters are estimated from the observations in the window. The popularity of these methods relies also on their simplicity. They also suffer from the similar drawbacks with historical simulation method, as they assume a constant distribution of returns and a non-time variant volatility. The most applied parametric methods for the calculation of VAR and Expected Shortfall are the normal and student's t distributions.

3.2.1 Normal Distribution Approach

The normal or Gaussian distribution approach was introduced by RiskMetrics (1996: 6). The approach is based on the assumption that returns in the rolling window are normally distributed. It is a very attractive method as it only requires the estimation of two independent parameters, a mean μ and a standard deviation σ, in every daily VAR and Expected Shortfall calculation. The normal distribution is also convenient as it produces straightforward formulas for the calculation of both risk measures. The corresponding formulas for the calculation of VAR and Expected Shortfall are presented in McNeil et al. (2015: 65, 70) as follows:

$$VAR_\alpha = \mu + \sigma\Phi^{-1}(\alpha) \qquad \text{(i)}$$

$$ES_a = \mu + \sigma\frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha} \qquad \text{(ii)}$$

Where α is the confidence level, Φ^{-1} is the standard normal variate, or inverse standard normal density function, corresponding to the chosen confidence level (e.g., $\Phi^{-1}(0{,}95) = 1{,}645$) and *φ(.)* is the value of the standard normal density function.

An assumption underlying the normal distribution approach is that market variables follow the normal distribution and confidence intervals for financial returns can be constructed after the calculation of μ and σ. For example there is a 5% probability that a financial return will move up by more than 1,645 standard deviations from its mean. In practice, though, financial variables are more likely to experience big moves than a normal distribution would suggest, as they present excess kurtosis (fat tails). The following method of student's t distribution partly accounts for this drawback as it more leptokurtic than the normal.

3.2.1 Student's t Distribution Approach

The student's t distribution approach for the calculation of VAR and Expected Shortfall captures extreme moves in financial returns more efficiently, as it exhibits more probability mass in the tails. Under this approach, a generalized t distribution is used, for which we have to specify the mean (μ), the standard deviation (σ) and the number of degrees of freedom (v). Following McNeil et al. (2015 : 66, 71), the corresponding formulas for VAR and Expected Shortfall calculation for the t-distribution approach are presented below:

$$VAR_{\alpha} = \mu + \sqrt{\frac{v-2}{v}}\sigma t_{v}^{-1}(\alpha) \qquad \text{(i)}$$

$$ES_{a} = \mu + \sqrt{\frac{v-2}{v}}\sigma\left(\frac{g_{v}(t_{v}^{-1}(\alpha))}{1-a}\right)\left(\frac{v+(t_{v}^{-1}(\alpha))^{2}}{v-1}\right) \qquad \text{(ii)}$$

where t_{v}^{-1} denotes the t inverse cumulative distribution function and g_v the density of the standard t.

As ν (degrees of freedom) gets large, the kurtosis approaches 3 and the distribution approximates the normal and as ν falls to 5, the kurtosis rises to 9. According to Dowd (2005: 160) we can approximately match ν to an empirical kurtosis by setting v as the integer closest to the value $\frac{4k-6}{k-3}$, where *k* is the kurtosis coefficient.

In the next figure we can see the comparison of a normal distribution with a heavy-tailed distribution of the t-locationscale distribution family that presents leptokurtosis.

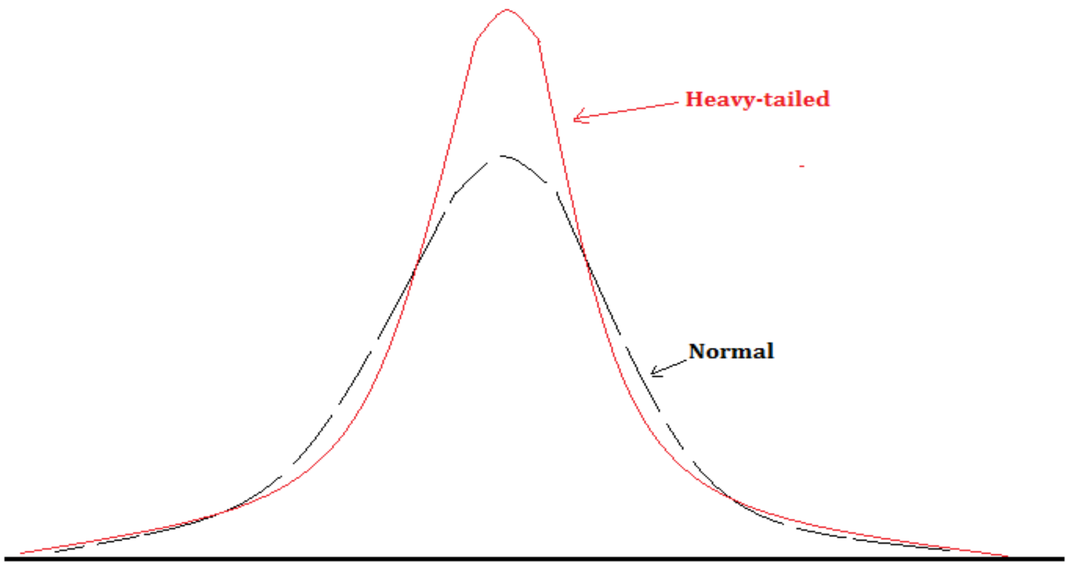

Figure 5. Comparison of normal distribution to a heavy-tailed distribution exhibiting leptokurtosis (Hull, 2012: 211).

Despite its advantage to handle excess kurtosis in financial returns, the t distribution approach has its own problems outlined by Dowd (2005: 160):

- It can produce misleadingly very conservative estimates of VAR and Expected Shortfall at periods of low market volatility.
- When used at very high or low confidence levels, it fails to deliver results equivalent to extreme value theory, a characteristic exhibited also by the normal distribution.
- The t distribution exhibits non stable properties (i.e. the sum of two or more random variables that follow the t-distribution is not necessarily t-distributed itself).

3.3 The Monte Carlo Simulation Approach

An alternative method to previous approaches for VAR and Expected Shortfall estimation is the Monte Carlo approach. Assuming that market variables x_i affecting the value of a portfolio follow a normal distribution, this method simulates the distribution of portfolio returns in the following way, described by Hull (2012 : 340):

1. The portfolio today is marked to market using the current value of market variables x_is.
2. We sample once from the multivariate normal distribution of Δx_is.
3. The sampled values of Δx_is are then used to revalue each market variable.
4. Portfolio is revalued according to the new estimates of market variables.
5. Portfolio value in step 4 is subtracted from that defined in step 1, to determine a sample ΔP.
6. We repeat this procedure many times in order to simulate a probability distribution of ΔP (for example 5.000 iterations).

The final step is to calculate VAR and Expected Shortfall from the simulated probability distribution of portfolio returns, using the correspondent quantile to the confidence level that is chosen.

Monte Carlo simulation approach has the advantage of more closely replicating the distribution of portfolio returns and thus produces more reliable estimates of VAR and Expected Shortfall. Its serious drawback is that it is computationally slow and the underlying assumption of multivariate normally distributed market variables does not account for fatter tails. This last drawback can be partly overcome by assuming that market variables follow a multivariate t distribution.

3.4 Methods Capturing Volatility Dynamics

All methods examined so far assume that financial returns in the estimation window follow a specific distribution with an unconditional constant volatility. An important characteristic of financial markets though is that volatilities of market factors tend to cluster, especially during stressed periods (Manganelli & Engle, 2001: 8). The inability of simple, but at the same time inefficient, methods of historical simulation and normal distribution to produce reliable VAR and Expected Shortfall estimates after the 1996 BCBS Amendment, led to the advent of academic research in the application of models capturing volatility dynamics in VAR and Expected Shortfall estimation. These studies were based on Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized

Autoregressive Conditional Heterscedasticity (GARCH) models introduced by Engle (1982) and Bollerslev (1986) respectively.

Financial returns (r_t) are said to follow an ARCH(q) process when:

$$y_t = \mu + \varepsilon_t$$
$$\varepsilon_t = \sigma_t z_t \qquad \text{where } z_t \sim i.i.d. \text{ with } E(z_t) = 0 \text{ and } Var(z_t) = 1$$
$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2$$

μ is the mean of returns and ε_t the innovation process. For the conditional variance to be positive, the parameters must satisfy: $\alpha_0>0$ and $\alpha_i \geq 0$ for i=1,...,q.

Empirical evidence has shown that a high q must be selected in order to estimate the conditional variance properly. This characteristic led Bollerslev (1986) to propose the generalized ARCH, or GARCH(p,q), model:

$$y_t = \mu + \varepsilon_t$$
$$\varepsilon_t = \sigma_t z_t \qquad \text{where } z_t \sim i.i.d. \text{ with } E(z_t) = 0 \text{ and } Var(z_t) = 1$$
$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$$

For the conditional variance to be positive, the parameters must satisfy: $\alpha_0>0$, $\alpha_i \geq 0$ for i=1,...,q and $b_j \geq 0$ for j=1,...,p.

A GARCH(p,q) model successfully captures several characteristics of financial returns such as thick tails and volatility clustering (Angelidis, Benos & Degiannakis, 2004: 3).

In the conditional mean equation of the above model, financial returns might be considered to be an AR(k) autoregressive process, so the GARCH(p,q) model would be transformed to an AR(k) GARCH(p,q) model:

$$y_t = \mu + \sum_{\theta=1}^{k} c_\theta y_{t-\theta} + \varepsilon_t$$
$$\varepsilon_t = \sigma_t z_t \qquad \text{where } z_t \sim i.i.d. \text{ with } E(z_t) = 0 \text{ and } Var(z_t) = 1$$
$$\sigma_t^2 = a_0 + \sum_{i=1}^{q} a_i \varepsilon_{t-q}^2 + \sum_{j=1}^{p} b_j \sigma_{t-j}^2$$

This specification allows for possible autocorrelation in daily returns.

In the above models, though, variance only depends on the magnitude and not the sign of ε_t, which somewhat at odds with the empirical behavior of financial returns (especially stock returns), where the 'leverage effect' might be present (Xekalaki & Degiannakis, 2010: 42). The 'leverage effect' was introduced by Black (1976). Black noticed that changes in stock returns often display a tendency to be negatively correlated with changes in volatility. This means that volatility tends to rise in response to bad news and fall in response to good news. The most popular model proposed to capture this asymmetry is Nelson's (1991) exponential GARCH or EGARCH(p,q). The model is plainly presented by Angelidis, Benos & Degiannakis (2004: 4) as follows:

$$y_t = \mu + \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim i.i.d. \text{ with } E(z_t) = 0 \text{ and } Var(z_t) = 1$$

$$\ln(\sigma_t^2) = a_0 + \sum_{i=1}^{q} \left(a_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right) + \sum_{j=1}^{p} \left(b_j \ln(\sigma_{t-j}^2)\right)$$

In contrast to GARCH, no restrictions are imposed on the parameters of the model estimation, since logarithmic transformation ensures the non-negativity of variance.

In such a model, the presence of leverage effect can be investigated by testing the hypothesis that $\gamma<0$.

The parameters of all the above models can be estimated by the maximum likelihood estimation method. Since parameters are to be estimated, models capturing volatility dynamics are classified, by many researchers, as a special part of parametric methods.

Many researchers have proposed the calculation of VAR and Expected Shortfall risk measures based on conditional variance models (Kellner & Rosch (2016), Righi & Caretta (2014), Degiannakis, Floros & Livada (2011) to name a few). They use the following formulas for the calculation of VAR and Expected Shortfall forecasts at time t:

$$VAR_{\alpha}^{t+1} = \mu_{t+1} + \sigma_{t+1} VAR_a(\varepsilon_{t+1})$$
$$ES_a^{t+1} = \mu_{t+1} + \sigma_{t+1} ES_a(\varepsilon_{t+1})$$

where $VAR_a(\varepsilon_{t+1})$and $ES_a(\varepsilon_{t+1})$are the forecasted VAR and Expected Shortfall of the innovation distribution respectively. As the innovation process is assumed to follow a specific distribution like standard normal or standardized student's t, we can substitute $VAR_a(\varepsilon_{t+1})$ and $ES_a(\varepsilon_{t+1})$ with $VAR\varepsilon_\alpha$ and $ES\varepsilon_\alpha$ respectively, with α representing the upper quantile of the innovation distribution that by assumption does not depend on t. With the use of a rolling window of certain width, the coefficients of a model are estimated each time t on the previous x observations (belonging to the window) and σ_{t+1} is estimated using one step ahead forecast.

We can observe that the forecasted risk measure at time t relies on the calculation of the forecasted mean, the forecasted standard deviation of returns from the conditional variance model and the upper quantile of the innovation distribution.

As for the innovation distribution it can be assumed to be the standard normal. The maximum likelihood estimation is based on this assumption. McNeal & Frey (2000: 5) propose a standardized t distribution with v>2 degrees of freedom, to account for fat tails. GARCH type models with t innovations can also be fitted by maximum likelihood and the parameter of v can be estimated.

In order to analyze the existence of correlation in volatility, which would indicate that a conditional variance model is appropriate to be adapted, Xekalaki & Degiannakis (2010: 35) propose the analysis of a proxy for variance, which is squared demeaned log-returns. A histogram of these returns would be indicative of time varying variance.

For a formal test of the above, Pindyck & Rubinfeld (1998: 496) propose the use of Box - Pierce statistic to assess if the k autocorrelations of squared log returns are statistically significant. The statistic is calculated as follows:

$$Q = T\sum_{\kappa=1}^{K} \rho_{\kappa}^{2}$$

, where T is the number of observations and ρ_κ is the k^{th} autocorrelation coefficient. This statistic is approximately distributed as a chi-square with K degrees of freedom. If we reject the null hypothesis at the 5% level of significance, we can be 95% certain that K autocorrelation coefficients are not all zero.

Berkowitz & O'Brien (2001) examined the statistical accuracy of large bank holding companies VAR forecasts in the period 1998-2000. They compared the VAR forecasts from banks' structural models with those from a standard GARCH model and concluded that the latter provided more accurate forecasts.

Kellner & Rosch (2016) fitted conditional variance models to returns from three indices and two exchange rates between 2001-2015, to produce one day ahead VAR and Expected Shortfall forecasts. Their results indicated that Expected Shortfall (at the 97,5% confidence level) is higher than VAR (at the 99% confidence level), which leads to higher capital requirements, something that is in line with the scope and intention of recent changes in Basel regulatory framework. They also concluded that heaviness in the tails of the innovation distribution can lead to a trade-off between capturing extreme events on one hand and the increase of model risk in Expected Shortfall estimation on the other.

Righi & Caretta (2015) analyzed data from equity, fixed income, exchange rates and commodities markets from 2000 to 2012. They tested the performance of non parametric, parametric and conditional variance models with fat-tailed innovation distributions on Expected Shortfall forecasts. They concluded that the non parametric historical simulation approach reacts slowly to the 2008 crisis and after that retains conservative forecasts. The conditional variance models, on the other hand, follow market fluctuations more efficiently, considering the crisis but recovering after it.

Degiannakis, Floros & Livada (2011) used data from mature (USA, UK, Germany) and emerging (Greece and Turkey) equity markets for the period 2004-2008. They produced VAR forecasts based on three conditional variance models for the period 2004-2007 and for the year 2008 separately, as the authors wanted to test the models in a small period of market turbulence. Overall, their findings suggest that conditional variance models gave satisfactory results before and during the highly volatile year of 2008.

Angelidis, Benos & Degiannakis (2003) conducted VAR forecasts for five equity index returns (CAC40, DAX30, FTSE100, NIKKEI225 and S&P) over a period of 1987 to 2002, using various conditional variance models. They conclude that innovation distributions allowing for fatter tails produce better results. Furthermore, for each index portfolio, a different model produces the most accurate forecasts.

Finally, Degiannakis, Floros & Dent (2013) tested various conditional variance models on their ability to forecast VAR and Expected Shortfall accurately and they reached the conclusion that long memory volatility models does not appear to outperform a short memory GARCH model, even for longer forecasting horizons.

Academic literature on VAR estimating techniques and their reliability has by far been more developed than Expected Shortfall equivalents. Various types of more sophisticated models like conditional volatility models have been recently examined for their accurate risk measure estimates. According to Angelidis, Degiannakis & Floros (2011: 446), the suitability of a model depends on the examined period and the market characteristics.

3.5 The Filtered Historical Simulation Approach

The most important drawback of historical simulation approach is that it considers past observations as a set of independently and identically distributed returns that can be applied to current asset prices and simulate their future returns (Barone-Adesi, Bourgoin & Giannopoulos, 1998: 100).

Filtered historical simulation is a hybrid, semi-parametric approach that combines the benefits of historical simulation with the power of conditional volatility models such as GARCH and circumvents the above drawback of historical simulation. It does so by bootstrapping from a distribution of independently and identically distributed standardized residuals which have been scaled by their conditional volatility.

The first step in the process is to fit a conditional variance model to the data. Barone-Adesi, Bourgoin & Giannopoulos (1998:100) use a GARCH model that allows for asymmetric behavior in the conditional variance equation. The specification they use is depicted below:

$$r_t = \mu + \varepsilon_t$$
$$\sigma_t^2 = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta\sigma_{t-1}^2$$

where γ is an additional term that allows for a surprise to have an asymmetric positive or negative effect on volatility.

The second step is to extract the fitted residuals and to use the model to forecast volatility for each day in the sample period. Then, standardized residual returns of the following form set the series of independently and identically distributed innovations:

$$z = {\widehat{\varepsilon}_t}\big/{\widehat{\sigma}_t}$$

If we assume a 1day VAR holding period, the third step involves bootstrapping with replacement from the set of standardized innovations. The independently and identically distributed property of z is important for bootstrapping, and allows us to avoid sampling from a population where observations are serially dependent. We then multiply a bootstrapped standardized innovation with the square root of our model forecast of tomorrow volatility in order to produce a forecast of tomorrow innovation:

$$\varepsilon_{t+1}^* = z\sqrt{\widehat{\sigma}_{t+1}^2}$$

Thus a simulated innovation is created by a randomly drawn (with replacement) standardized residual z, rescaled by next period's forecasted volatility, in order to reflect current market conditions. We then substitute this simulated innovation in the conditional variance equation of our model in order to create a forecasted portfolio return r_{t+1}.

In the fourth step we create a large number (M) of simulated returns by taking M drawings from the standardized residual return distribution.

Finally the last step involves the calculation of VAR and Expected Shortfall from the simulated distribution of portfolio returns. Giannopoulos & Tunaru (2005) proved that filtered historical simulation can be used for the estimation of Expected Shortfall and that this estimator is a coherent measure.

Filtered historical simulation can produce risk measures consistent with the current states of markets at any large confidence level (Barone-Adesi & Giannopoulos, 2002: 179). Its advantages allow it to dominate classic historical simulation approach easily and compete with models capturing volatility dynamics, as it simulates portfolio returns according to existing market conditions. Indeed, Nozari, Raei, Jahangiri & Bahramgiri (2010) used returns of European and Asian emerging market indices to compare the performance of selected VAR methods and concluded that the filtered historical simulation model outperformed a classic GARCH(1,1) model with a t-distribution innovation process.

BLANK PAGE

Chapter 4
Backtesting

As many banks adopt the internal models approach to estimate their market risk regulatory capital, a technique is needed for setting benchmarks for the evaluation of these models. Backtesting is a reality check that compares a bank's daily VAR measure with the subsequent daily portfolio profit or loss (BCBS, 1996b: 2). This is a test of how the method for calculating VAR would have performed if it had been used in the past (Hull, 2012: 197).

The importance of the backtesting procedure for the internal models approach was recognized by the Basel Committee on Banking Supervision under the publication of the backtesting supervisory framework (BCBS, 1996b) in the same year with the setting of the market risk capital requirements framework (BCBS, 1996a). In this way regulators were able to assess the efficiency of bank market risk models due to certain benchmarking criteria and intervene timely and accurately to prevent insufficient market risk capital coverage. Due to the Committee, backtesting acts like a controlling mechanism that sets certain limits to the internal models approach (BCBS, 1996b).

The drawbacks presented by VAR and the incorporation of Expected Shortfall risk measure in the Basel regulatory framework in 2012 (BCBS, 2012: 20), called for the development of a backtesting framework based on this measure. The discovery, though, that Expected Shortfall is not elicitable (Gneiting, 2011), sparked an academic debate as to whether it is backtestable. This is the reason why Basel Committee on Banking Supervision has retained VAR in its backtesting framework (BCBS, 2016: 95).

Acerbi & Szekely (2014) proved that Expected Shortfall's lack of elicitability property does not impede efficient backtesting. They also proposed three backtesting methods.

In this chapter various methods for VAR and Expected Shortfall backtests are presented. Regarding VAR, the Traffic Light test outlined by the Basel Committee on Banking Supervision is described. Subsequently, unconditional and conditional tests proposed by Kupiec (1995) and Christoffersen (1998) respectively are analyzed. Finally, two of the three Expected Shortfall backtesting methods proposed by Acerbi & Szekely (2014) are presented.

4.1 Backtesting Value at Risk

Value at Risk backtesting refers to the frequency of portfolio losses exceedances over VAR. If a daily loss exceeds the one day 99% VAR estimated under a certain model, this observation is called an exception. If the number of exceptions is statistically equal to the 1% worst case losses of the portfolio loss distribution over our data sample, then the model can be considered as accurate.

If we denote the daily loss of a portfolio as $x_{t,t+1}$, then we can define the 'hit' function as follows (Campbell, 2015: 3):

$$I_{t+1(a)} = \begin{cases} 1\_if\_x_{t,t+1} \geq VAR_t(a) \\ 0\_if\_x_{t,t+1} < VAR_t(a) \end{cases}$$

where α is the upper α% quantile of the daily portfolio loss distribution. The sequence of the hit function (e.g. 0,0,1,0,0,….,1) is the information criterion of the loss exceedances over VAR.

Mehta, Neukirchen, Pfetsch, & Poppensieker, (2012: 16) analyze the performance of VAR models based on the following evaluation criteria:

- Accuracy: The absolute difference between VAR estimate and daily portfolio actual loss. An accurate model traces daily portfolio return fluctuations closely, being highly reactive in market movements.
- Stability: The change in VAR from day to day should not be overly reactive to small short-term changes in market conditions.

- Outliers: The number of times in a given period that actual losses exceed VAR estimates.

The backtesting methods presented below examine the third criterion, which is the observed percentage of exceptions in comparison with the theoretical one.

4.1.1 Regulatory Framework – The Traffic Light Test

As previously mentioned, VAR backtesting regulatory framework was developed simultaneously with the market risk capital requirements standards by the Basel Committee on Banking Supervision in 1996. This framework is known as the 'Traffic Light Test' and has been retained by the Committee till today.

The primary reference point in the process is the use of the number of exceptions as the basis for appraising a bank's model. According to the Committee (BCBS, 1996b: 5), the appealing characteristic of this procedure is its simplicity and straightforwardness as well as its relatively few strong assumptions. The primary assumption in the process is the independent outcomes of the 'hit' function. More specifically, if we define α as the α% quantile of the return loss distribution above VAR level, then the 'hit' sequence $I_t(\alpha)$ is assumed to be independently and identically distributed as a Bernoulli random variable with probability α (Campbell, 2015: 4):

$$I_t(\alpha) \overset{i.d.d.}{\sim} B(\alpha)$$

Assuming a binomial distribution, with probability α indicating the theoretical percentage of exceedances over VAR (e.g. 1%), n indicating the number of days in the observation period and m the number of exceedances, the probability of the percentage of exceedances being greater that α can be found using the formula below:

$$\sum_{k=m}^{n} \frac{n!}{k!(n-k)!} a^k (1-a)^{n-k}$$

Considering a level of coverage of 99% (α=1%) and a sample of 250 observations, the number of exceptions and the corresponding cumulative probability, are presented in the following table:

Zone	Number of exceptions	multiplier	Cumulative probability
Green zone	0	1.50	8.11%
	1	1.50	28.58%
	2	1.50	54.32%
	3	1.50	75.81%
	4	1.50	89.22%
Yellow zone	5	1.70	95.88%
	6	1.76	98.63%
	7	1.83	99.60%
	8	1.88	99.89%
	9	1.92	99.97%
Red zone	10 or more	2.00	99.99%

Table 1. The Basel Committee on Banking Supervision Traffic Light test (BCBS, 2016: 77).

At the 5% confidence level we can see that a model exhibiting 5 or more exceptions is rejected from being accurate (at the green zone).

The range from 5 to 9 exceptions constitutes the yellow zone. In this zone the accuracy of a model is questionable. As the number of exceptions tends to reach the lower bound of 9, the accuracy is severely deteriorated.

Finally the red zone starts from the point of 10 or more exceptions. In this zone the framework dictates supervisory authorities to investigate the reasons of the poor model performance and require the bank to improve its accuracy.

In the table above we can observe that the multiplication factor is increased from 1.5 in the green zone to 2 in the red zone. This multiplication factor leads to greater market risk capital requirements as exceptions increase. According to the latest published

Standards of the Committee (BCBS, 2016: 64), the market risk capital charge is calculated as follows:

$$C_A = \max(IMCC_{t-1} + SES_{t-1}; m_c IMCC_{avg} + SES_{avg})$$

where IMCC is the aggregate capital charge for modellable risk factors and SES is the aggregate regulatory measure for K risk factors in model eligible desks that are non modellable.

It is worthwhile to mention that until the second consultative document of the Committee's 'Fundamental Review of the Trading Book' framework (BCBS, 2013), the multiplication factor for the green zone had been set to 3, scaling up to 4 in the red zone.

4.1.2 Unconditional Coverage Test – Kupiec Test

To test whether the probability of realizing a loss in excess of the reported $VAR_t(\alpha)$ must be precisely α X 100%, or that the cumulative probability of the hit sequence is equal to α $[Pr(I_{t+1}(\alpha)=1)=\alpha]$, Kupiec(1995) proposed a powerful two tailed test. Denoting α as the theoretical probability of an exception and m the number of exceptions observed in n trials, the following statistic should be distributed asymptotically as a chi-square distribution with one degree of freedom:

$$2\ln[(1-\frac{m}{n})^{n-m}(\frac{m}{n})^{m}] - 2\ln[(1-a)^{n-m}a^{m}] \sim x^2(1)$$

In the case where exceptions occur more frequently than the reported VAR suggest, then this would suggest that the VAR measure understates the actual level of risk. On the contrary, fewer exceptions than those suggested by α would be a signal of conservative VAR measure. The Kupiec's two tailed test of unconditional coverage reject the VAR model for either lower or greater percentage of failures than those implicit in the VAR confidence level.

4.1.3 Conditional Coverage Test – Christoffersen Test

The traffic light and unconditional coverage tests make the assumption that VAR violations are independent with each other. This means that two elements in the 'hit' sequence ($I_{t+g}(\alpha)$, $I_{t+k}(\alpha)$) do not exhibit bunching, requiring that previous VAR violations must not convey any information for an additional VAR violation. In practical terms, if volatility is low in some period and high in others, VAR forecasts should respond accurately to changing market conditions, not presenting clusters.

Christoffersen (1998: 845) proposed a test for the independence of VAR violations. We denote u_{ij} as the number of observations in which we go from a day we are in state i to a day we are in state j. State 0 is a day with no VAR violation and 1 is a day where there is a violation. The test of VAR violation independence can be conducted with the use of the following likelihood function which follows a chi distribution with 1 degree of freedom:

$$LR_{ind} = -2\ln[(1-\pi)^{u_{00}+u_{10}}\pi^{u_{01}+u_{11}}] + 2\ln[(1-\pi_{01})^{u_{00}}\pi_{01}^{u_{01}}(1-\pi_{11})^{u_{10}}\pi_{11}^{u_{11}}] \sim x^2(1)$$

where:

$$\pi = \frac{u_{01}+u_{11}}{u_{00}+u_{01}+u_{10}+u_{11}}$$

$$\pi_{01} = \frac{u_{01}}{u_{00}+u_{01}}$$

$$\pi_{11} = \frac{u_{11}}{u_{10}+u_{11}}$$

According to Christoffersen (1998: 846), we can test the combined hypothesis of unconditional coverage and correct conditional coverage using the following likelihood test statistic that follows a chi-square distribution with 2 degrees of freedom:

$$LR_{cc} = LR_{uc} + LR_{ind} \sim x^2(2)$$

where L_{uc} is the likelihood function used in the unconditional coverage test in the previous section. This approach enables us to test both coverage and independence hypothesis at the same time.

4.1.4 The Empirical Coverage Probability

Models which pass the conditional coverage test can be ranked according to an indicator of the relative frequency of violations (empirical coverage probability). The best model is the one with the lowest ratio between VAR violations and size of observations in the data sample window. A VAR measure must perform smaller number of violations than those implied by the VAR confidence level.

4.2 Backtesting Expected Shortfall

We have seen that Expected Shortfall was adopted in the Basel Committee on Banking Supervision market risk regulatory framework in 2012 (BCBS, 2012). Nevertheless, the Basel Committee preserved VAR as the measure to backtest in the common way. This was attributed to the scientific debate as to whether Expected Shortfall lack of elicitability would impede reliable backtesting. In addition to that, academic literature on expected shortfall backtesting had not been conclusive on certain methods.

In a pioneering article, Acerbi & Szekely (2014) concluded that elicitability of a risk measure is relevant only for model selection (i.e. comparison of different models forecasting the same process). From a regulatory point of interest, though, bank models are tested on an absolute validation basis (model testing), setting elicitability irrelevant for the backtesting process.

Furthermore, the authors proposed three non-parametric methods of backtesting Expected Shortfall which are distribution independent. In this section we present the first two tests, following from the representation of Expected Shortfall as a conditional and unconditional expectation respectively.

Lets denote X_t as a bank's portfolio loss in day t of a testing period (t=1,….,T), forecasted by a model predictive distribution P_t, also used to compute VAR and Expected Shortfall risk measures.

Test 1 of Acerbi & Szekely (2014) is derived from the conditional expectation representation of Expected Shortfall below:

$$ES_{a,t} = E(X_t | X_t > VAR_{a,t})$$

After some algebraic transformations we reach the following equation:

$$E[\frac{X_t}{ES_{a,t}} - 1 | X_t - VAR_{a,t} > 0] = 0$$

We now assume that VAR has already been tested and we want to separately test the magnitude of the verified violations against model predictions. Defining $I_t = (X_t - VAR_{a,t} > 0)$ as the indicator function of a VAR violation, the authors construct the following test statistic:

$$Z_1(\overrightarrow{X}) = \frac{\sum_{t=1}^{T} \frac{X_t I_t}{ES_{a,t}}}{N_T} - 1$$

where $N_T = \sum_{t=1}^{T} I_t > 0$

Under the null hypothesis H_0, the expected value of the test statistic is zero, whereas under the alternative hypothesis H_1 the expected value of the test statistic is positive. A possible acceptance of the alternative hypothesis would indicate risk underestimation.

$$E_{H_0}[Z_1 | N_T > 0] = 0$$
$$E_{H_1}[Z_1 | N_T > 0] > 0$$

From the above we can conclude that the conditional test has two parts: a VAR backtest and a standalone conditional test using the statistic Z_1. The conditional test accepts a model when these two tests accept it separately.

Test 2 follows from the representation of Expected Shortfall as an unconditional expectation of the following form:

$$ES_{a,t} = \mathrm{E}[\frac{X_t I_t}{1-a}]$$

From the above equation the authors construct the following statistic:

$$\mathrm{Z}_2(\overrightarrow{X}) = \frac{\sum_{t=1}^{T} \frac{X_t I_t}{ES_{a,t}}}{T(1-a)} - 1$$

The expected value of the test statistic under the null and alternative hypothesis is presented below:

$$E_{H_0}[Z_2] = 0$$
$$E_{H_1}[Z_2] > 0$$

In order to test the significance of the tests and accept or reject the models, M simulated scenarios of the N observations are made according to the model predictive distribution P_t ($X^S = (X_1^s, \ldots, X_t^s, \ldots, X_T^s)$ with $X_t^s \sim P_t$, $t = 1, \ldots, T$ and $s = 1, \ldots, M$). For each simulated scenario X^S we compute the test statistic of interest $Z_{1,2}^s$= $Z_{1,2}(X^s)$, s=1,…,M. The p-value is defined as the proportion of scenarios under which the simulated test statistic Z^s is larger than the test statistic evaluated through the observation period 1,…,T, defined as Z^{obs} ($Z^{obs} = Z(X_1, \ldots, X_T)$).

$$p = \frac{1}{M} \sum_{1}^{M} I(Z^s > Z^{obs})$$

where I is an indicator function with value 1 if $Z^s > Z^{obs}$ and zero otherwise. Defining p_{test} as the 1 minus the test confidence level, we reject the model if p< p_{test}.

According to the authors, critical values for Z_2 display remarkable stability across different distributional assumptions (eg. Normal or t with low degrees of freedom). This means that for the unconditional test, the above simulation procedure can be avoided.

BLANK PAGE

Chapter 5
Empirical Study Framework

In this part of the Thesis, the performance of various types of models in market risk estimation is evaluated through a stressed and a tranquil market condition period separately. The risk measures used in the analysis are VAR and Expected Shortfall, defined in the Basel Committee on Banking Supervision regulatory framework.

For the stressed market condition period, the US financial crisis, which started in 2007, was chosen. This crisis led large corporations such as Citigroup, UBS and Merrill Lynch to incur vast losses and Lehman Bothers to collapse. Banking regulatory framework was later accused of presenting serious inefficiencies that facilitated the escalation of events. Hull (2012: 132) considers regulatory arbitrage as the major inefficiency. Banks were actually induced to securitize mortgages, sell these to Special Purpose Vehicles, effectively setting them off their banking book and then buy created collateralized mortgage obligation tranches, moving these instruments to the trading book. This procedure significantly reduced their overall capital requirements. As referred in chapter 2, this issue was resolved in Basel 2.5 by the inclusion of an IRC (Incremental Risk Charge) in the market risk regulatory framework.

For the tranquil market condition period, the period following the US crisis was chosen. In this way, the comparison of the efficiency of certain models in accurately capturing different market conditions is made possible.

This chapter exhibits the purpose of the study, as well as the study limitations and methodology.

5.1 Purpose of the Study

This study aims at examining the efficiency of different models in estimating market risk in a stressed and a tranquil market condition period, using VAR and Expected Shortfall risk measures. Expected Shortfall is a risk measure officially introduced in the Basel Committee on Banking Supervision market risk regulatory framework in 2016 (BCBS, 2016) after two consultative documents in 2012 (BCBS, 2012) and 2013 (BCBS, 2013). Academic literature has recently begun to study the performance of market risk models using Expected Shortfall (e.g. Righi & Caretta (2015), Kellner & Rosch (2016)). Under its growing importance as risk measure capturing tail risk, this study uses Expected Shortfall in addition to VAR in order to examine the accuracy of market risk models using corresponding backtesting procedures, in stressed and tranquil market condition periods. The models examined are non parametric (Historical Simulation), parametric (Normal and t-distribution) and an EGARCH model capturing volatility dynamics. For the analysis, data of S&P index is used, simulating a univariate equity portfolio of a bank trading desk.

Historical Simulation has been criticized for slow reaction to fast changing market conditions (Pritsker (2005), Manganelli & Engle (2001)). Nevertheless, the vast majority of banking institutions adopt this method for estimating market risk capital requirements (EBA (2017), Perignon & Smith (2010)). In this study, the comparison of Historical Simulation VAR and Expected Shortfall estimation results with the equivalent of other models, especially during the US financial crisis, gives us insight into the level of possible market risk underestimation during this period.

The purpose of this study is to examine the extent at which certain market risk models and Expected Shortfall, as a newly introduced risk measure, could act as a preventing mechanism for upcoming crises in the banking industry that stem from market risk. The reliability of these models is examined with the use of VAR and Expected Shortfall backtesting techniques. Additionally, these models are also tested on their ability to track portfolio losses in a tranquil market condition period.

5.2 Study Limitations

Due to a great number of approaches proposed by academic literature for VAR estimation and the ongoing investigation in Expected Shortfall estimation methods, this study focus its analysis to certain models varying from simple and more applied (eg. Historical Simulation, Normal Distribution approach) to more sophisticated ones capturing heaviness in tail distribution and volatility clustering (EGARCH model).

Models based on Extreme Value Theory that estimate the tail of a distribution are not examined in this study. Extreme Value Theory is a field in risk management that focuses on the distinctiveness of high impact and low probability events and has recently gained considerable applicability in VAR and Expected Shortfall estimation methods (Harmantzis, Miao & Chien (2006), McNeil & Frey (2000), Nozari , Raei, Jahangiri & Bahramgiri (2010), to name a few studies).

Finally, to simplify the analysis, data is restricted in S&P equity index for two separate periods, indicating crisis and after crisis market conditions. This index consists of 500 US large company stocks, representing US equity market to a significant extent. As the study analysis is focused in the US financial crisis, the use of this index is intuitively justified as a simulation of a bank's equity trading desk.

5.3 Study Methodology

In this section, the methodology for the estimation of VAR and Expected Shortfall risk measures according to certain models is described. The methodology is based on certain assumptions made for the following factors affecting the estimation results:

- Forecast time horizon: A time horizon (holding period) of one day is selected for VAR and Expected Shortfall forecasts, consistent with the Basel Committee on Banking Supervision market risk regulatory framework.

- Rolling window: This is the subsample observation length used to estimate the model parameters. The rolling window used in this analysis is 250 daily observations, following the Basel Committee on Banking Supervision VAR backtesting procedure, which requires this length of the window for the purposes of backtesting (BCBS, 2016: 73). For example we estimate the VAR and Expected Shortfall for the first time on the 251st day based on the last 250 returns. The window is moved one day forward and using the data from day 2 to 251 we estimate VAR and ES for day 252.

- Confidence level: VAR and Expected Shortfall are estimated for each model using a confidence level of 97,5% and 99% for VAR and 97,5% for Expected Shortfall. According to the most recent market risk regulatory standards issued by the Basel Committee, banks are required to estimate for each business day and for each trading desk two daily VARs corresponding to a one tail 97,5% and 99% confidence level and a daily Expected Shortfall that corresponds to a 97,5% confidence level (BCBS, 2016: 56).

Apparently, the above assumptions are also used for VAR and Expected Shortfall backtesting procedures.

The analysis is conducted through MATLAB and Eviews.

Chapter 6
Empirical Results – VAR and Expected Shortfall Estimation

After presenting the scope, limitations and methodology of the Thesis, this chapter exhibits the data analysis and the VAR and Expected Shortfall estimation results of five types of models in the two separate periods chosen. Risk measure estimation procedure contains a certain assumption about the distribution and properties of portfolio returns (Manganelli & Engle, 2001: 7) and this can lead to important discrepancies among model results.

Using a time horizon of one trading day, a rolling window of 250 days and certain confidence levels (97,5% and 99% for VAR and 97,5% for Expected Shortfall), estimation results are presented for each model for the crisis and after crisis period. The chapter concludes with the comparison of model results that produces valuable findings with respect to significant market risk underestimation of certain models during the US financial crisis.

6.1 Data Analysis

Data consist of negative daily log returns of the S&P index in the periods 2006-2010 (crisis period) and 2011-2015 (after crisis period) and were retrieved from Yahoo Finance database. For the estimation of VAR and Expected Shortfall, we are interested in examining positive losses and not negative profits in a hypothetical profit and loss distribution, so daily negative log returns are estimated. In the next two figures, these returns are presented for the crisis and after crisis period.

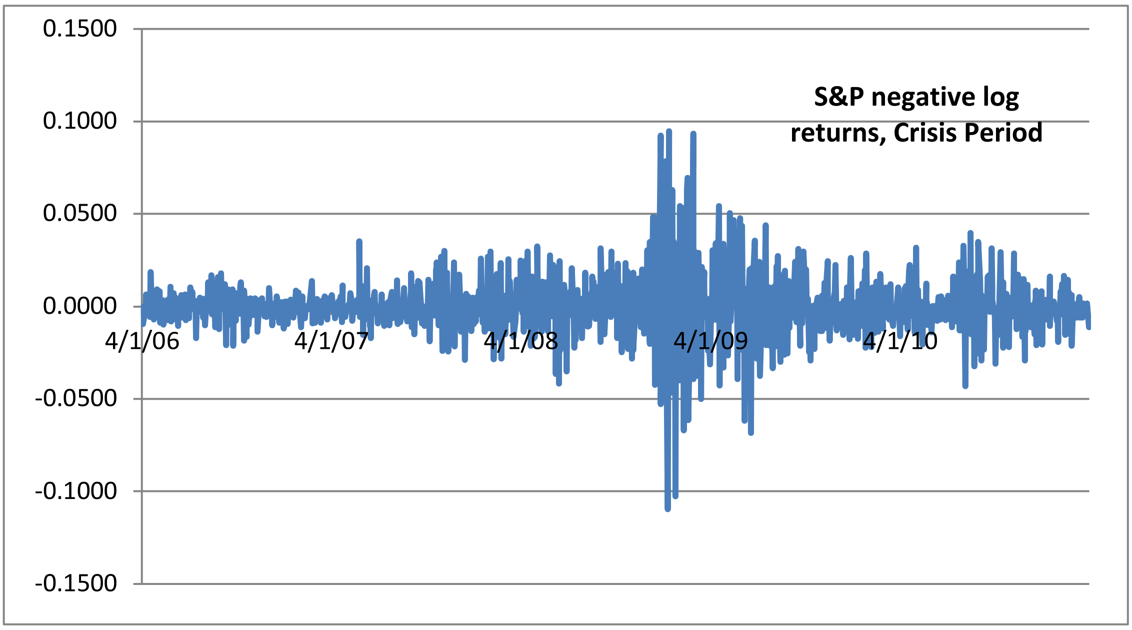

Figure 6. S&P negative log returns, Crisis Period.

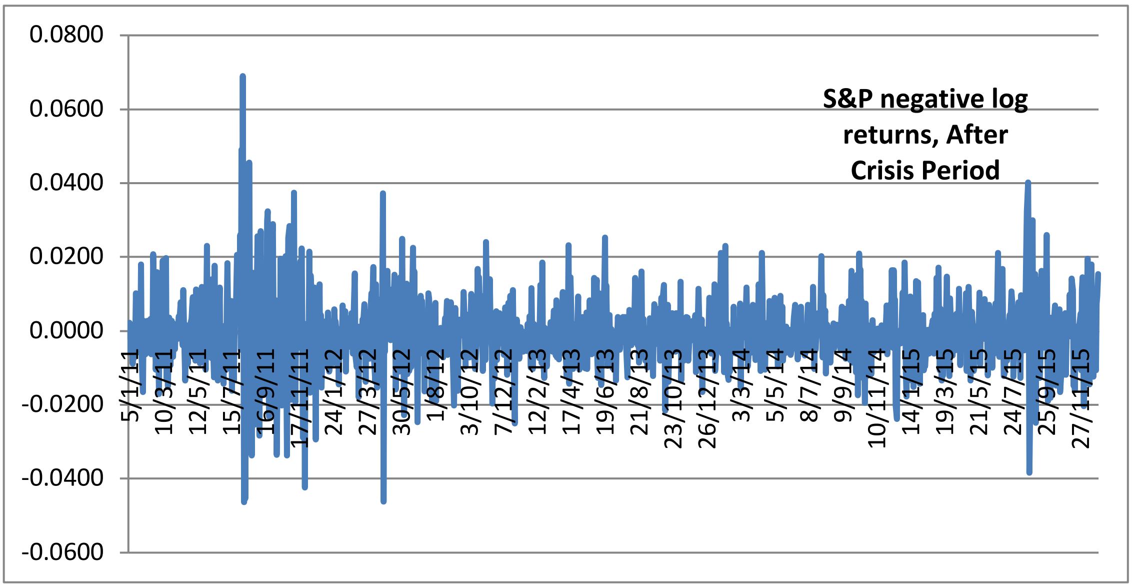

Figure 7. S&P negative log returns, After Crisis Period.

In figure 6 we can observe the peak of the crisis in mid year 2008, when it started to spiral out of control. In November 2008 the US Treasury Troubled Asset Relief Program was passed into law with the Emergency Economic Stabilization Act. Under this Program, US Treasury was given a large purchasing power in order to buy highly illiquid mortgage backed securities, thus reducing potential large losses for the financial institutions that owned them. S&P index return losses began to smooth out in the beginning of 2009.

In figure 7, the large return losses incurred in August 2011 were caused by stock prices fall in stock exchanges across United States, Europe, Asia and Middle East, due to fears of contagion of the European sovereign debt crisis. Not surprisingly, return losses in the after crisis period are of much lesser magnitude than in crisis period. In the next two figures, descriptive statistics for the S&P index return losses (negative log returns) for the two periods under study are exhibited.

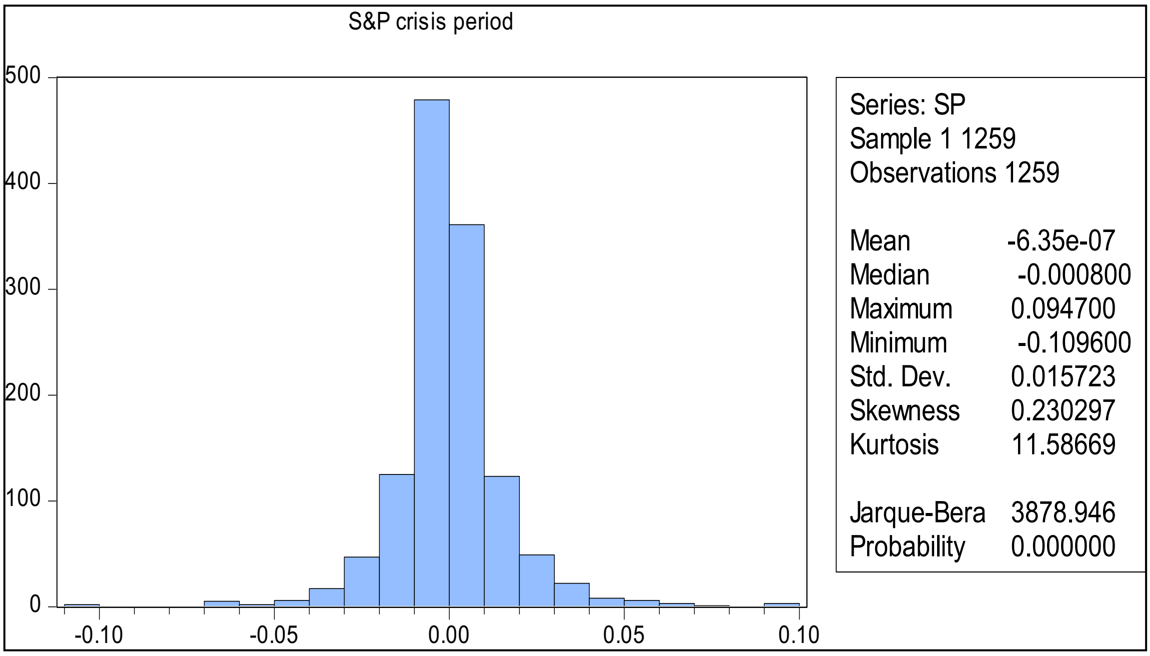

Figure 8. Descriptive statistics for S&P negative log returns, Crisis Period.

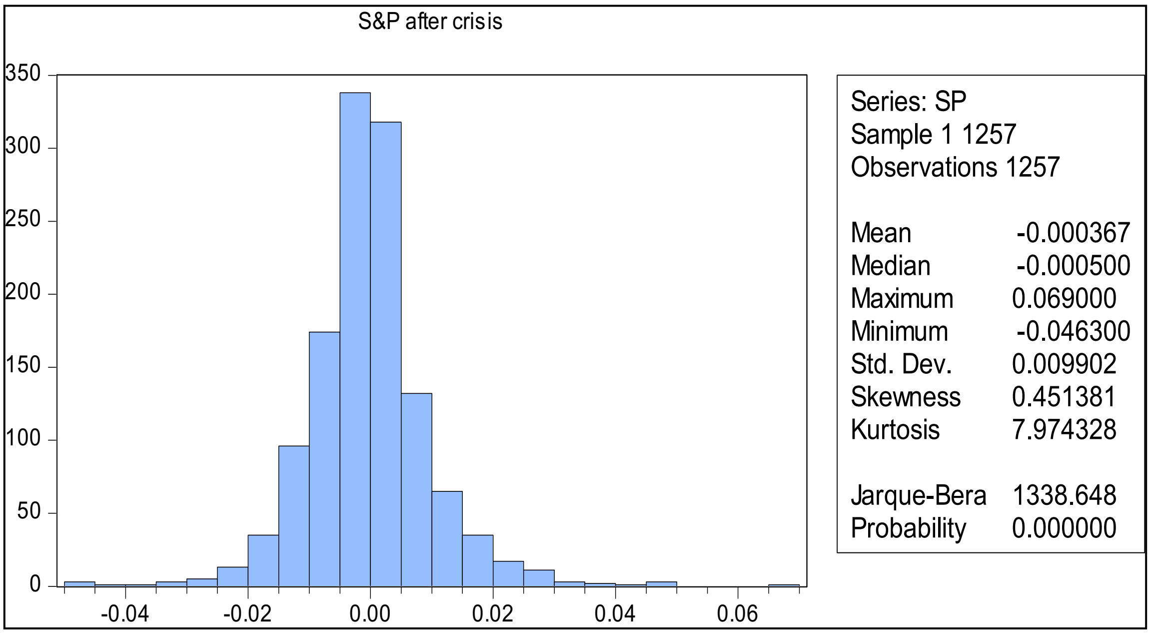

Figure 9. Descriptive statistics for S&P negative log returns, After Crisis Period.

The descriptive statistics show that mean is approximately equal to zero in both periods and distributions exhibit high kurtosis and are rightly skewed. Kurtosis coefficient is higher for return losses in crisis period, implying that large losses are more frequent. Additionally the Jarque-Bera statistic rejects the Normality condition in both periods. The heavier tail characteristic of return losses is depicted in the next quantile-quantile plots for both periods. It is obvious that the distribution is more leptokurtic in the crisis period.

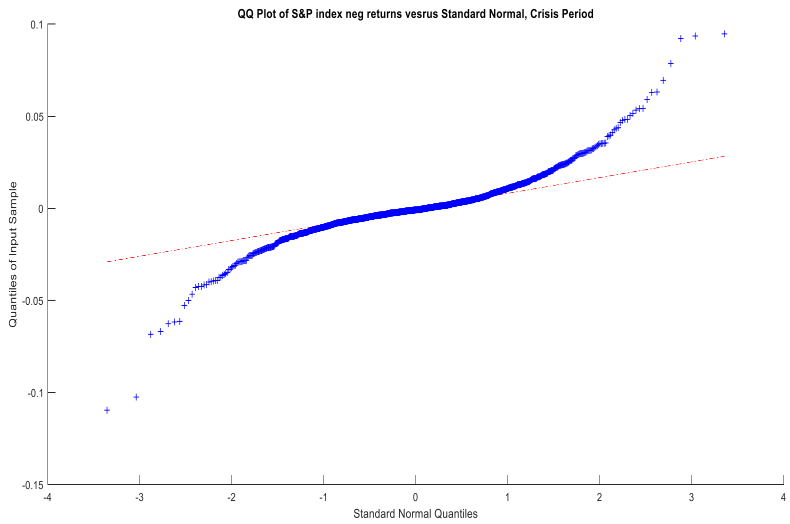

Figure 10. Quantile-Quantile plot, S&P index negative log returns, Crisis Period

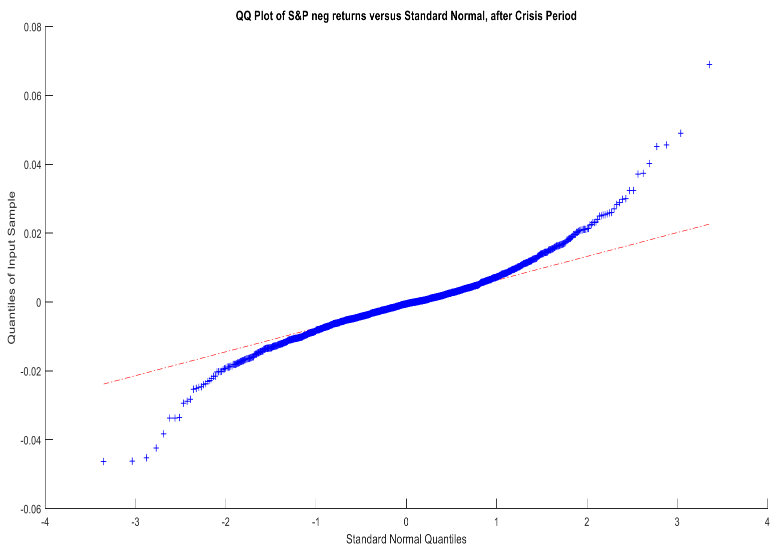

Figure 11. Quantile-Quantile plot, S&P index negative log returns, After Crisis Period

So far, return loss distributions for the two periods have presented two typical characteristic properties of financial returns: positive skewness and excess kyrtosis. One more characteristic to be examined is the autocorrelation of return volatilities, as these tend to cluster, especially during stressed market conditions (Manganelli & Engle, 2001). Squared log returns or squared demeaned log returns are popular proxies for daily volatility (Xekalaki & Degiannakis, 2010: 35).

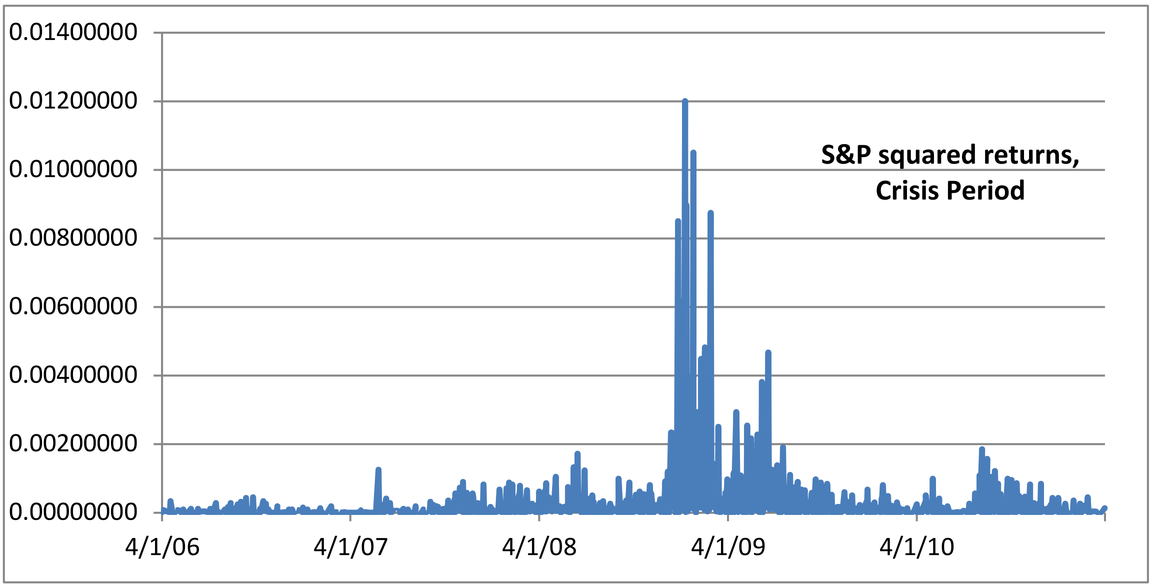

Figure 12. S&P squared returns, Crisis Period.

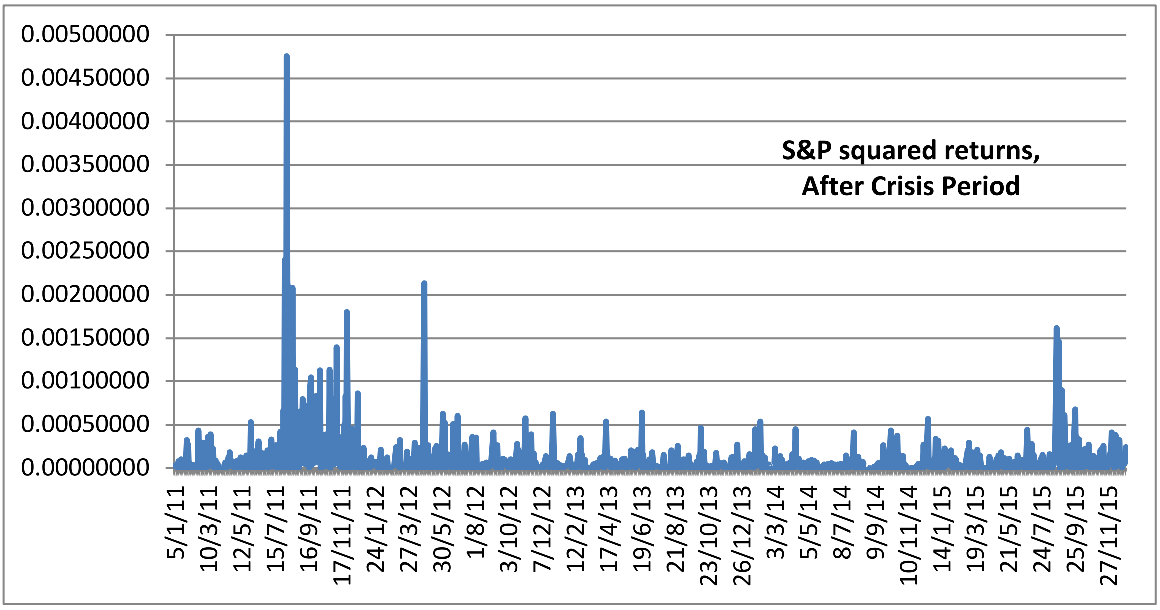

Figure 13. S&P squared returns, After Crisis Period.

We can observe that the volatility clustering property of returns is more apparent in the crisis period. Next we can test for the statistical significance of the squared return autocorrelations using the Box-Pierce statistic referred in chapter 3. In the next tables we can see that the null hypothesis of zero squared return autocorrelations is rejected due to values of the Box-Pierce statistic in both periods, implying an autoregressive pattern in daily volatilities.

S&P squared return correlogram, crisis period

Sample: 1 1259

Included observations: 1259

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|* |	|* |	**1**	0.189	0.189	45.048	0.000
|*** |	|*** |	**2**	0.388	0.365	234.71	0.000
|* |	| |	**3**	0.164	0.058	268.60	0.000
|** |	|* |	**4**	0.316	0.181	395.09	0.000
|** |	|** |	**5**	0.348	0.276	548.29	0.000
|** |	|* |	**6**	0.322	0.152	679.57	0.000
|** |	|* |	**7**	0.334	0.144	821.41	0.000
|** |	| |	**8**	0.220	0.020	883.04	0.000
|** |	|* |	**9**	0.324	0.108	1016.1	0.000
|** |	|* |	**10**	0.278	0.090	1114.0	0.000

Table 2. S&P squared return correlogram, Crisis Period.

S&P squared return corellogram, after crisis period

Sample: 1 1257

Included observations: 1257

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
|** |	|** |	**1**	0.301	0.301	113.99	0.000
|*** |	|*** |	**2**	0.425	0.368	341.88	0.000
|** |	|* |	**3**	0.269	0.098	433.02	0.000
|* |	| |	**4**	0.183	-0.047	475.19	0.000
|* |	| |	**5**	0.182	0.032	517.25	0.000
|* |	| |	**6**	0.163	0.071	550.97	0.000
|* |	| |	**7**	0.128	0.014	571.72	0.000
|** |	|* |	**8**	0.239	0.155	643.81	0.000
|* |	| |	**9**	0.093	-0.049	654.68	0.000
|* |	| |	**10**	0.159	0.000	686.63	0.000

Table 3. S&P squared return correlogram, After Crisis Period.

6.2 Crisis Period Models

In this section, the estimation results of four different models for daily VAR and Expected Shortfall during the crisis period are presented. Using 1259 observations of the S&P index for the period 2006-2010 and a rolling window of 250 trading days, the actual risk measure estimates are 1009. A confidence level of 97,5% for Expected Shortfall estimation and two confidence levels, 97,5% and 99%, for VAR estimation are used, following the recent Basel Committee market risk regulatory framework (BCBS, 2016: 56).

The models employed to simulate the loss function of S&P daily returns are Historical Simulation, Normal distribution, t distribution and an EGARCH model.

6.2.1 Historical Simulation

As presented in Chapter 3, this is a non parametric method that is based on the assumption that the actual distribution of portfolio losses in day t can be efficiently proxied by the empirical distribution of portfolio returns throughout the previous 250 return loss observations in the rolling window. The calculation of VAR and Expected Shortfall are based on this simplifying assumption. The estimation results for VAR are depicted in the next figure.

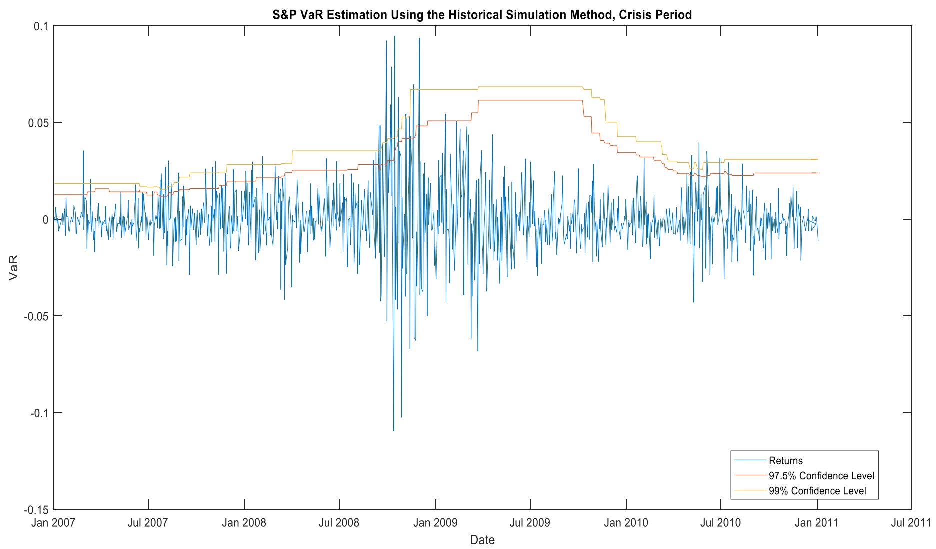

Figure 14. S&P VAR estimation under Historical Simulation method, Crisis Period.

We can observe that the Historical Simulation curve has a piecewise profile. The reason for this is that quantiles do not change for extended periods until extreme events occur. According to Perignon and Smith (2010: 367), 73% of banks that disclosed their VAR method used the Historical Simulation. More recently, in a benchmarking exercise by the European Banking Authority (EBA, 2017), 66% of participating banks were found to follow this methodological approach. Historical Simulation has been criticized for its inability to capture volatility clustering and react timely in fast changing market conditions. It is obvious that banks' market risk regulatory capital requirements were not accurately adjusted to restrict crisis escalating events.

Additionally, after the crisis peak in late 2008 and the subsequent introduction of the Troubled Asset Relief Program (TARP), VAR and Expected Shortfall estimates do not follow the reduction in return losses and volatility observed. Contrary to that, they follow a conservative pattern based on extreme observations in the last 250 day rolling window.

6.2.2 Normal Distribution

The normal or Gaussian distribution approach, is a parametric approach introduced by RiskMetrics (1996: 6) and is based on the assumption that return losses in the rolling window are normally distributed. Under this assumption the only parameters that have to be estimated for the risk measure estimation is distribution mean and standard deviation over the last 250 observations. Formulas for the derivation of VAR and Expected Shortfall for this method are straightforward, as already presented in Chapter 3. The only assumption we make is that mean return is zero for every risk measure calculation. The estimation results for VAR at 97,5% and 99% confidence levels are exhibited in the next figure.

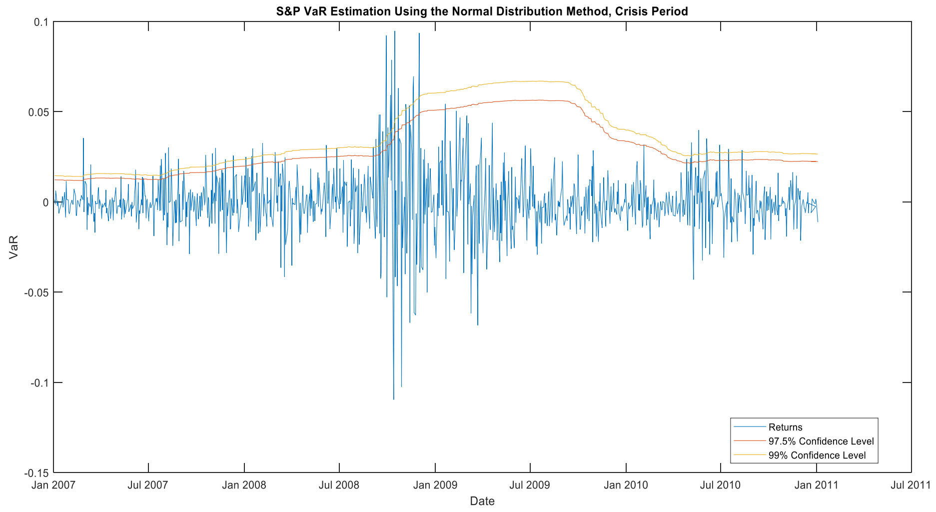

Figure 15. S&P VAR estimation under Normal Distribution method, Crisis Period.

Like the Historical Simulation approach, this method also seems to underestimate market risk during the crisis peak, as return losses surpass VAR estimates. Using standard deviation of the last 250 observations to calculate both risk measures, the Normal Distribution method presents the drawback of considering a stable volatility and moreover neglects any fat tailed return pattern. To partly overcome the second drawback, the t-distribution method is employed below.

6.2.3 Student's t Distribution.

The t-distribution approach is a parametric approach based on the properties of a t-distribution and allows for capturing more extreme losses than those implied by the normal distribution. A typical characteristic of financial returns is the excess than normal kurtosis of their distribution (fat tails). Defining the number of degrees of freedom in a t-distribution, we can adjust the kurtosis of the distribution to the observed data. In this Thesis the degrees of freedom ν are matched to an empirical kurtosis by setting ν as the integer closest to the value $\frac{4k-6}{k-3}$, where k is the kurtosis coefficient, following Dowd (2005: 160). For the crisis period this integer is 5.

Formulas for the calculation of VAR and Expected Shortfall for the t-distribution approach are also straightforward, like the normal distribution method, as presented in

Chapter 3. The estimation results for VAR calculation under t-distribution approach are exhibited in the figure below.

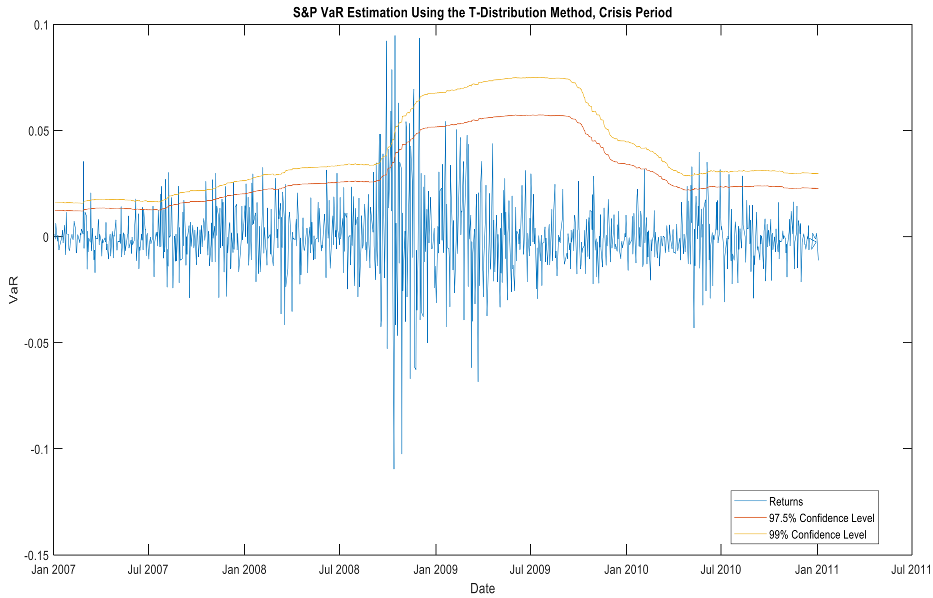

Figure 16. S&P VAR estimation under t Distribution method, Crisis Period.

The results show that VAR estimates account for extreme losses in the crisis peak period more efficiently than the two previous methods. Nevertheless, VAR violations are also of an important magnitude. After the crisis peak period, the model produces very conservative VAR estimates that follow a pattern of extreme returns and volatility observed in the previous 250 day period. These extreme values leave estimation results with a significant lag and this is the reason for the lack of risk measure accuracy under this approach.

6.2.4 EGARCH model

An important characteristic of financial return series is that its volatilities are not stable over time. Especially during stressed market conditions, return volatilities tend to cluster. The models examined so far do not account for this specific property. Instead they implicitly assume that volatility in the rolling window is stable.

In this section a model capturing volatility dynamics is employed. Specifically we adopt an EGARCH (1,1) model in order to capture thick tail and volatility clustering properties of S&P return losses. In addition to that, an EGARCH model also captures asymmetric effects in the conditional variance equation (Nelson, 1991: 354). The model specification used in the analysis is the following:

$$y_t = \varepsilon_t$$

$$\varepsilon_t = \sigma_t z_t \quad \text{where } z_t \sim i.i.d. \text{ with } E(z_t) = 0 \text{ and } Var(z_t) = 1$$

$$\ln(\sigma_t^2) = a_0 + (a_1 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}}) + b_1 \ln(\sigma_{t-1}^2)$$

More specifically, the innovation distribution is considered to be t-distributed with 5 degrees of freedom , to account for the fat tailed distribution property. The presence of leverage effect can be investigated by testing the hypothesis that $\gamma_1 > 0$, as our data consist of negative log returns.

First we proceed with the estimation of the model using all the observations in the crisis period (apart from the first 250 observations in year 2006). The results are the following:

S&P CRISIS PERIOD EGARCH MODEL EGARCH(1,1) Conditional Variance Model:			
Conditional Probability Distribution:			
Parameter	**Value**	**Standard Error**	**t-Statistic**
Constant	**-0.187897**	**0.0492172**	**-3.81772**
GARCH{1}	**0.978201**	**0.00568167**	**172.168**
ARCH{1}	**0.149686**	**0.0348218**	**4.29864**
Leverage{1}	**0.183574**	**0.0291854**	**6.28991**
DoF	**5**	**Fixed**	**Fixed**

Table 4. S&P Crisis Period EGARCH model estimation results

We can observe that all the coefficients are statistically significant. The leverage coefficient is positive. This coincides with the leverage effect theory of Black (1976), since our data consist of negative log returns. Consequently, a negative shock is implied by $\varepsilon_t > 0$ and volatility rises with respond to bad news.

The VAR and Expected Shortfall risk measure forecasts are estimated for each trading day according to the following formulas:

$$VAR_{\alpha}^{t+1} = \sigma_{t+1} VAR\,\varepsilon_a$$
$$ES_a^{t+1} = \sigma_{t+1} ES\varepsilon_a$$

where $VAR\,\varepsilon_a$ is the upper quantile of the innovation distribution that corresponds to a% confidence level and by definition does not depend on t. $ES\varepsilon_a$ is the equivalent Expected Shortfall of this distribution. σ_{t+1} is the EGARCH model forecast of next period volatility. This one day ahead volatility forecasts of the model are presented in the next figure.

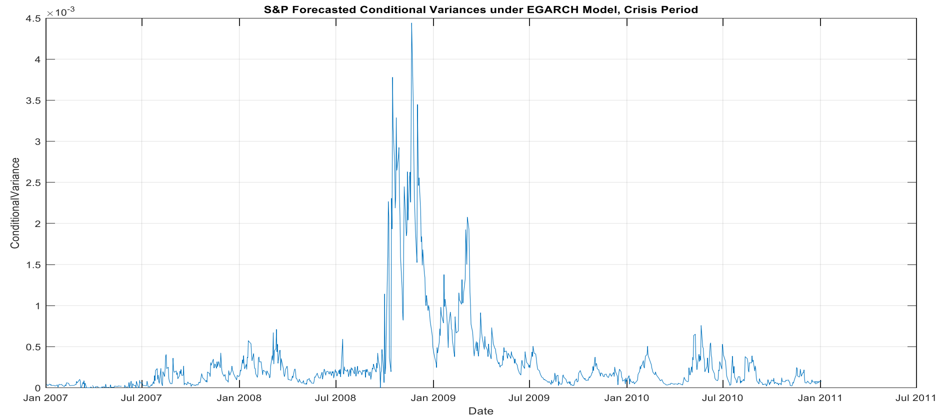

Figure 17. S&P forecasted conditional variances under EGARCH model, Crisis Period

From the above figure, volatility clustering in crisis period peak period is obvious. The estimation results for VAR according to the EGARCH model are presented below.

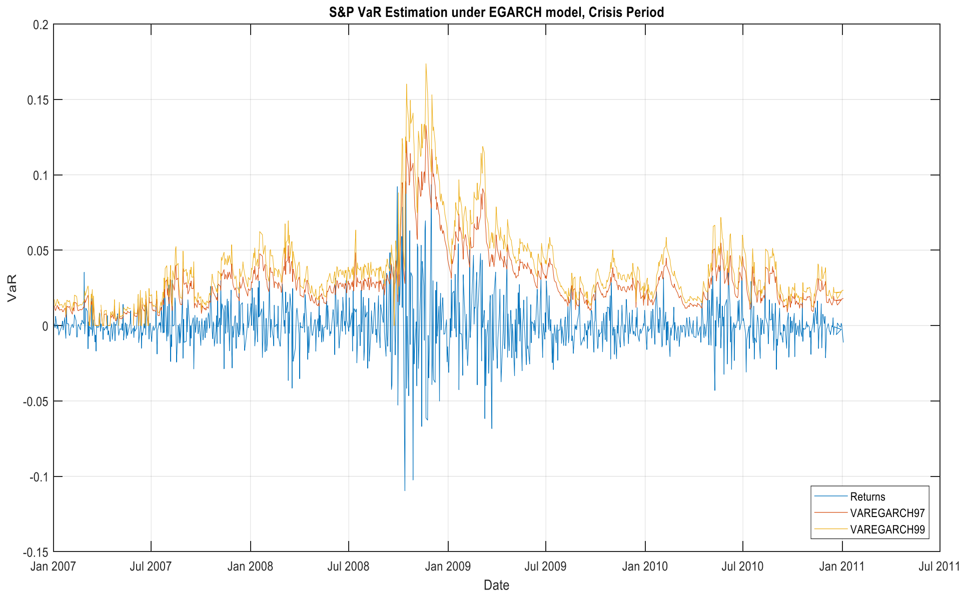

Figure 18. S&P VAR estimation under EGARCH model, Crisis Period.

VAR estimation under the EGARCH model seems to produce more accurate results than the previous models in our analysis. These estimates tend to track returns closely and account for changing volatility.

6.3 After Crisis Period Models

This section presents the results from VAR and Expected Shortfall estimation during the after crisis (tranquil) period. As we also want to examine the models for their efficiency to produce reliable risk measure estimates in a period with reduced market volatility and fewer extreme values, this section exhibits the level of model accuracy through equivalent plots. Data consists of daily negative log returns of the S&P index in period 2011-2015 (1257 observations). Using the defined rolling window of 250 trading days, 1007 VAR and Expected Shortfall estimates are produced for each model.

6.3.1 Historical Simulation

Being the approach most widely used for market risk estimation in the banking industry, the results of this model are very important for our analysis. In crisis period this method presented the serious drawback of following rather than confining the escalation of the crisis. The results of the VAR estimation in the after crisis period are presented in the following figure.

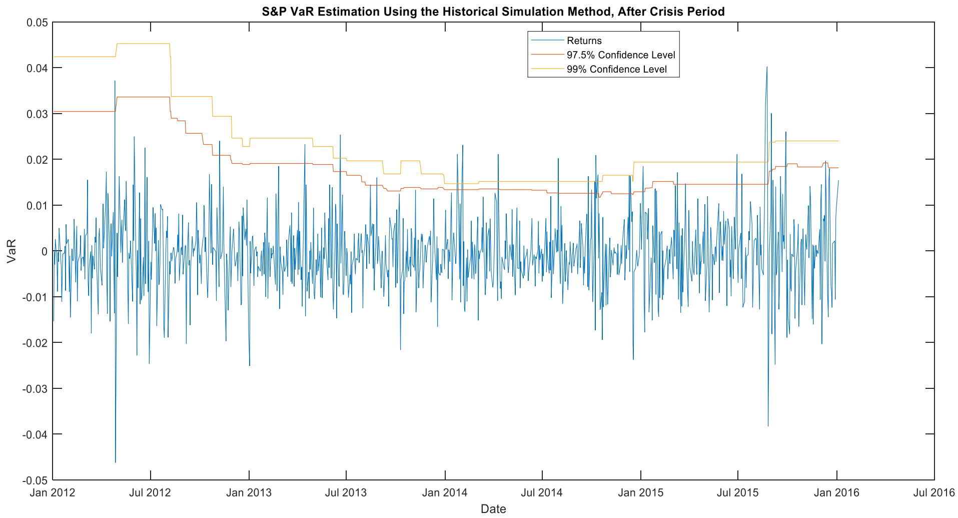

Figure 19. S&P VAR estimation under Historical Simulation method, After Crisis Period.

From the above figure we can observe that VAR violations under Historical Simulation are fewer in this period as there are less extreme return loss observations. VAR estimates from January 2012 to August 2012 clearly overestimate the underlying market risk, being influenced by the observations in the rolling window that belong to August 2011, a month with stressed conditions in stock markets in USA, Europe and Asia.

6.3.2 Normal Distribution

Normal distribution method VAR and Expected Shortfall estimates are based on the standard deviation of observations in the rolling window. Considering the reduced market volatility in after crisis period, we expect more reliable risk measure estimates in this period under this method. Next the VAR estimation results under the normal distribution method are presented.

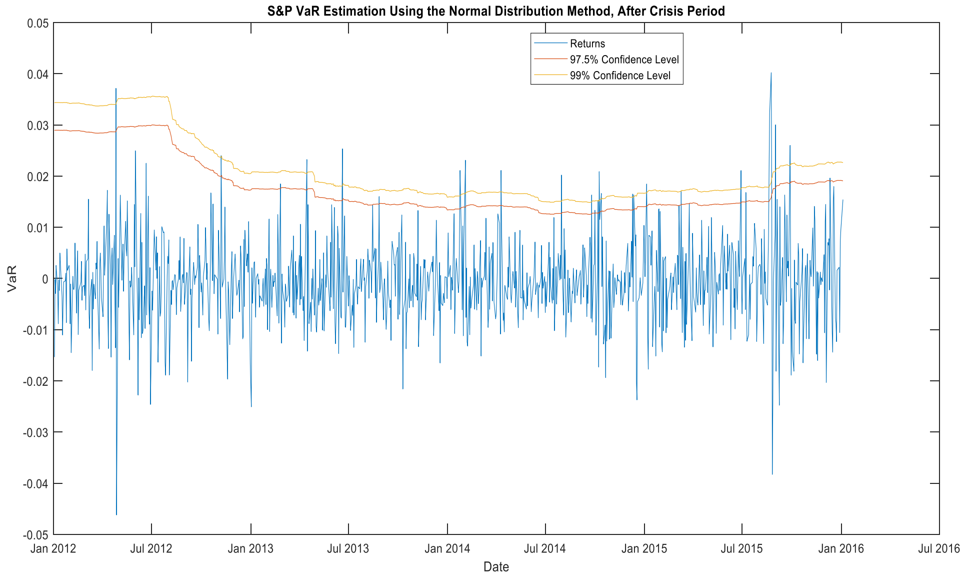

Figure 20. S&P VAR estimation under Normal Distribution method, After Crisis Period.

Clearly the VAR estimates in the beginning of 2012 are influenced by the increased volatility observed in August 2011, but this time to a lesser extent than their Historical Simulation equivalents. This method seems to produce more reliable estimates for the tranquil period than for the stressed one, where its late reaction to stressed conditions created many VAR breaks.

6.3.3 Student's t Distribution.

This method accounts for more extreme values in the tails and thus the excess kurtosis property of financial returns is captured to a considerable extent. In the data analysis section of the Thesis, the distribution of return losses in the after crisis period was presented to have lower kurtosis coefficient than in crisis period. Nevertheless, the excess than normal kurtosis is present in data and t-distribution must intuitively perform better than previous methods. Again the degrees of freedom ν are estimated according to Dowd (2005: 160) and defined to be 5.

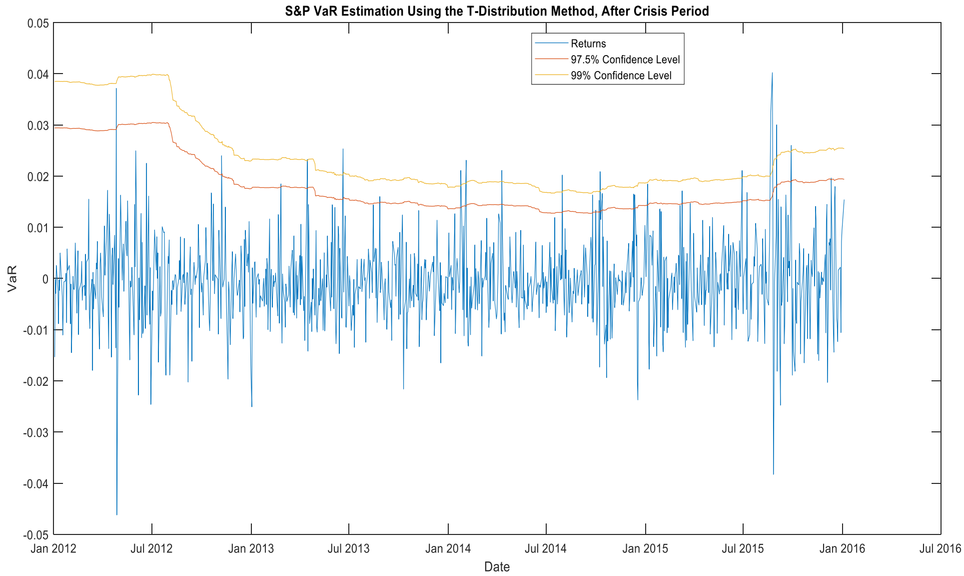

Figure 21. S&P VAR estimation under t Distribution method, After Crisis Period.

We can observe that VAR estimates are greater than the normal method in both confidence levels and VAR breaks at the 99% confidence level seem to be reduced by this method. The pattern of the t-distribution VAR estimation plot is similar to previous method plots in the first months of year 2012. It is also obvious that it performs much better in after crisis period than in crisis period.

6.3.4 EGARCH model

The EGARCH model efficiently captures the volatility clustering effect of our financial returns series and provides more accurate VAR estimates than other models in the crisis period, which means that tracks the fast changing market conditions more accurately. Next the estimation results of the EGARCH model using all the observations in after crisis period (except the first 250 observations in year 2011) are exhibited.

S&P AFTER CRISIS PERIOD EGARCH MODEL EGARCH(1,1) Conditional Variance Model:			
Conditional Probability Distribution:			
Parameter	Value	Standard Error	t-Statistic
Constant	-0.941507	0.13828	-6.80868
GARCH{1}	0.90414	0.0143389	62.8646
ARCH{1}	0.0514903	0.0430107	1.19715
Leverage{1}	0.383097	0.0368107	10.4072
DoF	5	Fixed	Fixed

Table 5. S&P After Crisis Period EGARCH model estimation results

We can observe that the ARCH coefficient is not statistically significant, possibly indicating that a conditional volatility model is not appropriate in a tranquil market condition period.

The one day ahead volatility forecasts of the model are presented in next figure.

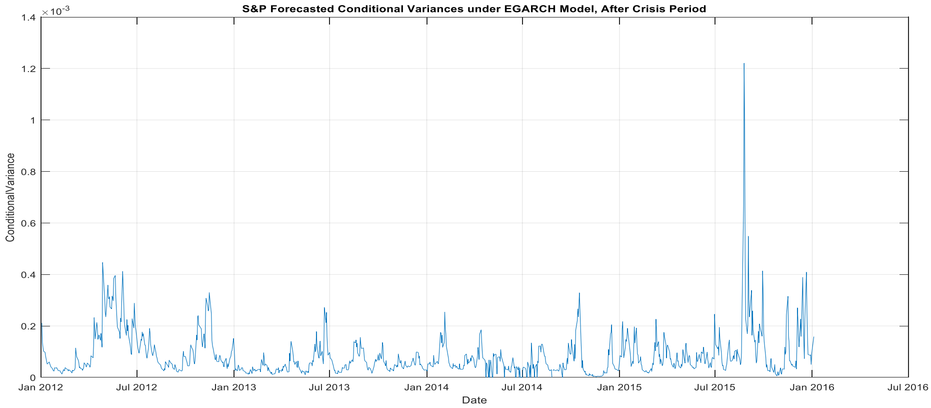

Figure 22. S&P forecasted conditional variances under EGARCH model, After Crisis Period

Indeed we observe that forecasted volatility is low, following our results from the data analysis section, where the plot of negative log returns indicates higher volatility only in August 2011, a month that is not included in forecasted volatilities (we use data from year 2012 onwards for the estimation results).

Next the VAR estimation results for the EGARCH model in the after crisis period are exhibited.

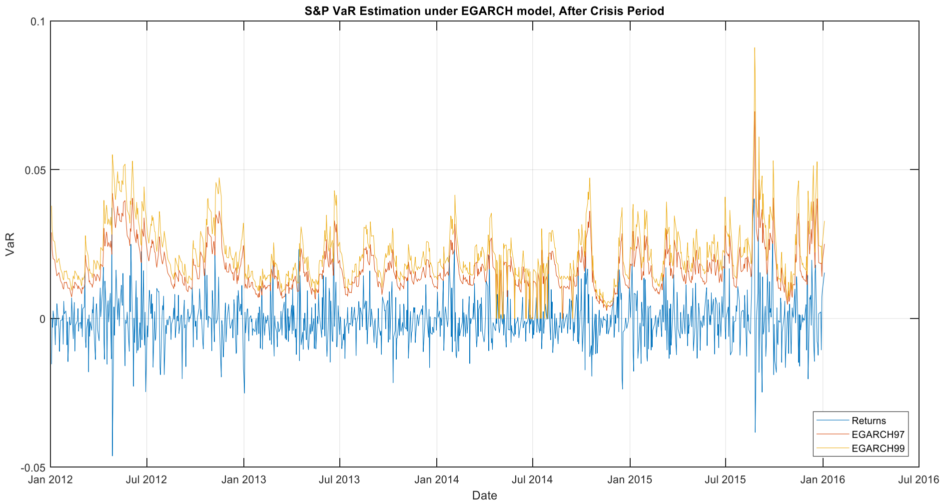

Figure 23. S&P VAR estimation under EGARCH model, After Crisis Period.

We can observe the VAR estimation accuracy again in this period, as VAR estimates closely track return series. But this time this comes at a cost of possibly many outliers, which are the number of times daily return losses exceed VAR estimates. Of course the results of this reliability analysis is presented in the next Chapter where backtesting results for VAR and Expected Shortfall gives as insight in which methods perform best and in which periods.

6.4 Comparison of VAR and Expected Shortfall estimation results

In this section the estimation results for VAR and Expected Shortfall for all models under analysis are presented in same plots in order to reach valuable conclusions for their accuracy in the crisis and after crisis period. These plots could indicate possible risk underestimation in the crisis period for some models, highlighting the need for the regulatory authorities to address this issue.

In the next two figures, VAR estimates at both confidence levels and for all models under analysis are presented for the crisis period.

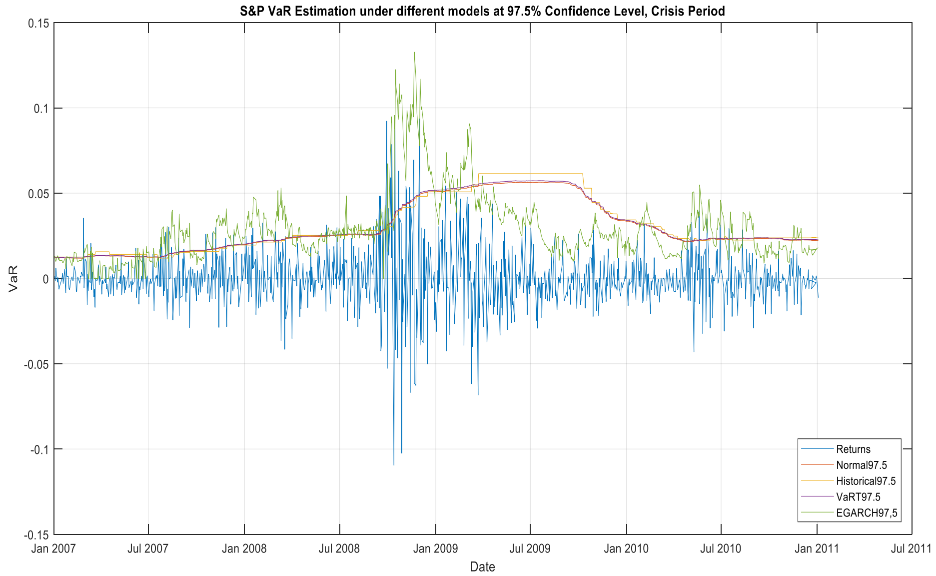

Figure 24. S&P VAR estimation under different models at 97,5% confidence level, Crisis Period.

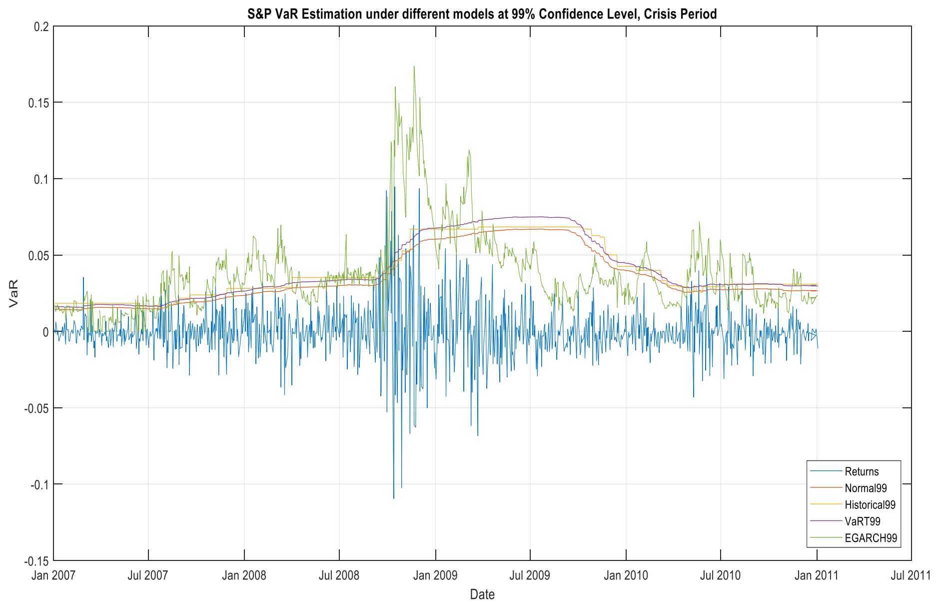

Figure 25. S&P VAR estimation under different models at 99% confidence level, Crisis Period.

Especially for the 99% confidence level, it is obvious that the EGARCH model outperforms its counterparts for the crisis period peak. The model accurately tracks the volatility clustering and high kurtosis effects and responds timely in stressed market conditions. Next the equivalent estimation results for the after crisis period are exhibited.

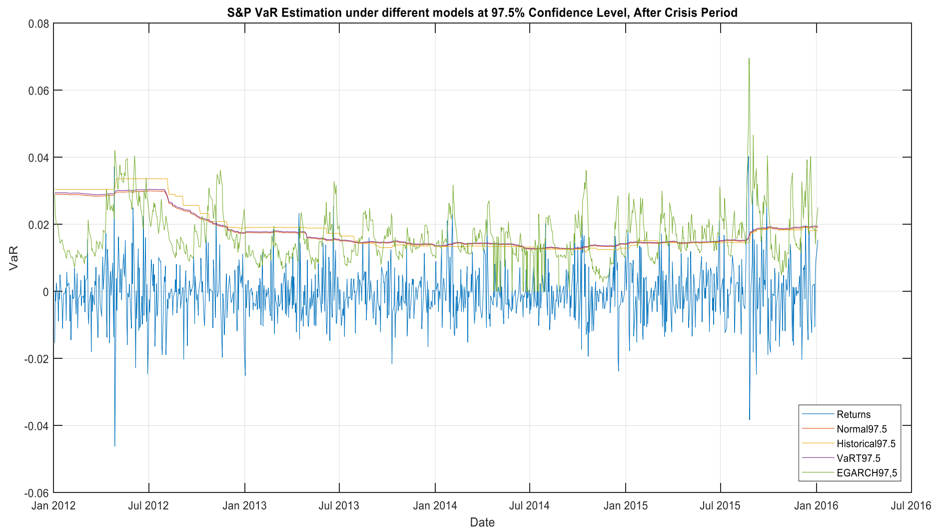

Figure 26. S&P VAR estimation under different models at 97,5% confidence level, After Crisis Period.

Figure 27. S&P VAR estimation under different models at 99% confidence level, After Crisis Period.

In the after crisis period, models other than EGARCH seem to perform better, as the low volatility market condition possibly leads to many VAR breaks under the later model. In the next two figures the Expected Shortfall estimation results for all models are presented for the crisis and after crisis periods. The Expected Shortfall estimates are conducted in the 97,5% confidence level, due to Basel Committee recent regulatory framework (BCBS, 2016: 56).

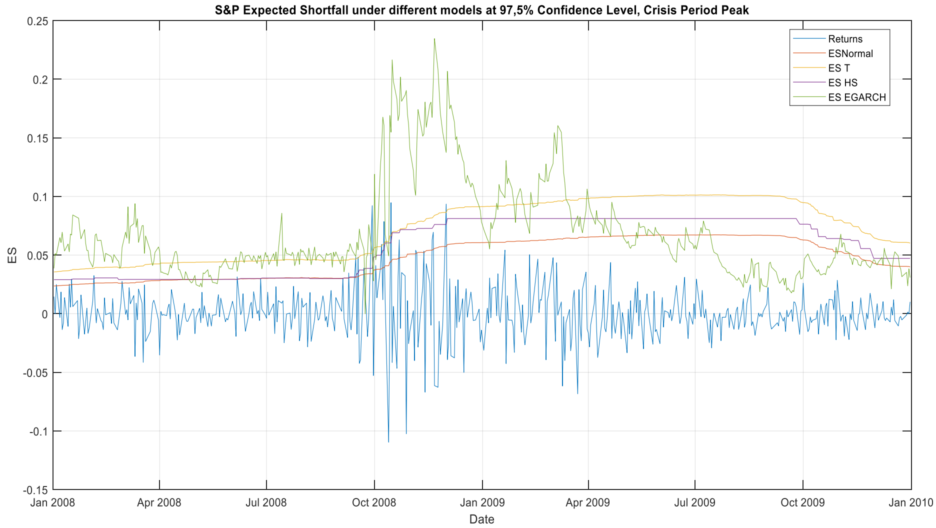

Figure 28. S&P Expected Shortfall estimation under different models at 97,5% confidence level, Crisis Period.

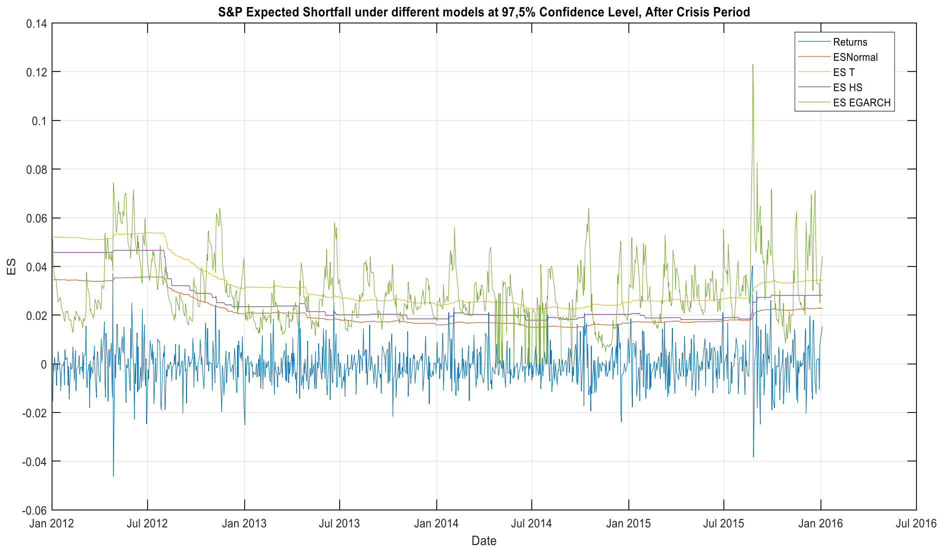

Figure 29. S&P Expected Shortfall estimation under different models at 97,5% confidence level, After Crisis Period.

Remarkably in the period of crisis peak, Expected Shortfall under EGARCH model produces estimates that are rarely breached and are higher than VAR equivalents, leading to increased market risk capital requirements. This is a very significant conclusion, as the new Basel Committee regulatory framework requires banks to estimate their market risk capital requirements under this risk measure.

Having analyzed the estimation results for VAR and Expected Shortfall under different models, the next step is to test their accuracy by a reliability test. This test is conducted through backtesting, a procedure presented for both risk measures in Chapter 4. In the next Chapter backtesting results are exhibited for VAR and Expected Shortfall in both periods under analysis.

Chapter 7
Empirical Results – Backtesting VAR and Expected Shortfall

Backtesting is a procedure for the examination of the reliability of a risk measure estimate. Since the late 1990s a variety of tests have been proposed for backtesting VAR, the most important of which were presented in Chapter 4. Under the Basel Committee on Banking Supervision market risk regulatory framework, the Traffic Light test is the one required by banks to perform. Expected Shortfall is a new measure that under the new regulatory market risk standards (BCBS, 2016) will substitute VAR in calculating capital requirements for market risk. Nevertheless, Basel Committee retained VAR as the risk measure for backtesting purposes. As analyzed in Chapter 4, this was possibly due to the ongoing discussion on Expected Shortfall lack of elicitability property and the subsequent scientific doubts about its backtesting potential. In a pioneering article though, Acerbi & Szekely (2014) proved that elicitability of a risk measure is not relevant to an absolute model validation procedure, which is backtesting. The authors also proposed three tests for Expected Shortfall backtesting.

In this Chapter the backtesting results for VAR and Expected Shortfall estimates of all models under analysis are presented for the crisis and after crisis periods separately. For VAR backtesting we focus on Traffic Light test, Kupiec's (1995) unconditional coverage test and Christoffersen's (1998) conditional coverage test. For Expected Shortfall backtesting, we use Test 2 of Acerbi & Szekely (2014: 4) with a small variation: As the test statistic Z_2 (analytically displayed in Chapter 4) is found by the authors to display remarkable stability across different distribution types (Acerbi & Szekely, 2014: 8), the critical values for the acceptance or rejection of a model are formed under the assumption that portfolio outcomes follow a standard normal distribution (Unconditional Normal test) or a t-student distribution with 3 degrees of freedom (Unconditional t test). With this way the simulation procedure for the calculation of critical values required by Test 2 is avoided, leading to a computationally easier solution.

This variation of Test 2 is proposed by MathWorks (2018).

Finally the reliability of a model in producing risk measures estimates is also tested for its accuracy. As backtesting is an informative criterion for the level of risk measure outliers in the examined period (i.e. the number of times the daily risk measure is exceeded by return losses), we also use the accuracy criterion to estimate the level of difference between daily risk measures and return losses (Mehta, Neukirchen, Pfetsch & Poppensieker (2012: 16). An accurate model closely tracks the return series pattern, avoiding risk overestimation and subsequent greater market risk capital requirements.

7.1 VAR Backtesting Results

In this section the reliability of VAR estimates is examined through various backtesting methods. Next the VAR backtesting results are exhibited for the crisis period (2007-2010).

S&P VaR Backtesting results under different models, Crisis Period

Portfolio ID	VaRID	VaR Level	Observed Level	Observations	Failures	Expected	Ratio	First Failure	Missing
"S&P"	Normal97	0,975	0,96432	1009	36	25,2250	1,4272	43	0
"S&P"	Historical97	0,975	0,96531	1009	35	25,2250	1,3875	43	0
"S&P"	VaRT97	0,975	0,96829	1009	32	25,2250	1,2686	43	0
"S&P"	EGARCH97	0,975	0.96421	1006	36	25,1500	1,4314	15	3
"S&P"	Normal99	0,99	0,97225	1009	28	10,0900	2,7750	43	0
"S&P"	Historical99	0,99	0,97919	1009	21	10,0900	2,0813	132	0
"S&P"	VaRT99	0,99	0,98117	1009	19	10,0900	1,8831	132	0
"S&P"	EGARCH99	0,99	0,97515	1006	25	10,06	2,4851	15	3

Portfolio ID	VaRID	VaR Level	TL	Bin	POF	TUFF	CC	CCI	TBF
"S&P"	Normal97	0,975	yellow	reject	reject	accept	accept	accept	reject
"S&P"	Historical97	0,975	yellow	reject	accept	accept	accept	accept	reject
"S&P"	VaRT97	0,975	green	accept	accept	accept	accept	accept	reject
"S&P"	EGARCH97	0,975	yellow	reject	reject	accept	reject	reject	reject
"S&P"	Normal99	0,99	red	reject	reject	accept	reject	accept	reject
"S&P"	Historical99	0,99	yellow	reject	reject	accept	reject	accept	reject
"S&P"	VaRT99	0,99	yellow	reject	reject	accept	reject	accept	reject
"S&P"	EGARCH99	0,99	red	reject	reject	accept	reject	reject	reject

Table 6. S&P VAR backtesting results, Crisis Period.

We can observe that the performance of VAR estimates is poor for all models, especially at the 99% confidence level, according to Traffic Lights test, Kupiec's (1995) Probability of Failures test and Christoffersen's (1998) Conditional Coverage test. As the VAR estimation results in Chapter 6 show that the EGARCH model has an exceptionally good performance in VAR estimation during the crisis peak period, producing few VAR breaks, we conduct a separate VAR backtesting analysis for the confined crisis peak period (2008-2009). For a total of 505 observations the results are the following.

S&P VaR Backtesting results under different models, Crisis Peak Period

Portfolio ID	VaRID	VaR Level	Observed Level	Observations	Failures	Expected	Ratio	First Failure	Missing
"S&P"	Normal97	0,975	0,96436	505	18	12,6250	1,4257	15	0
"S&P"	Historical97	0,975	0,96436	505	18	12,6250	1,4257	15	0
"S&P"	VaRT97	0,975	0,96832	505	16	12,6250	1,2673	48	0
"S&P"	EGARCH97	0,975	0,98614	505	7	12,6250	0,5545	62	0
"S&P"	Normal99	0,99	0,97426	505	13	5,0500	0,2574	48	0
"S&P"	Historical99	0,99	0,97822	505	11	5,0500	2,1782	48	0
"S&P"	VaRT99	0,99	0,98218	505	9	5,0500	1,7822	48	0
"S&P"	EGARCH99	0,99	0,99406	505	3	5,05	0,59406	189	0

Portfolio ID	VaRID	VaR Level	TL	Bin	POF	TUFF	CC	CCI	TBF
"S&P"	Normal97	0,975	green	accept	accept	accept	accept	accept	reject
"S&P"	Historical97	0,975	green	accept	accept	accept	accept	accept	reject
"S&P"	VaRT97	0,975	green	accept	accept	accept	accept	accept	reject
"S&P"	EGARCH97	0,975	green	accept	accept	accept	reject	accept	reject
"S&P"	Normal99	0,99	yellow	reject	reject	accept	reject	reject	reject
"S&P"	Historical99	0,99	yellow	reject	reject	accept	reject	accept	reject
"S&P"	VaRT99	0,99	yellow	accept	accept	accept	accept	accept	reject
"S&P"	EGARCH99	0,99	green	accept	accept	accept	accept	accept	accept

Table 7. S&P VAR backtesting results, Crisis Peak Period.

We can observe that during the crisis peak period the performance of EGARCH model is remarkably enhanced. This model passes the Traffic Light, Kupiec's POF and Christoffersen's CC tests. Finally it exhibits an empirical coverage ratio of 0.59406, significantly lower than the equivalent of its counterparts. Additionally, the model presents a high degree of accuracy, as exhibited in figure 18.

The t distribution model presents the second best performance, a result which is related to its ability to capture excess kurtosis.

After analyzing models VAR backtesting results for the crisis period, the equivalent results considering the after crisis period are presented in the next table.

S&P VaR Backtesting results under different models, After Crisis Period

Portfolio ID	VaRID	VaR Level	Observed Level	Observations	Failures	Expected	Ratio	First Failure	Missing
"S&P"	Normal97	0,975	0,98213	1007	18	25,1750	0,7150	81	0
"S&P"	Historical97	0,975	0,98113	1007	19	25,1750	0,7547	81	0
"S&P"	VaRT97	0,975	0,98411	1007	16	25,1750	0,6356	81	0
"S&P"	EGARCH97	0,975		1007	32	25,1750	1,2711	23	0
"S&P"	Normal99	0,99	0,98808	1007	12	10,0700	1,1917	81	0
"S&P"	Historical99	0,99	0,98908	1007	11	10,0700	1,0924	81	0
"S&P"	VaRT99	0,99	0,99007	1007	10	10,0700	0,9931	81	0
"S&P"	EGARCH99	0,99	0,98213	1007	18	10,0700	1,7875	23	0

Portfolio ID	VaRID	VaR Level	TL	Bin	POF	TUFF	CC	CCI	TBF
"S&P"	Normal97	0,975	green	accept	accept	accept	reject	reject	reject
"S&P"	Historical97	0,975	green	accept	accept	accept	accept	reject	reject
"S&P"	VaRT97	0,975	green	accept	reject	accept	reject	reject	reject
"S&P"	EGARCH97	0,975	green	accept	accept	accept	accept	accept	reject
"S&P"	Normal99	0,99	green	accept	accept	accept	reject	reject	reject
"S&P"	Historical99	0,99	green	accept	accept	accept	reject	reject	reject
"S&P"	VaRT99	0,99	green	accept	accept	accept	reject	reject	reject
"S&P"	EGARCH99	0,99	yellow	reject	reject	accept	reject	accept	reject

Table 8. S&P VAR backtesting results, After Crisis Period.

In this period we can observe that the model that performs best is the t distribution model. This model, at the 99% confidence level, passes the Traffic Light and Kupiec's POF tests but fails to pass the Christoffersen's CC test. The EGARCH model is the worst performer, implying that in low volatility tranquil periods this model is inefficient in producing reliable VAR estimates.

7.2 Expected Shortfall Backtesting Results

Having already analyzed the models' reliability of VAR estimates through backtesting, this section proceeds with the equivalent Expected Shortfall backtesting, making us capable of examining which models perform best in producing reliable estimates for both risk measures.

As already mentioned in the introduction of this Chapter, the Expected Shortfall backtesting procedure used in the analysis is a variation of Test 2 proposed by Acerbi & Szekely (2014: 4). For the same reasons as in the previous section we confine our analysis in the crisis peak period and the after crisis period. The results for Expected Shortfall backtesting for the crisis peak period are exhibited below.

S&P Expected Shortfall Backtesting results under different models, Crisis Period Peak										
Portfolio ID	**VaRID**	**VaR Level**	**Observed Level**	**Expected Severity**	**Observed Severity**	**Observations**	**Failures**	**Expected**	**Ratio**	**Missing**
"S&P"	Normal	0,975	0,96436	1,1928	1,5216	505	18	12,625	1,4257	0
"S&P"	Historical	0,975	0,96436	1,598	1,5984	505	18	12,625	1,4257	0
"S&P"	T 5	0,975	0,96832	1,7686	1,5605	505	16	12,625	1,2673	0
"S&P"	EGARCH	0,975	0,98614	1,7686	1,8721	505	7	12,625	0,5545	0

PortfolioID	VaRID	VaR Level	Unconditional Normal	Unconditional t
"S&P"	Normal	0,975	reject	reject
"S&P"	Historical	0,975	accept	accept
"S&P"	T 5	0,975	accept	accept
"S&P"	EGARCH	0,975	accept	accept

Table 9. S&P Expected Shortfall backtesting results, Crisis Peak Period.

The results show that all models except Normal distribution pass both Unconditional Normal and Unconditional t tests. Nevertheless, the EGARCH model presents the best empirical coverage ratio (0.5545). Next the equivalent results for the after crisis period are presented.

S&P Expected Shortfall Backtesting results under different models, After Crisis Period										
Portfolio ID	VaRID	VaRLevel	Observed Level	Expected Severity	Observed Severity	Observations	Failures	Expected	Ratio	Missing
"S&P"	Normal	0,975	0,98213	1,1928	1,3471	1007	18	25,175	0,715	0
"S&P"	Historical	0,975	0,98113	1,5535	1,4017	1007	19	25,175	0,75472	0
"S&P"	T 5	0,975	0,98411	1,7686	1,3679	1007	16	25,175	0,63555	0
"S&P"	EGARCH	0,975	0,96723	inf	inf	1007	33	25,175	1,3108	0

Portfolio ID	VaRID	VaRLevel	Unconditional Normal	Unconditional T
"S&P"	Normal	0,975	accept	accept
"S&P"	Historical	0,975	accept	accept
"S&P"	T 5	0,975	accept	accept
"S&P"	EGARCH	0,975	accept	accept

Table 10. S&P Expected Shortfall backtesting results, After Crisis Period.

The EGARCH model is clearly the worst performer, while the t distribution the best.

7.3 Comparison of Backtesting Results

According to the backtesting results of the two risk measures, the EGARCH model produces the most reliable VAR and Expected Shortfall estimates in the crisis peak period. This is clearly due to its ability to capture excess kurtosis, volatility clustering and asymmetries in the conditional variance equation (i.e. leverage effect).

Contrary to that, this model presents the worst performance in VAR and Expected Shortfall reliability of estimates in the after crisis period. The best performer in this period is the t distribution model, though at a cost of a reduced accuracy.

Chapter 8
Conclusions

Market risk is one of the most important risks faced by banking institutions. Basel Committee on Banking Supervision market risk regulatory framework has evolved from mid 1990s in an effort to address this kind of risk that stems from bank trading activities. Lessons from the past have contributed considerably in the formation of changes in regulatory standards.

This Thesis presented the evolution of the Basel Committee market risk regulatory framework and the challenges from a shift from a Value at Risk to an Expected Shortfall risk measure for the calculation of market risk capital requirements. According to the most recent Basel Committee market risk standards (BCBS, 2016: 4), national supervisors are expected to finalize implementation of this shift by January 2019. Thus the development of models that accurately estimate VAR and Expected Shortfall risk measures becomes a challenge for the banking industry.

In the theoretical part of the Thesis, various types of models were presented for VAR and Expected Shortfall estimation. More specifically, the non parametric method of Historical Simulation, parametric methods like the normal and t distributions, models of the GARCH family capturing volatility dynamics, the Monte Carlo simulation approach and finally the hybrid, semi parametric approach of Filtered Historical Simulation were analyzed in detail. VAR and Expected Shortfall backtesting techniques for model reliability checking were also exhibited.

The empirical part of the Thesis aimed at testing possible market risk underestimation under certain models, with the use of VAR and Expected Shortfall risk measures, in a crisis peak period like US financial crisis in 2008, which could lead to lower bank capital requirements and subsequent advent of systemic risk. In addition to that, a period of low market volatility was also chosen to test the model efficiency in an after crisis tranquil period. Four types of models were examined considering their VAR and Expected Shortfall estimates in these two market condition states separately using data of the S&P index, simulating a bank's equity trading portfolio.

The estimation and backtesting results showed that an EGARCH model capturing volatility dynamics, as well as asymmetries in the conditional variance equation, is more efficient in an acute crisis stage, where the crisis erupts. The adoption of other models that disregard these two properties of financial returns, like Historical Simulation, Normal Distribution or t Distribution produce poor results and underestimate growing market risk to a large extent. Additionally, the Expected Shortfall risk measure for the EGARCH model leads to higher level market risk capital requirements than VAR, based on its property to capture tail risk. The above results imply that the Expected Shortfall risk measure under a model capturing volatility dynamics for market risk would have acted as preventive or controlling mechanism for market risk systemic effects during the 2008 US financial crisis.

On the other hand, in a tranquil period with reduced market volatility, a t-distribution parametric model is found to perform best, but at a cost of a reduced accuracy (i.e. risk overestimation).

As the majority of banks today follow the simple but inefficient Historical Simulation approach (EBA, 2017: 31) for the calculation of their market risk capital requirements, the Basel Committee on Banking Supervision could consider requiring them to adopt more sophisticated models in periods of extreme market volatility. The results of the empirical part of the Thesis highlighted the level of failure of the Historical Simulation method to produce reliable risk measure estimates during such periods, as it follows rather than controls financial crises. Models accounting for volatility clustering effects and excess kurtosis in financial returns would prevent further escalation stages during a financial crisis period.

Appendix A
Matlab Code, Crisis Period

A.1 Matlab Variables

```
Returns1=SP_neg_ret;
DateReturns1=date1;
SampleSize=length(Returns1);
TestWindowStart=find(year(DateReturns1)==2007,1);
TestWindow=TestWindowStart:SampleSize;
EstimationWindowSize=250;
DatesTest1=DateReturns1(TestWindow);
ReturnsTest1=Returns1(TestWindow);
pVaR=[0.025 0.01];
```

A.2 Risk Measure Estimation under Different Models

A.2.1 Normal Distribution Model

```
Zscore=norminv(pVaR);
Normal97=zeros(length(TestWindow),1);
Normal99=zeros(length(TestWindow),1);
for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns1(EstimationWindow));
    Normal97(i)=-Zscore(1)*Sigma;
    Normal99(i)=-Zscore(2)*Sigma;
end
figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[Normal97 Normal99])
xlabel('Date')
ylabel('VaR')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the Normal Distribution Method, Crisis Period')

ESNormal97=zeros(length(TestWindow),1);

for t=TestWindow
```

```
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns1(EstimationWindow));
    ESNormal97(i)=Sigma*normpdf(norminv(0.975))./(1-0.975);
end
```

A.2.2 Historical Simulation Model

```
Historical97=zeros(length(TestWindow),1);
Historical99=zeros(length(TestWindow),1);
for t=TestWindow
  i=t-TestWindowStart+1;
  EstimationWindow=t-EstimationWindowSize:t-1;
  X1=Returns1(EstimationWindow);
  Historical97(i)=-quantile(X1,pVaR(1));
  Historical99(i)=-quantile(X1,pVaR(2));
end

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[Historical97 Historical99])
ylabel('VaR')
xlabel('Date')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the Historical Simulation Method, Crisis Period')

ESH97=zeros(length(TestWindow),1);
VARH97=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    N=length(Returns1(EstimationWindow));
    k=ceil(N*0.975);
    z=sort(Returns1(EstimationWindow));
    VARH97(i)=z(k);
if k<N
  ESH97(i)=((k-N*0.975)*z(k)+sum(z(k+1:N)))/(N*(1-0.975));
else
  ESH97(i)=z(k);
end
end
```

A.2.3 Student's t Distribution Model

```
DoF=5;
VaRT97=zeros(length(TestWindow),1);
VaRT99=zeros(length(TestWindow),1);
```

```
for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns1(EstimationWindow));
    VaRT97(i)=Sigma*sqrt((DoF-2)/DoF)*tinv(0.975,5);
    VaRT99(i)=Sigma*sqrt((DoF-2)/DoF)*tinv(0.99,5);
end

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[VaRT97 VaRT99]);
xlabel('Date')
ylabel('VaR')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the T-Distribution Method, Crisis Period')

EST97=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns1(EstimationWindow));
    ES_StandardT=(tpdf(tinv(0.975,DoF),DoF).*(DoF+tinv(0.975,DoF).^2)./((1-
0.975).*(DoF-1)));
    EST97(i)=Sigma*ES_StandardT;
end
```

A.2.4 EGARCH model

```
Mdl1=egarch('GARCH',NaN,'ARCH',NaN,'Leverage',NaN,'Distribution',struct('Name','t','D
oF',5));
fitMdl1=estimate(Mdl1,ReturnsTest1);

V=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    fitMdl1EW=estimate(Mdl1,Returns1(EstimationWindow),'Display','off');
    if i==1
      V(i)=forecast(fitMdl1EW,1,'Y0',Returns1(1:250),'V0',st0);
    else
      V(i)=forecast(fitMdl1EW,1,'Y0',Returns1(EstimationWindow));
    end
end

figure;
plot(DatesTest1,V)
ylabel('ConditionalVariance')
xlabel('Date')
title('S&P Forecasted Conditional Variances, Crisis Period')
```

```
grid on

VAREGARCH97=(sqrt(V))*(sqrt((DoF-2)/DoF))*tinv(0.975,5);
VAREGARCH99=(sqrt(V))*(sqrt((DoF-2)/DoF))*tinv(0.99,5);

ES_EGARCH97=(sqrt(V))*(ES_StandardT);

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,VAREGARCH97,DatesTest1,VAREGARCH99)
ylabel('VaR')
xlabel('Date')
legend('Returns','VAREGARCH97','VAREGARCH99','Location','SouthEast')
title('S&P VaR Estimation under EGARCH model, Crisis Period')
grid on
```

A.2.5 Comparison of Risk Measure Estimation Results under Different Models

```
figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,Normal97,DatesTest1,Historical97,DatesTest
1,VaRT97,DatesTest1,VAREGARCH97)
ylabel('VaR')
xlabel('Date')
legend('Returns','Normal97.5','Historical97.5','VaRT97.5','EGARCH97,5','Location','Sout
hEast')
title('S&P VaR Estimation under different models at 97.5% Confidence Level, Crisis
Period')
grid on

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,Normal99,DatesTest1,Historical99,DatesTest
1,VaRT99,DatesTest1,VAREGARCH99)
ylabel('VaR')
xlabel('Date')
legend('Returns','Normal99','Historical99','VaRT99','EGARCH99','Location','SouthEast')
title('S&P VaR Estimation under different models at 99% Confidence Level, Crisis
Period')
grid on

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,ESNormal97,DatesTest1,EST97,DatesTest1,ES
H97,DatesTest1,ES_EGARCH97);
ylabel('ES')
xlabel('Date')
legend({'Returns','ESNormal','ES T','ES HS','ES EGARCH'},'Location','Best')
title('S&P Expected Shortfall under different models at 97,5% Confidence Level, Crisis
Period')
grid on
```

A.3 Risk Measure Backtesting

```
vbt=varbacktest(ReturnsTest1,[Normal97 Historical97 VaRT97 VAREGARCH97
Normal99 Historical99 VaRT99
VAREGARCH99],'PortfolioID','S&P','VaRID',{'Normal97','Historical97','VaRT97','VAREGA
RCH97','Normal99','Historical99','VaRT99','VAREGARCH99'},'VaRLevel',[0.975 0.975
0.975 0.975 0.99 0.99 0.99 0.99]);
summary(vbt)
runtests(vbt)

ebt=esbacktest(ReturnsTest1,[Normal97 Historical97 VaRT97
VAREGARCH97],[ESNormal97 ESH97 EST97
ES_EGARCH97],'PortfolioID','S&P','VaRID',{'NORMAL','Historical','T
5','EGARCH'},'VaRLevel',[0.975 0.975 0.975 0.975]);
s=summary(ebt);
disp(s)

t=runtests(ebt);
disp(t)
```

BLANK PAGE

Appendix B
Matlab Code, After Crisis Period

B.1 Matlab Variables

```
Returns2=SP_neg_ret1;
DateReturns2=date1;
SampleSize=length(Returns2);
TestWindowStart=find(year(DateReturns2)==2012,1);
TestWindow=TestWindowStart:SampleSize;
EstimationWindowSize=250;
DatesTest1=DateReturns2(TestWindow);
ReturnsTest1=Returns2(TestWindow);
pVaR=[0.025 0.01];
```

B.2 Risk Measure Estimation under Different Models

B.2.1 Normal Distribution Model

```
Zscore=norminv(pVaR);
Normal97=zeros(length(TestWindow),1);
Normal99=zeros(length(TestWindow),1);
for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns2(EstimationWindow));
    Normal97(i)=-Zscore(1)*Sigma;
    Normal99(i)=-Zscore(2)*Sigma;
end
figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[Normal97 Normal99])
xlabel('Date')
ylabel('VaR')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the Normal Distribution Method, After Crisis Period')

ESNormal97=zeros(length(TestWindow),1);

for t=TestWindow
```

```
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns2(EstimationWindow));
    ESNormal97(i)=Sigma*normpdf(norminv(0.975))./(1-0.975);
end
```

B.2.2 Historical Simulation Model

```
Historical97=zeros(length(TestWindow),1);
Historical99=zeros(length(TestWindow),1);
for t=TestWindow
  i=t-TestWindowStart+1;
  EstimationWindow=t-EstimationWindowSize:t-1;
  X1=Returns2(EstimationWindow);
  Historical97(i)=-quantile(X1,pVaR(1));
  Historical99(i)=-quantile(X1,pVaR(2));
end

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[Historical97 Historical99])
ylabel('VaR')
xlabel('Date')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the Historical Simulation Method, After Crisis Period')

ESH97=zeros(length(TestWindow),1);
VARH97=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    N=length(Returns2(EstimationWindow));
    k=ceil(N*0.975);
    z=sort(Returns2(EstimationWindow));
    VARH97(i)=z(k);
if k<N
  ESH97(i)=((k-N*0.975)*z(k)+sum(z(k+1:N)))/(N*(1-0.975));
else
  ESH97(i)=z(k);
end
end
```

B.2.3 Student's t Distribution Model

```
DoF=5;
VaRT97=zeros(length(TestWindow),1);
VaRT99=zeros(length(TestWindow),1);
```

```
for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns2(EstimationWindow));
    VaRT97(i)=Sigma*sqrt((DoF-2)/DoF)*tinv(0.975,5);
    VaRT99(i)=Sigma*sqrt((DoF-2)/DoF)*tinv(0.99,5);
end

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,[VaRT97 VaRT99]);
xlabel('Date')
ylabel('VaR')
legend({'Returns','97.5% Confidence Level','99% Confidence Level'},'Location','Best')
title('S&P VaR Estimation Using the T-Distribution Method, After Crisis Period')

EST97=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    Sigma=std(Returns2(EstimationWindow));
    ES_StandardT=(tpdf(tinv(0.975,DoF),DoF).*(DoF+tinv(0.975,DoF).^2)./((1-
0.975).*(DoF-1)));
    EST97(i)=Sigma*ES_StandardT;
end
```

B.2.4 EGARCH model

```
Mdl1=egarch('GARCH',NaN,'ARCH',NaN,'Leverage',NaN,'Distribution',struct('Name','t','D
oF',5));
fitMdl1=estimate(Mdl1,ReturnsTest1);
V0=infer(fitMdl1,ReturnsTest1);

 st0=var(Returns2(1:250));

V=zeros(length(TestWindow),1);

for t=TestWindow
    i=t-TestWindowStart+1;
    EstimationWindow=t-EstimationWindowSize:t-1;
    fitMdl1EW=estimate(Mdl1,Returns2(EstimationWindow),'Display','off');
    if i==1
      V(i)=forecast(fitMdl1EW,1,'Y0',Returns2(1:250),'V0',st0);
    else
      V(i)=forecast(fitMdl1EW,1,'Y0',Returns2(EstimationWindow));
    end
end
```

```
figure;
plot(DatesTest1,V)
ylabel('ConditionalVariance')
xlabel('Date')
title('S&P Forecasted Conditional Variances under EGARCH Model, After Crisis Period')
grid on

VAREGARCH97=(sqrt(V))*(sqrt((DoF-2)/DoF))*tinv(0.975,5);
VAREGARCH99=(sqrt(V))*(sqrt((DoF-2)/DoF))*tinv(0.99,5);

ES_EGARCH97=(sqrt(V))*(ES_StandardT);

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,VAREGARCH97,DatesTest1,VAREGARCH99)
ylabel('VaR')
xlabel('Date')
legend('Returns','EGARCH97','EGARCH99','Location','SouthEast')
title('S&P VaR Estimation under EGARCH model, After Crisis Period')
grid on
```

B.2.5 Comparison of Risk Measure Estimation Results Under Different Models

```
figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,Normal97,DatesTest1,Historical97,DatesTest
1,VaRT97,DatesTest1,VAREGARCH97)
ylabel('VaR')
xlabel('Date')
legend('Returns','Normal97.5','Historical97.5','VaRT97.5','EGARCH97,5','Location','Sout
hEast')
title('S&P VaR Estimation under different models at 97.5% Confidence Level, After Crisis
Period')
grid on

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,Normal99,DatesTest1,Historical99,DatesTest
1,VaRT99,DatesTest1,VAREGARCH99)
ylabel('VaR')
xlabel('Date')
legend('Returns','Normal99','Historical99','VaRT99','EGARCH99','Location','SouthEast')
title('S&P VaR Estimation under different models at 99% Confidence Level, After Crisis
Period')
grid on

figure;
plot(DatesTest1,ReturnsTest1,DatesTest1,ESNormal97,DatesTest1,EST97,DatesTest1,ES
H97,DatesTest1,ES_EGARCH97);
ylabel('ES')
xlabel('Date')
```

```
legend({'Returns','ESNormal','ES T','ES HS','ES EGARCH'},'Location','Best')
title('S&P Expected Shortfall under different models at 97,5% Confidence Level, After
Crisis Period')
grid on
```

B.3 Risk Measure Backtesting

```
vbt=varbacktest(ReturnsTest1,[Normal97 Historical97 VaRT97 VAREGARCH97
Normal99 Historical99 VaRT99
VAREGARCH99],'PortfolioID','S&P','VaRID',{'Normal97','Historical97','VaRT97','VAREGA
RCH97','Normal99','Historical99','VaRT99','VAREGARCH99'},'VaRLevel',[0.975 0.975
0.975 0.975 0.99 0.99 0.99 0.99]);
summary(vbt)
runtests(vbt)

ebt=esbacktest(ReturnsTest1,[Normal97 Historical97 VaRT97
VAREGARCH97],[ESNormal97 ESH97 EST97
ES_EGARCH97],'PortfolioID','S&P','VaRID',{'NORMAL97','Historical97','VaRT97','VAREG
ARCH97'},'VaRLevel',[0.975 0.975 0.975 0.975]);
s=summary(ebt);
disp(s)

t=runtests(ebt);
disp(t)
```

BLANK PAGE

Bibliography

Acerbi, C. & Tasche, D. (2001) Expected Shortfall: a Natural Coherent Alternative to Value at Risk, *Economic Notes*, 31(2), 379-388.

Acerbi, C. & Szekely, B. (2014) Backtesting Expected Shortfall. *MSCI Inc.*

Angelidis, T., Benos, A. & Degiannakis, S. (2004) The Use of GARCH Models in VAR Estimation, *Statistical Methodology*, 1, 105-128.

Artzner, P., Delbaen, F. & Eber, J.M. & Heath, D. (1999) Coherent Measures of Risk, *Mathematical Finance*, 9(3), 203-228.

Barone-Adesi, G., Bourgoin, F. & Giannopoulos, K. (1998) Don't Look Back, *RISK*, 11, 100-104.

Barone-Adesi, G. & Giannopoulos, K. (2002) Non-parametric VAR Techniques; Myths and Realities, *Economic Notes*, 30, 167-181.

Basel Committee on Banking Supervision. (1996a) Amendment to the Capital Accord to incorporate Market Risks, Bank for International Settlements, Basel.

Basel Committee on Banking Supervision (1996b) Supervisory Framework for the use of 'Backtesting' in Conjunction with the Internal Models Approach to Market Risk Capital Requirements, Basel.

Basel Committee on Banking Supervision, (2011) Revisions to the Basel II Market Risk Framework, Bank for International Settlements, Basel.

Basel Committee on Banking Supervision. (2012) Fundamental Review of the Trading Book, Bank for International Settlements, Basel.

Basel Committee on Banking Supervision. (2013) Fundamental Review of the Trading Book: A Revised Market Risk Framework, Bank for International Settlements, Basel.

Basel Committee on Banking Supervision. (2016) Minimum Capital Requirements for Market Risk, Bank for International Settlements, Basel.

Berkowitz, J. & O'Brien, j. (2002) How Accurate are Value-at-Risk Models at Commercial Banks?, *The Journal of Finance*, 57(3), 1093-1111.

Berkowitz, J., Christoffersen, P. & Pelletier, D. (2009) Evaluating Value at Risk Models with Desk-Level Data, *Management Science*, 1-15.

Black, F. (1976). Studies of Stock Market Volatility Changes. Proceedings of the *American Statistical Association, Business and Economic Statistics Section*, 177-181.

Bollerslev, T (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, 31, 177-181.

Boudoukh, J., M. Richardson and R. Whitelaw (1998) The Best of Both Worlds, *Stern School of Business*, 1-13.

Campbell, S. (2005) A Review of Backtesting and Backetsting procedures, *Finance and Economics Discussion Series, Division of Research & Statistics and Monetary Affairs, Federal Reserve Board*, 1-23.

Christoffersen, P. (1998) Evaluating interval forecasts, *International Economic Review*, 39(4), 841-862.

Danielsson, J. (2011) *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk, with Implementation in R and MATLAB.* Chichester: Wiley.

Danielsson, J. & Zhou, C. (2016) Why Risk is so hard to measure, *Systemic Risk discussion paper 36, London School of Economics*, 1-29.

Degiannakis, S., Floros, C. & Dent, P. (2012) Forecasting Value-at-Risk and Expected Shortfall using Fractionally Integrated Models of Conditional Volatility: International Evidence, *International Review of Financial Analysis*, 27, 21-33.

Degiannakis, S., Floros, C. & Livada, A. (2012) Evaluating Value at Risk Models Before and After the Financial Crisis of 2008: International Evidence, *Managerial Finance*, 38(4), 436-452.

Dowd, K. (2005) *Measuring Market Risk*. Chichester: Wiley.

Embrechts, P., Resnick, S. & Samorodnitsky, G. (1999) Extreme Value Theory as a Risk Management Tool, *North American Actuarial Journal*, 3, 30-41.

Emmer, S., Kratz, M. & Tasche, D. (2015) What is the Best Risk Measure in Practice? A Comparison of Standard Measures, *Journal of Risk*, 18(2), 31-60.

Engle, R. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation. *Econometrica*, 50, 987-1008.

European Banking Authority (2017) Results from the 2016 Market Risk Benchmarking Exercise

Giannopoulos, K. & Tunaru, R. (2005) Coherent Risk Measures under Filtered Historical Simulation, *Journal of Banking & Finance,* 29, 979-996.

Gurrola, P. & Murphy, D. (2015) Filtered Historical Simulation Value-at-Risk Models and their competitors, *Bank of England, Working paper No 525*, 1-30.

Gneiting, T. (2011) Making and Evaluating Point Forecasts. *Journal of the American Statistical Association*, 106, 746-762.

Harmantzis, F., Miao, L. & Chien, Y. (2006) Empirical Study of Value-at-Risk and Expected Shortfall Models with Heavy Tails, *Journal of Risk Finance*, 7(2), 117-135.

Hull, J. (2012) *Risk Management and Financial Institutions*. Hoboken, New Jersey: Wiley Finance.

Jorion, P. (1996) Measuring the Risk in Value at Risk, *Financial Analysis Journal*, 52(6), 47-56.

Kellner, R. & Rosch, D. (2016) Quantifying Market Risk with Value-at-Risk or Expected Shortfall? - Consequences for Capital Requirements and Model Risk. *Journal of Economic Dynamics and Control,* 68, 45-63.

Kupiec, P. (1995) Techniques for Verifying the Accuracy of Risk Management Methods, *Journal of Derivatives*, 3, 73-84.

Manganelli, S. & Engle, R. (2001) Value at Risk Models in Finance, *European Central Bank, Working Paper Series (No 75)*, 1-40.

Mathworks (2018) Overview of Expected Shortfall Backtesting. Retrieved from http://www.mathworks.com/help/risk/examples/expected-shortfall-estimation-and backtesting.html.

McNeil, A., Frey, R. & Embrechts, P. (2015) *Quantitative Risk Management*. Princeton: Princeton University Press.

McNeil, A. & Frey, R. (2000) Estimation of Tail-Related Risk Measures for Heteroscedastic Time Series: An Extreme Value Approach, *Journal of Empirical Finance*, 7, 271-300.

Mehta, A., Neukirchen, M., Pfetsch, S. & Poppensieker, T. (2012) Managing Market Risk: Today and Tomorrow. https://www.mckinsey.com/business-functions/risk/our-insights/managing-market-risk-today-and-tomorrow (accessed 03-12-2017)

Ming, J. (2014) Measuring Market Risk under the Basel Accords: VAR, Stressed VAR and Expected Shortfall, *The IEB International Journal of Finance,* 8, 184-201.

Nadarajah, S., Zhang, B. & Chan, S. (2013) Estimation methods for expected shortfall, *Quantitative Finance*, 14(2), 271-291.

Nelson, D. (1991) Conditional Heteroscedasticity in Asset Returns: A New Approach, *Econometrica,* 59(2), 347-370.

Nozari, M., Raei, S.M., Jahangiri, P. & Bahramgiri, M. (2010) A Comparison of Heavy-Tailed VAR Estimates and Filtered Historical Simulation: Evidence from Emerging Markets, *International Review of Business Research Papers*, 6(4), 347-359.

Perignon, C. & Smith, D. (2009). The Level and Quality if Value-at-Risk Disclosure by Commercial Banks, *Journal of Banking & Finance*, 34, 362-377.

Pindyck, R. & Rubinfeld, D. (1998) *Econometric Models and Economic Forecasts,* Boston: McGraw-Hill.

Pritsker, M. (2006) The Hidden Dangers of Historical Simulation, *Journal of Banking & Finance*, 30 (2), 561-582.

Righi, M. & Ceretta, P. (2014) A Comparison of Expected Shortfall Estimation Models. *Journal of Economics and Business,* 78, 14-47.

RiskMetrics (1996) Technical Document, 4th Edition, *Morgan Guaranty Trust Company,* 1-284.

Saunders, A. & Cornett, M. (2003) *Financial Institutions Management: A Risk Management Approach*. International Edition. Boston: McGraw-Hill.

Xekalaki, E. & Degiannakis, S. (2010) *ARCH Models for Financial Applications.* Chichester: Wiley.

Yamai, Y. & Yoshiba, T. (2005) Value-at-Risk versus Expected Shortfall: a Practical Perspective, *Journal of Banking & Finance*, 29, 997-1015.