Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER THESIS

Value at Risk: GARCH vs. Stochastic Volatility Models: Empirical Study

Author: **Bc. Viktória Tesárová**

Supervisor: **PhDr. Petr Gapko**

Academic Year: **2011/2012**

Declaration of Authorship

The author hereby declares that she compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 30, 2012

Signature

Acknowledgments

I would like to thank my supervisor PhDr. Petr Gapko for successful guidance and valuable comments while writing my master thesis. I would like to thank also PhDr. Boril Šopov, MSc., LL.M. for his valuable help with empirical estimation of parameters via Efficient Importance Sampling procedure.

Abstract

The thesis compares GARCH volatility models and Stochastic Volatility (SV) models with Student's t distributed errors and its empirical forecasting performance of Value at Risk on five stock price indices: S&P, NASDAQ Composite, CAC, DAX and FTSE. It introduces in details the problem of SV models Maximum Likelihood examinations and suggests the newly developed approach of Efficient Importance Sampling (EIS). EIS is a procedure that provides an accurate Monte Carlo evaluation of likelihood function which depends upon high-dimensional numerical integrals.

Comparison analysis is divided into in-sample and out-of-sample forecasting performance and evaluated using standard statistical probability backtestig methods as conditional and unconditional coverage.

Based on empirical analysis thesis shows that SV models can perform at least as good as GARCH models if not superior in forecasting volatility and parametric VaR.

JEL Classification	F12, C22, C52, C53, G15
Keywords	VaR, GARCH, Stochastic Volatility, backtesting, conditional coverage, unconditional coverage

Author's e-mail	`viki.tesar@gmail.com`
Supervisor's e-mail	`petr.gapko@seznam.cz`

Abstrakt

Práca porovnáva GARCH modely volatility a modely Stochastickej volatility so študentovým t rozdelením a ich empirickú schopnosť predpovedania Value at Risk na piatich akciových indexoch: S&P, NASDAQ Composite, CAC, DAX a FTSE. Detailne predstavuje problém vyrátania metódy maximálnej vierohodnosti pre Stochastickú volatilitu a navrhuje nedávno vyvinutú metódu tzv. Efficient Importance Sampling. Táto metóda poskytuje veľmi primerané Monte Carlo odhady vierohodnostnej funkcie, ktoré sú závislé na numerických integráloch vysokého rádu.

Komparatívna analýza je rozdelená na predpovedací výkon v prvom období zo vzorky a v druhom období mimo vzorku. Tie sú vyhodnotené na základe štandardných štatistických a pravdepodobnostných backtestových metódach ako je tzv. podmienený a nepodmienený coverage.

Na základe empirickej analýzy táto práca ukazuje, že SV modely môžu fungovať aspoň tak dobre ako GARCH modely, ak nie k nim byť nadradené pri predpovedaní volatility a následne parametrického Value at Risk.

Klasifikácia JEL	F12, C22, C52, C53, G15
Kľúčové slová	VaR, GARCH, Stochastická volatilita, backtestové metódy, podmienený coverage, nepodmienený coverage

E-mail autora	viki.tesar@gmail.com
E-mail vedúceho práce	petr.gapko@seznam.cz

Contents

List of Tables

List of Figures

Acronyms

ACF Autocorrelation Function

ADF Augmented Dickey-Fuller test

AR Autoregressive

ARCH Autoregresive Conditional Heteroskedasticity

ARSV Autoregressive Stochastic Volatility

EIS Efficient Importance Sampling

EMM Efficient Method of Moment

GARCH General Autoregresive Conditional Heteroskedasticity

GMM Generalized Method of Moments

HMSE Heteroskedasticity-adjusted Mean Square Error

KPSS Kwiatkowski-Phillips-Schmidt-Shin test

MAD Mean Absolute Deviation

MCMC Markov chain Monte Carlo

ML Maximum Likelihood

MLAE Mean Log-absolute Error

MM Method ofMmoments

MSE Mean Square Error

QML Quasi-Maximum Likelihood

RMSE Root Mean Equare Error

SMM Simulated Method of Moments

SV Stochastic Volatility

TUFF Time until First Failure

VaR Value at Risk

Master Thesis Proposal

Author	Bc. Viktória Tesárová
Supervisor	PhDr. Petr Gapko
Proposed topic	Value at Risk: GARCH vs. Stochastic Volatility Models: Empirical Study

Topic characteristics Value at Risk (VaR) has over time evolved to one of the most popular comprehensive tools used to estimate exposure to market risks. VaR claims the maximum loss of portfolio, expressed in its units, with certain probability during given period. It works with the distribution of loss and profit. With zero mean only standard deviation of the loss matters. A time horizon and a confidence level are chosen and a cumulative distribution function is assumed.

Because volatility is a key input to VaR models, the characterization of asset or portfolio volatility is of great importance when implementing and testing VaR models. The correct choice of volatility model is one of the most important factors in determining the effectiveness of VaR. Volatility modeling is nowadays dominated by three families of models: the Conditional Volatility ARCH/GARCH models developed by Engle (1982) and Bollerslev (1986), respectively; Stochastic Volatility (SV) models, which specifies a stochastic process for volatility, first introduced by Taylor (1982); and Realized Volatility (RV) models. This paper will consider first two methods of estimating volatility, while GARCH and SV are two competing, well-known, often-used models to explain volatility of financial series.

ARCH/GARCH models have subsequently led to a huge family of autoregressive conditional volatility models. Its popularity is attributed to the fact of easy to implement, bringing great results and having large ability to capture several stylized facts of financial returns, such as time-varying volatility, persistence and clustering of volatility, and asymmetric reactions to positive and negative shocks of equal magnitude. The ARCH/GARCH family proved to be a rich framework and many different extensions and generalizations of the initial ARCH/GARCH models have been proposed.

SV models have been until nowadays extremely time consuming to estimate. But it is not longer a case since the strong evolution of simulation based econometric methods in last years. Problem of difficult estimation is handled with lot of algorithms developed recently: Generalized Methods of Moments, the Quasi Maximum Likelihood method, Simulated Maximum Likelihood technique, the Markov Chain Monte Carlo method. The idea behind the family of SV models is that the volatility is driven by a latent process representing the flow of price relevant information

Both GARCH and SV models take account the important volatility clustering of financial returns. But the main difference is that in SV model the volatility is a latent variable with unexpected noise, while in the GARCH model, the volatility one period ahead is observable given todays information.

However, VaR models are useful only if they predict future risks accurately. In order to evaluate the quality of the VaR estimates, the models should always be backtested with appropriate methods. Backtesting is a statistical procedure where actual profits and losses are systematically compared to corresponding VaR estimates. The objective of this paper will be the theoretical and empirical comparison and evaluation of GARCH and SV models for forecasting of VaR. Empirical part will be applied on 4 different western European stock indices

Hypotheses

1. SV is better (less forecast errors) than GARCH in estimating asset volatility

2. SV model captures more aspects of volatility than the GARCH model due to its sources of variability

3. SV model is more flexible than the GARCH model in the sense that it is able to generate series with properties more in compliance with the properties often observed in real financial time series.

4. Best model for volatility forecasting also is the best model for VaR forecasting.

Methodology The paper will start with theoretical introduction to risk management measure VaR and its different types of modeling. We consider important aspect in VaR modeling: volatility modeling methods. We discuss all the three used nowadays. Conditional variances, stochastic volatility and realized volatility methods. More in details we devote to GARCH family and SV family. Here we examine the importance of the choice of distribution. We will use three different distributions: the normal distribution, the student t distribution and the skewed t distribution and consider heaviness of theirs tails. Coming to the empirical part we run GARCH models and SV models with some its extensions on our data sample and we forecast the volatility and VaR, respectively. GARCH will be estimated using QMLE (quasi-maximum likelihood estimate) and SV through Kalman Filter or Markov Chain Monte Carlo method. Good VaR estimates should have correct unconditional coverage: fraction of VaR violation should be equal to the nominal coverage probability. Therefore VaR forecasts are evaluated using the LR Kupiec test for correct unconditional coverage, and a Markov test for correct conditional coverage. But it is implicitly assumed that Var is independent, while VaR estimates should have independence. VaR violations should be spread out over the sample and not come in clusters. So another test will be used to check for the independent VaR violations, proposed by Christoffersen (1998).

Outline

1. Risk management - VaR
2. VaR forecasting
3. Comparison of GARCH vs SV methods
4. Empirical analysis
5. Backtesting of models
6. Conclusion

Core bibliography

1. Awartani, B. M. A. & W. Corradi (2003): "Predicting the Volatility of the S&P-500 Index via GARCH Models: The Role of Asymmetries." *University of Exeter*.

2. Blomstrom, M. & A. Kokko (2003): "The Economics of Foreign Direct Investment Incentives.", NBER *Working Papers 9489*, National Bureau of Economic Research, Inc.

3. Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **31**.

4. Bollerslev, T., K. F. Kroner & R. Y. Chou (1992): "RCH Modeling in Finance: A Review of the Theory and Empirical Evidence." *Journal of Econometrics* **52**, **April**.

5. Christoffersen, F. E., .J. Hahn, & A. Inoue (2001): *"Testing and Comparing Value at Risk Measures."* Working Papers - 03: CIRANO.

6. Eberlein, E., J. Kallsen, & J. Kristen (2002): *"TRisk Management Based on Stochastic Volatility." Institut fr Mathematische Stochastik* University of Freiburg.

7. Engle, R. F. (2003): "Risk and Volatility: Econometric Models and Financial Practice, Nobel Lecture." New York University, Department of Finance, December 8.

8. Engle, R. F., S. M. Focardi, & F. J. Fabozzi (2007): *"ARCH/GARCH Models in Applied Financial Econometrics."*, http://pages.stern.nyu.edu/ rengle/ARCHGARCH.pdf.

   Engle, R. F. & S. Manganelli (2001): "Value at Risk Models in Finance." *European Central Bank*, Working Paper No. 75, ISSN 1651 1810.

   Giot, P. & S. Laurent (2003): "Modelling Daily Value at Risk Using Realized Volatility and ARCH Types Models." *Journal of Empirical Finance*, May.

9. Jimnez-Martn, J., M. McAleer, & T. Prez-Amaral (2009): *"The Ten Commandments for Managing Value-at-Risk Under the Basel II Accord." ECO2008-06091/ECON*, March.

10. Larsson, O. (2005): "Forecasting Volatility and Value at Risk: Stochastic Volatility vs GARCH." November 1.

11. Loddo, A. (2006): "Bayesian Analysis of Multivariate Stochastic Volatilityand Dynamic Models." *University of Missouri-Columbi*, August.

Author　　　　Supervisor

Chapter 1

Introduction

Behavior of financial time series is a hot topic considering unexpected extreme events of price movements. Recent global financial crises 2008 as well as the former market crash in October 1987 and many others are perpetual triggers to further search for more efficient and exact predicting instruments. All of these crisis are characterized by the substantial increase in market volatility. Volatility in finance is a variable with ability to characterize the variations of returns in time. The point of the interest is therefore highly precise estimate of this volatility and its subsequent prediction. By virtue of its fundamentals it is also a cynosure of risk management and its measures for risk evaluations.

One of the most widely known measure in use of risk management departments across industry is the popular Value at Risk (VaR). For financial institutions it is also required by market risk framework accord Basel II to report 10-day ahead VaR at 95% level. VaR is simply defined as a measure of potential loss of some risky value associated with general market movements over a defined period of time with a given confidence interval.

It is now obvious that by our thesis we will contribute to works that study volatility models and consequently evaluate them in purpose of VaR predictions. Considering a bulk of aforementioned literature we find out that there is still not much written and examined on topic of Stochastic Volatility (SV) models, firstly introduced by Taylor (1982). That is why we choose it as center point of our thesis. Contrary to Stochastic Volatility models there are very popular and nowadays widely employed General Autoregressive Conditional Heteroskedasticity (GARCH) models established by Bollerslev (1986). To deviate from the majority of GARCH models analysis we therefore choose comparison of GARCH models with SV models and evaluate them in sence of parametric VaR predicting ability. From the group of SV models we chose autoregressive SV model of order one and collaterally GARCH(1,1) both with Student's t distributions.

SV models and GARCH models are very similar and well established in financial econometrics. Both distinguish of time-varying and persistent volatility as well as of the leptokurtosis in financial return series (Liesenfeld & Richard 2003). Generally, excepting the calculation methods, SV is treated as adequate substitution for GARCH models. However, the advantage of SV models is also the ability to account for leverage effect.

Literature discussing SV is more or less analyzing the Gaussian Stochastic Volatility. The one discussing GARCH models with its modifications and extensions is very far-reaching. From the theoretical papers comparing these two models are worth noticing, e.g, Carnero *et al.* (2004), Gerlach & Tuyl (2006) and Bai *et al.* (2003). These paper will be also the fundamentals for our thesis.

The reason why SV models did not gain on popularity as GARCH models did is simply in its estimation procedure. While GARCH is easily estimated via Maximum Likelihood Estimator, calculation of likelihood value for SV model is very complicated and time consuming. Here it comes to the need of integrations trough high-dimensional integrals. This cannot be done analytically and some more complex numerical methods are required. The main reason is that SV volatility enters the model nonlinearly, as log-volatility and is explained, except its lagged variable, by random error process. So the SV model has two random error terms[1] and can be more flexible in explaining price movements. Nevertheless this implies that volatility in SV is latent process that cannot be measurable.

For estimation of SV model many procedures have been proposed. From methods of moments, trough maximum likelihood to Monte Carlo Markov Chain simulations. In our thesis we adopt a relatively new procedure of Efficient Importance Sampling (EIS) Monte Carlo (MC) based on paper by Liesenfeld & Richard (2003). Importance sampling is a tool which provides an estimate of a high-dimensional numerical integrals and so on the value of likelihood function. It is ideally suited for SV differential equations computations.

Aim of this paper is to find out whether SV models with t-distribution can be alternative or superior to GARCH-t models in predicting risk measures. These two models when calculated, will be compared to each other for five different world indices in empirical part of this thesis. As we already men-

[1] GARCH has only one error term in return equation.

tioned, we calculate VaR measure from estimated volatilities and subsequently we test these risk measures by backtesting procedures. We choose particular backtests based on our data analysis. The most decision power is ascribed to unconditional and conditional coverage.

The thesis is organized as follows. In Chapter 2 we introduce the risk management popular measure Value at Risk together with its many different approaches of calculations. We go in details trough parametric method, the one used for our computations. In Chapter 3 we analyze volatility forecasting and review SV and GARCH models. Here we also describe common Maximum Likelihood Estimator in a nutshell and EIS procedure in details. 4th Chapter is dedicated to introduction of backtesting methods used for evaluation. Finally, Chapter 5 is an empirical study of two competing models with resume of results. Conclusion of whole work is in the last but not least Chapter 6.

Chapter 2

Value at Risk

The concept of value at risk (VaR) emerged as the one way of measuring the uncertainty in the future value of different financial instruments that a financial institution faces. Its fundamental purpose is to summarize the potential for deviations from a target or expected value and to define a loss that we are sure at some confidence level $(1-\alpha)$ will not be exceeded over some period of time t. VaR gives us possibility to easily capture this risk by one formula. Three major components of risk that financial institution faces[1] are besides market risk also credit risk and operational risk. As there are several sources of risk, in our thesis we consider only market risk, while VaR is generally preferred approach for measuring it. However the concept is also applicable to other types of risk.

Reasons to be one of the most popular comprehensive tools in this area can be summarized as follows: It is easy to implement and understand; it simply claims the maximum possible loss with some chosen probability; is measurable at any level, from an individual trade or portfolio, up to a single enterprise-wide VaR measure covering all the risks in the firm as a whole; it is expressed in its units during given period; is a universal metric that applies to all activities and to all types of risk and can be compared across markets and different exposures; and last, but not least, it is used for risk assessment as well as for setting margin requirements.

Following the lead from main regulators and large international banks during the mid-1990s, almost all financial institutions nowadays use some form of VaR as a risk metric (Alexander 2009). As popularity of VaR rises its criticism widely spreads in recent years, also encouraged by global financial crisis 2008 and many modifications of fundamental VaR have been proposed so far.

[1] In fact, it faces many more risks as e.g. systemic risk, pension risk, concentration risk, strategic risk, reputation risk, liquidity risk, legal risk.

Between classical modifications of VaR belongs mean-conditional VaR (CVaR) introduced by Rockafellar & Uryasev (2000) that exclusively incorporate the negative tail risk while it focuses not only on frequency but also on the size of losses when extreme event occurs and can involve discreetness. Good comparison study on VaR vs. CVaR provides Sarykalin *et al.* (2008). Another modification is implementation of Cornish-Fisher Expansion (CFVaR). This method takes into account the higher skewness and kurtosis and is used when dealing with returns wit not normal distribution. For more details see Maillard (2012). Extreme Value Theory(EVT) is a risk capital measure similar to VaR, but it accounts only for the tails of distribution where extreme values lies. VaR looks at whole distribution, see e.g. McNeil & Frey (2000).

2.1 Brief History of VaR

It is known that history of risk measurement VaR is traced since 1980s, when firms began to quickly develop, rise and became more complex. All they were in need of suitable management and measures for handling their internal risk. Thorough, the mathematics that underlie VaR were largely developed already in the context of portfolio theory i.e. by Leavens (1945), Markowitz (1952) and Roy (1952), even their efforts were led towards a different ends, concretely to devising optimal portfolios for equity investors (Damodaran 2007).

However, the main trigger for the first VaR theory boost is assigned to stock market crash in 1987. Later there was also contribution in the rest of the world occurring as Mexican (1994-1995), Asian (1997-1998) and Japanese (1990) crises. Those years are also known for the increase of the off-balance-sheet products and disastrous derivatives trading techniques e.g. Long-Term Capital Management in 1998 and Orange County in 1994. All these affairs lead to a great need of sophisticated risk management between and within banks and financial institutions. The thing of matter is also the extending volume of trading portfolios with its growing volatilities. The first regulatory measures that evoke VaR, were initiated in 1980, when the Securities Exchange Commission (SEC) tied the capital requirements of financial service firms to the losses that would be incurred, with 95% confidence over a thirty-day interval, in different security classes. At that time historical returns were used to compute these potential losses.

Latter regulatory frameworks for reliable financial risk management and regulation of financial institutions are Basel Capital Accords. Beginning with

Basel I (1988), Basel II (2004) with its many updates by the Basel Committee on Banking Supervision (BCBS) under the Bank of International Settlements (BIS) or Capital Adequacy Directive 1 (93/6/EEC) and (98/31/EEC) of the European Economic Community. BASEL II requires reporting 10-day VaR at 95% level. In 2010-2011 there was agreed BASEL III, as a result of recent financial crises. It strengthens bank capital requirements and introduces new regulatory requirements on bank liquidity and bank leverage. It should be brought into effect in 2013.

At about the same time in 1994 J.P.Morgan releases its RiskMetrics^{RM} system, where the term of VaR is used. System allows the access of public to data on the variances of and covariances across various security and asset classes, that it had used internally for almost a decade to manage risk, and allowed software-makers to develop a new software to measure risk. (Damodaran 2007)

Since then, as we already mentioned, VaR has gained lot of popularity between banks, regulators, financial institutions and lately also between non-financial service firms as commodity and energy merchants, and other trading organizations.

2.2 Mathematical Definition of VaR

According to Jorion (2006) for a given portfolio, probability and time horizon, VaR is defined as a threshold value such that the probability that the mark-to-market loss on the portfolio over the given time horizon exceeds this value (assuming normal markets and no trading in the portfolio) is the given probability level. In other words, VaR is measure of potential loss of risky value associated with general market movements over a defined period of time for a given confidence interval. In statistics, VaR represents a one-side quantile from a return distribution observed over a particular time horizon for a defined time period. When estimating VaR, we work with the distribution of loss and profit and two basic parameters, mentioned confidence level and time horizon. Important issue is to find most suitable cumulative distribution function (CDF) for given financial returns. As we will mention later, they are usually not normal distributed while they feature fat tails, higher kurtosis, etc..

For mathematical definition of VaR we use one by McNeil *et al.* (2005):

Definition 2.1 (Value at Risk). Given some confidence level $\alpha \in (0, 1)$. The *Value-at-Risk* VaR of our portfolio at the confidence level α is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1-\alpha)$. Formally,

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1-\alpha\} = \inf\{l \in :F_L(l) \geq \alpha\}. \tag{2.1}$$

First equation is definition of VaR and second presents the probability distribution of profit and loss, where $F_L(l)$ is its cumulative distribution function. The minus sign presents VaR only as a positive value.

As we already mentioned, VaR is a one-side quantile function, represented by an inverse function to CDF. To be precise, we define quantile function as:

Definition 2.2 (Generalized Inverse and Quantile Function).

(i) Given some increasing function $T : \mathbb{R} \to \mathbb{R}$, the *generalized inverse* function of T is defined by $T^{\leftarrow}(y) := \inf\{x \in :T(x) \geqslant y\}$, where we use the convention that the infimum of an empty set is ∞.

(ii) Given some df F, the generalized inverse $F^{\leftarrow}$ is called the *quantile function* of F. For $\alpha \in (0,1)$ the α-quantile of F is given by

$$q_\alpha(F) := F^{\leftarrow}(\alpha) = \inf\{x \in :F(x) \geqslant y\}. \tag{2.2}$$

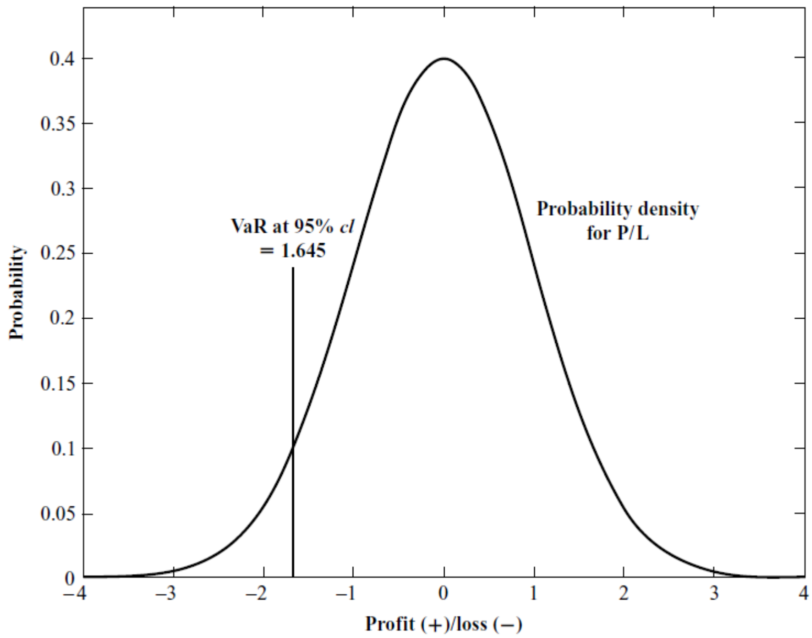
If F is continuous and strictly increasing, then $q_\alpha(F) = F^{-1}(\alpha)$, where F^{-1} is the *ordinary inverse* of F. Figure 2.1 graphically shows the 95% VaR for standardized normally distributed profit and loss.

2.3 VaR Approaches

Popularity of VaR and its many forms cause that there are known many different approaches for VaR calculations. There is no single universal method to obtain perfectly suitable result, but they all follow a common general structure and result from the characteristics of financial data. This structure can be summarized in three points: first to mark-to-market the portfolio, second to estimate the distribution of portfolio returns and last to compute the VaR of this portfolio. For more details see Engle & Manganelli (2004). We will introduce briefly the main methods that are classified into three broad categories:

- Parametric methods (the one we are going to use in our thesis)
- Nonparametric methods (usually simulations)

Figure 2.1: 95% VaR of normal distributed returns



Source: Dowd (2005).

- Semi-parametric methods

Parametric method involve GARCH and SV methods as well as RiskMetrics^{RM} method, that is actually one case of Exponentially Weighted Moving Average (EWMA) method. This is the approach we are going to use in our thesis. These methods do make assumptions about the distribution of returns, which is their weaker point. Also to provide accurate estimations, they need data sets with enough observations in order to estimate parameters. Advantages of these methods are concerned to be that they provide straightforward VaR formulas for calculations.

The connection between our analyzed SV and GARCH volatility models and α-percentage value at risk at time t is straightforward

$$VaR_{\alpha,t} = \mu_t + h_t q_\alpha(F), \tag{2.3}$$

where μ_t and h_t are mean and standard deviation of given loss distribution F , respectively, and $q_\alpha(F)$ is the α-quantile of given distribution. h_t will be the point of our interest. The goal of this thesis will be the comparative study of SV and GARCH volatility estimation in VaR performance. We will discuss it in more details in next chapter. Here is also crucial the right choice of financial returns distribution. It is widely known that time series returns does not behave according to normal distribution. Their known characteristics, as fat-tailness and leptokurtosis, are better described with Students-t, or skewed-t distribution, as well as with generalized error distribution (GED). In this thesis we consider Students-t distribution that has cumulative distribution function (cdf)

$$p = F(x|\nu) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \frac{1}{\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}} \mathrm{dt}, \tag{2.4}$$

where ν denotes the degrees of freedom (for $\nu \to \infty$, the t-distribution approaches the normal distribution) and the gamma function is $\Gamma(\nu) = \int_0^\infty e^{-x} x^{\nu-1} dx$. Degree of freedom represents a parameter to be estimated. This distribution is symmetric around zero and for $\nu > 4$ the conditional kurtosis equals $3(\nu-2)(\nu-4)^{-1}$.

Nonparametric method is for example very known and easy to implement Historical Simulation (HS), where no distributional assumption are needed. It reduces the risk measure estimation problem to one-dimensional problem and uses a simple concept of rolling windows. But as it is also unconditional method and has one implicit assumption, that distribution of portfolio return doest change over time, this method does not provide very creditable results. On the other hand Alexander (2009) claims that HS method is becoming more and more popular because of its limitations, that are actually advantages. He states that historical VaR does not have to make an assumption about the form of the distribution and dependencies of the risk factors are inferred directly from historical observations which still include the dynamic behavior of risk factors in a natural and realistic manner. These estimates may also present predictable jumps, due to the discreteness of extreme returns.

Monte Carlo method is rather general name for any approach based on simulations of an explicit parametric model for risk factor changes. It is an extremely flexible tool that has numerous applications to finance. Can be either conditional or unconditional according to the explicit model is dynamic time series or static distributional model. This method is generally considered as very accurate VaR estimator, however, the calculations are very time-consuming as all simulations. It is used as method of 'last resort' when analytical methods fail. We will use such a simulations in our thesis when calculating SV volatility.

Semi-parametric method are recently developed and they gain on popularity. They are hybrid approaches, arising from criticism of particular VaR methods. One appreciable approach is combination of RiskMetrics^{RM} and historical simulation methodologies. It differs by applying exponentially declining weights to past returns of the portfolio.

2.4 Criticism & Limitations

VaR criticism captures wide area of its characteristics from very different points of view. Common drawback of many risk measures and is not unique for VaR, is that their estimates can be subject to errors. VaR systems can be subject to model risk, the risk of errors arising from inappropriate assumptions on which model was based, or implementation risk, the risk of errors arising from the way in which the systems are implemented.

More peculiar drawbacks are analyzed e.i. in Dowd (2005) or McNeil *et al.* (2005). One is that it does not give us any information about how much we can loose if certain probability is over-crossed, i.e. when loss really occurs. It tells us only the bottom bound of the loss and the probability with which it wont occur during certain period.

Another limitation is that VaR doesn't satisfy the risk diversification principal and has poor aggregation properties. Mathematically, it is not a coherent risk measure as Artzner *et al.* (1999) firstly showed in his work, that VaR of the diversified portfolio is larger than the VaR of the undiversified portfolio.

Often highlighted limitation of VaR is that it disobey the feature of sub-additivity. A risk measure $\rho(.)$ is said to be sub-additive if the measured risk of the sum of positions A and B is less than or equal to the sum of the measured risks of the individual positions considered on their own, $\rho(A + B) \leq \rho(A) + \rho(B)$. It means that cumulating individual risks does not increase overall risk, what is highly desirable from a risk measure.

Based on this criticism alternative risk measures have been developed as Expected Shortfall (ES) and Extreme Value Theory (EVT), demonstrably superior. ES, also called Conditional VaR, is the expected loss conditional that the loss is greater than some q percentile of the loss distribution, certain VaR. It is clear that ES is more precise on the tails of distribution than normal VaR.

On the other hand, EVT is more recent approach in economics, even it was already widely used in other sciences. This approach considers only extreme risk events that are rare and have catastrophic impact. This approach does not need any assumptions about return distribution. It considers only distribution of tails and can fit on this tail any probability distribution.

The question arises why to bother with VaR predictions when other nowadays methods are superior in many different manners. As Dowd (2005) answers: *"Part of the answer is that there will be a need to measure VaR for as long as there is a demand for VaR itself: if someone wants the number, then someone has to measure it, and whether they should want the number in the first place is another matter."* Another reason is that VaR is simply the quantile of some distribution what is still desirable when calculating other risk measures, e.g. ES.

Chapter 3

Modeling Volatility

The main characteristic of any financial asset is its return that is considered to be a random variable. This random variable can be partly described with assets volatility, which describes the spread of outcomes of this variable and can be used to value a market risk.[1]

It is now widely agreed that financial asset return volatilities and correlations (henceforth volatilities) are time varying, with persistent dynamics. (Andersen *et al.* 2007). Even some believes that volatility as a random future event is unpredictable, there are some stylized facts about financial volatility, that can disaffirm this thinking. Return series are generally found to have their marginal distributions leptokurtic. They feature fatter tails than normal distribution and excess kurtosis. Very familiar pattern that volatility exhibits is volatility clustering. As noted by Mandelbrot (1972), large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes, meaning that while returns themselves are uncorrelated the absolute and squared returns display a positive, significant and slowly decaying autocorrelation function. Fact, that volatility tends to increase when prices decrease (price movements are negatively correlated with volatility) is called leverage effect. This has been documented for the first time by Black (1976). There are also known patters that volatilities of different time series are mutually correlated and there are co-movements of these volatilities. These correlations are explained by multivariate group of models.

Modeling volatility is crucial issue to risk management and always been of a big importance, since volatility is considered to be a measure of risk. Banks and

[1] Although volatility is related to risk, it is not exactly the same. Risk is uncertainty of a negative outcome of some event (e.g. stock return); volatility measures the spread of outcomes (Ladokhin 2009).

many other financial institutions assess risks by VaR models based on volatility estimates as it was explained in previous chapter. In other words, modeling and forecasting the covariance structure of asset returns is very crucial for managing risk and relevant models (Tersvirta 2009).

In this chapter we will analyze and compare two, on the first sight, very similar models for volatility predictions. We consider popular General Autoregressive Conditional Heteroskedasticity model against less known Autoregressive Stochastic Volatility model. Both for modeling time varying volatility, capable to capture the volatility clustering. Further, we will use these models for modeling volatility and VaR, respectively, and we analyze their performance in the sense of VaR predictions.

Nevertheless, todays widely used approach for modeling volatility is concept of Realized Volatility (RV) that takes advantage of intra day information and analysis the high-frequency data. This method is quickly developing in recent years due to extensive databases and booming evolution of computer sciences. It is confirmed by many works that RV can outperform GARCH volatility approaches while it provides far more accurate calculations and has the advantage of being a non-parametric approach. However, this is not subject of our thesis it can be analyzed in further research.

3.1 Overview: GARCH vs. SV

When talking about modeling volatility two similar approaches were developed synchronously among many others, but only one of these gained the popularity, especially, when it comes to empirical examinations.

Since 1982, when Engle (1982) introduced his first Autoregressive Conditional Heteroskedasticity (ARCH) model and consecutively his student Bollerslev (1986) upgraded it to the General Autoregressive Conditional Heteroskedasticity (GARCH) model, these methods attracted plenty of attention. Lot of work has been dedicated to research these practices, many extensions of the ARCH/GARCH models have been developed and the method is widely used nowadays.

Collaterally, in 1986 another method for modeling volatility has been suggested, called the Autoregressive Stochastic Volatility (ARSV) or simply Stochastic Volatility (SV) model by Taylor (1986). It is at large considered to be a successful alternative to the class of ARCH models in distinguishing of the time-varying and persistent volatility as well as of the leptokurtosis in financial return series (Liesenfeld & Richard 2003). Even the distinct advantage of stochastic volatility models is the incorporation of leverage effect, at least at the univariate level (McAleer 2009) and also the possibility to capture the main empirical properties often observed in daily series of financial returns in a more appropriate way (for more details, see Carnero *et al.* (2004)), there is still much less literature on this topic. The main reason of being not such a favorite method is its complexity and difficulty of estimations, mainly its computations. Basically, the trouble stems from the core of the SV model, where volatility is modeled as an unobserved latent variable and while it assumes two error processes compared to the GARCH that assumes only one.

Our analysis will be based on several papers trying to provide the comparison of SV and ARCH/GARCH family of models, mentioned subsequently.

The theoretical examination of GARCH and SV comparison provide Carnero *et al.* (2001). They show that SV models better explain the excess kurtosis, low first order autocorrelation and high persistence of volatility. They also show that SV model is less dependent on the choice of returns distribution. So *et al.* (1999) compare the predictive performance of both models analyzing returns of five exchange rates. They conclude that both models have similar performance in terms of the Mean Square Prediction Error and Mean Absolute Prediction Error. Only in two of the all series considered, the ARSV predictions of volatility outperform the GARCH predictions. Carnero *et al.* (2004) shows the different relationship between kurtosis, persistence of shocks to volatility, and first-order autocorrelation of squares in GARCH and ARSV models. While parameters explaining these moments are closely linked in GARCH model, they can be modeled independently in SV model and better represent observed patterns of financial time series. Study of Yu (2002) examined nine univariate models to forecast volatility of index NZSE 40[2] and considers SV model to be superior according to three different asymmetric loss function and Root Mean Equare Error (RMSE) testing. In the paper Mapa *et al.* (2010) they compare basic GARCH volatility forecasts with one from SV model, computed trough Kalman Filter or Markov chain Monte Carlo (MCMC) method. They conclude that SV model captures more aspects of volatility than GARCH model due to its sources of variability and produces lower forecast errors. Bai *et al.* (2003) study volatility clustering and conditional non-normality and its implications

[2]New Zeland index, one of the least regulated economies and freest share markets in the world

to leptokurtosis for both GARCH and ARSV models. They compare this two models using kurtosis autocorrelation relationship. Gerlach & Tuyl (2006) also confront the empirical fit of both model and presents MCMC and importance sampling techniques for volatility estimation.

According to reviewed literature we assume, that ARSV models are adequate substitutes to GARCH models, when omitting their calculation difficulties, if not superior. In our thesis we will try to testify this affirmation in our empirical part. However, as noticed by Hafner & Preminger (2010), the two models are non-nested and this can complicate the model comparison. Models can be rejected or accepted against each other and no one can be superior.

3.2 Stochastic Volatility Models

The origins of SV models are diminish. We can say they date back 40 years when Clark (1973) introduced Bochner's (1949) time-changed Brownian motion into financial economics. Initiatory published paper considering real volatility clustering and introducing bases of SV models is considered to be the one by Taylor (1982). Later in Taylor (1986) he expands his theory and defines the first SV model in context of Financial Econometrics[3]. The key feature of SV models, the possibility to deal with the leverage effects of financial time series, is discussed afterwards in paper of Hull & White (1987).

In our thesis we will stem from the later work of Taylor (1986) where the basic model of SV is defined. It is also called log-normal autoregressive stochastic volatility ARSV(1) model, while he suggests to model the logarithm of volatility as AR(1) process. This model is used to account for the well documented autoregressive behavior in the volatility of financial return series and represents an alternative to the ARCH/GARCH models (Liesenfeld & Jung 1997). To demark the family of SV models, we can say, that these models have the volatility driven by a latent process representing the flow of price relevant information. The corresponding log-normal ARSV model is given by

$$r_t = \sqrt{h_t}\epsilon_t \tag{3.1}$$

$$\ln h_t = \gamma + \delta \ln h_{t-1} + \rho\eta_t, \tag{3.2}$$

where r_t is return observed at time t, $\{\epsilon_t\}$ and $\{\eta_t\}$ are mutually independent and identically distributed (i.i.d.) random variables, and at the same time $\{\eta_t\}$

[3]Taylor (1982) and Taylor (1986) did not discussed leverage effect, so far

is standard normal. Assumption of $\{\eta_t\}$ normality is validated by Andersen *et al.* (2003) who show that log-volatility process can be well approximated by a Normal distribution. $\{\epsilon_t\}$ in the basic Taylor's model is also standard normal, but as we already mentioned, in this thesis we will work with Student's t distribution of returns, $\{\epsilon_t\} \sim t_\nu$. Vector $\theta = (\gamma, \delta, \rho)' \in \Theta$ of parameters is to be estimated, where δ measures the persistence of the latent volatility process $\ln h_t$, while ρ measures the standard deviation of volatility shocks. For $|\delta| < 1$ returns r_t are strictly stationary, ρ is assumed to be greater than zero. And as it was mentioned before, the log-volatility $\ln h_t$ as latent variable is unobservable, contrary to the GARCH model volatility.

Autocorrelation function (ACF) of squared returns, variance and kurtosis of returns can be expressed as a function of parameters vector θ. $\phi(\tau, \theta)$, $\psi(\theta)$ and $\kappa(\theta)$ are ACF, variance and kurtosis, respectively, and for the Gaussian ARSV(1) model holds

$$\phi(\tau, \theta) = \frac{\exp\left(\frac{\rho^2}{1-\delta^2}\delta^\tau\right) - 1}{\kappa - 1}, \tau \geq 1 \tag{3.3}$$

$$\psi(\theta) = \exp\left(\frac{\gamma}{1-\delta} + \frac{\rho^2}{2(1-\delta^2)}\right) \tag{3.4}$$

$$\kappa(\theta) = 3\exp\left(\frac{\rho^2}{1-\delta^2}\right). \tag{3.5}$$

The latent volatility h_t enters the SV model nonlinearly, which leads to a likelihood function that depends on high-dimensional integrals, from where the problem of computations results. Many papers concentrate on finding the most appropriate way how to overcome the hassle of estimating the stochastic volatility parameters. The paper of Broto & Ruiz (2004) provides very clear review of so far known methods for estimations. They divide these methods onto the two main groups. First group is based on statistical properties of r_t and second group uses linearized SV model and subsequently the estimation is based on logarithm of squared returns. In our thesis we will be interested only in the first group.

Estimating SV parameters based on r_t

To present at least the basics of these estimators we start with the simplest procedures called Method ofMmoments (MM), also used by Taylor (1986) in his early works. Later, Melino & Turnbull (1990) apply Generalized Method of Moments (GMM) and coincidently, at the same time Duffie & Singleton (1990) propose the Simulated Method of Moments (SMM). These methods are generally considered to be easy to implement but have poor finite sample properties. The GMM is based on convergence of selected sample moments to their unconditional expected values and the later one simply replaces analytical moments by simulated processes.

Maximum Likelihood (ML) methods are considered to be better while their efficiency is superior with respect to MM. In order to derive likelihood from the equation, the vector of unobserved volatilities has to be integrated and this is the already mentioned problem. Harvey *et al.* (1994) and Ruiz (1994) suggest potential Quasi-Maximum Likelihood (QML) approach but do not solve the computational difficulty.

MCMC estimators are firstly and independently proposed by Jacquier *et al.* (1994), Kim *et al.* (1998) and have been followed by many other authors in the context of SV parameters estimations. Jacquier *et al.* (1994) also show that it is more efficient method that both QML and MM estimators. Concurrently the numerical methods based on important sampling and MCMC procedure are developing and the first authors that used these procedures are, for example, Shephard *et al.* (1995) and Sandmann & Koopman (1998). The big advantage of this importance sampling over MCMC algorithm is that it is less computationally demanding and avoids convergence problems (Broto & Ruiz 2004). In our thesis we are going to follow the method used in Liesenfeld & Richard (2003), the ML estimation which is based on the Efficient Importance Sampling (EIS). This technique was firstly proposed by Richard & Zhang (1996) for computing high-dimensional integrals. Recently elaborated summarizing work in details on EIS is presented by Richard & Zhang (2007).

An optional method is an estimation procedure by means of auxiliary model. One known method is Efficient Method of Moment (EMM), used by Gallant *et al.* (1997) or Jiang & Van Der Sluis (1999), and another is Indirect Inference method proposed by Gourieroux *et al.* (1993), They also show that these two estimators are asymptotically equivalent. The idea is in the right choice of an auxiliary model that is easy to estimate. It is clear that efficiency of these methods strongly depends on the choice of auxiliary model and the main drawback is the expensiveness in computational terms.

In the following subsection we will introduce the technique of EIS and its implementation to SV model and its extensions.

3.2.1 Efficient Importance Sampling

EIS is relatively new method developed by Richard & Zhang (1996), Liesenfeld & Richard (2003) and Richard & Zhang (2007). Essentially, it is a strategy to build a sampling density containing a huge number of parameters and ideally suited to provide extremely accurate estimates of some high dimensional integrals (Pastorello & Rossi 2010). For explanation of EIS we rewrite the SV model so that we replace $\ln h_t$ with λ_t, that will represent log volatility. Then we have

$$r_t = \beta \exp(\frac{\lambda_t}{2})\epsilon_t \quad \{\epsilon_t\} \sim t_\nu \tag{3.6}$$

$$\lambda_t = \delta\lambda_{t-1} + \rho\eta_t, \tag{3.7}$$

where ν are degrees of freedom and $t = 1...T$, $\{\eta_t\} \sim (0, 1)$ and β is a scale parameter that removes the necessity of including a constant term γ in the log-volatility equation. According to Liesenfeld & Richard (2003) we denote r_t an n-dimensional vector of observable random variable and sequence $R = \{r_t\}_{t=1}^T$, along with λ_t a q-dimensional vector of latent variable and sequence $\Lambda = \{\lambda_t\}_{t=1}^T$. We let $f(R, \Lambda; \theta)$ to represent a joint density of R and Λ, then we can write the likelihood associated with observable R as

$$L(\theta; R) = \int f(R, \Lambda; \theta) \mathrm{d}\Lambda. \tag{3.8}$$

This can be factorized to

$$L(\theta; R) = \int \prod_{t=1}^{T} f(r_t, \lambda_t | \Lambda_{t-1}, R_{t-1}, \theta) \mathrm{d}\Lambda \tag{3.9}$$

and the right side of the equation can be further factorized to

$$f(r_t, \lambda_t|\Lambda_{t-1}, R_{t-1}, \theta) = g(r_t|\lambda_t, R_{t-1}, \theta)p(\lambda_t|\Lambda_{t-1}, R_{t-1}\theta), \tag{3.10}$$

where $g(.)$ denotes the conditional density of r_t given (λ_t, R_{t-1})associated with ϵ_t in (3.6) and $p(.)$ the conditional density of λ_t given (Λ_{t-1}, R_{t-1}) associated with AR process of the latent variable.

The problem of SV models estimations is obvious from the likelihood equation (3.9) where we have to integrate over the latent factor a T dimensional integral. This can not be solved analytically, so the numerical simulation method is required. But the natural MC estimator is highly inefficient in this case, according to the dramatic increase of the MC sampling variance with the sample size T, plus it ignores the relation of observation of R and the underlying latent process λ_t. For more details see Danielsson & Richard (1993).

Here comes the EIS technique to solve this struggle. It searches for such a sequence of samplers which takes into account that r_t's are conveying the sample information on λ_t's. Therefore, we denote a sequence of auxiliary samples, $\{m(\lambda_t|\Lambda_{t-1}, a_t)\}_{t=1}^T$ indexed by $A = \{a_t\}_{t=1}^T$, the auxiliary parameters. Now for any A the integral in likelihood equation (3.9) can be rewritten as

$$L(\theta; R) = \int \prod_{t=1}^{T} \left[\frac{f(r_t, \lambda_t|\Lambda_{t-1}, R_{t-1}, \theta)}{m(\lambda_t|\Lambda_{t-1}, a_t)} \right] \prod_{t=1}^{T} m(\lambda_t|\Lambda_{t-1}, a_t) \mathrm{d}\Lambda \tag{3.11}$$

with its corresponding importance sampling MC estimate

$$\tilde{L}_N(\theta; R, A) = \frac{1}{N} \sum_{i=1}^{N} \left\{ \prod_{t=1}^{T} \left[\frac{f\big(r_t, \tilde{\lambda}_t^{(i)}(a_t)|\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}), R_{t-1}, \theta\big)}{m\big(\tilde{\lambda}_t^{(i)}(a_t)|\tilde{\Lambda}_{t-1}^{(i)}(a_{t-1}), a_t\big)} \right] \right\}, \tag{3.12}$$

where $\{\tilde{\lambda}_t^{(i)}(a_t)\}_{t=1}^T$ is the trajectory of importance sampler m.

EIS tries to minimize the MC sampling variance of (3.12) by choosing such a values of auxiliary parameters A that will provide a good match of product in nominator and denominator of equation. While it is a high-dimensional minimization problem, when solving for all relevant θ's, Liesenfeld & Richard (2003) breaks it down to separate low-dimensional subproblems. They draw up a functional approximation $k(\Lambda_t; a_t)$ for the density $f(r_t, \lambda_t|\Lambda_{t-1}, R_{t-1}, \theta)$, called density kernel for m, that will be analytically integrable with respect to λ_t as

$$m(\lambda_t|\Lambda_{t-1}, a_t = \frac{k(\Lambda_t; a_t)}{\chi(\Lambda_{t-1}, a_t)}, \text{ where } \chi(\Lambda_{t-1}, a_t) = \int k(\Lambda_t; a_t) \mathrm{d}\lambda_t. \tag{3.13}$$

Now finding parameters of auxiliary sampler is solving low-dimensional least squares problem

$$\hat{a}_t(\theta) = \arg\min_{a_t} \sum_{i=1}^{N} \Big\{ \ln \Big[f\big(r_t, \tilde{\lambda}_t^{(i)}(\theta)|\tilde{\Lambda}_{t-1}^{(i)}(\theta), R_{t-1}, \theta\big) . \chi\big(\tilde{\Lambda}_t^{(i)}(\theta); \hat{a}_{t+1}(\theta)\big) \Big] \\ - c_t - \ln k(\tilde{\Lambda}_t^{(i)}(\theta); a_t) \Big\}^2, \tag{3.14}$$

for $t : T \to 1$, based on natural samplers.

Finally, the EIS estimate of likelihood in (3.12) is obtained by substituting

$\{\hat{a}_t(\theta)\}_{t=1}^T$ for $\{a_t(\theta)\}_{t=1}^T$. ML-EIS estimates of θ are obtained by maximizing importance sampling MC estimate of likelihood in (3.12) with respect to θ, using an iterative numerical optimizer. Estimates of θ are obtained by maximizing $\tilde{L}_N(\theta; R, A)$, (3.12), with respect to θ, using an iterative numerical optimizer. As Liesenfeld & Richard (2003) declare, EIS turns out to be provider of very flexible and numerically extremely efficient procedure for the likelihood evaluation oh high-dimensional nonlinear latent variable models such as SV.

In our case of standard normal η_t and Student's t ϵ_t, the densities g, p will be given by

$$g(r_t|\lambda_t, \theta) \propto \exp(-\lambda_t/2)\left\{1 + \frac{r_t^2 \exp(-\lambda_t)}{\nu - 2}\right\}^{-(\nu+1)/2}, \nu > 2, \quad (3.15)$$

$$p(\lambda_t|\lambda_{t-1}, \theta) \propto \exp\left\{ - \frac{1}{1\rho^2}(\lambda_t - \delta\lambda_{t-1})^2\right\}, \quad (3.16)$$

where multiplicative factors not depending on λ_t are omitted. To calculate the EIS estimate of likelihood we first use the natural samplers p to draw N trajectories. Then we use these draws to solve the sequence of least-squares problems in (3.14) and finally we use the EIS samplers $\{m(\lambda_t|\Lambda_{t-1}, \tilde{a}_t(\theta))\}_{t=1}^T$ to draw N trajectories $\{\tilde{\lambda}_t^{(i)}(\tilde{a}_t(\theta))\}_{t=1}^T$ from which EIS estimate of likelihood is calculated via (3.12).

Filtering, Smoothing and Predictions

As the parameters of SV are estimate, we can get estimate of λ_t in two forms. An estimate based on all observations up to, and possibly including, the one at time t is called filtered $\lambda_{t|t-1}$. Contrary, an estimate based on all observations in the sample including those which came after time t is called smoothed $\lambda_{t|t}$. (Ghysels *et al.* 1996). For a one-step-ahead volatility prediction for time t we then calculate

$$\hat{\beta} \exp(E(\hat{\lambda}_{t|t-1}/2)). \quad (3.17)$$

3.3 GARCH models

General Autoregressive Conditional Heteroskedasticity model was introduced in the seminal work of Bollerslev (1986) as generalization of ARCH(p,q) model, that has has shown to need high q in order to obtain estimates of conditional variance properly. This extension allows the past conditional variance to be considered in the present conditional variance estimate and extends the AR process equation to an ARMA process equation. Independently to Taylor (1982) another non-linear models considering volatility clustering were introduced. In other word, these models allow for heteroskedasticity and are able to model this changing conditional volatility in time. AS these models are very popular we are not going to analyze them in such details as SV models. Here we provide the fundamentals of the basic GARCH(p,q) model, that is Bollerslev (1986) is defined as

$$\epsilon_t = \sqrt{h_t} z_t \tag{3.18}$$

$$\text{where} \qquad \epsilon_t | \mathcal{F}_{t-1} \sim \mathrm{N}(0, h_t), \tag{3.19}$$

$$h_t = a_0 + \textstyle\sum_{i=1}^{q} a_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} b_j h_{t-j}, \tag{3.20}$$

where ϵ_t is mean-corrected return $\epsilon_t = y_t - \mu$, $h_t > 0$ is conditional variance of ϵ_t (conditional on the information available in t) and z_t is i.i.d. standard normally distributed random variable, $p \geq 0$ and $q \geq 0$. Vector $\theta_2 = (a_0, a_i, b_j) \in \Theta_2$ are parameters to be estimated. To ensure that conditional variance of ϵ_t is stationary, $a_0 > 0$, $a_i \geq 0$ for $i = 1, 2...q$, $b_j \geq 0$ for $j = 1, 2...p$ and $\sum_{i=1}^{max(p,q)} (a_i + b_j) < 1$ have to hold. Last constraint is necessary and sufficient condition for the second-order stationarity of the Bollerslev (1986) model. It implies that the unconditional variance of ϵ_t is finite, whereas its conditional variance h_t^2 evolves over time.[4] Following Theorem 3.3.1 summarizes all previous restrictions of wide-sense stationarity.

Theorem 3.3.1 (Bollerslev (1986))**.** *The GARCH(p,q) process as defined in* (3.18) *is wide-sense stationary with* $E(\epsilon_t) = 0$*,* $var(\epsilon_t) = \frac{a_0}{\sum_{i=1}^{q} a_i + \sum_{i=1}^{q} b_i}$ *and* $cov(\epsilon_t, \epsilon_s) = 0$ *for* $t \neq s$ *if and only if* $\sum_{i=1}^{q} a_i + \sum_{i=1}^{q} b_i < 1$.

For the GARCH(1,1) model with parameters $\theta_2 = (a_0, a, b) \in \Theta_2$ we can write ACF, variance and kurtosis, $\phi(\tau, \theta_2)$, $\psi(\theta_2)$ and $\kappa(\theta_2)$, respectively, as follows

$$\phi(1, \theta_2) = \frac{a(1 - b(a+b))}{1 - b^2 - 2ab} \tag{3.21}$$

$$\phi(\tau, \theta_2) = (a+b)\phi(\tau - 1), \quad \tau \geq 2 \tag{3.22}$$

$$\psi(\theta_2) = a_0/(1 - a - b) \tag{3.23}$$

$$\kappa(\theta_2) = 3 + \frac{6a^2}{1 - 3a^2 - 2ab - b^2}. \tag{3.24}$$

[4] The addition to mentioned ARCH model is that GARCH allows the conditional variance to be modeled by past values of itself in addition to the past shocks.

The special case of GARCH(p,q) is the Exponentially Weighted Moving Average (EWMA). This models is authorized and used by RiskMetrics^{TM} company for VaR calculations. The variance h_t is modeled as exponentially declining process $h_t = ah_{t-1} + (1-a)\epsilon_{t-1}^2$, the special case of GARCH(1,1). RiskMetrics^{TM} has standardized wights of $a = 0.94$ for daily data and go 75 data points backward in their estimation (Angelidis *et al.* 2004). Between preferred extensions of GARCH(p,q) belongs EGARCH(p,q) by Nelson (1991), TGARCH(p,q) by Zakoian (1994), APARCH(p,q) by Ding *et al.* (1993), FYGARCH(p,d,q) by Baillie *et al.* (1996) and many others. EGARCH and TGARCH account for missing asymmetry due to some extra parameters. Later one allows a response of volatility to good and bad news with different coefficients. APARCH and FY-GARCH encompass asymmetry, leverage effect and also the long memory property. We will not examine these models neither compare them with SV. This will be left for the further research together with extensions of SV.

When talking about comparison of two models, the main difference compared to SV models is that volatility in GARCH models is measurable with respect to the information set in $t-1$ (Hafner & Preminger 2010). The basic GARCH model is symmetric in the sense that positive shocks has the same impact on volatility as negative and does not account for leverage effect. And the huge difference important for practical usage is the estimation of GARCH models that is much less complicated than for SV models. Usually it is done by Maximum Likelihood Estimator, or Quasi-maximum Likelihood Estimator for non-normal innovations either when the conditional distribution is not perfectly known. The QMLE estimates are in general less precise than those from MLE, provided that the z_t are indeed Gaussian (Fabozzi *et al.* 2008). We will analyze it in details in following subsection 3.25.

As we already noted, for our thesis we will assume only Student's t distribution of $\{\epsilon_t\} \sim t_\nu$ in (3.18), where $\nu > 2$ are degrees of freedom and as $\nu \to \infty$ Student's t distribution approximates to Normal Distribution. If $\{\epsilon_t\} \sim t_\nu$, so the random variable z_t will have standardized Student's t distribution with $\nu > 2$. Formula of density function of this distribution is in (2.4).

3.3.1 Maximum Likelihood Estimator

Maximum Likelihood Estimation (MLE) is commonly used approach for estimating GARCH models, which provides an asymptotically efficient estimates.

Under the assumption of independently and identically distributed standardized innovations, z_t, and its density function $D(z_t;\nu)$ the log-likelihood function of $\{y_t(\theta)\}$ for a sample of T observations is defined as

$$\mathbb{L}_T(\{y_t\};\theta) = \sum_{t=1}^{T} \left[\ln\left[D(z_t(\theta);\nu)\right] - \frac{1}{2}\ln\left(h_t(\theta)\right)\right] \tag{3.25}$$

where θ is parameters vector that ought to be estimated for the conditional mean, variance and density function. This estimates of $\hat{\theta}$ we obtain by numerically maximizing the log-likelihood function, (3.25), using iterative optimization methods.[5] For our Student's t distribution we will have

$$\begin{aligned} \mathbb{L}_T(\{y_t\};\theta) = \quad & T \quad \left[\ln\Gamma\left(\frac{\nu+1}{2} - ln\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln[\pi(\nu-2)]\right]\right. \\ - \quad & \frac{1}{2}\sum_{t=1}^{T}\left[ln(h_t) + (1+\nu)\ln\left(1+\frac{z_t^2}{\nu-2}\right)\right], \end{aligned} \tag{3.26}$$

where all parameters are defined as in model GARCH(p,q) in (3.18). For more details and other distributions see Angelidis *et al.* (2004).

[5]Marquardt algorithm written by Marquardt (1963) in software Matlab

Chapter 4

Backtesting the Accuracy of VaR Models

For the verification or rejection of our assumptions about VaR predicting performance of the two models that are stated at the beginning of this paper, with certain confidence, we will use different quantitative, mostly statistical, methods called backtesting. However, the important part is also the data analysis and graphical analysis of observed returns with its corresponding VaRs, backtesting methods will give us the means to conclude the accuracy of our different VaR models and their performances. Under the word of accuracy we will understand how well does the model predicts a particular percentile or the entire distribution of returns, how well does it capture possible losses on given confidence level and how well does it predict the size and frequency of possible losses. There exists no uniform backtesting method with correct reliable results. The task is to choose such an appropriate method that balances out its performance in measuring power, its drawbacks and its appropriateness for our data. In this section we will introduce three promoted wide groups of statistical backtesting procedures known nowadays.

Since late 1990's there have been published many papers and studies that propose variety of backtesting methods and analyze which method is more appropriate. We will try to go over this literature and highlight the important ones. In details we describe only those methods that will be used in our empirical part. The choice of such a methods will be done based on summarization of pros and cons of each method in combination with the data analysis and the most suitable ones will be chosen. Quick overview of such an analysis is described below.

Preliminary Data Analysis

As we mentioned before, the indispensable part of the verification is also the analysis of the data. Our data should be, first, cleaned from all components that are not directly related to current or earlier market risk-taking. When we have the "cleaned" data, we can now plot the realized and forecast values and examine the obvious statistics. Very helpful is the construction of backtesting charts where observations are superimposed with corresponding VaR forecasts, histograms of distributions and quantile-quantile (QQ) charts of realized against forecast distributions of returns. On these graphs the behavior of the outliers, exceptions and violations of VaR is nicely visually performed and we are able to create a preliminary image of performance of our models. According to Dowd (2002) there are few common facts that we can use for interpretation in graphical analysis:

- A relatively high or low number of cases when losses are exceeding the VaR line indicates that our risk measure is likely to be too low, high respectively.
- If the VaR line seems to be too flat or smooth, probably our model is not adjusting adequately.
- Rapid changes in VaR line suggest considerable changes in volatilities or in the way Var is estimated.
- Also the major differences between high and low exceptions can indicate the biasedness of the risk measure.

4.1 Statistical Backtesting

All statistical tests are based on hypothesis-testing paradigm. This consists of specifying the null hypothesis and possible alternative hypothesis, selecting the significance level on which the null can be rejected or accepted and estimating the probability, p-value, associated with null being "true". Naturally we accept the null hypothesis if its p-value exceeds the significance level and otherwise.

When testing the hypothesis in statistics, we can make two types of errors that can be very costly for risk management: Type I occurs when falsely rejecting an accurate model and Type II error occurs when we fail to reject (incorrectly accept) the wrong model. The choice of significance level should be well consider, because it can weight out the likelihoods of occurring Type I or II errors. The higher significance level, more likely we are to accept null, what indicates lower likelihood of rejecting the true model (Type I) but higher likelihood of accepting a false model (Type II) and vice versa. A test can be said to be reliable if it is likely to avoid both types of error when used with an appropriate significance level (Dowd 2002). In our thesis we will work with widely adopted significance level $(1-\alpha) = 5\%$ for testing the hypothesis.

In the following sections we will discuss three main groups of backtesting methods. The most popular and easy to implement test group is the one investigating the sequences of VaR violations. For example the binomial backtesting proposed by Kupiec (1995), also the only current approach in Basel Accord and interval forecast backtesting approach of Christoffersen (1998). Second group are statistical backtest of VaRs at multiple confidence levels or also called density forecasts backtesting methods. These methods were firstly proposed by Crnkovic & Drachman (1997) and Diebold *et al.* (1998a) and later modified by Berkowitz (2001). The basic idea of this group is that conditional and unconditional property of an accurate VaR model should hold over all confidence levels at the same time. The last group, that we will present, is not based on hypothesis testing, instead, the forecast distribution transforms what might happen in the future into probability forecasts (Lopez 1999). Here the alternative evaluation method is proposed and its called probability forecast backtesting. Loss function, score function and benchmark are determined and it allows also for ranking the alternative models and diagnostic purposes. We use this method to decide between two models when other test results seem to be the same.

On the Basel II Backtesting Requirements

- Ultimate goal of Basel II Framework is to promote adequate capitalization of banks and to encourage improvements in risk management.

- Banks are allowed from Basel II to compute capital requirements on their own, using IRB (Internal Rating Based) approach.

- Backtesting must be be done over the longer look back period and on other confidence intervals than 99% interval required.

- "Traffic light" approach to backtesting is the only assessment of VaR accuracy prescribed in the current regulatory framework.

- "Traffic Light" approach by Basel Committee on Banking Supervision 1996: The three zones have been delineated and their boundaries chosen in order to balance the type I and type II error (Green zone, Yellow zone, Red zone).
- The results of these backtests are used by supervisors to assess the risk models, and to determine the multiplier (or hysteria) factor to be applied: if the number of exceptions during the previous 250 days is less than 5, then the multiplier is 3; if the number of exceptions is 5, the multiplier is 3.40, and so forth; and 10 or more exceptions warrant a multiplier of 4.
- Backtesting must be performed daily.
- Banks must identify the number of days when trading losses, if any exceed the VaR.

4.2 Exceedance-Based Methods

The idea of these tests is to focus on numbers and frequencies of exceedance observations. Let define them as $x_{t,t+1}$ realized during the certain time interval. They are also called the tail losses and stay for the violations of forecast VaRs by realized returns. We obtain them when we define a hit sequence I_{t+1} as:

$$I_{t+1}(\alpha) = \begin{cases} 1 & \text{if } x_{t,t+1} \leq -VaR_t(\alpha), \\ 0 & \text{if } x_{t,t+1} > -VaR_t(\alpha). \end{cases} \tag{4.1}$$

Now we are able to check whether our hit sequence I_{t+1} satisfies two properties of exceedance-based backtests: unconditional coverage property (number of violations follows a binomial distribution) and independence property (violations do not come in clusters). In principle, a particular VaR model could result in a hit sequence that satisfies the unconditional coverage property but not the independence property and in reverse. Only hit sequences that satisfy both properties can be described as evidence of an accurate VaR model. Each property characterizes a different dimension of an accurate VaR model (Campbell 2005).

4.2.1 Unconditional Coverage

Testing unconditional coverage also called binomial backtesting method proposed by Kupiec (1995) is focused exclusively on this one property. It tests whether the observed frequency of tail losses that exceed VaR is consistent with frequency predicted by the model. Under the null hypothesis (model is accurate) the frequency of violations follow binomial distribution with probability p, where p is tail probability, also $p = (1 - \alpha)$. Given x violations and n observations we calculate the probability of this case:

$$\Pr(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x} \tag{4.2}$$

Our model predicts np violations under $H_0 : p = (1 - \alpha)$ and alternative hypothesis is specified as $H_1 : p > (1 - \alpha)$. We will accept the null if test p-value will be larger than certain significance level. The most suitable test for comparing a theoretical and realized value is the likelihood ratio test LR_{POF}. This test computes a test statistic for each number of realized exceptions. Also related approach is the time-to-first-tail-loss test (or TUFF test) under which the probability of a tail loss is p, the probability of observing the first tail loss in period T is $p(1-p)^{T-1}$, and the probability of observing the first tail loss by period T is $1 - (1-p)^T$ which obeys a geometric distribution. This test is inferior to binomial backtest and can be used as supplementary to other tests.

If we have sufficiently large number of observations n we can approximate our binomial distribution of x with approximately normal with mean np and variance $np(1-p)$. It implies that variable $z = (x - np)/\sqrt{np(1-p)}$ is distributed as standard normal. Here we can simply test whether variable "z" has $N \sim (0, 1)$ distribution. The "z" statistic is actually the Wald variant of the likelihood ratio statistic proposed by Kupiec (1995). One potential advantage of the Wald test over the likelihood ratio test is that it is well-defined in the case that no VaR violations occur (Campbell 2005).

This test is easy to understand and requires knowledge of easily obtainable variables, however, it has its shortcomings:

- Low ability to identify incorrect models when having small sample sizes.
- Ignores the independence property of violations.
- Discards the volume of singe violations.

4.2.2 Independence

Besides the binomial distribution of frequency of violations we should test also for the independence. Independence property restrict any temporal pattern in the series, i.e., the probability of the next observation being an exceedance should be independent of whether any previous observation was an exceedance or not (Dowd 2008). Simple test for independence is a *runs test* and more sophisticated tests are suggested by Engle & Manganelli (2004)[1] or Christoffersen & Pelletier (2004)[2]. Other very popular test will be introduced below and its author is Christoffersen (1998).

Independence tests provide very important informations about the predicted risk measures but they have one principal disadvantage. They all have the assertion that any accurate VaR measure will result in a series of independent hits. Accordingly, any test of the independence property must fully describe the way in which violations of the independence property may arise[3]. Intuitively, the violations of the independence property which are not related to the set of anomalies defined by the test will not be systematically detected by the test (Campbell 2005).

4.2.3 Conditional Coverage

For testing the property of conditional coverage we will use likelihood ratio tests as suggests the author of this backtesting method Christoffersen (1998). To form a complete test he uses an LR test of correct unconditional coverage, an LR test of independence, and an LR test that combines the two.

Given x/n observed violations and predicted probability of violations p, under the null of correct unconditional coverage, the test statistic

$$LR_{UC} = -2\ln\big[(1-p)^{n-x}p^x\big] + 2\ln\big[(1-\frac{x}{n})^{n-x}(\frac{x}{n})^x\big] \tag{4.3}$$

has distribution $\chi^2(1)$.

Moving further, Christoffersen (1998) tests independence against an explicit first-order Markov chain alternative. He considers a binary first-order Markov

[1] They regress the hit sequence I_{t+1} against possible explanatory variables and subsequently test for joint insignificance of the these variables.

[2] Idea of this test is that VaR violations should not exhibit any kind of, so called, duration dependence. It means that the amount of time that elapses between violations should be independent of the amount of time that has elapsed since the last violation

[3] The alternative hypothesis must be completely specified

chain, $\{I_t\}$, with transition probability matrix

$$\mathbf{\Pi} = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix}$$

where $\pi_{ij} = \Pr(I_t = j \mid I_{t-1} = i)$ and where n_{ij} is the number of observations with value i followed by j[4].

Then he considers the output sequence, $\{I_t\}$, from an interval model, estimates a first-order Markov chain model on the sequence and tests the hypothesis that the sequence is independent, having that Π_2[5] corresponds to independence[6].

Under the null the LR test of independence

$$\begin{aligned} LR_{ind} = &-2\ln\left[(1-\hat{\pi_2})^{n_{00}+n_{10}}\hat{\pi_2}^{n_{01}+n_{11}}\right] \\ &+ 2\ln\left[(1-\hat{\pi_{01}})^{n_{00}}\hat{\pi_{01}}^{n_{10}}(1-\hat{\pi_{11}})^{n_{10}}\hat{\pi_{11}}^{n_{11}}\right] \end{aligned} \tag{4.4}$$

will have $\chi^2(1)$ distribution where n_{ij} is the number of observation with state i followed by j.

Now we combine the two hypothesis of correct unconditional coverage (4.3) and independence (4.4) to one, the hypothesis of unconditional coverage. The test statistic

$$LR_{cc} = LR_{uc} + LR_{ind} \tag{4.5}$$

is distributed as $\chi^2(2)$. If the tested model is accurate, then the violations should be Bernoulli variables[7].

Advantages of this method are that it is easy to implement and understand and it can identify the source of failure, while at the same time it enables to test both coverage and independence properties. Drawbacks of this method is that

- While joint tests have the property that they will eventually detect a VaR measure which violates either of these properties, this comes at the

[4]The approximate likelihood function for this process is $L(\Pi_1; I_1, ...I_T) = (1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}$, where parameters are $\hat{\mathbf{\Pi}}_\mathbf{1} = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix}$.

[5]$\mathbf{\Pi_2} = \begin{bmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{bmatrix}$

[6]The likelihood under the null becomes $L(\Pi_2; I_1, ...I_T) = (1-\pi_2)^{(n_{00}+n_{10})}\pi_2^{(n_{01}+n_{11})}$ and the Maximum Likelihood estimate is $\hat{\Pi_2} = \hat{\pi_2} = (n_{01}+n_{11})/(n_{00}+n_{10}+n_{01}+n_{11})$ Christoffersen (1998).

[7]Bernoulli variables are components of Bernoulli process and are identical and independent

expense of a decreased ability to detect a VaR measure which only violates one of the two properties (Campbell 2005).

- The test does not include the information about the accuracy of model on other confidence levels than determined by VaR measure.
- It also discards the information of magnitude of potential losses.

4.3 Density Forecast Backtesting Methods

Al the tests mentioned so far focus only on one confidence level, but the two properties of unconditional coverage and independence should hold over all confidence levels at the same time. Main authors that contributed to these methods of testing are Crnkovic & Drachman (1997) and Diebold *et al.* (1998a) and later extension by Berkowitz (2001).

Diebold *et al.* (1998a) furthermore adopt less formal, but more revealing, graphical methods. They suggest visual assessment using obvious graphical tools, a density estimate. Simple histograms - regards unconditional uniformity - and simple correlograms - regard evaluation whether a random variable is iid.

Testing Uniformity

Both Diebold *et al.* (1998a) and Crnkovic & Drachman (1997) suggest that if two properties should hold for any level of α that it can be formalized in following manner. Realized values of variables whose density is being forecast should be mapped to their probability integral transformation[8] (PIT) or forecast cumulative density values (the Rosenblatt transformation[9]). Rosenblatt (1952) defines the transformation

$$\hat{U}_t = \int_{-\infty}^{y_t} \hat{f}(u)du = \hat{F}_t(y_t) \tag{4.6}$$

[8] Methods are based on the relationship between the data generating process, $f_t(y_t)$, and the sequence of density forecasts, $p_t(y_t)$, as related through the probability integral transform, z_t, of the realization of the process taken with respect to the density forecast. The probability integral transform is simply the cumulative density function corresponding to the density $p_t(y_t)$ evaluated at y_t (Diebold *et al.* 1998a).

[9] Rosenblatt (1952) described a transformation mapping a k-variate random vector with a continuous distribution to one with a uniform distribution on the k-dimensional hypercube (Brockwell 2007)

where y_t is ex post realized profit or loss and $\hat{f}(.)$ is the ex ante forecast loss density. He also showed that if $\hat{F}_t(.)$ is correctly defined, then $\hat{U}$ is iid and distributed uniformly on (0,1)

$$\hat{U}_t \overset{i.i.d.}{\sim} U(0,1). \tag{4.7}$$

Now the task of validating the Var model accuracy boils down to task of testing for uniformity and independence. Here we can apply conventional tests as Kolmogorov-Smirnov (KS) test for uniformity, where test statistic D[10],

$$D = \max \left| F(y_t) - \hat{F}(y_t) \right|,$$

is compared to the relevant critical value and the null is accepted or rejected accordingly. The KS statistic is easy to calculate but has its drawback that it tends to be much more sensitive around the median value of the distribution and less sensitive around the extremes what is crucial for our case. Alternative test is Kuiper test with test statistic D^*[11],

$$D^* = \max \left| F(y_t) - \hat{F}(y_t) \right| + \left| \hat{F}(y_t) - F(y_t) \right|,$$

that is much more sensitive to deviations on the tails. However, Crnkovic & Drachman (1997) also report that the Kupier test statistic is very data-intensive. Another option is to use chi-squared goodness-of-fit test, χ^2, suggested by Diebold *et al.* (1998b).

Berkowitz Transformation and Testing Standard Normality

While testing the iid uniform distribution hypothesis can be technically demanding, Berkowitz (2001) suggests to transform the classified observations, $\hat{U}_t$, to standard normal, $\hat{Z}_t$. We obtain them by applying an inverse cumulative distribution function, Φ^{-1}

$$\hat{Z}_t = \Phi^{-1}(\hat{U}_t) = \Phi^{-1}\left(\int_{-\infty}^{y_t} \hat{f}_t(u)du \right) = \Phi^{-1}\big(\hat{F}_t(y_t)\big).$$

[10] KS test statistic represents the maximum distance between the predicted cumulative density $F(y)$ and the empirical cumulative density $\hat{F}(y)$

[11] Kuiper test statistic is the sum of the maximum amount by which each distribution exceeds the other.

Now we can apply more powerful statistical tool as data follow standard normal distribution under the null. Berkowitz (2001) suggest to nest the null hypothesis within the first-order autoregressive process with a possible different mean and variance and construct the LR test statistic that under null is distributed as $\chi^2(3)$. The drawback of this method is that it focuses only on the first two moments of the distribution and has very little detecting power in the higher moments.

Another way is to divide the test to two subgroups of tests. One is to test whether $\hat{Z}_t$ is $N(0,1)$ assuming it is iid. Here we can use several tests. If $\hat{Z}_t$ is standard normal, then it should have zero mean, $\mu = 0$, variance of 1, $\sigma^2 = 1$, zero skew and kurtosis of 3. If we assume $\hat{Z}_t$ is iid, we can apply z-test or t-test to test $\mu = 0$, variance ratio test for variance predictions and Jarque-Bera test for skewness and kurtosis predictions. JB test can be regarded as test of normality itself and has very good power properties.

Second group tests whether $\hat{Z}_t$ is iid. Common one are runs tests, binary regression tests or we can estimate the autocorrelation structure of $\hat{Z}_t$ observations or fit an ARMA process to them. All the parameters in an autocorrelation function o r an ARMA process should be insignificant, and we can test for their insignificance using standard tests such as Box-Pierse Q test (Dowd 2008). Or having enough data, we can also use BDS (Brock, Dechert and Scheinkman) test.

4.4 Probability Forecast Backtesting Methods

This method is alternative to the previous two, and it does not examine the behavior of the hit sequence, (4.1), neither test any hypothesis. This method is able to provide the information about magnitude of exceedance, to give every model a score in term of some function and to label the model that most closely approximate the true data. Lopez (1999) suggests the backtesting method based on binomial loss function, defined in 4.8. It is called second, size-adjusted, loss function and measures how well certain model predicts losses when they occur and describes observed exceedances of VaR.

$$L(VaR_t(\alpha), x_{t,t+1}) = \begin{cases} 1 + \big(x_t - VaR_t(\alpha)\big)^2 & \text{if } x_t \leq -VaR_t(\alpha) \\ 0 & \text{if } x_t > -VaR_t(\alpha) \end{cases} \tag{4.8}$$

, where x_t is the actual realized return and $VaR_t(\alpha)$ is the corresponding calculated VaR of in-sample subset, or forecast VaR for time period t based on the information set available at time $t-1$ for out-of-sample subset. Here the task is to determine what would be expected when the reported VaR accurately reflects underlying risk. Lopez (1999) therefore defines the sample average loss,

$$\hat{L} = \frac{1}{T}\sum_{t=1}^{T} L\big(VaR(\alpha), x_{t,t+1}\big), \tag{4.9}$$

. But the degree to which the observed average loss is still consistent with an accurate VaR model has to be still assessed. Here comes the informational burden associated with determining whether the average loss, $\hat{L}$, is "too large relative to what would be expected" and so it is necessary to understand the stochastic behavior of the loss function, $L(VaR_t(\alpha), x_t)$ (Campbell 2005). What is more, the mentioned loss function does not penalize possible overestimation of the VaR.

In our thesis we will use loss functions that are more commonly used in practice for evaluation of both in-sample and out-of-sample forecasts. As mentioned by Lopez (2001), it is not outright which loss function is more appropriate than other. So we will use six different accurate statistics or loss functions as decision criteria. As we are interested in accurate forecasting of VaR a high-level of volatility will be our cynosure. It implies Mean Square Error (MSE),Mean Absolute Deviation (MAD), Mean Log-absolute Error (MLAE), Heteroskedasticity-adjusted Mean Square Error (HMSE), R2LOG and QLIKE[12] are appropriate criterion for our thesis.

$$\text{MSE:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a}(\hat{\sigma}_t^2 - VaR_{t|t-1})^2 \tag{4.10}$$

$$\text{MAD:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a}|\hat{\sigma}_t^2 - VaR_{t|t-1}| \tag{4.11}$$

$$\text{MLAE:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a}\log|\hat{\sigma}_t^2 - VaR_{t|t-1}| \tag{4.12}$$

$$\text{HMSE:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a}\Big(1 - \frac{\hat{\sigma}_t^2}{VaR_{t|t-1}}\Big)^2 \tag{4.13}$$

[12]The metric QLIKE is the loss implied by a Gaussian likelihood, while the R2LOG loss function penalizes volatility forecasts asymmetrically in low volatility and high volatility periods.(Lorde & Moore 2008)

$$\text{R2LOG:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a} \log\left[\frac{\hat{\sigma}_t^2}{VaR_{t|t-1}}\right]^2 \tag{4.14}$$

$$\text{QLIKE:} \qquad L(\hat{\sigma}_t^2, VaR_{t|t-1}) = \frac{1}{a}\sum_{t=1}^{a} \left(\log h_t + \frac{\hat{\sigma}_t^2}{VaR_{t|t-1}}\right) \tag{4.15}$$

where a is the number of predicting data. All functions or statistics are based on comparing predicted Value at Risk $VaR_{t|t-1}$ with true realized variance σ_t^2, that is hardly observable. That is why conditionally unbiased volatility proxy $\hat{\sigma}_t^2$ is used instead. Patton (2011) shows that squared returns ${r_t}^2$ are good volatility proxy of realized variance under three different distributional assumption of r_t: Student's t (0, ${\sigma_t}^2$, ν); N (0, ${\sigma_t}^2$) ; F_t (0, ${\sigma_t}^2$), where F_t is unspecified distribution with referenced mean and variance. In all cases it is clear that $E_{t-1}(r_t^2) = \sigma_t^2$. Good review of most of them provides above mentioned Patton (2011), Wei *et al.* (2010) and Lopez (2001).

4.5 Use of Backtests for Empirical Analysis

Based on the characterization of individual backtesting methods we choose few, that will be directive for our conclusions. We will widely apply exceedance-based methods and as the most conclusive we consider conditional coverage. It will be examined also separately and we calculate statistics for its individual parts independence and conditional coverage. As supplementary we use also TUFF test. For more precise analysis we state after what period the first violation occurred and failure rate of particular model. Probability forecast backtesting method will be call on when we wont be able to decide on previous results about models performance.

Chapter 5

Empirical Research

In this chapter we are going to fit the above mentioned models and methods to calculate VaR on real datasets and evaluate theirs performance. VaR calculations are done by parametric method mentioned earlier in Chapter 2 both by GARCH and SV volatility estimates. Evaluation of their performance will be done by statistics values of several backtests, which are described in Chapter 4. Analysis will be executed separately on two subsets of data, in-sample and out-of-sample. Later one provides more precise assessment of performance of certain models, while it is calculated on other sample then subsequently evaluated.

The chapter is organized as follows. First we provide analysis of data that we are going to work with. Secondly, we use our two groups of models to make volatility prediction for both subsamples with its diagnostics. In the third section we calculate VaR and backtest the results. Last section is dedicated to overall performance of modes and theirs summarized results. We try to conclude whether our assumption that SV models with t-distribution can be a good alternative if not superior to GARCH-t models in predicting VaR measure holds.

5.1 Data Analysis

We are going to analyze dataset that contains of 5 different financial indices obtained from Yahoo Finance. Each series involves 1000 daily observations of closing prices that are divided into two subsets. Indices S&P, NASDAQ, DAX, CAC and FTSE are market-value wighted indices and measure change in market price whilst DAX considers also the dividends payment and reflects the effect of dividend reinvestment. This can bias our analysis of DAX index as prices can contain a discontinuous jumps not caused by actual price movements.

- S&P 500 (GSPC): 5/14/08 - 7/18/11 - 5/1/12
- NASDAQ Composite (IXIC): 5/14/08 - 7/18/11 - 5/1/12
- FTSE 100 (FTSE): 5/19/08 - 7/19/11 - 5/1/12
- CAC 40 (FCHI): 6/10/08 - 7/21/11 - 4/30/12
- DAX 30 (GDAXI): 6/9/08 - 7/20/11 - 4/30/12

These daily closing prices S_t, at time t, are in our purpose transformed to daily continuously compounded returns $r_t = ln(S_t/S_{t-1})$. For general intention of our thesis and to refrain of regional characteristics these worldwide known indices were chosen, two American and three European. As long as the trading dates are not internationally standardized the beginning date of all samples vary across indices. This does not constitute a problem for us while we do not combine indices into portfolio or we do not study their relative correlations.

The important part of choosing suitable datasets is to opt its appropriate length. Although, generally holds the longer period, the better results, we had to satisfy ourselves with only less then 3-year periods. The reason is the heftiness of empirical calculations of parameters when numerical examining SV by EIS.

As we already mentioned, for more precise evaluation of models we divide series into two sub datasets; in-sample (first 800 observations) and out-of-sample (remaining 200 observations). Former one will be used for parameters, likelihood and volatility estimations and then both will be used for evaluation of its performance. For out-of-sample subset parameters of certain model will be recalculated every 20 observations and volatility will be obtained as one-day-ahead forecast. This will enable us to make better decisions about models performances by reason that the situation is happening analogously as in the real-life context. Here we are able to test the predictive power of the model on 200 real observations backward.

In the Figure 5.1 we can see closing prices of all indices. Dashed vertical line at 800th observation divide our sample into in-sample subset and out-of-sample subset. Indices behave alike, as the beginning of our observable period is mid of year 2008 when all the prices have declining tendencies due to contemporary economic and financial crises. Only after 200 observation the market firstly reinforce and prices start to rise again. The whole sample is quite volatile for all indices by virtue of impact of mentioned crises which consequences are fading away very slowly. From the Figure we are not able to capture any stronger correlations between indices. Although, there is similar evolution of FTSE and DAX index for the first subset and for out-sample there seems to be some relations between FTSE and CAC, even CAC index has much lower prices (the whole index is moved down by constant).

Figure 5.1: Closing Prices

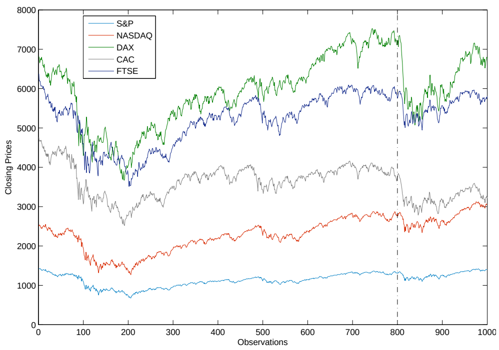
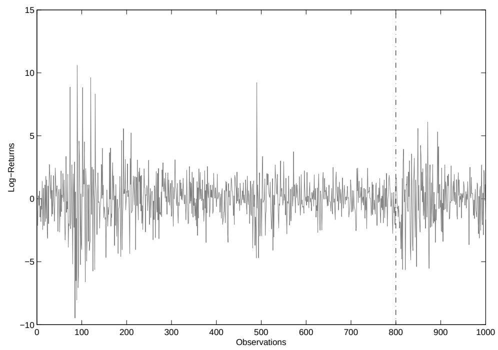


Figure 5.2: CAC Log-Returns



Source: author's computations in Matlab.

What we are going to work with and analyze are aforementioned log-returns of closing prices and their properties. Since the evolution of closing prices is non-stationary, log-returns usually transform the series to stationary. Figure 5.2 shows Log-Returns of CAC index for the whole period. We chose example of CAC in behalf of it seems to be average volatile from all five indices. On this plot we can compare volatility of both subsets. In-sample subset seems to reach higher volatilities and volatility clustering is nicely observable. Later subset is generally more volatile taking as a whole, even it does not reach such a high volatility. Volatility clustering is also obvious, though, in its weaker form.

Now and then, we are going to study sub-samples and their characteristics only separately. Daily Log-Returns are depicted in Figure B.1 in appendix. For all five indices in main period the volatility clustering is very clear while one period is concentrated around year 2009 and second, less volatile, in mid 2010. So, from the graphical analysis we can deduce the heteroskedasticity of this sub-sets. Further Figures B.2 and B.3, histograms of log-returns and theirs qq-plots against normal distribution, respectively, imply the non-normality of log-returns distribution. This is validated also by descriptive statistics of series in Table A.1. Kurtosis of each series is approximately three times higher than 3 what is kurtosis of normal distribution. Negative skewness is present in S&P, NASDAQ and FTSE indices around -0.2 and positive skewness in remaining two around 0.28. Aforementioned together with histograms, B.2, confirms the assumption about fat-tails of financial time series and leptokurtic distribution. The normality of series is tested also by Jarque-Bera test statistic and we can reject the null of normality at all significance levels for each series. Minimum and maximum for all indices are very similar and fall in the claims of highly volatile period. We investigated stationarity of our series by two different tests. Familiar Augmented Dickey-Fuller test (ADF) rejects the unit root for all series. As well, the less known but stricter Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) cannot reject null hypothesis of trend stationarity for all five series. From both tests the stationarity of in-sample subset is presumed.

Compared with in-sample, out-of-sample subset will have quite different characteristics. From the first sight, the main odds will be that series quite appear like a normal distribution. The one reasonable explanation is that the subset has only 200 observation what is not sufficiently long for stylized facts of financial volatilities to become evident and for us to make a precise judgments. Moreover, volatility clustering is seem-able very barely as depicted in Figure B.6 of log-returns. Looking at histograms and qq-plots of our indices, B.7 and B.8 respectively, we see that normal distribution fit is much closer to realized returns that was the case of in-sample subset. Table of descriptive statistics A.2 move towards our normality presumptions. It is confirmed for CAC and DAX index for which Jarque-Bera test null cannot be rejected at 5% significance level and for DAX neither at 10% level. Their kurtosis is very close to 3 and have very small negative skewness. S&P, NASDAQ, FTSE indices have kurtosis around 5, are also negatively skewed and the fat-tailness is present. For all indices ADF and KPSS tests affirm stationarity.

Based on the former survey we chose for our research the assumption of Student's t distribution of errors. This distribution is symmetric and fat-tailed. Parameter ν, degrees of freedom, will be estimated by both models for both subsamples and for all indices. These estimates will also affirm that some indices have very close to normal distribution. ν will be assigned by significantly high numbers, occasionally around 20, 30 and 40 degrees of freedom.

The analysis of dependence structure in log-returns nods to our previous conclusions. Autocorrelation Function (ACF) B.4 and B.9 of in-sample and out-of-sample subsets, respectively, does not indicate strong or considerable autocorrelation up to 30th lag, except the first lag. As financial time series are often characterized by high persistence in squared observations so it is in our graphs of squared returns ACF, B.5 and B.10, where there is only very slow decay.

5.2 Models Application

Hereafter we apply our two competing models to estimate parameters and volatilities of in-sample and out-of-sample datasets. To compare Stochastic Volatility with GARCH models we chose basic form of both models, meaning SVAR(1) and GARCH(1,1). The selection of GARCH(1,1) is supported, first, by our own calculations and comparison of Akaike information criteria for GARCH(p,q), where $p = \{1, 2\}$ and $q = \{1, 2\}$ as presented in Table A.3. Even GARCH(2,2) has for two indices lower AIC, we consider the parsimony rule and choose simpler model. Second, by the common practice and favorable results of many research papers and third, regarding the selection of ARSV(p). As a consequence of SV's difficult numerical calculations the only option to consider was the ARSV(1) model for our thesis.

As we conclude in the above section the in-sample sets of all five indices are fat-tailed and leptokurtic we will assume Student's t distribution of errors. For GARCH(1,1)-t we estimate parameters $\hat{a}_0$, $\hat{a}$, $\hat{b}$ and $\hat{\nu}$ by maximum likelihood estimator as described in Section 3.3. All calculations from now on are executed in software Matlab. The results of ML for GARCH-t model are presented in Table 5.1. Each parameter has significant estimates and degrees of freedom are approximately around value of 8.5 for every index. Immediately we are able to calculate GARCH volatility as $\sqrt{\hat{h}_t}$.

For SVAR(1) calculations we estimate parameters $\hat{\beta}$, $\hat{\delta}$, $\hat{\rho}$, $\hat{\nu}$ with efficient importance sampling Monte Carlo numerical method as explained in details in Section 3.2. ML-EIS estimations of parameters of SV-t are presented in Table 5.2 below. Also for this model all parameters are significant where the slightest significance inhere ρ estimates. However it is still sufficient at 95% confidence level for all indices. Degrees of freedom are higher than for GARCH model. Estimates move around 40 degrees of freedom what indicate more Gaussian shape of distribution.

Table 5.1: ML estimation results for GARCH(1,1)-t model

	$\hat{a_0}$	$\hat{a}$	$\hat{b}$	$\hat{\nu}$	LL
S&P	0.0183 (1.0824e-04)	0.1022 (4.9855e-04)	0.8956 (3.8595e-04)	6.7005 (3.5287)	-1361.21
NASDAQ	0.0228 (1.6447e-04)	0.0896 (4.2269e-04)	0.9042 (3.8833e-04)	8.7955 (8.5439)	-1426.69
DAX	0.0255 (2.2700e-04)	0.0727 (3.7611e-04)	0.9180 (4.1700e-04)	8.5403 (8.6266)	-1423.89
CAC	0.0483 (5.4949e-04)	0.0801 (4.9187e-04)	0.9034 (5.8381e-04)	9.3977 (9.4592)	-1488.89
FTSE	0.0296x (9.4592)	0.0825 (3.9920e-04)	0.9029 (4.6025e-04)	11.0852 (18.9197)	-1350.65

Standard errors are calculated in brackets.

Source: author's computations.

Table 5.2: ML-EIS estimation results for ARSV(1)-t model

	$\hat{\beta}$	$\hat{\delta}$	$\hat{\rho}$	$1/\hat{\nu}$
S&P	1.6732 (2.0509e-05)	0.9982 (4.2116e-04)	0.2350 (0.0250)	0.0279 (0.0958)
NASDAQ	1.6732 (2.2295e-05)	0.9982 (3.7560e-04)	0.2003 (0.0097)	0.0309 (0.2083)
DAX	1.6731 (2.3240e-05)	0.9977 (4.0517e-04)	0.1778 (0.0143)	0.0183 (0.0826)
CAC	1.6731 (1.8172e-05)	0.9985 (3.0210e-04)	0.1593 (0.0108)	O.0207 (0.0833)
FTSE	1.6731 (1.9528e-05)	0.9982 (2.4680e-04)	0.1794 (0.0078)	0.0245) (0.0973

MC numerical standard errors are in brackets.

Source: author's computations.

In this model we will be working only with filtered estimate[1] of $\hat{\lambda}_t$ and we get volatility as $\hat{\beta}\exp(E(\hat{\lambda}_t/2))$. When the parameters of models are estimated and particular volatilities predicted we can forecast the out-of-sample subset. For the first 20 observations we do one-day-ahead forecast for time (t) where we use parameters estimated on in-sample subset and returns observed at $(t-1)$. For next 20 observations we recalculate parameters and again do one-day-ahead forecast with one-day back observations. This procedure is repeated ten times

[1] Existence of smoothed estimates of lambda, for more details see Section 3.2.1.

until we have predicted 200th observation.

Estimated volatilities of particular indices for in-sample set are depicted in Figure B.11. We can see that volatilities calculated by SV are over the GARCH volatilities for all indices. The only time that SV volatility meets GARCH volatility predictions is when volatility is at its peaks, i.e. year 2008-2009 the period of global financial crises strong impact. But as the volatility is unobservable we can say nothing about the fits of these two models to the data. Therefore we take a quick look on analysis of normalized residuals z^* and their squares.

Diagnostics for SV-t model are summarized in Table 5.3 and plotted in Figure B.12 with its qq-plots against N(0,1) in B.13. If model is correctly specified distribution of z^* should be standard normal. We provide values of skewness and kurtosis, Kolmogorov-Smirnov z-statistics $KS(z^*)$ and Ljung-Box Q-statistics $Q_{20}(z^*)$, $Q_{20}(z^{2*})$ with its particular p-values. For all indices is residual kurtosis around 3 what is kurtosis of normal distribution and negative skewness just very slight. Q-statistic of residuals indicates that they exhibit no autocorrelations up to 20th lag, although squared residuals exhibit serial correlations and suggest a failure to account for some dynamics. On the other hand, standard normal distribution is tested by $KS(z^*)$ and for all series it cannot be rejected that residuals are standard normal at 1% significance level.

Table 5.3: Diagnostics for the SV(1)-t model normalized residuals

	Skewness	Kurtosis	$Q_{20}(z^*)$	$Q_{20}(z^{2*})$	$KS(z^*)$
S&P	-0.2751	3.1904	20.50 (0.4270)	195.46 (0.0000)	0.0520 (0.0253)
NASDAQ	-0.2444	3.2551	17.53 (0.6180)	255.88 (0.0000)	0.0513 (0.0285)
DAX	-0.0482	3.5684	20.67 (0.4163)	204.62 (0.0000)	0.0435 (0.0937)
CAC	-0.0003	3.5895	20.09 (0.4520)	151.82 (0.0000)	0.0289 (0.5056)
FTSE	-0.1119	3.2551	29.70 (0.0749)	314.97 (0.0000)	0.0351 (0.2709)

p-values are calculated in brackets.

Source: author's computations.

Diagnostics for GARCH model normalized residuals are in Table 5.4. Figures of plotted residuals and their qq-plots, B.14 and B.15 respectively, are presented in appendix. Correspondingly to SV diagnosis, values of kurtosis are very close to 3 and skewness differs weakly from zero. Although in the case of GARCH residuals is the skewness higher. Ljung-Box Q-statistic for both normalized residuals and its squared values indicates no residual autocorrelations in not even one case. On the other hand, $KS(z^*)$ test rejects standard normal distribution in case of S&P and NASDAQ index on all significant levels. For the rest three indices null cannot be rejected on 1% level.

Table 5.4: Diagnostics for the GARCH(1,1)-t model normalized residuals

	Skewness	Kurtosis	$Q_{20}(z^*)$	$Q_{20}(z^{2*})$	$KS(z^*)$
S&P	-0.4357	3.6922	19.03 (0.5200)	6.7005 (0.0707)	0.0762 (0.0001)
NASDAQ	-0.4186	3.7004	14.11 (0.8248)	8.7955 (0.2787)	0.0591 (0.0071)
DAX	-0.1217	3.5854	14.61 (0.7984)	8.5403 (0.1712)	0.0516 (0.0273)
CAC	0.0296	3.9474	13.70 (0.8456)	9.3977 (0.3342)	0.0389 (0.1738)
FTSE	-0.1177	3.6063	14.82 (0.7868)	11.0852 (0.2302)	0.0439 (0.0987)

p-values are calculated in brackets.

Source: author's computations.

Nevertheless, the point of our interest is to decide which model has better predicting performance of VaR measure and therefore we move on to VaR predictions and its subsequent backtesting.

5.3 VaR Calculations and Backtesting

This part is the aim of our empirical study. Here we use above calculated volatilities of both sub-samples to calculate one-day-ahead parametric VaR predictions for three confidence levels $\alpha = \{90\%, 95\%, 99\%\}$. This calculation method is illustrated in Chapter 2. For each index separately we compare the results from both models and based on backtesting procedure we will define better performing one for in-sample and out-of-sample subsets. We will consider out-of-sample results as more conclusive while as it was mentioned, testing data out-of-sample gives us more real-life conclusions. Summarization of these results is provided in next section.

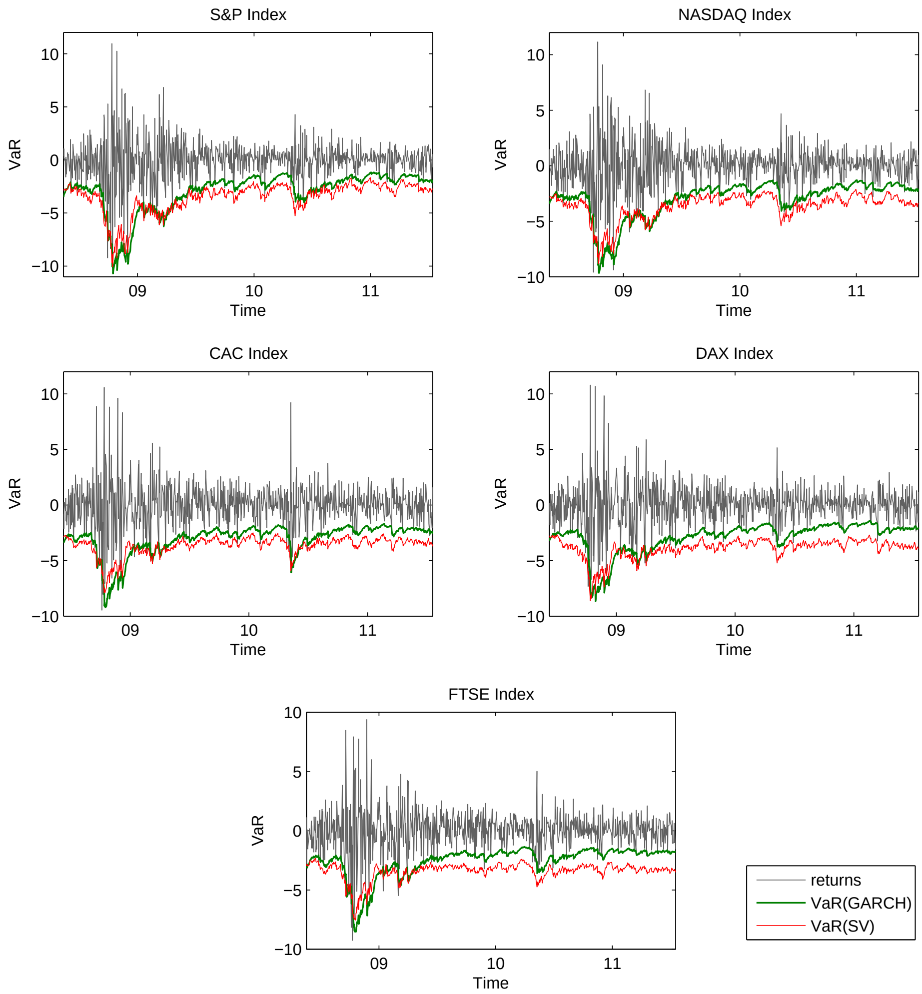
In-Sample

For graphical illustration we calculated one-day-ahead 95% VaR and plotted it in Figure 5.3. Result seems to be in accord of the volatility estimates plotted in Figure B.11 when SV was mainly above the GARCH volatility. As we consider only long position of VaR, the one calculated from SV model is in most under the VaR calculated by GARCH model. From the first sight it looks that VaR(SV) overstates the risk and so it is more safer. But for the financial institutions this can be very costly to keep correspondent reserves. Vice versa, in the period of crises VaR(SV) seems to less overvalue the negative returns and particular risk of losses than VaR(GARCH). To valuate it more accurately we have to use backtesting methods described in Chapter 4. Every index has it own table of particular backtesting results in appendix.

We analyze first S&P 500 index which results are presented in Table A.4 starting with 90% confidence interval. This one-day-ahead VaR(GARCH) and VaR(SV) are both in the Red zone of traffic light approach. According to Basel II framework, this would indicate the need of recalculations of both models and the use of particular multiplicators to reestimate them for more secure ones. Considering the percentage of failure the SV model has only 3.13% what can be little strict when calculating VaR at 90% confidence level. On the other hand, VaR by GARCH has 8.38% failure rate, presumably more accurate. But the accuracy of an ex-post loss exceeding probability of VaR is the subject of unconditional coverage. For the most crucial tests we therefore consider unconditional and conditional coverage and independence tests. Conditional coverage contains the likelihood ratio of unconditional coverage and independence so it evaluates model more generally. 90% VaR is fairly rejected by unconditional and conditional coverage for both methods of calculations while ex-post exceedances of VaR are not consistent with its α coverage rate. As a supplementary test we have used TUFF test that assumes the first failure to occur in $1/(1-\alpha)$ days. Same for GARCH and SV models TUFF test together with independence test accept the models.

For 95% VaR the 4.5% failure rate of VaR(GARCH) will be more accurate what is also accepted by unconditional coverage. But according to Basel II Backtesting Framework the model should be recalculated as long it is still in

Figure 5.3: 95% VaR Calculations: In-Sample



Source: author's computations in Matlab.

the Red zone. Comparing the results of all backtests, VaR(GARCH) appears to be superior for this index and confidence level. GARCH model is accepted by all four tests based on their p-values. VaR(SV) is again as for the 90% level rejected by unconditional and conditional coverage. It has failure rate of 0.05% and can be too strict and expensive, however it is recommended by Basel II traffic light approach where it is in the Green zone.

Last but not least, 99% VaR has reversed conclusion. SV model is accepted by all four test whereas GARCH model is accepted only by independence test. Its failure rate seems to be accurate, 0.38%, but is not promoted by unconditional test. In favor of this model is Basel II framework for which is the model in the Green zone. But closer look on VaR(SV) reveals that model has not even one violation and this can cause problems when evaluating it. In this case our backtests assume one violation instead of zero to avoid calculation errors and one violation is just approved by every test.

For index NASDAQ, backtests presented in Table A.5, and 90% VaR seems to be very accurate model calculated by stochastic volatility. It is accepted by all four tests, failure rate seems to be adequate and first violation comes after common 17 periods. However, it is again in the Red zone. But this framework takes into account the 90% confidence level what is not in accord of its manner in itself. VaR(GARCH) model is here rejected by conditional and unconditional coverage and is also in the Red zone.

95% VaR which was earlier displayed in Figure 5.3 is accepted for GARCH calculated model by all four tests and for model calculated by SV only by independence test. It is apparent from the plot how VaR(SV) overrates the risk what is for $\alpha = 95\%$ unnecessarily too much. But this come in favor of Basel II framework and the SV is in the Green zone, meaning expensive but safe. Big difference is also in the period when first exceedance occurs what is 17 periods for GARCH and 88 for SV. It is quite big contrast and VaR(SV) may seem to be more reliable model from the beginning. Though at large this tells us nothing about its general performance and so this is the task for TUFF test which rejects the model, too.

For most rigid 99% VaR is the overstating risk by SV model calculations suitable and so it is validated by all four tests p-value. However we are facing again the problem of no violations and cannot evaluate it reliably. VaR(GARCH) is for this index accepted only by TUFF test and independence test.

Next two indices, DAX presented in Table A.6 and FTSE presented in Table A.8, has very similar results and we analyze them together. Their 90% VaR of both SV and GARCH models is accepted only by TUFF test and independence test. First violation comes already after second period for FTSE index and tenth in DAX. All four fall into the Red zone. Percentage of failure of GARCH model is around 9% and seems to be accurate, however it is not affirmed by unconditional and conditional test for either index. Unconditional coverage as a part of conditional coverage rejects the model due to inequality between probability of ex-post loss exceedance and defined rate α.

Similarly to previous indices 95% VaR performs very well also for DAX and FTSE when volatility was estimated by GARCH model and is little overstated when predicted by SV. Latter VaR is accepted by TUFF , independence test, has only 0.38% and 0.75% percentage of failure, respectively, and is in the Green zone. VaR(GARCH) has accurate 6% of failure for both indices.

As desired stochastic volatility VaR is very suitable for 99% confidence predictions and adhere the high demand of safeness. It is accepted by all four tests in case of DAX when again no violations were occurred. For FTSE it is, on the other hand, accepted for the first time only by independence test. Here the first violation occurred after 99 periods and failure rate is 0.13% but was not supported by backtests. GARCH model is at this confidence VaR performing also very poorly. It is accepted by independence test for both indices and by conditional coverage for FTSE what is more desirable. All four models are in the Green zone.

For the last fifth index CAC, results presented in table A.7, common conclusion holds. SV models is performing well for 90% VaR and GARCH model for 95% VaR. Complements to them are accepted only in two cases of independence and TUFF test. According to Basel II only one model is in the Green zone, the 95% VaR(SV). For the most strict risk measure are models performing same and both are rejected by unconditional and conditional coverage tests. Very similar is also the percentage of failure around value of 0.2 and first violation after approximately 90 periods.

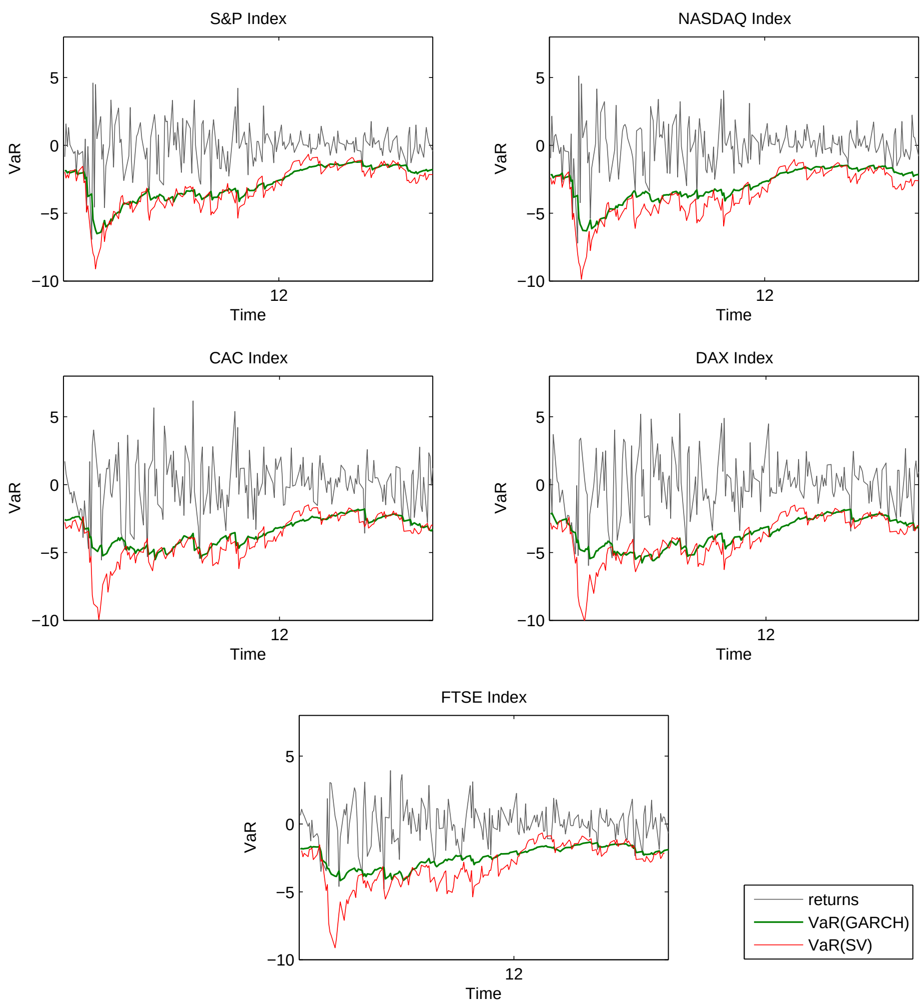
For all indices the results were constitutive. Models that were calculated based on volatility estimated by GARCH perform well for 95% confidence interval of one-day ahead Value at Risk. Contrary, for more strict confidence interval of 99% Value at Risk, models calculated from volatility estimated by SV seem to be very accurate, even it does not have to hold in cases with no violations. For 90% VaR, SV model comes out as superior for two of five indices.

And still there is another methodology of backtesting as it is described in Chapter 4.4, not based on testing hypothesis. Loss functions are the mean to rank the model, while they take into account the size of the losses. We employed six different loss functions. Based on them it can help us conclude which methods is more suitable. MSE, MAD and MLAE for all five indices acquire values in favor of GARCH model, although HMSE states for SV model. This diversification is caused by different evaluation of VaR violations.

Out-of-Sample

Out-of-sample subset analysis was carried out as comparison of one-day ahead forecasts with involved observed returns. Plots of return observations against both calculated 95% VaR models are in the Figure 5.4. It is clear that SV copy the financial returns movements up and down more precisely than GARCH which is flatter. But also it is obvious how again SV overstates the risk more than GARCH model, what mainly holds for FTSE and NASDAQ index,.

Figure 5.4: 95% VaR Calculations: Out-of-Sample



Source: author's computations in Matlab.

Looking at outcomes from backtests of out-of-sample data, we will get more realistic view on performance of two models than from in-sample. We are actually comparing the predictive ability of models with reality. All results are summed again in five tables consisting of VaR with $\alpha = \{90\%, 95\%, 99\%\}$ confidence level and its loss functions according to particular index.

S&P index is summarized in Table A.9. For the first 90% VaR we have very good results for both models. All four tests accept both models and the only difference is the one regarding the Basel II framework. VaR(GARCH) is in the Red zone whilst VaR(GARCH) is in the Yellow zone. We consider Yellow zone as very appropriate middle way as it is neither very expensive either not so risky. Percentage of failure for the former model is 7.54% and the other model only 2.51%.

However, for the left two VaRs of 95% and 99% are results contradictory. All tests accept GARCH model for convenient 95% confidence interval and the model is also in the Yellow zone. SV model is in the Green zone and can be too strict for given level. It is also rejected by unconditional and conditional coverage. As VaR(SV) was always little overstate the risk, result for 99% VaR are very pleasant. Model is accepted by all four tests and there is zero failure rate. This means no violations but very high reserves. Model by GARCH is at this level rejected by coverage tests.

For index NASDAQ the results come out bit better for SV computed models as presented in Table A.10. Thorough for 90% VaR is this SV model rejected by unconditional coverage, for the rest levels of VaR it is accepted by all four tests. The problem is again that they exhibits no exceedances and this results do not have to be reliable. On the other hand, GARCH model is still very good for 90% and 95% VaR and actually it is also in the Yellow zone for the latter one. 99% VaR(GARCH) is rejected commonly by unconditional and conditional coverage.

Table A.11 displays DAX evaluations and it is clear from the first sight that GARCH models do not perform good as before. For the less strict VaR the percentage of failure is 11.6% what can be over the feasible level. It is also rejected by independence test and conditional coverage test. For the same confidence VaR(SV) performs similarly. It is rejected by same tests but at least failure rate is 3% less than GARCH. For 95% VaR GARCH performs well as usually however is in the Red zone. Otherwise, SV is rejected by unconditional test and belongs to Yellow zone. Comparing the last 99% VaRs, SV model seems to be perfectly applicable but again there is a problem that risk is overstated and it embodies no exceedances. We can just hardly conclude about the accuracy of the model. GARCH for this case has also very little failure rate of only 0.5%

but it is appropriate for the strict risk management tool. Worst thing is that it is rejected by three of four tests, except independence.

The result of CAC index are in Table A.12. SV calculated model for the index performs very well for 90% and 95% level. For both it is accepted by all tests and additionally, the latter one is in the Yellow zone. The case of the strictest VaR is rejected by unconditional and conditional coverage for both models. For VaRs calculated via GARCH, the best performing one is on 95% confidence level. Feasible is also 90% VaR where it is rejected only by unconditional coverage. The worst performing from GARCH models is for the index the 99% VaR, rejected by all tests except independence.

For FTSE is better accuracy of SV more distinctive. Last index has its results in Table A.13 and for 95% VaR, SV is superior while it is accepted by all four tests and fall into Yellow zone. Contra to this, GARCH model is rejected by independence test but performs also adequately. For 90% VaR, SV model outperforms GARCH while it is accepted expect TUFF and independence test in addition by conditional coverage. 99% VaR has the worst results for both models and is rejected by conditional and unconditional coverage and for VaR(SV) model also by TUFF test.

5.4 Results Summary

Looking at overall VaR predicting performance of both models we can make some interesting conclusions. For in-sample subset the VaR calculated trough GARCH model is accepted by all four tests in case of 95% confidence level and there is one case when it is accepted only by two backtests, but one of them is conclusive conditional coverage. Considering Stochastic Volatility calculations of VaR, by all four tests there was accepted three times 99% VaR and two times 90% VaR. It is obvious that as it was noticed trough analysis, VaR(SV) overstates risk more than GARCH models, however it is also more flexible and copy the evolution of returns in time more resembling way compared to GARCH calculated risk measure that is much more flatter curve. This is also confirmed by six loss statistics that measure the difference between observed returns and calculated VaRs.

For out-of-sample subset of our 200 observations were GARCH VaR models accepted by all four tests four times on 95% confidence level and two times on 90% level. On the other hand, SV VaR models were accepted by all backtests once for 90% VaR, three times for 95% VaR and four times for 99% VaR. By three tests out of four were GARCH models accepted once at 90% and once at 95% confidence interval. For SV models it was two times for 90% VaR and once for 95%. Both models were accepted by two tests on 90% level out of one it was conditional coverage test, that we consider as the most important one. Here it seems that VaR(SV) performs a lot better at the strictest level of confidence. Nevertheless, we have to remember that not even once there was any exceedance and the calculations of statistics are zeros then. And also our calculations show acceptance by all four tests and do not have to be definitely right. Except this, are both models very comparable while GARCH performs little more accurate for 95% VaR and SV model for 90% VaR. Here the overstating of risk is not so significant as the forecast is calculated from the lagged value of real observation for each day ahead.

According to Basel II regulatory framework, were models 90%, 95% and 99% VaR in Red, Green, Green zones for all SV in-sample models and Red, Red, Green zones for all GARCH in-sample models, respectively. As it was already mentioned Green zone is although very safe, but very expensive for financial institutions, indeed. For out-of-sample subset were for this three different confidence levels VaR models also in Yellow zone. Stochastic volatility models were two times in Yellow, Green, Green and three times in Red, Yellow, Green zones and GARCH models two times in Red, Yellow, Green and three times in Red, Red, Green zones. Again SV(VaR) comes out as preferable in compliance with Basel II which first place merit is safeness.

We can now confirm our assumption that SV models perform at least as good as GARCH models in predictive performance of VaR measures. In our analysis they were superior to GARCH models as well as they were inferior. It has depended on what confidence level was calculated for what index. DAX index is quoted as total return and is affected by divided payouts and results might be distorted. For 99% VaR we could not conclude many results, because the overstating of risk caused zero exceedances. In this case our backtests replace zero with one violations which passes all tests on highly significant levels.

The problem that stays unsolved remains the comparison of SV and GARCH models performance when taking into account the calculation heftiness of both models. This is space for further research when some decision rule should be set to compare the two models in all of their aspects. One of the papers trying to do so is work of Hafner & Preminger (2010).

Chapter 6

Conclusion

This thesis compares two competing volatility models that are able to predict time-varying volatility considering its clustering. Subsequently, it evaluates them in accordance with theirs predicting performance of risk measure Value at Risk. For our analysis we chose widely used and known GARCH model with Student's t distribution contra to less popular Autoregressive Stochastic Volatility model, also t distributed. We did not analyze models with Normal distribution in behalf of two reasons. First, many papers, that are mentioned in our thesis and serve as the base for our research, compare just right models with Gaussian distribution. Second, distribution of financial returns is generally no longer assumed to be normal, as it ordinarily inherent fat tails and leptokurtosis.

The aim of this thesis was to show that SV models can in aforementioned purpose perform at least as good as GARCH models and perhaps to be superior. The main challenge was in choosing of the suitable process for SV's likelihood calculations where analytical methods are insufficient. In our thesis we propose method of Efficient Importance Sampling Monte Carlo Simulations based on paper of Richard & Zhang (2007). This provides us estimates of high-dimensional integrals necessary for SV parameters estimates.

Paper is divided into two parts, theoretical and empirical. In the first part we offer background of both volatility models, its calculation methods and used distribution. Likewise we introduce Value at Risk measure and its subsequent backtesting methods. In the empirical part we run our analysis on the five world-wide stock indices splitted into two sub-analysis. One of in-sample data set and other for more precise and real-life results on out-of-sample subset.

Our empirical analysis confirms our assumptions about SV models and their performance. From the backtests they come out as comparable with GARCH models and for stricter confidence level of 99% VaR even superior. SV VaRs curves are unambiguously more capable to imitate the shape of indices returns than GARCH models. They also incorporate the leverage effect at least at the univariate level. On the other hand, they have tendency to overstate risk what can be uselessly expensive for financial institutions.

However, considering calculation heftiness of both models, new questions about adequacy of both models arise. Does the estimation process of SV model worths its results? It could be reevaluated with application of some decision rule based on how much better the results should be to redeem SV models. Then, there are many extensions of both models. EGARCH, TGARCH, APARCH, etc. are capable to incorporate volatility properties as asymmetry, leverage effect, long memory, respectively, and are still computationally simply compared to SV. But SV has also its own extensions with which its accuracy rises apart from multivariate extensions of both models. This could be proposes to further research in this topics.

Bibliography

Alexander, C. (2009): *Market Risk Analysis, Value at Risk Models*, volume 4. Wiley.

Andersen, T., T. Bollerslev, P. Christoffersen, & F. Diebold (2007): "Practical volatility and correlation modeling for financial market risk management."

Andersen, T., T. Bollerslev, F. Diebold, & P. Labys (2003): "Modeling and forecasting realized volatility." *Econometrica* **71(2)**: pp. 579–625.

Angelidis, T., A. Benos, & S. Degiannakis (2004): "The use of garch models in var estimation." *Statistical Methodology* **1(1)**: pp. 105–128.

Artzner, P., F. Delbaen, J. Eber, & D. Heath (1999): "Coherent measures of risk." *Mathematical finance* **9(3)**: pp. 203–228.

Bai, X., J. Russell, & G. Tiao (2003): "Kurtosis of garch and stochastic volatility models with non-normal innovations." *Journal of econometrics* **114(2)**: pp. 349–360.

Baillie, R., T. Bollerslev, & H. Mikkelsen (1996): "Fractionally integrated generalized autoregressive conditional heteroskedasticity." *Journal of econometrics* **74(1)**: pp. 3–30.

Berkowitz, J. (2001): "Testing density forcasts with applications to risk management." *Journal of Business and Economic Statistics* **19(4)**: pp. 465–474.

Black, F. (1976): "Studies in stock price volatility changes." *Proceedings of the 1976 meetings of the business and economics statistics section, American Statistical Association* pp. 177–181.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of econometrics* **31(3)**: pp. 307–327.

Brockwell, A. E. (2007): "Universal residuals: a multivariate transformation." *Stat Probab* **77(14)**: pp. 1473–1478.

Broto, C. & E. Ruiz (2004): "Estimation methods for stochastic volatility models: a survey." *Journal of Economic Surveys* **18(5)**: pp. 613–649.

Campbell, S. D. (2005): "A review of backtesting and backtesting procedures." *Finance and Economics Discussion Series* .

Carnero, M., D. Peña, & E. Ruiz (2001): "Is stochastic volatility more flexible than garch?" *Nº.: UC3M Working Papers. Statistics and Econometrics 2001-05* .

Carnero, M., D. Peña, & E. Ruiz (2004): "Persistence and kurtosis in garch and stochastic volatility models." *Journal of Financial Econometrics* **2(2)**: pp. 319–342.

Christoffersen, P. F. (1998): "Evaluating interval forecasts." *International Economic Review* **39(4)**: pp. 841–862.

Christoffersen, P. F. & D. Pelletier (2004): "Backtesting value-at-risk: A duration-based approach." *Journal of Financial Econometrics* **2(1)**: pp. 84–108.

Clark, P. (1973): "A subordinated stochastic process model with finite variance for speculative prices." *Econometrica: journal of the Econometric Society* pp. 135–155.

Crnkovic, C. & J. Drachman (1997): ""quality control" in var: Understanding and applying value-at-risk." *Risk Publications* .

Damodaran, A. (2007): *Strategic risk taking: a framework for risk management.* Wharton School Publishing.

Danielsson, J. & J. Richard (1993): "Accelerated gaussian importance sampler with application to dynamic latent variable models." *Journal of Applied Econometrics* **8(S1)**: pp. S153–S173.

DIEBOLD, F. X., T. GUNTHER, & A. TAY (1998a): "Evaluating density forecasts with applications to financial risk management." *International Economic Review* **39(4)**: pp. 863–883.

DIEBOLD, F. X., A. TAY, & K. F. WALLIS (1998b): "Evaluating density forecasts of inflation: The survey of professional forecasters." *Statistics* **No. 6228**: pp. 76–90.

DING, Z., C. GRANGER, & R. ENGLE (1993): "A long memory property of stock market returns and a new model." *Journal of empirical finance* **1(1)**: pp. 83–106.

DOWD, K. (2002): *Measuring Market Risk*. John Wiley & Sons Ltd.

DOWD, K. (2005): *Measuring market risk*, volume 2. John Wiley & Sons Hoboken^ eNJ NJ.

DOWD, K. (2008): *Backtesting Market Risk Models*, volume 3. Wiley: Hoboken, NJ.

DUFFIE, D. & K. SINGLETON (1990): "Simulated moments estimation of markov models of asset prices." *Technical report*, National Bureau of Economic Research.

ENGLE, R. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica: Journal of the Econometric Society* pp. 987–1007.

ENGLE, R. F. & S. MANGANELLI (2004): "Caviar: Conditional autoregressive value-at-risk by regression quantiles." *Journal of Business and Economic Statistics* **22(4)**: pp. 367–381.

FABOZZI, F., S. FOCARDI, S. RACHEV, T. JAŠIĆ, & S. MITTNIK (2008): "Financial econometrics: From basics to advanced modelling techniques." *EKONOMSKI PREGLED* **59(1-2)**: pp. 94–98.

GALLANT, A., D. HSIEH, & G. TAUCHEN (1997): "Estimation of stochastic volatility models with diagnostics." *Journal of Econometrics* **81(1)**: pp. 159–192.

GERLACH, R. & F. TUYL (2006): "Mcmc methods for comparing stochastic volatility and garch models." *International Journal of Forecasting* **22(1)**: pp. 91–107.

GHYSELS, E., A. HARVEY, E. RENAULT, & U. DE MONTRÉAL. CENTRE DE RECHERCHE ET DÉVELOPPEMENT EN ÉCONOMIQUE (1996): *Stochastic volatility.* Université de Montréal, Centre de recherche et développement en économique.

GOURIEROUX, C., A. MONFORT, & E. RENAULT (1993): "Indirect inference." *Journal of applied econometrics* **8(S1)**: pp. S85–S118.

HAFNER, C. & A. PREMINGER (2010): "Deciding between garch and stochastic volatility via strong decision rules." *Journal of Statistical Planning and Inference* **140(3)**: pp. 791–805.

HARVEY, A., E. RUIZ, & N. SHEPHARD (1994): "Multivariate stochastic variance models." *The Review of Economic Studies* **61(2)**: pp. 247–264.

HULL, J. & A. WHITE (1987): "The pricing of options on assets with stochastic volatilities." *Journal of finance* pp. 281–300.

JACQUIER, E., N. POLSON, & P. ROSSI (1994): "Bayesian analysis of stochastic volatility models (with discussion)." *Journal of business and Economic Statistics* **12**: pp. 371–389.

JIANG, G. & P. VAN DER SLUIS (1999): "Index option pricing models with stochastic volatility and stochastic interest rates." *European Finance Review* **3(3)**: pp. 273–310.

JORION, P. (2006): "Value at risk: the new benchmark for managing financial risk." *New York* .

KIM, S., N. SHEPHARD, & S. CHIB (1998): "Stochastic volatility: likelihood inference and comparison with arch models." *The Review of Economic Studies* **65(3)**: pp. 361–393.

KUPIEC, P. (1995): "Techniques for verifying the accuracy of risk management models." *Journal of Derivatives* **4(3)**: pp. 73–84.

LADOKHIN, S. (2009): "Forecasting volatility in the stock market." *VU University Amsterdam, Business Mathematics and Informatics* .

LEAVENS, D. (1945): "Diversification of planning." *Trusts and Estates* **80**: pp. 469–473.

LIESENFELD, R. & R. JUNG (1997): "Stochastic volatility models: conditional normality versus heavy-tailed distributions." .

LIESENFELD, R. & J. RICHARD (2003): "Univariate and multivariate stochastic volatility models: estimation and diagnostics." *Journal of empirical finance* **10(4)**: pp. 505–531.

LOPEZ, J. (2001): "Evaluating the predictive accuracy of volatility models." *Journal of Forecasting* **20(2)**: pp. 87–109.

LOPEZ, J. A. (1999): "Regulatory evaluation of value at risk models." *Federal Reserve Bank of San Francisco* .

LORDE, T. & W. MOORE (2008): "Modeling and forecasting the volatility of long-stay tourist arrivals." *Tourism Analysis* **13(1)**: pp. 43–51.

MAILLARD, D. (2012): "A users guide to the cornish fisher expansion." .

MANDELBROT, B. (1972): "Certain speculative prices(1963)." *The Journal of Business* **45(4)**: pp. 542–543.

MAPA, D., H. MERCADER, Mazhiel, & P. TOLENTINO, Joy (2010): "GARCH vs. SV in Forecasting Asset Volatility: Do We Need Another Volatility Models?" *11th National Convention on Statistics (NCS)* .

MARKOWITZ, H. (1952): "Portfolio selection." *The journal of finance* **7(1)**: pp. 77–91.

MARQUARDT, D. (1963): "An algorithm for least-squares estimation of nonlinear parameters." *Journal of the society for Industrial and Applied Mathematics* **11(2)**: pp. 431–441.

MCALEER, M. (2009): "The ten commandments for optimizing value-at-risk and daily capital charges." *Journal of Economic Surveys* **23(5)**: pp. 831–849.

MCNEIL, A. & R. FREY (2000): "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach." *Journal of empirical finance* **7(3)**: pp. 271–300.

MCNEIL, A., R. FREY, & P. EMBRECHTS (2005): *Quantitative risk management: Concepts, techniques and tools.* Princeton Univ Pr.

MELINO, A. & S. TURNBULL (1990): "Pricing foreign currency options with stochastic volatility." *Journal of Econometrics* **45(1-2)**: pp. 239–265.

NELSON, D. (1991): "Conditional heteroskedasticity in asset returns: a new approach." *Econometrica: Journal of the Econometric Society* pp. 347–370.

PASTORELLO, S. & E. ROSSI (2010): "Efficient importance sampling maximum likelihood estimation of stochastic differential equations." *Computational Statistics & Data Analysis* **54(11)**: pp. 2753–2762.

PATTON, A. (2011): "Volatility forecast comparison using imperfect volatility proxies." *Journal of Econometrics* **160(1)**: pp. 246–256.

RICHARD, J. & W. ZHANG (1996): "Econometric modelling of uk house prices using accelerated importance sampling*." *Oxford Bulletin of Economics and Statistics* **58(4)**: pp. 601–613.

RICHARD, J. & W. ZHANG (2007): "Efficient high-dimensional importance sampling." *Journal of Econometrics* **141(2)**: pp. 1385–1411.

ROCKAFELLAR, R. & S. URYASEV (2000): "Optimization of conditional value-at-risk." *Journal of risk* **2**: pp. 21–42.

ROSENBLATT, M. (1952): "Remarks on a multivariate transformation." *The Annals of Mathematical Statistics* **23(3)**: pp. 470–472.

ROY, A. (1952): "Safety first and the holding of assets." *Econometrica: Journal of the Econometric Society* pp. 431–449.

RUIZ, E. (1994): "Quasi-maximum likelihood estimation of stochastic volatility models." *Journal of econometrics* **63(1)**: pp. 289–306.

SANDMANN, G. & S. KOOPMAN (1998): "Estimation of stochastic volatility models via monte carlo maximum likelihood." *Journal of Econometrics* **87(2)**: pp. 271–301.

SARYKALIN, S., G. SERRAINO, & S. URYASEV (2008): "Value-at-risk vs. conditional value-at-risk in risk management and optimization." In "Tutorials in operations research, INFORMS Annual Meeting," .

SHEPHARD, N., M. PITT, & N. COLLEGE (1995): "Likelihood analysis of non-gaussian parameter-driven models." *Economics Discussion Paper-Nuffield College Oxford* .

So, M., K. Lam, & W. Li (1999): "Forecasting exchange rate volatility using autoregressive random variance model." *Applied Financial Economics* **9(6)**: pp. 583–591.

Taylor, S. J. (1982): "Financial returns modelled by the product of two stochastic processes - a study of daily sugar prices 1961-79." In "Time Series Analysis: Theory and Practise," volume 1, pp. 203–226. Amsterdam: North Holland.

Taylor, S. J. (1986): "Modelling financial time series." *Chichester* .

Tersvirta, T. (2009): "An introduction to univariate garch models." In "Handbook of Financial Time Series," pp. 17–42. Springer Berlin Heidelberg.

Wei, Y., Y. Wang, & D. Huang (2010): "Forecasting crude oil market volatility: Further evidence using garch-class models." *Energy Economics* **32(6)**: pp. 1477–1484.

Yu, J. (2002): "Forecasting volatility in the new zealand stock market." *Applied Financial Economics* **12(3)**: pp. 193–202.

Zakoian, J. (1994): "Threshold heteroskedastic models." *Journal of Economic Dynamics and control* **18(5)**: pp. 931–955.

Appendix A

Tables

Table A.1: Descriptive Statistics: In-Sample Subset

Statistics	S&P	NASDAQ	CAC	DAX	FTSE
min	-9.4695	-9.5877	-9.4715	-7.3355	-9.2646
max	10.9572	11.1594	10.5946	10.7975	9.3842
median	0.1025	0.1258	0.0027	0.0439	0.0311
mean	-0.0080	0.0140	-0.0307	0.0069	-0.0114
var	3.3392	3.4873	3.4973	3.0598	2.6716
st.dev.	1.8273	1.8674	1.8701	1.7492	1.6345
skewness	-0.2156	-0.1668	0.2640	0.3110	-0.0345
kurtosis	9.7463	8.2820	8.7458	9.4753	9.5109
Jarque-Bera test	1	1	1	1	1
p-value	0.0010	0.0010	0.0010	0.0010	0.0010
test statistic	1523.30	933.68	1109.76	1410.53	1413.21
ADF test	1	1	1	1	1
p-value	0.0010	0.0010	0.0010	0.0010	0.0010
test statistic	-31.90	-31.23	-29.70	-28.75	-29.32
KPSS test	0	0	0	0	0
p-value	0.1000	0.1000	0.1000	0.1000	0.1000
test statistic	0.0824	0.0947	0.0780	0.0824	0.1019

Source: author's computations.

Table A.2: Descriptive Statistics: Out-on-Sample Subset

Statistics	S&P	NASDAQ	CAC	DAX	FTSE
min	-6.8958	-7.1489	-5.6346	-5.9947	-4.7792
max	4.6317	5.1592	6.0891	5.2104	3.9414
median	0.0927	0.0599	-0.0255	-0.0388	-0.0019
mean	0.0330	0.0447	-0.0779	-0.0309	0.0051
var	2.4825	2.8543	4.0564	4.2552	2.0574
st.dev.	1.5756	1.6895	2.0140	2.0628	1.4344
skewness	-0.5295	-0.4361	-0.1356	-0.1164	-0.2808
kurtosis	5.5663	5.2605	3.7644	3.4704	4.0015
Jarque-Bera test	1	1	0	0	1
p-value	0.0010	0.0010	0.0536	0.2643	0.0120
test statistic	64.2282	48.9221	5.4821	2.2955	10.9850
ADF test	1	1	1	1	1
p-value	0.0010	0.0010	0.0010	0.0010	0.0010
test statistic	-16.0225	-15.5921	-13.1309	-12.4469	-12.7127
KPSS test	0	0	0	0	0
p-value	0.1000	0.1000	0.1000	0.1000	0.1000
test statistic	0.0478	0.0457	0.0908	0.1102	0.0529

Source: author's computations.

Table A.3: Akaike Information Criteria for GARCH(p,q)

(p,q)	S&P	NASDAQ	CAC	DAX	FTSE
(1,1)	2736.3	2866.7	2989.7	2860.1	2711.5
(1,2)	2740.3	2870.7	2993.7	2864.1	2715.5
(2,1)	2738.4	2868.5	2990.0	2861.8	2712.5
(2,2)	2727.4	2862.5	2989.7	2853.8	2709.5

Source: author's computations.

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 17 periods			1. violation	after 17 periods			1. violation	after 96 periods		
PoF	8.38%			PoF	4.50%			PoF	0.38%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.0264	3.841	0.871	TUFF	0.0264	3.841	0.871	TUFF	4.6189	3.841	0.032
Uncond.	16.090	3.841	0.000	Uncond.	0.4351	3.841	0.510	Uncond.	60.231	3.841	0.000
Indep.	3.5430	3.841	0.060	Indep.	0.2996	3.841	0.584	Indep.	0.0000	3.841	1.000
Cond.	19.633	5.991	0.000	Cond.	0.7347	5.991	0.693	Cond.	60.231	5.991	0.000
MSE = 2001.42		MAD = 1.8076		MLAE = 0.0871		HMSE = 1.9399		R2LOG = -2.3932		QLIKE = 0.3234	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 17 periods			1. violation	after 17 periods			1. violation	after 500 periods		
PoF	3.13%			PoF	0.05%			PoF	0.00%		
Basel II	in the Red zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.0264	3.841	0.871	TUFF	0.0264	3.841	0.871	TUFF	0.0000	3.841	1.000
Uncond.	6.7939	3.841	0.009	Uncond.	55.258	3.841	0.000	Uncond.	0.0000	3.841	1.000
Indep.	0.0597	3.841	0.807	Indep.	0.0000	3.841	0.1000	Indep.	0.0000	3.841	1.000
Cond.	6.8536	5.991	0.032	Cond.	55.258	5.991	0.000	Cond.	0.0000	5.991	1.000
MSE = 3495.15		MAD = 2.1805		MLAE = 0.431		HMSE = 1.3827		R2LOG = -1.7913		QLIKE = 0.6799	

Table A.4: VaR Backtesting Results for S&P Index: In-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 17 periods			1. violation	after 17 periods			1. violation	after 88 periods		
PoF	9.00%			PoF	4.25%			PoF	0.63%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.0264	3.841	0.871	TUFF	0.0264	3.841	0.871	TUFF	3.9732	3.841	0.046
Uncond.	22.008	3.841	0.000	Uncond.	0.9960	3.841	0.318	Uncond.	50.793	3.841	0.000
Indep.	8.1768	3.841	0.004	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	1.000
Cond.	30.185	5.991	0.000	Cond.	0.9960	5.991	0.608	Cond.	50.793	5.991	0.000
MSE = 2114.58		MAD = 1.8711		MLAE = 0.1380		HMSE = 1.9333		R2LOG = -2.3833		QLIKE = 0.4054	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 17 periods			1. violation	after 88 periods			1. violation	after 500 periods		
PoF	4.00%			PoF	0.63%			PoF	0.00%		
Basel II	in the Red zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.0264	3.841	0.871	TUFF	3.9732	3.841	0.046	TUFF	0.0000	3.841	1.000
Uncond.	1.8027	3.841	0.179	Uncond.	50.793	3.841	0.000	Uncond.	0.0000	3.841	1.000
Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.100	Indep.	0.0000	3.841	1.000
Cond.	1.8027	5.991	0.406	Cond.	50.793	5.991	0.000	Cond.	0.0000	5.991	1.000
MSE = 3651.08		MAD = 2.2506		MLAE = 0.4651		HMSE = 1.4109		R2LOG = -1.7993		QLIKE = 0.7439	

Table A.5: VaR Backtesting Results for NASDAQ Index: In-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 10 periods			1. violation	after 48 periods			1. violation	after 86 periods		
PoF	8.50%			PoF	6.00%			PoF	0.25%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.4131	3.841	0.520	TUFF	1.0916	3.841	0.296	TUFF	3.8143	3.841	0.051
Uncond.	17.210	3.841	0.000	Uncond.	0.6856	3.841	0.408	Uncond.	65.886	3.841	0.000
Indep.	0.7261	3.841	0.394	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.999
Cond.	17.936	5.991	0.000	Cond.	0.6856	5.991	0.710	Cond.	65.886	5.991	0.000
MSE = 1979.77		MAD =1.8406		MLAE = 0.1978		HMSE = 1.9629		R2LOG = Inf		QLIKE = 0.3867	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 14 periods			1. violation	after 86 periods			1. violation	after 500 periods		
PoF	2.63%			PoF	0.38%			PoF	0.00%		
Basel II	in the Red zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	1.8027	5.991	0.406	TUFF	3.8143	3.841	0.051	TUFF	0.0000	3.841	1.000
Uncond.	11.408	3.841	0.001	Uncond.	60.231	3.841	0.000	Uncond.	0.0000	3.841	1.000
Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.100	Indep.	0.0000	3.841	1.000
Cond.	11.408	5.991	0.003	Cond.	60.231	5.991	0.000	Cond.	0.0000	5.991	1.000
MSE = 4332.89		MAD = 2.4302		MLAE = 0.6241		HMSE = 1.3620		R2LOG = Inf		QLIKE = 0.8333	

Table A.6: VaR Backtesting Results for DAX Index: In-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 13 periods			1. violation	after 24 periods			1. violation	after 96 periods		
PoF	9.25%			PoF	4.50%			PoF	0.25%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.1716	3.841	0.679	TUFF	0.0371	3.841	0.847	TUFF	3.7352	3.841	0.053
Uncond.	24.592	3.841	0.000	Uncond.	0.4351	3.841	0.510	Uncond.	65.886	3.841	0.000
Indep.	0.2233	3.841	0.637	Indep.	0.2996	3.841	0.584	Indep.	0.0000	3.841	0.999
Cond.	24.815	5.991	0.000	Cond.	0.7347	5.991	0.693	Cond.	65.886	5.991	0.000
MSE = 2368.68		MAD = 1.9949		MLAE = 0.2646		HMSE = 2.0263		R2LOG = -2.7735		QLIKE = 0.4384	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 24 periods			1. violation	after 85 periods			1. violation	after 85 periods		
PoF	3.75%			PoF	1.00%			PoF	0.13%		
Basel II	in the Red zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.0371	3.841	0.847	TUFF	3.7352	3.841	0.053	TUFF	3.7352	3.841	0.053
Uncond.	2.8701	3.841	0.090	Uncond.	39.578	3.841	0.000	Uncond.	72.590	3.841	0.000
Indep.	0.0159	3.841	0.900	Indep.	0.0000	3.841	1.000	Indep.	0.0005	3.841	0.983
Cond.	2.8860	5.991	0.236	Cond.	39.578	5.991	0.000	Cond.	72.591	5.991	0.000
MSE = 3928.42		MAD = 2.3579		MLAE = 0.5321		HMSE = 1.5196		R2LOG = -2.2757		QLIKE = 0.7472	

Table A.7: VaR Backtesting Results for CAC Index: In-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 2 periods			1. violation	after 17 periods			1. violation	after 17 periods		
PoF	9.88%			PoF	6.00%			PoF	6.00%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	3.3215	3.841	0.068	TUFF	3.3215	3.841	0.068	TUFF	4.8648	3.841	0.027
Uncond.	31.566	3.841	0.000	Uncond.	1.3677	3.841	0.242	Uncond.	43.011	3.841	0.000
Indep.	0.2537	3.841	0.614	Indep.	0.1151	3.841	0.734	Indep.	0.0000	3.841	1.000
Cond.	31.820	5.991	0.000	Cond.	1.4828	5.991	0.476	Cond.	0.0000	3.841	1.000
MSE = 1682.36		MAD = 1.6837		MLAE = 0.0953		HMSE = 1.9987		R2LOG = -3.0750		QLIKE = 0.2757	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 2 periods			1. violation	after 94 periods			1. violation	after 99 periods		
PoF	3.00%			PoF	0.75%			PoF	0.13%		
Basel II	in the Red zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	3.3215	3.841	0.068	TUFF	4.4561	3.841	0.035	TUFF	4.8648	3.841	0.027
Uncond.	7.8149	3.841	0.005	Uncond.	46.733	3.841	0.000	Uncond.	72.590	3.841	0.000
Indep.	0.1031	3.841	0.748	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	1.000
Cond.	7.9180	5.991	0.019	Cond.	46.733	5.991	0.000	Cond.	72.591	5.991	0.000
MSE = 3347.47		MAD = 2.1551		MLAE = 0.4919		HMSE = 1.4355		R2LOG = -2.3593		QLIKE = 0.6882	

Table A.8: VaR Backtesting Results for FTSE Index: In-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 8 periods			1. violation	after 8 periods			1. violation	after 14 periods		
PoF	7.54%			PoF	3.52%			PoF	1.01%		
Basel II	in the Red zone			Basel II	in the Yellow zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.6812	3.841	0.409	TUFF	0.6812	3.841	0.409	TUFF	0.1202	3.841	0.729
Uncond.	2.3504	3.841	0.125	Uncond.	1.0225	3.841	0.312	Uncond.	9.812	3.841	0.002
Indep.	0.6483	3.841	0.421	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.997
Cond.	2.9987	5.991	0.223	Cond.	1.0225	5.991	0.600	Cond.	9.812	5.991	0.007
MSE = 439.5639		MAD = 1.6725		MLAE =0.0758		HMSE = 2.0053		R2LOG = -2.2123		QLIKE = 0.2809	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 8 periods			1. violation	after 14 periods			1. violation	after 500 periods		
PoF	2.51%			PoF	0.50%			PoF	0.00%		
Basel II	in the Yellow zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.6812	3.841	0.409	TUFF	0.1202	3.841	0.729	TUFF	0.0000	3.841	1.000
Uncond.	3.1471	3.841	0.076	Uncond.	13.722	3.841	0.000	Uncond.	0.0000	3.841	1.000
Indep.	0.0000	3.841	1.000	Indep.	0.0019	3.841	0.965	Indep.	0.0000	3.841	1.000
Cond.	3.1471	5.991	0.207	Cond.	13.724	5.991	0.001	Cond.	0.0000	5.991	1.000
MSE = 694.8683		MAD = 1.9145		MLAE = 0.1987		HMSE = 1.4422		R2LOG = -2.0891		QLIKE = 0.4839	

Table A.9: VaR Backtesting Results for S&P Index: Out-of-Sample

90% VaR(GARCH)				95% VaR(GARCH)				99% VaR(GARCH)			
1. violation	after 8 periods			1. violation	after 8 periods			1. violation	after 14 periods		
PoF	7.04%			PoF	3.52%			PoF	1.01%		
Basel II	in the Red zone			Basel II	in the Yellow zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.6812	3.841	0.409	TUFF	0.6812	3.841	0.409	TUFF	0.1202	3.841	0.729
Uncond.	1.5490	3.841	0.213	Uncond.	1.0225	3.841	0.312	Uncond.	9.812	3.841	0.002
Indep.	0.0001	3.841	0.991	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.997
Cond.	1.5491	5.991	0.461	Cond.	1.0225	5.991	0.600	Cond.	9.812	5.991	0.007
MSE = 510.4345		MAD = 1.8274		MLAE = 0.1871		HMSE = 1.9920		R2LOG = -2.8268		QLIKE = 0.3871	

90% VaR(SV)				95% VaR(SV)				99% VaR(SV)			
1. violation	after 8 periods			1. violation	after 500 periods			1. violation	after 500 periods		
PoF	2.01%			PoF	0.00%			PoF	0.00%		
Basel II	in the Yellow zone			Basel II	in the Green zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.6812	3.841	0.409	TUFF	0.0000	3.841	1.000	TUFF	0.0000	3.841	1.000
Uncond.	4.7951	3.841	0.029	Uncond.	0.0000	3.841	1.000	Uncond.	0.0000	3.841	1.000
Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	1.000
Cond.	4.7951	5.991	0.091	Cond.	0.0000	5.991	1.000	Cond.	0.0000	5.991	1.000
MSE = 850.1778		MAD =2.1294		MLAE = 0.3379		HMSE = 1.4193		R2LOG = -2.6324		QLIKE = 0.6298	

Table A.10: VaR Backtesting Results for NASDAQ Index: Out-of-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 2 periods			1. violation	after 2 periods			1. violation	after 163 periods		
PoF	11.06%			PoF	6.03%			PoF	0.50%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	3.3215	3.841	0.068	TUFF	3.3215	3.841	0.068	TUFF	10.429	3.841	0.001
Uncond.	3.3215	3.841	0.068	Uncond.	0.4183	3.841	0.518	Uncond.	13.722	3.841	0.000
Indep.	5.1237	3.841	0.024	Indep.	0.0000	3.841	1.000	Indep.	0.0019	3.841	0.965
Cond.	16.721	5.991	0.000	Cond.	0.4183	5.991	0.811	Cond.	13.724	5.991	0.001
MSE =760.1525		MAD = 2.2777		MLAE = 0.3653		HMSE =2.1151		R2LOG = -3.1942		QLIKE = 0.6165	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 9 periods			1. violation	after 34 periods			1. violation	after 500 periods		
PoF	8.04%			PoF	2.01%			PoF	0.00%		
Basel II	in the Red zone			Basel II	in the Yellow zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.5332	3.841	0.465	TUFF	0.3538	3.841	0.552	TUFF	0.0000	3.841	1.000
Uncond.	3.2962	3.841	0.069	Uncond.	4.7951	3.841	0.029	Uncond.	0.0000	3.841	1.000
Indep.	8.2419	3.841	0.004	Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	1.000
Cond.	11.538	5.991	0.003	Cond.	4.7951	5.991	0.091	Cond.	0.0000	5.991	1.000
MSE = 1110.2578		MAD = 2.5812		MLAE = 0.5416		HMSE = 1.6921		R2LOG = -3.0077		QLIKE = 0.7899	

Table A.11: VaR Backtesting Results for DAX Index: Out-of-Sample

90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 8 periods			1. violation	after 11 periods			1. violation	after 163 periods		
PoF	9.05%			PoF	5.03%			PoF	0.50%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.6812	3.841	0.409	TUFF	0.3153	3.841	0.574	TUFF	10.429	3.841	0.001
Uncond.	5.5885	3.841	0.018	Uncond.	0.0003	3.841	0.987	Uncond.	13.722	3.841	0.000
Indep.	0.0925	3.841	0.761	Indep.	0.0000	3.841	1.000	Indep.	0.0019	3.841	0.965
Cond.	5.6810	5.991	0.058	Cond.	0.0003	5.991	1.000	Cond.	13.724	5.991	0.001
MSE = 716.5550		MAD = 2.2037		MLAE = 0.3991		HMSE = 2.1100		R2LOG =-3.1444		QLIKE =0.5889	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 11 periods			1. violation	after 11 periods			1. violation	after 11 periods		
PoF	6.53%			PoF	3.52%			PoF	0.50%		
Basel II	in the Red zone			Basel II	in the Yellow zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.3153	3.841	0.574	TUFF	0.3153	3.841	0.574	TUFF	0.3153	3.841	0.574
Uncond.	10.901	3.841	0.342	Uncond.	1.0225	3.841	0.312	Uncond.	13.722	3.841	0.000
Indep.	0.0000	3.841	1.000	Indep.	0.0000	3.841	0.100	Indep.	0.0019	3.841	0.965
Cond.	0.9013	5.991	0.637	Cond.	1.0225	5.991	0.600	Cond.	13.724	5.991	0.001
MSE = 1107.7787		MAD = 2.5579		MLAE = 0.5336		HMSE = 1.7474		R2LOG = -2.9192		QLIKE = 0.7738	

Table A.12: VaR Backtesting Results for CAC Index: Out-of-Sample

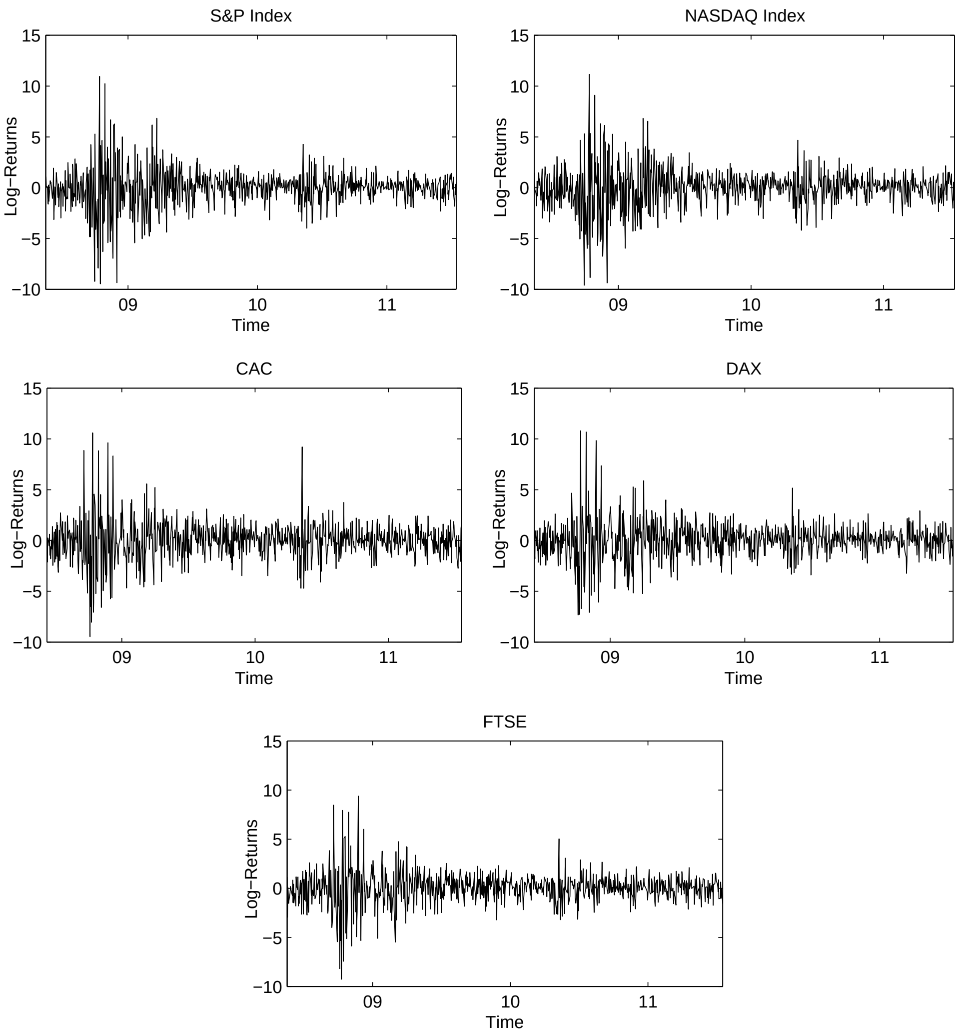
90% VaR(GARCH)				**95% VaR(GARCH)**				**99% VaR(GARCH)**			
1. violation	after 12 periods			1. violation	after 12 periods			1. violation	after 13 periods		
PoF	9.05%			PoF	6.03%			PoF	0.50%		
Basel II	in the Red zone			Basel II	in the Red zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.2359	3.841	0.627	TUFF	0.2359	3.841	0.627	TUFF	0.1716	3.841	0.679
Uncond.	5.5885	3.841	0.018	Uncond.	0.4183	3.841	0.518	Uncond.	13.722	3.841	0.000
Indep.	1.1547	3.841	0.283	Indep.	4.9715	3.841	0.026	Indep.	0.0019	3.841	0.965
Cond.	6.7433	5.991	0.034	Cond.	5.3898	5.991	0.068	Cond.	13.724	5.991	0.001
MSE = 380.6693		MAD = 1.5863		MLAE = 0.0538		HMSE = 2.0427		R2LOG = -3.1913		QLIKE = 0.2436	
90% VaR(SV)				**95% VaR(SV)**				**99% VaR(SV)**			
1. violation	after 12 periods			1. violation	after 13 periods			1. violation	after 183 periods		
PoF	8.54%			PoF	2.51%			PoF	0.50%		
Basel II	in the Red zone			Basel II	in the Yellow zone			Basel II	in the Green zone		
	Value	chi-sq	p-value		Value	chi-sq	p-value		Value	chi-sq	p-value
TUFF	0.2359	3.841	0.627	TUFF	0.1716	3.841	0.679	TUFF	12.249	3.841	0.000
Uncond.	4.3780	3.841	0.036	Uncond.	3.1471	3.841	0.076	Uncond.	13.722	3.841	0.000
Indep.	1.5722	3.841	0.210	Indep.	2.7292	3.841	0.099	Indep.	0.0019	3.841	0.965
Cond.	5.9502	5.991	0.051	Cond.	5.8764	5.991	0.053	Cond.	13.724	5.991	0.001
MSE = 697.9614		MAD = 1.9689		MLAE =0.2999		HMSE = 1.7229		R2LOG = -2.8157		QLIKE = 0.4534	

Table A.13: VaR Backtesting Results for FTSE Index: Out-of-Sample

Appendix B

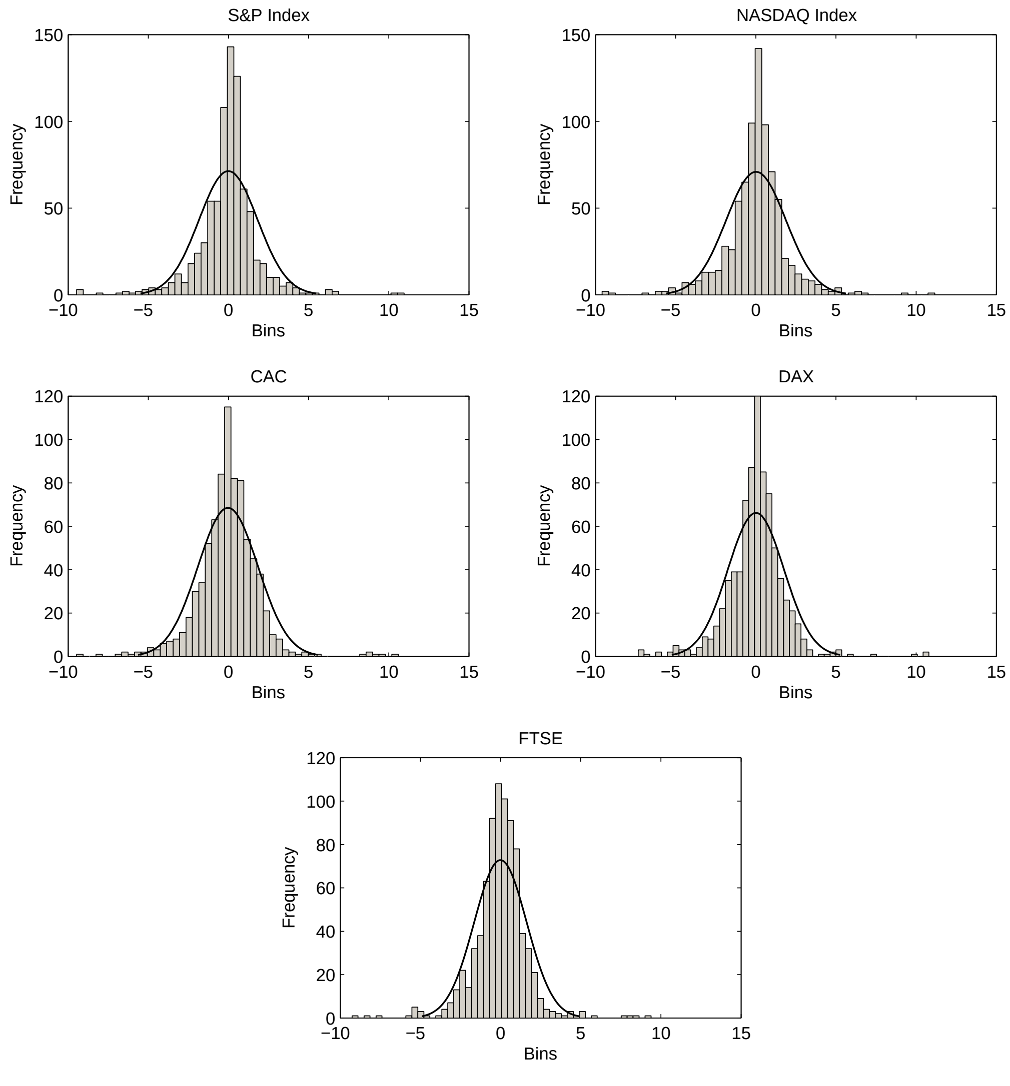
Figures

Figure B.1: Log-Returns of Indices: In-Sample



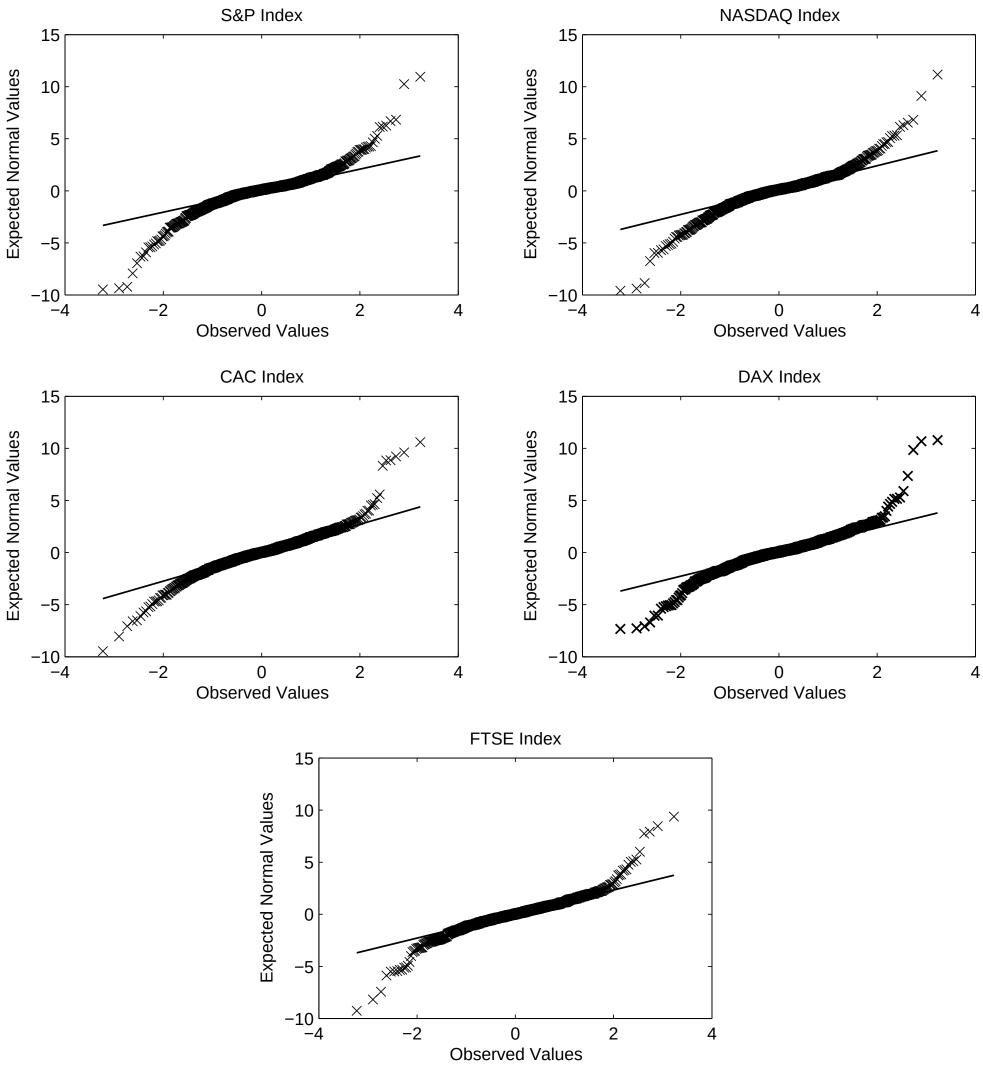
Source: author's computations in Matlab.

Figure B.2: Histograms of Log-Returns with a fitted normal distribution: In-Sample



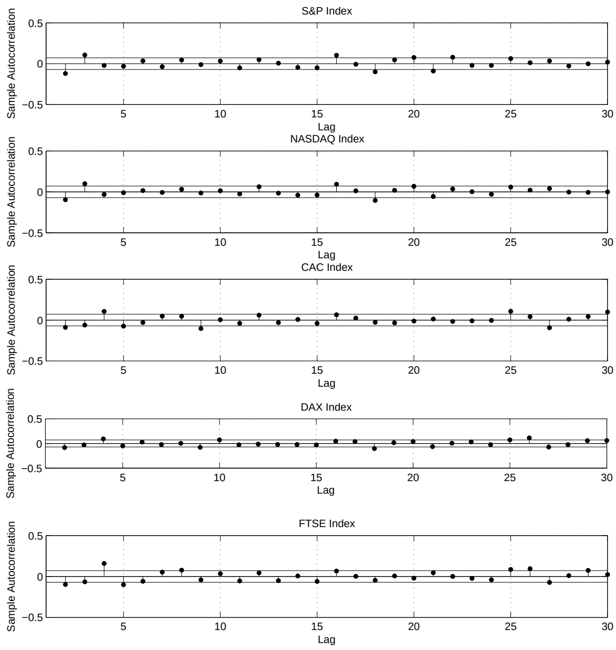
Source: author's computations in Matlab.

Figure B.3: QQ Plots: Empirical cdf compared with theoretical normal cdf: In-Sample



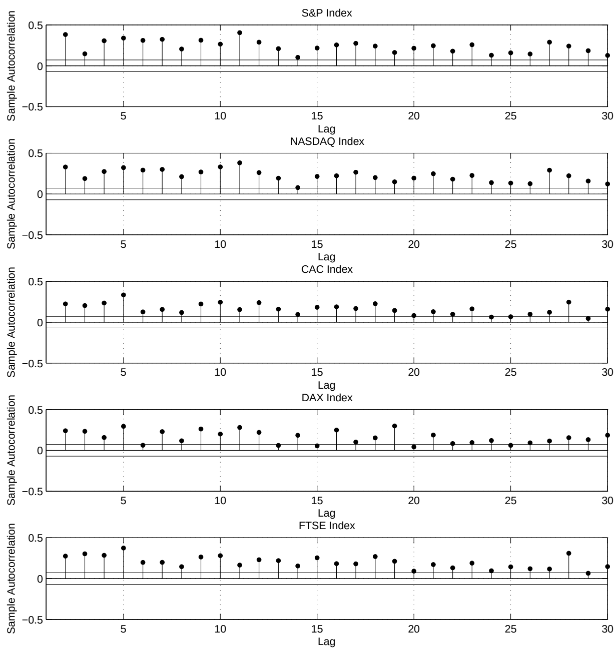
Source: author's computations in Matlab.

Figure B.4: Sample ACF for Returns: In-Sample



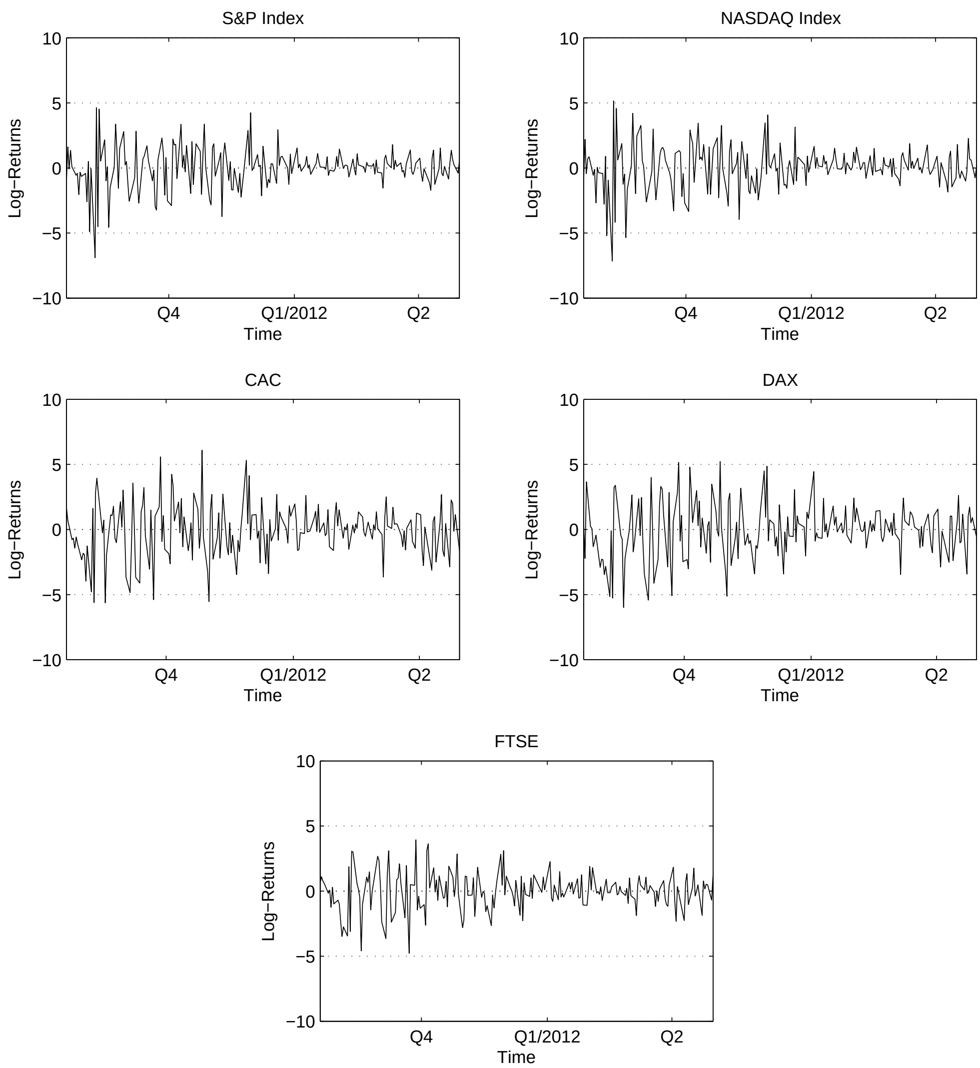
Source: author's computations in Matlab.

Figure B.5: Sample ACF for Squared Returns: In-Sample



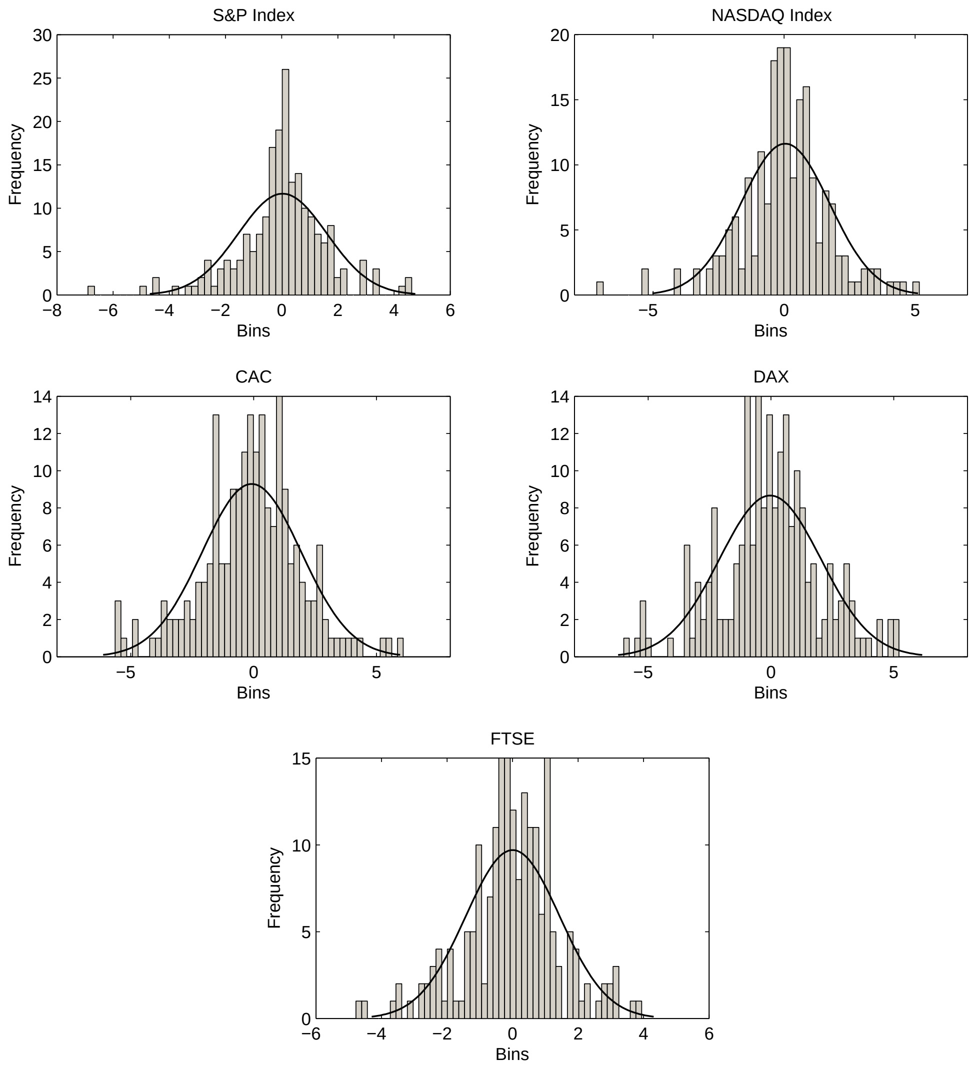
Source: author's computations in Matlab.

Figure B.6: Log-Returns of Indices: Out-of-Sample



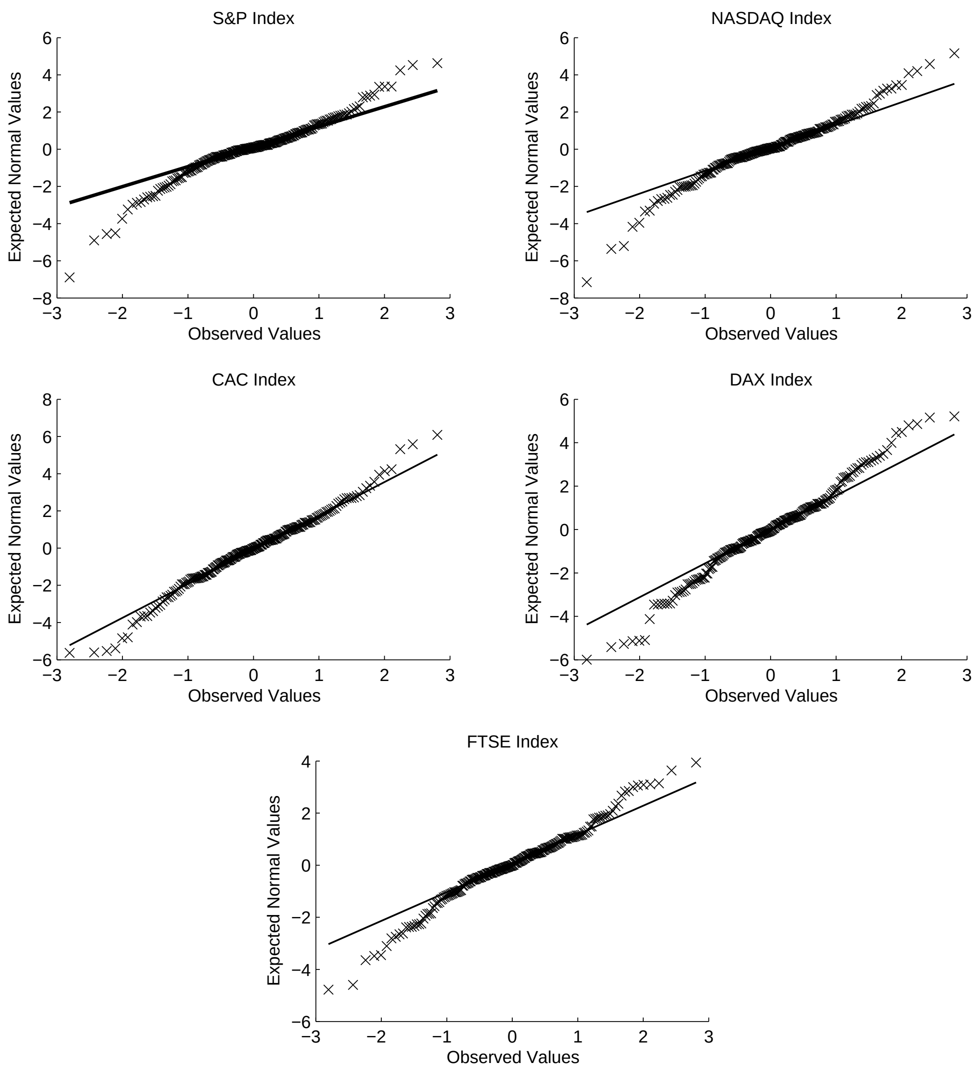
Source: author's computations in Matlab.

Figure B.7: Histograms of Log-Returns with a fitted normal distribution: Out-of-Sample



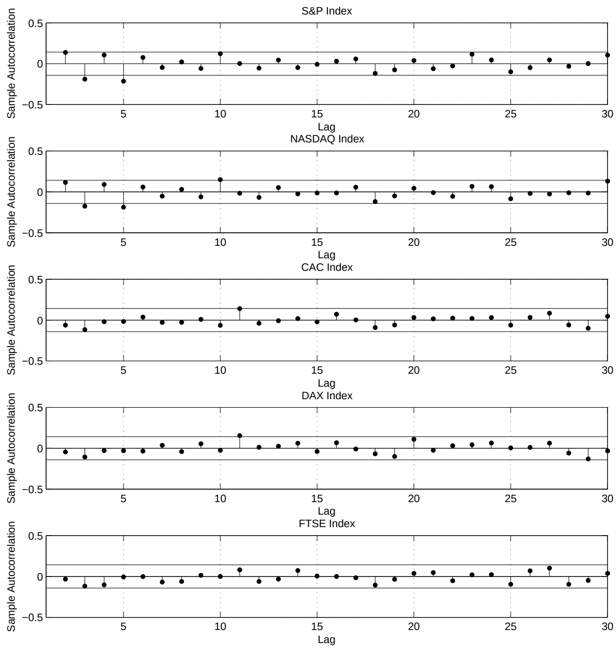
Source: author's computations in Matlab.

Figure B.8: QQ Plots: Empirical cdf compared with theoretical normal cdf: Out-of-Sample



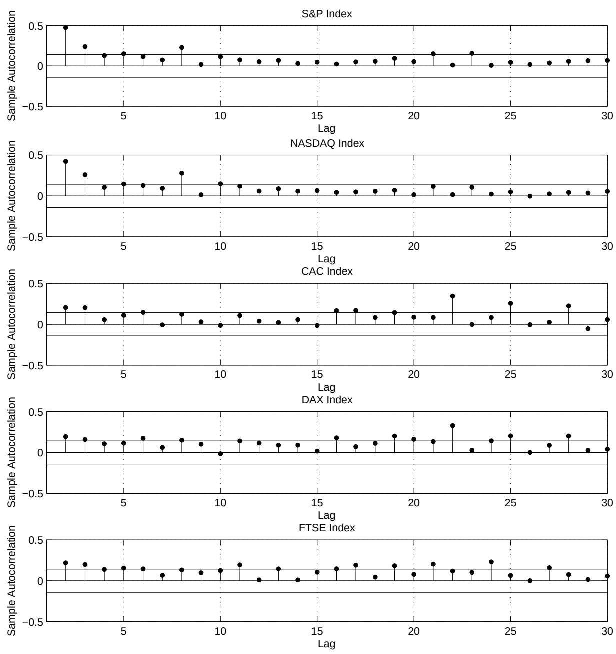
Source: author's computations in Matlab.

Figure B.9: Sample ACF for Returns: Out-of-Sample



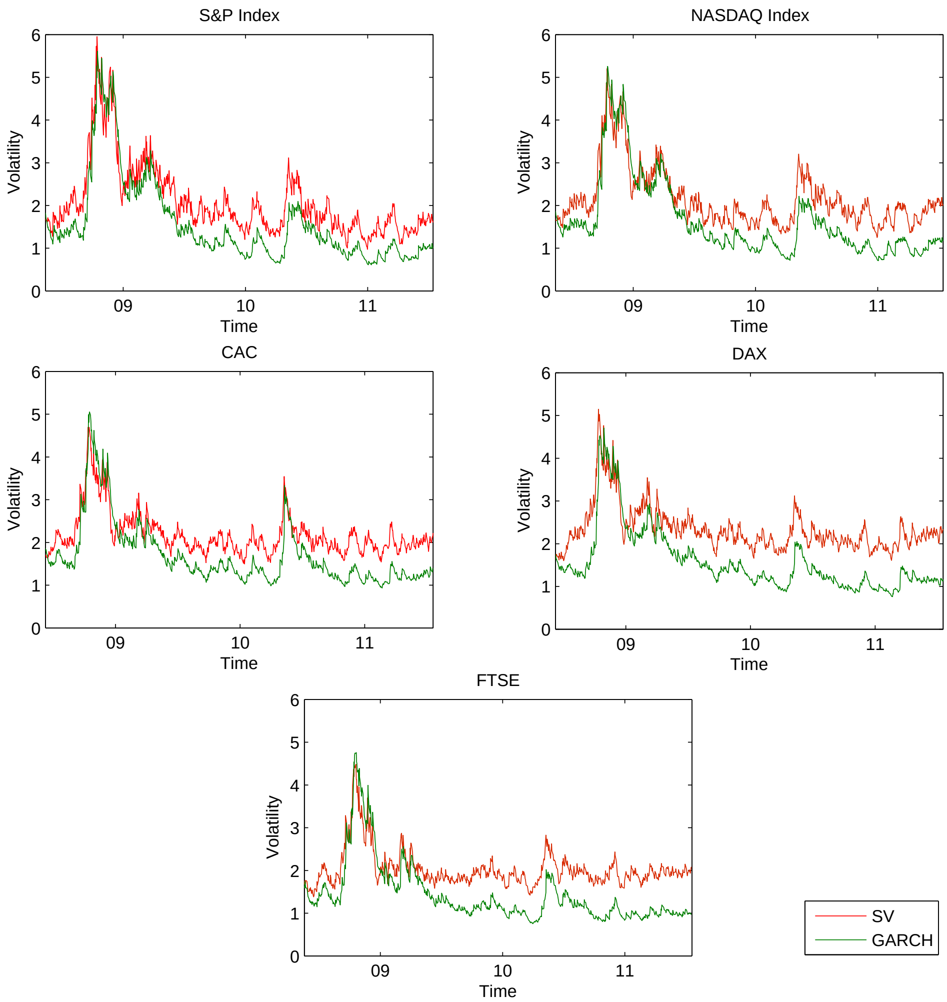
Source: author's computations in Matlab.

Figure B.10: Sample ACF for Squared Returns: In-Sample



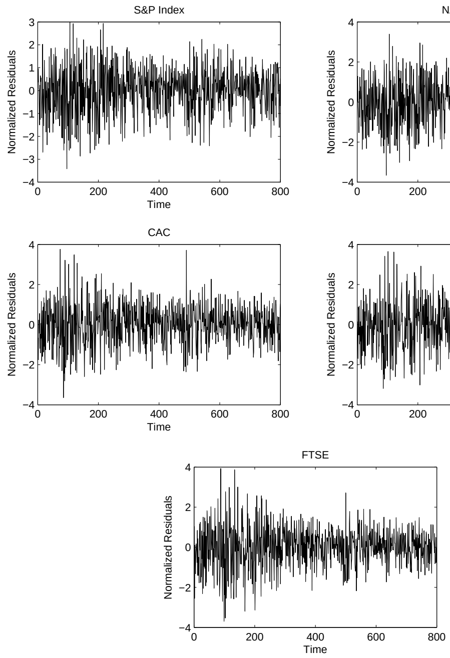
Source: author's computations in Matlab.

Figure B.11: Volatility Predictions: In-Sample



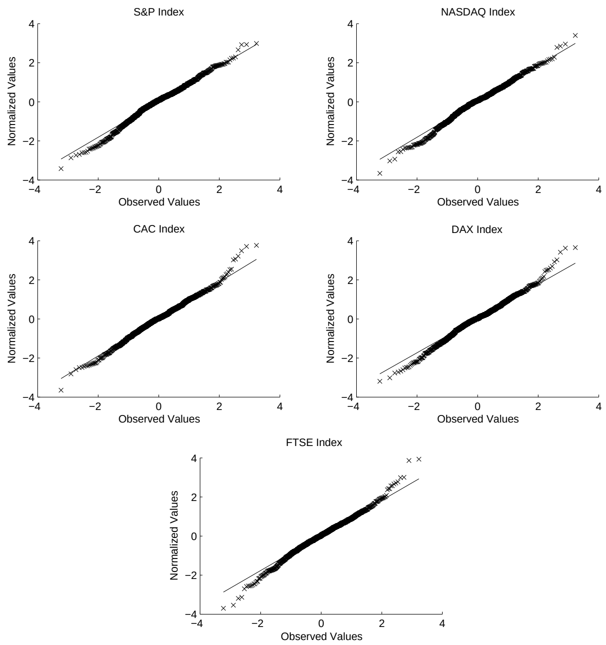
Source: author's computations in Matlab.

Figure B.12: Normalized Residuals of SV-t model



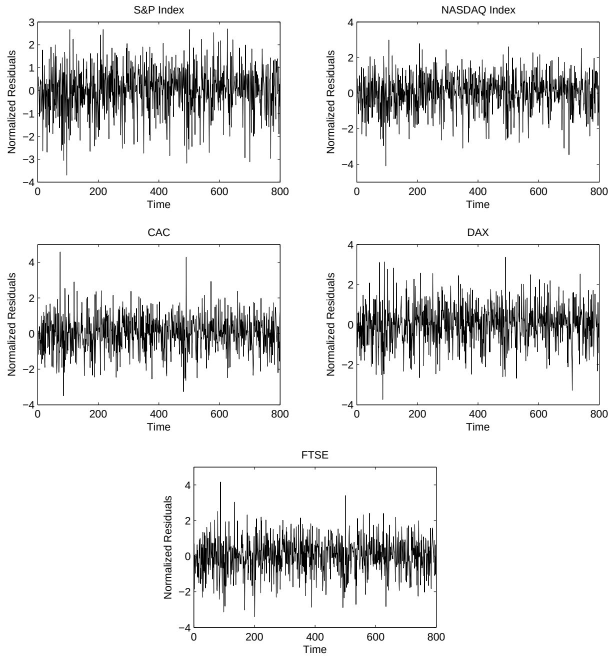
Source: author's computations in Matlab.

Figure B.13: QQ plots of normalized residuals: SV-t model



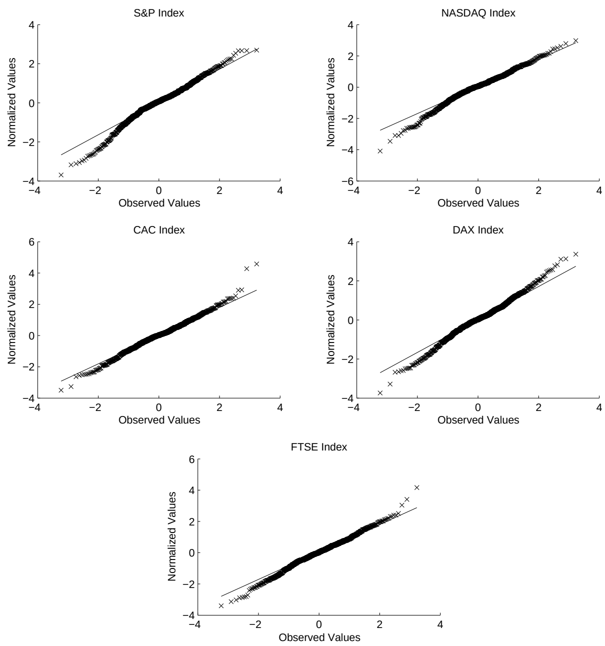
Source: author's computations in Matlab.

Figure B.14: Normalized Residuals of GARCH-t model



Source: author's computations in Matlab.

Figure B.15: QQ plots of normalized residuals: GARCH-t model



Source: author's computations in Matlab.