Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

MASTER THESIS

Modeling Conditional Quantiles of Central European Stock Market Returns

Author: **Bc. Diana Burdová**

Supervisor: **PhDr. Jozef Baruník, Ph.D.**

Academic Year: **2013/2014**

Declaration of Authorship

The author hereby declares that he 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, January 5, 2014

Signature

Acknowledgments

I am honestly grateful to my supervisor, Jozef Baruník, whose guidance, encouragement and valuable suggestions helped me complete this thesis.

I would like to thank my family for supporting me during my studies and Štefan for his patience, love and support.

Abstract

Most of the literature on Value at Risk concentrates on the unconditional nonparametric or parametric approach to VaR estimation and much less on the direct modeling of conditional quantiles. This thesis focuses on the direct conditional VaR modeling, using the flexible quantile regression and hence imposing no restrictions on the return distribution. We apply semiparametric Conditional Autoregressive Value at Risk (CAViaR) models that allow time-variation of the conditional distribution of returns and also different time-variation for different quantiles on four stock price indices: Czech PX, Hungarian BUX, German DAX and U.S. S&P 500. The objective is to investigate how the introduction of dynamics impacts VaR accuracy. The main contribution lies firstly in the primary application of this approach on Central European stock market and secondly in the fact that we investigate the impact on VaR accuracy during the pre-crisis period and also the period covering the global financial crisis. Our results show that CAViaR models perform very well in describing the evolution of the quantiles, both in absolute terms and relative to the benchmark parametric models. Not only do they provide generally a better fit, they are also able to produce accurate forecasts. CAViaR models may be therefore used as a suitable tool for VaR estimation in practical risk management.

JEL Classification C14, C22, C52, G11, G17

Keywords VaR, GARCH, CAViaR, conditional quantiles, quantile regression, backtesting

Author's e-mail burdova.diana@gmail.com

Supervisor's e-mail barunik@utia.cas.cz

Abstrakt

Prevažná časť literatúry na tému Value at Risk (VaR) sa zameriava na nepodmienené neparametrické alebo parametrické prístupy k jeho odhadovaniu, oveľa menšia časť na priame modelovanie podmienených kvantilov. Táto práca sa sústreďuje na priame modelovanie podmieneného VaRu, za pomoci flexibilnej kvantilovej regresie, a teda nekladie žiadne obmedzenia na rozdelenie výnosov. Na štyri cenové indexy, a to český PX, maďarský BUX, nemecký DAX a americký S&P 500, aplikujeme semiparametrické podmienené autoregresné Value at Risk (CAViaR) modely, ktoré umožňujú variáciu podmieneného rozdelenia výnosov v čase a takisto rôznu časovú variáciu pre rôzne kvantily. Hlavným cieľom práce je skúmať ako zavedenie dynamiky ovplyvňuje presnosť VaR odhadov. Hlavný prínos práce spočíva v tom, že sa jedná o prvú aplikáciu tohto prístupu na stredoeurópsky akciový trh a po druhé, že skúmame vplyv na presnosť VaR odhadov v období pred krízou a takisto počas krízy. Výsledky dokazujú, že CAViaR modely veľmi dobre popisujú vývoj kvantilov v čase, či už z hľadiska absolútneho alebo relatívneho v porovnaní s parametrickými modelmi. Nielen že poskytujú všeobecne lepšie odhady, ale prinášajú aj presné predpovede. Tieto modely preto môžu slúžiť ako vhodný nástroj na odhadovanie VaRu pri praktickom riadení rizík.

Klasifikace JEL	C14, C22, C52, G11, G17
Klíčová slova	VaR, GARCH, CAViaR, podmienené kvantily, kvantilová regresia, backtestové metódy

E-mail autora	`burdova.diana@gmail.com`
E-mail vedoucího práce	`barunik@utia.cas.cz`

Contents

List of Tables

List of Figures

Acronyms

ADF Augmented Dickey-Fuller test

ARCH Autoregressive Conditional Heteroskedasticity

CAViaR Conditional Autoregressive Value at Risk

CDF Cumulative Distribution Function

DM Diebold-Mariano test

DQ Dynamic Quantile test

GARCH Generalized Autoregressive Conditional Heteroskedasticity

MLE Maximum Likelihood Estimation

VaR Value at Risk

Master Thesis Proposal

Author	Bc. Diana Burdová
Supervisor	PhDr. Jozef Baruník, Ph.D.
Proposed topic	Modeling Conditional Quantiles of Central European Stock Market Returns

Topic characteristics Effective risk management is a very important issue for financial institutions, their investment decisions and regulation. The necessity of accurate risk measures has also been stressed by the recent global financial crisis. The exposure to market risk of a financial institution is measured by value at risk (VaR). It is a standard tool, which represents the maximum potential loss of the asset or portfolio value within a given time period at a given confidence level (usually 1 % or 5 %). In other words, VaR is a particular quantile of future portfolio values conditional on current information. Although it is widely used due to its conceptual simplicity, it is necessary to investigate value at risk dynamically as the volatility of assets exhibits strong dynamics. The nature of risks and also the distribution of portfolio returns change over time; therefore there is a need for models accounting for time-varying conditional quantiles. The aim of my thesis will be to examine whether dynamic modeling of VaR on Central European stock market provides more satisfactory results. To do this, I would like to apply conditional autoregressive value at risk (CAViaR) proposed by Engle and Manganelli (2004) and evaluate its predictive performance in comparison with other value at risk models. To my best knowledge, there has not been published any academic work focusing on applying this model on CEE market so far.

Hypotheses

1. Dynamic modeling of value at risk using CAViaR provides us with more accurate estimates than methods based on static modeling.

2. Introduction of dynamics to VaR modeling using CAViaR improves its performance in the crisis period.

3. Dynamics of quantiles of CEE markets is different from dynamics of US markets quantiles.

Methodology The main method used in my thesis will be the model known as conditional autoregressive value at risk (CAViaR). As opposed to many VaR estimation methods, CAViaR is based on direct modeling of quantile instead of modeling the whole distribution of returns and then recovering its quantile. Engle and Manganelli (2004) propose a conditional autoregressive quantile specification due to volatility clustering of stock market returns, which results in autocorrelated distribution. The unknown parameters of the CAViaR specifications are estimated using regression quantile framework introduced by Koenker and Bassett (1978). In my thesis, different specifications of VaR and CAViaR approaches will be applied on a portfolio of CEE and US stocks for various confidence levels. To investigate their performance under different circumstances, I will focus on three different periods: before-crisis period, crisis period and after-crisis period.

Outline

1. Introduction
2. Literature Overview
3. Value at Risk Models and Methods
4. Conditional Autoregressive Value at Risk Model (CAViaR)
5. Forecast Evaluation
6. Data Overview
7. Empirical Application and Discussion of Results - Comparison of VaR and CAViaR Estimation
8. Conclusion

Core bibliography

1. Chen, C. W. S., Gerlach, R., Hwang, B.B.K. & M. McAleer (2012): "Forecasting Value-at-Risk using nonlinear quantiles and the intra-day range." *International Journal of Forecasting* **28**: pp. 557–574.

2. Chernozhukov, V. & L. Umantsev (2001): "Conditional Value-at-Risk: Aspects of modeling and estimation." *Empirical Economics* **26**: pp. 271–293.

3. Engle, R. F. & S. Manganelli (2004): "CAViAR: conditional autoregressive value at risk by regression quantiles." *Journal of Business and Economic Statistics* **22**: pp. 367–381.

4. Engle, R. F. & S. Manganelli (2001): "Value at Risk models in finance." *European Central Bank, Working paper no. 75*

5. Jeon, J. & J. W. Taylor (2011): "Using CAViaR models with implied volatility for Value-at-Risk estimation."

6. Koenket, R. & G. Bassett (1978): "Regression quantiles." *Econometrica* **1**: pp. 33–50.

7. Taylor, J. W. (2001): "Estimating Value at Risk and expected shortfall using expectiles." *Journal of Financial Econometrics* **6**: pp. 231–252.

Author

Supervisor

Chapter 1

Introduction

Financial institutions are exposed to various kinds of risk that reflect the level of uncertainty regarding the future returns. In this thesis we concentrate on the market risk, which represents the potential loss due to unexpected price movements and changes in market conditions, hence significant especially for traders. Substantial increase in financial uncertainty in the 1990s has led to intensive study of price volatility (as a measure of riskiness) in the stock markets and search for precise and more efficient tools for its estimation. Also recent global financial crisis of 2008/2009 has stressed that effective risk management and accurate risk measures are of great importance for financial institutions, as these are used for their key decisions and should therefore respond properly to the changing nature of risks and overall market development.

The exposure to market risk is measured by Value at Risk (VaR), which has become a standard tool among risk managers, financial analysts and regulators, such as the Basel Committee on Banking Supervision via The Basel II Accord. VaR represents the potential loss in the asset or portfolio value over a given time horizon, for a given confidence level. It was pioneered by J.P. Morgan in 1996 as the RiskMetrics system and has gained on popularity mostly due to its simple concept. Various methods have been since developed for its calculation. Most of the literature focuses either on the methods that model VaR in an unconditional sense or parametric (location-scale) methods that retrieve VaR from conditional volatility forecasts. Even though the latter model the volatility conditionally and hence capture the phenomenon of volatility clustering, they are restricted to a certain distribution of return innovations that is not assumed to be changing over time, which is their main drawback. These methods involve mostly the well-known generalized autoregressive conditional heteroskedastic

(GARCH) volatility models of Engle (1982) and Bollerslev (1986) with various error distributions and their subsequent modifications and have been in fact widely applied in the sense of VaR estimation among researchers, for example by Angelidis *et al.* (2004), Kuester *et al.* (2006), Bams *et al.* (2005) or Mittnik *et al.* (2000). However, much less literature concentrates on the direct modeling of conditional quantiles. This is actually crucial, since VaR represents a particular quantile of future portfolio returns conditional on current information. Inspired by the stylized facts about financial return data and the changing nature of return distributions, semiparametric methods based on the flexible quantile regression (see Koenker & Bassett (1978), Chernozhukov & Umantsev (2001)) have been proposed, such as Conditional Autoregressive Value at Risk (CAViaR) of Engle & Manganelli (2004) and its recent extensions of Jeon & Taylor (2013), Gerlach *et al.* (2011) or Chen *et al.* (2012)). This approach is very attractive since it aims to derive VaR directly, imposes no distributional restrictions and hence allows time-variation of the conditional return distribution and also different time-variation for different quantiles.

The objective of this thesis is to investigate how the introduction of dynamics to VaR estimation using the quantile regression impacts VaR accuracy. We apply semiparametric Conditional Autoregressive Value at Risk (CAViaR) of Engle & Manganelli (2004) and model conditional quantiles directly on four stock price indices: Czech PX, Hungarian BUX, German DAX and U.S. S&P 500. Few empirical applications of this method focus mainly on American and Asian stock markets. In this respect, our first contribution to the existent literature lies in the application of this approach on Central European stock market indices. To the best of our knowledge, this study presents the primary results for this region. Another contribution lies in the fact that we investigate the impact on VaR accuracy during the pre-crisis period and also the period covering the global financial crisis. In our empirical analysis, we also compare this approach to the parametric GARCH approach that serves as the benchmark.

Our results for four stock price indices between January 2003 and December 2012 suggest that the dynamic approach seems to improve VaR accuracy. As a group, CAViaR models perform very well in describing the evolution of the quantiles, both in absolute terms and relative to the benchmark parametric models. On one hand, they generally provide a better fit, also further in the left tail and even in the crisis period, and on the other hand are able to produce accurate forecasts as well.

The rest of the thesis is structured as follows. Chapter 2 provides an overview of Value at Risk concept and various approaches to its calculation, their advantages and shortcomings. Chapter 3 introduces CAViaR and describes its different specifications in detail. Chapter 4 is dedicated to the forecast evaluation and backtesting methods. In Chapter 5 we present an empirical application of GARCH and CAViaR models with the summary of results. Chapter 6 concludes.

Chapter 2

Value at Risk

Starting with the theory of Value at Risk (VaR), in this chapter we first provide the reader with its basic notion from practical and statistical point of view. Further, we continue with approaches to its calculation and describe in detail those commonly used, pointing at their advantages as well as limitations. Finally, we present briefly several shortcomings of VaR in general as a risk measure.

Foundations of the Value at Risk approach as a measurement of market risk go back to 1990s. This concept, which is very easily understood and implemented in practice, was developed to evaluate the cost of positions in terms of risk and therefore enable managament of many financial institutions to allocate the underlying risk effectively. It serves not only financial and risk managers, but also regulators. As a part of The Basel II Accord, VaR methodology is used by the Basel Committee on Banking Supervision (1996) at the Bank for International Settlements for capital requirements of various financial institutions and investment firms in order to cover their exposures to market risk. The objective is clear: estimate the underlying risk as precisely as possible in order to avoid maintaining too high or low levels of capital, which could eventually lead to very inaccurate investment or other decisions. Both underestimation or overestimation of the underlying risk could potentially result in substantial losses, whereas in case of the latter the loss refers to the opportunity costs, to be more precise. VaR was first implemented by Morgan Corporation in 1996, known as RiskMetrics system.

Jorion (1996) formally defines VaR as *'the worst expected loss over a target horizon within a given confidence interval'*. Nice feature of this measure is that the information about the exposure to risk and the associated probability is summarized in just a single number, which is expressed in the same units as the bottom line (for example returns on indices, commodities or foreign exchange rates).

We want to find VaR such that

$$\Pr\left(y_t < -VaR_t|\mathcal{F}_t\right) = \alpha, \tag{2.1}$$

where y_t denotes return of an asset or portfolio at time $t = 1, \ldots, T$, $\mathcal{F}_t$ represents the information set available at time t and α is the probability level referring to the risk.[1] The purpose of VaR measure is to define an upper bound on losses, thus forecast a value each period that will be exceeded with only a small probability α, which is usually 1 % or 5 %.

From statistical point of view, VaR is understood as a particular quantile of future portfolio returns conditional on current information. Formally, given some probability $\alpha \in (0, 1)$ and conditional on the information available at time t, VaR for time $t + h$ at confidence level $1 - \alpha$ is defined as the negative α-quantile of the conditional return distribution, that is

$$VaR_{t+h,\alpha} = -q_\alpha(y_{t+h}|\mathcal{F}_t) = -\inf\left\{x \in \mathcal{R} : Pr(y_{t+h} \leq x|\mathcal{F}_t) \geq \alpha\right\} \tag{2.2}$$

where $q(\cdot)$ denotes the quantile function (the inverse cumulative distribution function (CDF) of y_t).[2]

According to Engle & Manganelli (2004), appropriate method for VaR calculation should fulfil certain features that can be summarized in the following points:

- providing a formula for VaR_t calculation as a function of variables known at time $t - 1$
- providing a set of unknown parameters that have to be estimated
- providing a procedure for estimation of these unknown parameters (loss function and some optimization algorithm)
- providing a test for the estimate evaluation.

Developing a suitable model for VaR forecasting is still an ongoing and very challenging research. In the further text we go through several approaches to VaR calculation, as well as their stronger and weaker points.

[1] The confidence level is therefore $1 - \alpha$.

[2] For more rigorous background see for example McNeil *et al.* (2005).

2.1 VaR approaches

Since the concept of VaR has become very popular over the last 20 years, it has also become a subject of intestive research among academics. Thus, there exist many different approaches for its computation, but there is no universal method providing perfectly satisfactory results. Despite certain distinctions between them, accoding to Engle & Manganelli (2001), they all share three common features: (1) daily marking-to-market of the portfolio, (2) estimating the distribution of portfolio returns and finally (3) computing the VaR of this portfolio. The second point here is related to the estimation of possible price changes of the asset or portfolio and is in fact the reason why there are many different approches to VaR forecasting. All of these methods are designed to capture and incorporate some or all of the well known characteristics of financial data, which have been summarized by Mandelbrot (1963) and Fama (1965). These include non-normality of return distribution, as returns commonly exhibit heavier tails and higher kurtosis (therefore their distribution is rather leptokurtic), negative skeweness and significant autocorrelation of squared returns, which results in volatility clustering (volatility of assets evolves over time, it is stable in the short period, but can change in the long period).

From the extensive literature concentrating on VaR and its calculation methods, in this thesis we use the classification of the existing VaR models according to Engle & Manganelli (2001), since they consider also recently developed and less common methodologies. They divide these methods into three broad categories:

- Nonparametric methods
- Parametric methods
- Semiparametric methods

We describe each category in the following subsections.

2.1.1 Nonparametric methods

Nonparametric methods involve for example Historical Simulation (HS), Hybrid model or Monte Carlo simulation. All of these methods are relatively non-restrictive.

HS is a very common method substantially simplifying VaR calculations as there are no assumptions about the distribution of the portfolio returns. It is based on a simple concept of rolling windows. This means that a window consisting of observations of the most recent periods is chosen and current VaR is estimated as the quantile of the empirical distribution of historical returns from this window. By moving the window one observation forward, one obtains VaR forecast for the following day. Any return within the window is equally likely, but returns outside this window have zero probability of occuring. Implicit assumption about the return distribution is that it does not change over time. Hence, the risk of the portfolio is related to its historical experience. There are several problems referring to this method. One of them is inconsistency of the empirical quantile estimator (to be consistent, the size of the window has to go to infinity). Another one is related to the size (length) of the moving window. On one hand, the number of observations in the window has to be large enough to avoid large sampling errors, but on the other hand not too large in order to adapt accurately to dynamics in the true distribution. By choosing too large window we risk combining the periods with high and low volatility and consequently obtaining biased VaR estimates. Another issue associated with this method is the effect of extreme observations. VaR estimates may be jumping upward whenever a large observation is included in the window and downward when it drops out of the rolling window. Therefore this method is not cosidered to provide very reliable results, especially for the extreme quantiles (Engle & Manganelli (2001)).

To overcome some of these drawbacks, Boudoukh *et al.* (1998) proposed the hybrid method, which is basically a combination of historical simulation and RiskMetrics. In comparison with HS, where each observation has the same weight, this method applies exponentially declining weights to the past portfolio returns.

Monte Carlo simulation is a general and very flexible method based on simulations, widely applicable in finance, but more demanding. To compute VaR with this method, return processes are simulated according to a certain type of distribution and VaR is then calculated using a particular quantile from these simulated processes.

2.1.2 Parametric methods

These include RiskMetrics by J.P.Morgan & Reuters (1996), GARCH models of Bollerslev (1986), or Stochastic Volatility (SV) methods of Taylor (1986). The basic intuition behind these models is the parametrization of the behavior of prices. This means that we assume a certain type of the distribution of returns and the model of dynamics (conditional volatility forecast), which are then employed together to estimate conditional quantiles. The advantage of the parametric approach is that under each model, this is done simply by multiplying the estimated standard deviation and a particular quantile of the error distribution from the corresponding model of dynamics. Formally, the general formula for the parametric one-step-ahead VaR at confidence level $(1-\alpha)$ has the following form:

$$VaR_t = \mu + \sqrt{h_t} q_\alpha(D), \tag{2.3}$$

where μ is the mean of the distribution D, $\sqrt{h_t}$ its standard deviation and $q_\alpha(D)$ is the α-quantile of this distribution. When applying all sorts of parametric models, the task is to estimate $\sqrt{h_t}$ and then retrieve VaR_t at confidence level $1-\alpha$.

To compute variance and consequently the standard deviation under RiskMetrics approach, Exponentially Weighted Moving Average (EWMA) method is used. First, we consider series of returns y_t for $t = 1, ..., T$, such that

$$y_t = z_t \sqrt{h_t} \qquad z_t \overset{iid}{\sim} N(0,1), \tag{2.4}$$

where h_t is the conditional variance of y_t. If the conditional mean of y_t is assumed to be 0, the variance equation is defined as

$$h_t = \lambda h_{t-1} + (1-\lambda)\, y_{t-1}^2 \tag{2.5}$$

where λ is usually set to 0.94 for daily data.[3] EWMA is a special case of GARCH. Since GARCH models are a benchmark in our comparative study, we discuss them separately in more detail in the further text of this chapter.

Focusing on the GARCH type models in the sense of VaR estimation, apart from the specification of the variance equation, these rely also on the assumption that the errors are i.i.d and require specification of their distribution (such as Gaussian, Student-t, skewed Student-t or Generalized error distribution). After choosing a certain type of error distribution it is possible to estimate the unknown parameters. When applying this approach, several problems may arise, such as the misspecification of the variance equation or the distribution used for log-likelihood function, or the fact that the errors may not be

[3] For more details see J.P.Morgan & Reuters (1996).

i.i.d. Their performance can be improved by avoiding the assumption about the normality of errors and employing alternative distributions, which are more consistent with the empirical findings about the behavior of financial returns.

In general, the normality assumption (in GARCH or RiskMetrics) leads to underestimation of the risk, as has been shown in vast empirical research. GARCH type models with various probability distributions in VaR estimation have been applied for example by Angelidis *et al.* (2004), who find that leptokurtic distributions provide more accurate one-step-ahead VaR forecasts. According to Kuester *et al.* (2006), the performance of parametric VaR may be considerably improved by employing the skewed Student-t distribution that accounts for both skeweness and heavier tails. Another empirical applications involve for example Bams *et al.* (2005) or Mittnik *et al.* (2000).

To demonstrate that the choice of the distribution of the standardised residuals is of great importance when estimating VaR parametrically, in this thesis we consider two different distributions: normal (Gaussian) distribution, which serves as a benchmark and Student-t distribution, which is more in accord with the characteristics of financial returns we already discussed. The cumulative distribution function (CDF) of Student-t distribution has the following form:

$$D(x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}}\mathrm{d}t, \tag{2.6}$$

where ν denotes the degrees of freedom - a parameter that describes the heaviness of tails and needs to be estimated. $\Gamma(\cdot)$ denotes the gamma function, where $\Gamma(\nu) = \int_0^{\infty} x^{\nu-1}e^{-x}\mathrm{d}x$ for $\nu > 0$.The distribution is symmetric around zero, bell-shaped and with growing ν approaches the normal distribution with mean 0 and variance 1.

GARCH models

Now we turn to the description of the well known and very popular approach to volatility modeling, Generalized Autoregressive Conditional Heteroskedasticity (GARCH), which we apply to obtain volatility forecast in our empirical research. GARCH model was proposed by Bollerslev (1986) as the extension of the ARCH process introduced by Engle (1982). As opposed to conventional econometric models, the ARCH process does not consider a constant conditional variance, but allows the conditional variance to be time-varying (as a function of past errors), while the unconditional variance remains constant.

However, based on the empirical evidence, relatively high order of lags in the conditional variance equation of ARCH(q) process was needed for accurate variance estimation.[4] This led to the search for a model that would have on one hand reasonable lag structure and on the other hand relatively long memory. Whereas in the ARCH(q) process the conditional variance is specified only as a linear function of past sample variances, GARCH models were developed to consider also the effect of the past conditional variance on the current conditional variance by using its own lagged values as some sort of '*adaptive mechanism*'. The extension of the ARCH class models to GARCH type models corresponds to the extension of an AR process to an ARMA process.

As the name already suggests, GARCH models allow for heteroskedasticity in the data and enable modeling of changing variance. However, one possible drawback that is worth mentioning is that the basic GARCH model does not allow for leverage effect. This means that both positive and negative returns have the same effect on volatility, since only the magnitude of return matters, not the sign. To capture this feature and also other empirically found properties of the financial data, numerous modifications of the original GARCH model and the variance equation have been proposed since the pioneering works of Engle (1982) and Bollerslev (1986). Extensions accounting for the different response to positive and negative returns and asymmetry in the data include for example Threshold GARCH (TGARCH) of Zakoian (1994), Exponential GARCH (EGARCH) of Nelson (1991) or Non-linear asymmetric GARCH (NGARCH) of Engle & Ng (1993), whereas Asymmetric power ARCH (A-PARCH) of Ding *et al.* (1993) and Fractionally integrated GARCH (FIGARCH) by Baillie *et al.* (1996) address also a long memory property of stock market returns.

Having introduced the background of GARCH type models, we now proceed with the fundamentals of the basic GARCH(p,q) process. Following Bollerslev (1986), ϵ_t denotes a real-valued discrete time stochastic process and $\mathcal{F}_t$ denotes the information available at time t. GARCH(p,q) process is then defined as:

$$\epsilon_t = \sqrt{h_t} z_t \tag{2.7}$$

$$\epsilon_t | \mathcal{F}_{t-1} \sim N(0, h_t) \tag{2.8}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}, \tag{2.9}$$

[4]See Bollerslev (1986) for the list of empirical applications.

where

$$p \geq 0, \quad q \geq 0$$

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad i = 1, \ldots, q$$

$$\beta_j \geq 0, \quad j = 1, \ldots, p.$$

In this context, ϵ_t represents mean corrected return $\epsilon_t = y_t - \mu$, h_t is its conditional variance (positive and time-varying) and z_t is assumed to be i.i.d. random variable. Conditions $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1, ..., q$ and $\beta_j \geq 0$ for $j = 1, ..., p$ ensure the stationarity of conditional variance.

Restriction

$$\sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) < 1 \tag{2.10}$$

ensures the second-order stationarity of GARCH(p,q) process, which means that the unconditional variance is finite and the conditional variance h_t changes over time. The conditions for wide-sense stationarity are summarized in the following proposition:

Proposition 2.1. *(Bollerslev (1986)) The GARCH (p,q) process as defined in* (2.7), (2.8) *and* (2.9) *is wide-sense stationary with* $E(\epsilon_t) = 0$, $var(\epsilon_t) = \alpha_0/(1 - \sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i))$ *and* $cov(\epsilon_t, \epsilon_s) = 0$ *for* $t \neq s$ *iff* $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$.[5]

In GARCH(p,q) model, α_0, α_i and β_j are the parameters to be estimated. If $p = 0$ we obtain the ARCH(q) process and if also $q = 0$, ϵ_t is white noise. The restricted version of GARCH(p,q) model, in which $\sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) = 1$, is known as Integrated GARCH (IGARCH). The restriction property implies that there is a unit-root in the process indicating that past squared shocks are persistent. Obviously, Exponentially Weighted Moving Average (EWMA) model, which we described in Subsection 2.1.2, is a special case of GARCH model as it corresponds to the Integrated GARCH(1,1).

Estimation of the GARCH regression model

(G)ARCH regression models are often estimated using Maximum Likelihood Estimator (MLE). Generally, assuming i.i.d. innovations z_t with density func-

[5] For proof see the appendix in the paper of Bollerslev (1986).

tion $D(z_t;\nu)$[6], the log-likelihood function of $y_t(\theta)$ for $t = 1,\ldots,T$ has the following form

$$l_T\left(\{y_t\};\theta\right) = \sum_{t=1}^{T}\left[\ln\left[D\left(z_t\left(\theta\right);\nu\right)\right] - \frac{1}{2}\ln\left(h_t\left(\theta\right)\right)\right], \tag{2.11}$$

where θ is a vector of parameters to be estimated for the conditional mean, variance and density function, and $z_t(\theta) = \frac{\epsilon_t(\theta)}{\sqrt{h_t(\theta)}}$. Estimated parameters $\widehat{\theta}$ are obtained as a solution to the maximization of the log-likelihood function given by Equation 2.11 with respect to parameters. As we assume normal and Student-t distribution in this thesis, we provide the reader with the form of the log-likelihood function for this particular distributions. For the normal distribution of z_t, the log-likelihood function takes on the form

$$l_T\left(\{y_t\};\theta\right) = -\frac{1}{2}\left[T\ln(2\pi) + \sum_{t=1}^{T} z_t^2 + \sum_{t=1}^{T}\ln(h_t)\right], \tag{2.12}$$

whereas for the t-distributed z_t the log-likelihood function is as follows

$$\begin{aligned} l_T\left(\{y_t\};\theta\right) = T&\left[\ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln\left[\pi\left(\nu-2\right)\right]\right] \\ &- \frac{1}{2}\sum_{t=1}^{T}\left[\ln\left(h_t\right) + \left(1+\nu\right)\ln\left(1+\frac{z_t^2}{\nu-2}\right)\right]. \end{aligned} \tag{2.13}$$

More details regarding other distributions can be found in Angelidis *et al.* (2004).

2.1.3 Semiparametric methods

Semiparametric methods are relatively recently introduced methodology and include for example applications of Extreme Value Theory (EVT) (see for example Danielsson & de Vries (2000)) that concentrates on the tail of the return distribution, rather than the whole distribution and is generally more suitable for very low (extreme) probability levels.[7] Further, these methods involve also quasi-maximum likelihood GARCH models (see Bollerslev & Woolridge (1992) or McNeil & Frey (2000)) and methods based on quantile regression (see Cher-

[6] While Engle (1982) assumed normal distribution of z_t in ARCH(q) process, Bollerslev (1986) introduced t-distribution with $\nu > 2$ in GARCH(p,q).

[7] For rigorous review of advantages and disadvantages of EVT see for example Diebold *et al.* (1998).

nozhukov & Umantsev (2001)), such as the CAViAR models proposed by Engle & Manganelli (2004). As the center point of this thesis is to examine the performance of the CAViaR models, we give them special attention and describe them in detail in Chapter 4.

2.2 Shortcomigs of VaR as a risk measure

Having mentioned several shortcomings of individual methodologies commonly used for VaR calculation in the previous section, we now briefly mention some drawbacks of VaR as a risk measure.

These drawbacks are nicely summarized and described for example in Dowd (2002). We start with the ones that refer not only to VaR individually, but also to other risk measures in general. Obviously, VaR estimates can often be subject to errors, which stem especially from relying on inaccurate assumptions within the models that are implemented for risk calculation, or the implementation itself, which can result in quite different and imprecise results of similar models. Turning to those that are characteristic for VaR itself, we continue with three main drawbacks. The first one is that VaR is uninformative of tail losses, meaning that it provides information only about the potential expected loss and the associated probability with which this will not be exceeded (sets an upper bound on losses), but it does not specify how much we lose (lower bound), when the violation of VaR (tail event) really occurs. Hence, this may create a deceptive image about the relative riskiness of various positions with the same VaR at some confidence level. However, this problem may be possibly resolved by estimating VaR at various confidence levels, especially high levels. Another drawback refers to the portfolio diversification. When looking at VaR, the diversification of risks can be noticeably discouraged, as the VaR of the diversified portfolio is greater than the VaR of undiversified portfolio. As pointed out by Artzner *et al.* (1999), probably the most striking limitation is that VaR is not a coherent risk measure, since it does not satisfy the property of sub-additivity. Formally, a risk measure $\rho(\cdot)$ is sub-additive if

$$\rho(A+B) \leq \rho(A) + \rho(B), \tag{2.14}$$

i.e. the risk of the sum of positions A and B is less than or equal to the sum of the risk of the individual positions alone. This property implies that the overall risk does not increase when the individual risks are aggregated. This is generally desirable for any risk measure, especially because of the reporting conservatism. A measure with this feature always gives an overestimate of aggregated risk.

Nevertheless, VaR is still a common and popular measure used in risk management. We disregard the above mentioned limitations, since the scope of this thesis is not to assess the suitability of VaR measure as such, but to assess the accuracy of methodologies developed for its calculation. Moreover, in case of conditional VaR estimation, which is our main focus, the last drawback does not concern us. In the further text, we proceed with the theoretical background of conditional, dynamic approach to VaR modeling.

Chapter 3

CAViaR

Having described widely used methods for VaR estimation in previous chapter, either in an unconditional sense or with the restriction to a certain distribution, we now turn to the theoretical framework that imposes no distributional restrictions and comfortably enables us to model VaR conditionally, since VaR represents a particular quantile of future portfolio returns *conditional* on current information.

To allow dynamics in the VaR modeling, Engle & Manganelli (2004) propose a new approach to quantile estimation, the so-called Conditional Autoregressive Value at Risk (CAViaR). Contrary to the methods that are based on modeling of the whole distribution and then recovering its quantile, CAViaR aims to derive the evolution of the quantile directly. This is a very strong point of these models, as they allow time-variation of the conditional distribution of returns and also different time-variation for different quantiles. The approach is inspired by the well known fact that volatilities of market returns exhibit certain dynamics and clustering over time, meaning that they possess autocorrelated structure. Since VaR is linked to the standard deviation of the distribution of returns, it should be designed to follow a similar pattern. To do so, it is reasonable to use some sort of autoregressive process, for example conditional autoregressive specification as proposed by Engle & Manganelli (2004), which we present in the following section.

3.1 CAViaR specifications

Following Engle & Manganelli (2004), we start with several assumptions and notations: y_t denotes return of an asset or portfolio at time $t = 1, \dots, T$, α

denotes the probability level associated with VaR, $\mathbf{x}_t$ a vector of variables observable at time t (which represent the information set at time t) and θ_α a vector of unknown parameters to be estimated. Then the time t α-quantile of the distribution of returns formed at time $t-1$ is denoted as $q_t(\theta) \equiv q_t(\mathbf{x}_{t-1}, \theta_\alpha)$. The α subscript has been dropped for simplicity reasons.

Generally, CAViaR specification is then defined as follows:

$$q_t(\theta) = \beta_0 + \sum_{i=1}^{s} \beta_i q_{t-i}(\theta) + \sum_{j=s+1}^{r+s+1} \beta_j l(\mathbf{x}_{t-j}), \tag{3.1}$$

where l is a function of a finite number of lagged values of observable variables and θ is a vector of βs to be estimated and has the dimension of $p = s + r + 1$. Vector $\mathbf{x}_{t-j}$ is usually chosen as a vector of lagged returns. The second term in Equation 3.1, $\sum_{i=1}^{s} \beta_i q_{t-i}(\theta)$, is autoregressive and enables smooth transition of the quantile over time. Function $l(\mathbf{x}_{t-j})$ is there to tie the quatile $q_t(\theta)$ to observable variables in the information set.

The authors suggest that the dependence of VaR on lagged returns $|y_{t-1}|$ could be symmetrical. That is, if the return y_{t-1} reaches very negative value, one could anticipate that the VaR will increase, because the probability of bad day happening again increases. But on the other hand, very positive returns could increase VaR as well (as suggested by volatility models).

As the main objective is to specify the l function in various alternatives, Engle & Manganelli (2004) propose four different CAViaR specifications:

1. Adaptive

$$q_t(\theta) = q_{t-1}(\theta) + \beta_1 \left\{ \left[1 + \exp\left(G\left[y_{t-1} - q_{t-1}(\theta)\right]\right)\right]^{-1} - \alpha \right\} \tag{3.2}$$

2. Symmetric Absolute Value (SAV)

$$q_t(\theta) = \beta_0 + \beta_1 q_{t-1}(\theta) + \beta_2 |y_{t-1}| \tag{3.3}$$

3. Asymmetric Slope (AS)

$$q_t(\theta) = \beta_0 + \beta_1 q_{t-1}(\theta) + \beta_2 (y_{t-1})^+ + \beta_3 (y_{t-1})^- \tag{3.4}$$

where $(y_{t-1})^+ = y_t I(y_t \geq 0)$ and $(y_{t-1})^- = y_t I(y_t < 0)$.

4. Indirect GARCH(1,1)

$$q_t(\theta) = (\beta_0 + \beta_1 q_{t-1}^2(\theta) + \beta_2 y_{t-1}^2)^{1/2} \tag{3.5}$$

In Adaptive model, G is a positive finite number and as it approaches infinity, the last term in Equation 3.2 converges to $\beta_1[I(y_{t-1} \leq q_{t-1}(\theta)) - \alpha]$, where $I(\cdot)$ is the indicator function. The idea of this model specification is to increase VaR immediately as it is exceeded and decrease it moderately in the cases when it is not exceeded. This leads to the reduction of the probability of observing sequence of hits, but on the other hand ensures that some hits will occur. As the authors point out, the model does not distinguish the magnitude by which VaR is exceeded. When it is exceeded by a large margin, VaR is simply increased by the same amount as it would be exceeded by a small margin.

The following three models are similar in structure to GARCH models. In these the coefficient of the lagged VaR is not constrained to equal 1, which means that they are mean-reverting. As the name already suggests, Symmetric Absolute Value model responds symmetrically to past returns. On the other hand, Asymmetric slope model allows for leverage effect of past returns. In the fourth specification, Indirect GARCH(1,1), the response to past returns is again symmetrical. This model is correctly specified if the data follow a true GARCH(1,1) process with i.i.d. errors. Contrary to typical GARCH(1,1), this model differs in estimation technique. It is not estimated using MLE, but quantile regression which we describe in Section 3.2. Overall, the important feature is that compared to GARCH models, CAViaR models are more general as they may be employed also in the cases when the error distributions are not i.i.d., or volatilities or error distributions are subject to a change.

Recently, there have been some contributions made in the area of CAViaR that appear to improve its predictive ability. Jeon & Taylor (2013) extend these models by utilizing implied volatility, Yu *et al.* (2010) propose threshold and mixture type indirect-GARCH CAViaR models, while Gerlach *et al.* (2011) propose nonlinear threshold CAViaR model. Chen *et al.* (2012) introduce range-based CAViaR models that incorporate intra-day high-low price range data. Huang *et al.* (2010) propose the so called index-exciting CAViAR model, in which the parameters are allowed to be driven by the market index return as these are designed as its time-varying functions.

In this thesis, we employ approach of Engle & Manganelli (2004). Similarly to Huang *et al.* (2010) or Barunik & Zikes (2013), in our empirical application on Central European stock market data we estimate two among four different CAViaR specifications, Symmetric Absolute Value (SAV) that adapts VaR to the size of the returns and Asymmetric Slope (AS) that allows also for the leverage effect of returns.

3.2 Regression quantiles

In this section we focus on the regression quantile framework of Koenker & Bassett (1978) that is used to estimate the unknown parameters in the CAViaR models.[1] The aim is to minimize the regression quantile loss function as we proceed to show formally.

Following Engle & Manganelli (2004), we first consider a sample of return observations y_t for $t = 1, \ldots, T$ that are generated by the model:

$$y_t = \mathbf{x}_t^{'}\theta^0 + \epsilon_{\alpha t} \qquad Q_\alpha(\epsilon_{\alpha t}|\mathbf{x}_t) = 0, \tag{3.6}$$

where $\mathbf{x}_t$ is a vector of p regressors and $Q_\alpha(\epsilon_{\alpha t}|\mathbf{x}_t)$ is the α-quantile of $\epsilon_{\alpha t}$ conditional on $\mathbf{x}_t$. We also define $q_t(\theta) \equiv \mathbf{x}_t\theta$. Within this framework, the α-regression quantile is then determined by any $\hat{\theta}$ that solves the following minimization problem:

$$\min_\theta \frac{1}{T}\sum_{t=1}^{T} \left[\alpha - I\left(y_t < q_t(\theta)\right)\right]\left[y_t - q_t(\theta)\right], \tag{3.7}$$

where $I(\cdot)$ is the indicator function.

Over the years, regression quantile framework has been extended variously. Some extensions are based on the different assumptions about the errors, for example Koenker & Bassett (1982) consider heteroskedastic errors, Portnoy (1991) nonstationary dependent errors. Also, a number of alternative linear models have been developed, such as time series models, simultaneous equation models, censored regression models or the most recent autoregressive quantiles.[2] Furthermore, there are some extensions regarding the non-linear regression quantiles, as well as in the context of time series (for example White (1994),

[1]See Koenker & Bassett (1978) for the extension of a sample quantile concept to a linear regression model. To see how a linear quantile regression is applied to VaR estimation, we refer the reader to Chernozhukov & Umantsev (2001).

[2]See Engle & Manganelli (2004) for the list of the relevant literature.

who provides proofs that the regression quantile estimates for the i.i.d. and stationary cases are consistent).

We point out that Engle & Manganelli (2004) consider a non-linear regression quantile estimator $\hat{\theta}$. The advantage of their regression quantile model is that there is no need for specification of the whole distribution of the errors, the only requirement is to specify the quantile process properly. However, the authors note that even if this specification is not correct, the minimization procedure still provides estimates satisfying the Kullback-Leibler Information Criterion. In their paper, they also establish the consistency and asymptotical normality of this estimator, provide a consistent estimator of the variance-covariance matrix and derive the asymptotic distribution of the estimator, which is necessary for hypothesis testing of the quantile models.[3]

[3]For technicalities, such as theorems and proofs, we refer the interested reader to the original paper of Engle & Manganelli (2004).

Chapter 4

Forecast evaluation

Crucial part when estimating and forecasting various economic variables is generally model validation and forecast evaluation. Value at Risk is no exception, therefore we dedicate this chapter to the backtesting methods that serve to assess the accuracy of its estimates and forecasts. Assessing correctness of VaR is actually a huge part of the literature on this topic. Over the years, various tests have been developed to assess its predictive accuracy. Nice summary of numerous backtesting and forecast evaluation procedures can be found for example in Campbell (2007), Dowd (2002) or Berkowitz *et al.* (2011). Commonly used statistical backtests for VaR accuracy employ transformation of VaR and realized returns, such as sequence of violations, and test for certain properties of these transformations. Many of these tests are developed in an unconditional setting, but since we focus on conditional VaR modeling in this thesis, it is important to assess if VaR models are correctly specified and perform well in a conditional sense. From various procedures we therefore focus not only on unconditional, but also conditional methods that serve to assess the absolute performance and lead us to either accept or reject considered models. To find out which models perform best, we also employ methods that serve to assess the relative performance of competing models.

4.1 Evaluation of absolute performance

First, to evaluate the absolute performance of the models under study, we use the unconditional coverage test of Kupiec (1995), test of independence and conditional coverage of Christoffersen (1998) and dynamic quantile (DQ) test of Engle & Manganelli (2004).

Before we go further, let us first define violation of VaR forecast by realized return (also called tail loss) as $I(y_t < -VaR_t)$, where $I(\cdot)$ is the indicator function. $\{I(y_t < -VaR_t)\}_{t=1}^{T} = \{I_t\}_{t=1}^{T}$ then denotes sequence of violations, where T is the number of observations. As suggested by Christoffersen (1998), accurate VaR model should result in such a sequence of violations that satisfies both of the following properties:

1. proportion (percentage) of violations is close to the risk level α, formally: $Pr(I_t = 1) = E[I_t] = \alpha$
2. violations do not exhibit clustering (they are independent).

Test of unconditional coverage

Unconditional Coverage test proposed by Kupiec (1995), also known as Proportion of Failures (PoF)[1] test, focuses on the first property we mention above. It tests for the correct number of violations - whether the actual number of violations is consistent with the frequency predicted by the model. Under the null hypothesis that the model is accurate (consistent with the data), the number of violations follows a binomial distribution with probability α:

$$Pr\left(x|n, \alpha\right) = \binom{n}{x} \alpha^{x} (1-\alpha)^{n-x}, \tag{4.1}$$

where x is the number of VaR violations, n is the total number of observations and α is the frequency of violations (tail losses) predicted by the model. Note that $\alpha = (1 - confidence\ level)$. The null hypothesis is tested using likelihood-ratio test statistic LR_{uc}. Under the null hypothesis $H_0 : \alpha = \frac{x}{n}$, the model predicts $n\alpha$ violations and for LR_{uc} it holds that

$$LR_{uc} = -2ln\left[(1-\alpha)^{n-x}\alpha^{x}\right] + 2ln\left[(1-\frac{x}{n})^{n-x}(\frac{x}{n})^{x}\right] \overset{asy}{\sim} \chi^2_{(1)}. \tag{4.2}$$

The unconditional coverage test rejects the null hypothesis if the proportion of VaR violations statistically differs from the risk level α.

Test of independence

For a particular VaR model to be accurate, it is important that this is not only correct on average (satisfies the unconditional coverage property), but

[1]PoF is defined as a number of violations x to the total number of observations n.

also accounts for temporal volatility dependence, meaning that it results in an independet sequence of violations. Independence can be tested by various tests, for example runs test or Ljung-Box test of Ljung & Box (1978). Under the null hypothesis, the probability of the next observation violating VaR is not influenced by the previous VaR violations. To test independence, Christoffersen (1998) employs a two-state Markov process and models I_t as a binary first-order Markov chain, with transition probability matrix

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

where $\pi_{ij} = Pr(I_t = j|I_{t-1} = i)$, i and j refer to states of violations/non-violations. Under the null hypothesis of independence, for the likelihood test statistic LR_{ind} it holds that

$$LR_{ind} = 2ln\left[\frac{(1-\widehat{\pi}_{01})^{n_{00}}\widehat{\pi}_{01}^{n_{01}}(1-\widehat{\pi}_{11})^{n_{10}}\widehat{\pi}_{11}^{n_{11}}}{(1-\widehat{\pi})^{n_{00}+n_{10}}\widehat{\pi}^{n_{01}+n_{11}}}\right] \overset{asy}{\sim} \chi^2_{(1)}, \tag{4.3}$$

where $n_{ij} = \sum_{t=1}^{n} \mathcal{I}(I_t = i|I_{t-1} = j)$ denotes the number of transitions from state i to state j and

$$\widehat{\pi}_{01} = \frac{n_{01}}{n_{00}+n_{01}} \;,\; \widehat{\pi}_{11} = \frac{n_{11}}{n_{10}+n_{11}} \text{ and } \widehat{\pi} = \frac{n_{01}+n_{11}}{n_{00}+n_{01}+n_{10}+n_{11}}.$$

Test of conditional coverage

Since Equation 4.3 does not account for correct unconditional coverage, Christoffersen (1998) proposes to combine Equation 4.2 and Equation 4.3 to test for a joint property $E[I_t|\mathcal{F}_{t-1}] = \alpha$, which is referred to as correct conditional coverage. Formally, under the null hypothesis:

$$LR_{cc} = LR_{uc} + LR_{ind} \overset{asy}{\sim} \chi^2_{(2)}. \tag{4.4}$$

In the thesis, we calculate the test statistics for both unconditional coverage and independence separately and jointly as conditional coverage. The reason is to detect which property is violated, if any, as the joint test has a reduced statistical power in detecting violation of a particular property.

Dynamic quantile test

Since the independence test of Christoffersen (1998) is limited to the first-order dependence, Engle & Manganelli (2004) propose a linear regression test for correct dynamic specification, the so-called Dynamic Quantile (DQ) test. This test is more strict as it enables to test for higher-order dependences in the sequence of violations. Its main advantage is the ability to incorporate a wide range of alternative specifications by making a simple extension. Following Engle & Manganelli (2004), we first define the $Hit_t(\theta)$ variable as de-meaned indicator function. Formally,

$$Hit_t(\theta) = I(y_t < q_t(\theta)) - \alpha, \tag{4.5}$$

where $Hit_t(\theta)$ takes on value $(1-\alpha)$ when y_t exceeds the quantile and value $-\alpha$ otherwise. The expected value of $Hit_t(\theta)$ equals zero as well as the conditional expectation of $Hit_t(\theta)$ given an information set at time $t-1$, based on the definition of the quantile equation. This property indicates that $Hit_t(\theta)$ has to be uncorrelated with its own lags and also with $q_t(\theta)$ and its expected value has to be zero. Every property mentioned above being satisfied, the hits will not be autocorrelated, the number of violations will be correct and there will be no measurement error.

Engle & Manganelli (2004) suggest to regress $Hit_t(\theta)$ on a set of explanatory variables, for example in the following form

$$Hit_t(\theta) = \alpha_0 + \sum_{i=1}^{m} w_i Hit_{t-i}(\theta) + w_{m+1} q_t(\theta) + u_t, \tag{4.6}$$

where w_i for $i = 1, .., m+1$ are the parameters to be estimated and $u_t = -\alpha$ with probability $1-\alpha$ or $1-\alpha$ with probability α. Following Kuester *et al.* (2006) notation, we have

$$\mathbf{H} - \alpha\mathbf{1} = \mathbf{X}\mathbf{w} + \mathbf{u}, \tag{4.7}$$

where $\mathbf{1}$ is a vector of ones and $w_0 = \alpha_0 - \alpha$.

Under the null hypothesis, $\alpha_0 = \alpha$, $w_i = 0$ for $i = 1, .., m+1$ and

$$\widehat{\mathbf{w}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}' \left(\mathbf{H} - \alpha\mathbf{1}\right) \overset{asy}{\sim} N\left(0, (\mathbf{X}'\mathbf{X})^{-1}\alpha(1-\alpha)\right), \tag{4.8}$$

from which Engle & Manganelli (2004) derive the test statistic

$$DQ = \frac{\widehat{\mathbf{w}}'\mathbf{X}'\mathbf{X}\widehat{\mathbf{w}}}{\alpha(1-\alpha)} \overset{asy}{\sim} \chi^2_{(m+2)}. \tag{4.9}$$

Engle & Manganelli (2004) derive two separate test statistics: in-sample DQ test and out-of-sample DQ test.[2] As suggested, the first one can serve for selection of a particular CAViaR model among various alternative specifications and the latter can be utilized by regulators or risk managers in order to assess the accuracy of the VaR estimates provided by financial institutions. The test is not dependent on the procedure used for estimation, it requires only a series of VaR forecasts and series of actual returns to obtain the result, so it is easily implemented in practice.

Berkowitz *et al.* (2011) modify the DQ test of Engle & Manganelli (2004) and refer to this test as the CAViaR test of Engle and Manganelli. They define the $Hit_t(\theta)$ as

$$Hit_t(\theta) = I(y_t < q_t(\theta)) \tag{4.10}$$

where $Hit_t(\theta)$ is equal to 1 when the quantile is exceeded and 0 otherwise. $Hit_t(\theta)|\mathcal{F}_{t-1} \overset{iid}{\sim} Ber(\alpha)$ if the conditional quantiles are correctly specified. As $Hit_t(\theta)$ is a binary variable, Berkowitz *et al.* (2011) suggest to estimate the following m-th order autoregression as a logit model

$$Hit_t(\theta) = c + \sum_{i=1}^{m} \beta_{1i} Hit_{t-i}(\theta) + \sum_{i=1}^{m} \beta_{2i} q_\alpha(y_{t-i+1}|\mathcal{F}_{t-i}) + u_t \tag{4.11}$$

where u_t is assumed to have logistic distribution. The authors set $m = 1$. Under the null hypothesis, β coefficients are zero and $Pr(Hit_t(\theta) = 1) = e^c/(1+e^c) = \alpha$. To test for the null hypothesis, the likelihood-ratio is used.

In our empirical analysis, we consider the original version of Engle & Manganelli (2004) to make our results comparable to the literature.

4.2 Evaluation of relative performance

Further, to compare the performance of competing models, we focus on the loss function as suggested by Giacomini & Komunjer (2005). Precisely, they employ so called 'tick' loss function $\mathcal{T}_\alpha$, which is the asymmetric linear loss

[2] For the detailed derivations of in-sample and out-of-sample distributions of the DQ test see their original paper.

function of order α defined as

$$\mathcal{T}_{\alpha}(e_{t+1}) = (\alpha - I(e_{t+1} < 0))\, e_{t+1}, \tag{4.12}$$

where $e_{t+1} = y_{t+1} - q_{\alpha}(y_{t+1}|\mathcal{F}_t)$ and $I(\cdot)$ denotes the indicator function. This particular form of the loss function is chosen as the object of the interest in their paper is the conditional α-quantile of the distribution of y_{t+1}. This is indeed the case of this thesis, therefore we make use of this functional form as well.

To test for equal predictive ability of competing models, we use Diebold & Mariano (1995) test, which is based on the loss differential. The loss differential is defined as

$$d = e_1^2 - e_2^2, \tag{4.13}$$

where e_1 is a vector of the forecast errors from the first model and e_2 is a vector of the forecast errors from the second model. In this thesis, as forecast errors we use loss as defined in Equation 4.12. In each case, we use the loss from the benchmark model as e_2. Under the null hypothesis $H_0 : E[d] = 0$, for the test statistic it holds that

$$DM = \frac{\overline{d}}{\sqrt{\frac{var(d)}{n}}} \overset{asy}{\sim} N(0,1), \tag{4.14}$$

where n is the number of observations, $\overline{d}$ and $var(d)$ are the unconditional mean and the unconditional variance of d respectively. In case of the multi-step-ahead forecasts, Diebold-Mariano test uses Newey-West estimator of the variance to account for autocorrelation of forecast errors. However, since we work only with the one-step-ahead forecasts, this does not concern us. Having concluded the theoretical part of this thesis, we now continue with the empirical research.

Chapter 5

Empirical application

After introducing the underlying theory of Value at Risk in previous chapters, we proceed with the most essential part of this thesis, our own empirical research. In this chapter we present empirical application of semiparametric CAViaR models and parametric GARCH models to calculate VaR on Central European, U.S. and German stock market data and consequently evaluate their performance using several backtesting methods, which we described in Chapter 4. In our empirical research, we investigate whether the introduction of dynamics to VaR modeling using CAViaR approach provides us with more accurate results, thus if its performance is superior to the commonly used parametric approach, focusing also on the crisis period.

We examine the performance of these approaches during two different time periods, the period before the crisis and the other covering the time of global financial crisis (GFC). In each period, the analysis is performed on two subsets of data, in-sample and out-of-sample, in order to obtain as reliable assessment of the performance of the models under study as possible. In all cases, we estimate VaR for two confidence levels, 95% and 99% as these are the most common in literature and also in practice.

It is important to point out that for both the parametric and the semiparametric method of VaR calculation, we consider static forecasting approach instead of rolling sample approach, which means that we keep the parameters of models constant throughout the period for which they are estimated.[1] The reason behind the static approach is very simple: appropriate model for fore-

[1] Our argumentation and choice of the estimation period length of 4 years and forecast period of one year is similar to Huang *et al.* (2010), who use the rolling sample of 1000 days and keep the parameters constant for the following 250 days. In our case, accounting for different characteristics of the data within the total time span of 10 years (2003 - 2012) is to some extent resolved by dividing the data into two different time periods.

casting should be designed in such a way that it is able to capture the dynamics of the data for some time before it is necessary to re-estimate it, which means that its parameters should not change very quickly and the model should be able to deliver satisfactory out-of-sample results. In practice, it is highly desired that the models do not need to be re-estimated very often as it may be very demanding, time-consuming or not feasible, especially when considering large portfolios. This is indeed the case of CAViaR models, which we investigate in this thesis.

This chapter is organized as follows. In the first section, we provide the reader with a thorough analysis of the corresponding market data. Secondly, we perform estimation procedures - we fit the GARCH model with normally and t-distributed errors to forecast volatility in both periods and after that we calculate VaR parametrically. As for the semiparametric CAViaR models, we calculate VaR directly. Further, we perform backtesting to evaluate the competing approaches. In the last section, we summarize the results and discuss the overall performance of our investigated models.

5.1 Data analysis

We will analyze the stock market data consisting of Central European, German and U.S. indices. For the purpose of this thesis, we use four different value-weighted indices: PX, BUX, DAX and S&P 500 representing Czech, Hungarian, German and U.S. stock markets respectively. The closing prices of the first two, representing Central European stock market, were obtained from Prague Stock Exchange and Budapest Stock Exchange. The closing prices of DAX and SP500, which serve as benchmarks, were obtained from Yahoo Finance.

For each index, the dataset is divided into two subsets, the first one covers the period of 2003 - 2007 and the second one the period of 2008 - 2012. Furthermore, each period is then split into two subsamples: in-sample consisting of observations from the first four years in the sample, and out-of-sample consisting of the remaining observations that cover the last year in each period. Overall, that gives us around 1000 observations in the in-sample and approximately 250 observations in the out-of-sample. In-sample will be used for estimations of unknown parameters for all models in question and also for forecast evaluation, while out-of-sample will be used just for forecast evaluation. The backtesting results referring to the out-of-sample will therefore be more conclusive and decisive about the investigated models.

In the first subset, models will be estimated on the stable period. The second subset will demonstrate how and if the time of economic turmoil impacts the performance of the models as this one covers the period of the 08/09 global financial crisis. This division will also allow us to investige the differences and regional characteristics across the stock markets under study.

First, we use the closing prices P_t to construct the series of daily percentage retuns as 100 times the difference of log of the prices: $y_t = 100 \times \ln(P_t/P_{t-1})$. These returns will then be employed for in-sample estimation of all VaR models under study. Plots of all closing prices and the corresponding percentage log returns can be found in Appendix B. As we can see on the plots, the second period is generally more volatile compared to the first one, due to the 08/09 GFC. Volatility clustering is clearly observable, especially around year 2009. This is also supported by descriptive statistics of log returns of all indices, which we present in Table 5.1 and Table 5.2 for the full samples. Descriptive statistics for in-sample and out-of-sample separately within each period can be found in Appendix A.

Table 5.1: Descriptive Statistics: full sample, first period

Statistics	PX	BUX	DAX	S&P 500
Observations	1256	1253	1273	1257
Mean	0.1084	0.0956	0.0750	0.0381
Median	0.1662	0.1285	0.12025	0.0808
Std. dev.	1.0801	1.2935	1.1954	0.8285
Minimum	-6.1250	-5.6027	-6.3360	-3.5867
Maximum	7.0482	4.8660	6.6446	3.4814
Skewness	-0.6145	-0.2203	-0.1935	-0.1749
Ex. kurtosis	4.5678	1.1008	3.2352	1.6893
Jarque-Bera test				
p-value	5.36335e-255	1.15079e-016	5.27328e-123	1.41377e-034
test statistic	1170.9600	73.4018	563.1110	155.8830
ADF test				
p-value	2.125e-011	5.641e-013	3.819e-012	3.6e-011
test statistic	-7.6804	-8.1610	-7.9113	-7.6082

Source: Author's computations.

In general, the second period exhibits higher maximum and minimum values for all indices, which results in higher standard deviations throughout the subset. Compared to the first period, standard deviations of indices are approximately twice higher, which indicates that the stocks are much more volatile

Table 5.2: Descriptive Statistics: full sample, second period

Statistics	PX	BUX	DAX	S&P 500
Observations	1257	1253	1281	1258
Mean	-0.0441	-0.0282	-0.0034	-0.0012
Median	-0.0088	-0.0353	0.0411	0.0702
Std. dev.	1.9007	2.0126	1.7242	1.6596
Minimum	-16.1850	-12.6490	-7.4335	-9.4695
Maximum	12.3600	13.1780	10.7970	10.9570
Skewness	-0.4311	-0.0025	0.1293	-0.2449
Ex. kurtosis	11.1420	5.5617	5.0155	6.9764
Jarque-Bera test				
p-value	0	0	4.71304e-293	0
test statistic	6540.95	1614.93	1346.21	2563.71
ADF test				
p-value	6.463e-005	1.248e-006	1.768e-007	1.038e-007
test statistic	-5.2234	-5.9870	-6.3250	-6.4136

Source: Author's computations.

here. According to unconditional standard deviation, S&P 500 appears to be generally less volatile than Central European indices.

In the first period, the returns of all indices are negatively skewed. That is also the case of the second period with the exception of DAX index, where the skewness is only slightly positive. Excess kurtosis always suggests higher peaks for all indices, especially in the second period, where its values are much higher compared to the first period (for PX more than twice higher, for S&P 500 more than three times higher). This adds to the assumption about non-normality of our time series in both time periods. Formally, we test the data for normality using the Jarque-Bera (JB) test. Based on the results, normality of daily returns is strongly rejected at all significance levels for all the indices in the first and also in the second period. Therefore, we can conclude that the distribution of daily returns is rather leptokurtic, which is in accord with the stylized facts about the financial data. As we work with the time series in our analysis, it is important to test the data for stationarity as well. This was done using Augmented Dickey-Fuller (ADF) test. Based on the p-values, the null hypothesis of unit root is rejected for all indices in both periods, therefore we consider all of our series to be stationary. Having described our stock market data, we proceed with the estimation procedures and VaR calculation.

5.2 Estimation and VaR calculation

Firstly, to estimate VaR parametrically, we estimate two GARCH models for volatility for all indices in both periods - GARCH with normally distributed errors and GARCH with Student t-distributed errors, to which we also refer in our analysis only as GARCH and GARCH-t respectively. As for the functional form of GARCH(p,q), we use the simplest model - GARCH (1,1) in all cases. Our decision is supported by Akaike Information Criteria (AIC), which we present in Appendix A. Even though the AIC is not the lowest in all cases for GARCH(1,1), the differences among various combinations of p and q, where $p, q = \{1, 2\}$, are insignificant. Therefore, we consider the parsimony rule and estimate the simplest model possible. For both GARCH(1,1) and GARCH(1,1)-t, the estimates were obtained using MLE. In case of GARCH with normal errors, we estimate α_0, α_1 and β_1 and for GARCH with t-distributed errors we report also estimate of the degrees of freedom ν. The results of the estimations for the first period are presented in Table 5.3 and for the second one in Table 5.4.

Apart from the constant, all the parameters are always significant at 5% significance level and very similar for both models within each period. Volatility clustering seems to be well captured. Based on the estimated parameter for lagged conditional variance, it is obvious that this term has the strongest impact on volatility forecast as the value of β_1 ranges approximately from 0.83 to 0.93 across indices in the first period and from 0.82 to 0.9 in the second period. As for the estimated degrees of freedom of error distribution in GARCH-t, in the first period the values of ν vary from 5 to 15 for PX, BUX and DAX, which supports the assumption about the heavy tailedness of the distribution. S&P 500 is presented with the value of 31, which is closest to the normal distribution among all indices. This could suggest that in this particular case there might not be a distinct difference between GARCH and GARCH-t in the sense of VaR estimation. In the second period, the estimated values of ν are approximately around 8 for PX, BUX and DAX, suggesting a bit heavier tails for BUX and DAX compared to the pre-crisis period. Interestingly, for PX this result suggests that the distribution is a bit closer to normal here as the value of ν was previously somewhat lower. The result for S&P 500 presents quite a different picture compared to that obtained in the first period, since the value of ν is considerably lower here ($\nu = 6$), which indicates much heavier tails in this time horizon. In general, the results regarding the degrees of freedom are not surprising, due to the presence of more extreme observations, especially

Table 5.3: MLE parameter estimates for GARCH and GARCH-t, first period

The left-hand side panel reports results for GARCH(1,1) with normally distributed errors (GARCH) and the right-hand side panel reports results for GARCH(1,1) with t-distributed errors (GARCH-t). Estimated parameters are reported with standard errors (s.e) and p-values. Significant coefficients at 5 % confidence level are formatted in bold. For each model we report the value of log-likelihood function (LL) and Bayesian Information Criterion (BIC).

	GARCH				**GARCH-t**			
	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
α_0	**0.0639**	**0.0714**	**0.0173**	0.0079	**0.0645**	**0.0706**	**0.0141**	0.0075
s.e.	0.0211	0.0239	0.0070	0.0042	0.0204	0.0241	0.0067	0.0042
p-value	0.0025	0.0029	0.0136	0.0608	0.0016	0.0035	0.0346	0.0776
α_1	**0.1098**	**0.0771**	**0.0767**	**0.0499**	**0.1076**	**0.0721**	**0.0744**	**0.0493**
s.e.	0.0290	0.0168	0.0155	0.0123	0.0255	0.0162	0.0152	0.0124
p-value	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	**0.8372**	**0.8819**	**0.9086**	**0.9347**	**0.8389**	**0.8869**	**0.9144**	**0.9362**
s.e.	0.0293	0.0219	0.0165	0.0156	0.0299	0.0229	0.0157	0.0156
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ν					**5.2100**	**15.9940**	**10.4029**	31.3391
s.e.					0.7607	6.9305	3.2816	27.3235
p-value					0.0000	0.0210	0.0015	0.2514
LL	-1432.99	-1671.11	-1497.83	-1110.15	-1392.53	-1667.60	-1492.11	-1109.60
BIC	2893.65	3369.88	3023.39	2247.95	2819.64	3369.79	3018.87	2253.77

Source: Author's computations.

Table 5.4: MLE parameter estimates for GARCH and GARCH-t, second period

The left-hand side panel reports results for GARCH(1,1) with normally distributed errors (GARCH) and the right-hand side panel reports results for GARCH(1,1) with t-distributed errors (GARCH-t). Estimated parameters are reported with standard errors (s.e) and p-values. Significant coefficients at 5 % confidence level are formatted in bold. For each model we report the value of log-likelihood function (LL) and Bayesian Information Criterion (BIC).

	GARCH				**GARCH-t**			
	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
α_0	**0.0426**	**0.0936**	**0.0280**	**0.0266**	**0.0477**	**0.0669**	**0.0300**	0.0186
s.e.	0.0207	0.0465	0.0136	0.0109	0.0205	0.0333	0.0152	0.0106
p-value	0.0393	0.0442	0.0387	0.0150	0.0198	0.0443	0.0473	0.0811
α_1	**0.1732**	**0.1222**	**0.0889**	**0.1103**	**0.1454**	**0.1122**	**0.0861**	**0.1051**
s.e.	0.0353	0.0304	0.0200	0.0201	0.0315	0.0270	0.0196	0.0205
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	**0.8281**	**0.8591**	**0.9043**	**0.8831**	**0.8469**	**0.8761**	**0.9073**	**0.8952**
s.e.	0.0285	0.0310	0.0193	0.0167	0.0284	0.0269	0.0187	0.0167
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ν					**8.6265**	**8.7602**	**7.9015**	**6.1212**
s.e.					1.8819	2.2041	2.1129	1.2244
p-value					0.0000	0.0000	0.0002	0.0000
LL	-1876.02	-2057.24	-1909.56	-1775.83	-1863.95	-2045.36	-1898.55	-1761.51
BIC	3779.71	4142.15	3846.87	3579.33	3762.48	4125.31	3831.78	3557.60

Source: Author's computations.

around year 2009.

Further, we report the value of log-likelihood function (LL) and Bayesian Information Criterion (BIC), which allow us to assess the relative fit of GARCH and GARCH-t. Our results suggest that GARCH-t is uniformly a better fit, for all the indices in both first and second period. The only exception is observed for S&P 500 in the first period, where GARCH appears to a better fit.

Having estimated the parameters and assessed the relative fit of the models, we proceed with the volatility forecast. As we have already mentioned at the beginning of this chapter, for both periods we do not allow the parameters to change over time and use the estimates for the corresponding period. After obtaining the standard deviation forecasts from both GARCH and GARCH-t, we calculatee one-step-ahead parametric VaR for both 95% and 99% confidence levels, according to Equation 2.3 provided in Chapter 2.

Turning to the semiparametric CAViaR models, we employ the estimation procedure of Engle & Manganelli (2004) and focus on modeling VaR as 1% and 5% quantile of the return distribution directly.[2] To compute series of one-step-ahead VaR forecasts, we follow Engle & Manganelli (2004) and set $q_1(\theta)$ to the empirical quantile of the first 300 observations. In case of SAV, we estimate the constant β_1, the lagged conditional quantile coefficient β_2 and the lagged absolute return parameter β_3. For AS, we also report an estimate of the lagged negative return coefficient β_4. The estimated parameters for both specifications for the first period are presented in Table 5.5 and Table 5.6 for the second period, where the upper panel reports the results for 1% quantile and the lower panel reports the results for 5% quantile.

In analogy with Engle & Manganelli (2004) or Kouretas & Zarangas (2005), we obtain qualitatively similar results regarding the coefficient of autoregressive term, β_2. The parameter is close to one and highly statistically significant in case of 99% and also 95% VaR, across all indices in both time periods, which suggests that the well known fact about volatility clustering is also revelant when focusing on the tails of the return distribution. The value of the parameter ranges from 0.84 to 0.96 for both 99% and 95% VaR in the first period, somewhat lower values are only demostrated in case of AS for PX. In the second period, the results come out very similar, only the range is slightly wider, the coefficient value varies between 0.8 and 0.97 across indices, model specifications and confidence levels. The conditional quantile parameter exhibits much higher values compared to the parameters corresponding to the other

[2]We are gratuful to Simone Manganelli for providing the Matlab code.

Table 5.5: Estimation results for CAViaR, first period

The left-hand side panel reports results for SAV and the right-hand side panel reports results for AS. Estimated parameters are reported with standard errors (s.e) and p-values, RQ is the value of regression quantile objective function. Significant coefficients at 5 % confidence level are formatted in bold.

	Symmetric Absolute Value				**Asymmetric Slope**			
	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
99% VaR								
β_1	0.3245	0.2107	0.0621	**0.0687**	**1.1073**	-0.0322	0.0592	0.0190
s.e.	0.3536	0.2108	0.0545	0.0154	0.2724	0.1012	0.0415	0.0199
p-value	0.1794	0.1588	0.1274	0.0000	0.0000	0.3751	0.0768	0.1697
β_2	**0.8496**	**0.8779**	**0.9624**	**0.9400**	**0.5630**	**0.8876**	**0.9307**	**0.9495**
s.e.	0.1302	0.0878	0.0475	0.0212	0.1830	0.0278	0.0369	0.0264
p-value	0.0000	0.0000	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000
β_3	0.3999	**0.2592**	0.0988	**0.1865**	**-0.5592**	**0.4513**	0.1394	0.1241
s.e.	0.2525	0.0704	0.0944	0.0532	0.0610	0.1653	0.1119	0.0831
p-value	0.0566	0.0001	0.1477	0.0002	0.0000	0.0032	0.1063	0.0678
β_4					1.2000	**0.3089**	**0.1841**	**0.1179**
s.e.					1.3709	0.1296	0.0930	0.0611
p-value					0.1907	0.0086	0.0238	0.0268
RQ	40.4641	40.9730	34.3095	19.9587	34.7222	40.0865	34.0075	20.4480
95% VaR								
β_1	**0.0900**	**0.1684**	**0.0550**	**0.1449**	0.3114	0.1635	**0.0365**	0.0166
s.e.	0.0361	0.0564	0.0169	0.0372	0.1947	0.0996	0.0132	0.0229
p-value	0.0063	0.0014	0.0006	0.0000	0.0548	0.0504	0.0030	0.2347
β_2	**0.8889**	**0.8571**	**0.9383**	**0.8483**	**0.7135**	**0.8162**	**0.9505**	**0.9665**
s.e.	0.0364	0.0512	0.0249	0.0553	0.1336	0.0860	0.0224	0.0391
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_3	**0.2022**	**0.1968**	**0.1162**	**0.1862**	-0.0712	0.1227	-0.0189	0.0107
s.e.	0.0832	0.0580	0.0573	0.0742	0.1097	0.0964	0.0429	0.0579
p-value	0.0076	0.0003	0.0212	0.0060	0.2582	0.1016	0.3296	0.4267
β_4					**0.5567**	**0.2890**	**0.1566**	**0.0705**
s.e.					0.2139	0.1145	0.0568	0.0347
p-value					0.0046	0.0058	0.0029	0.0210
RQ	129.6511	139.4371	132.6544	79.2464	123.9350	137.8244	129.9458	79.3107

Source: Author's computations.

Table 5.6: Estimation results for CAViaR, second period

The left-hand side panel reports results for SAV and the right-hand side panel reports results for AS. Estimated parameters are reported with standard errors (s.e) and p-values, RQ is the value of regression quantile objective function. Significant coefficients at 5 % confidence level are formatted in bold.

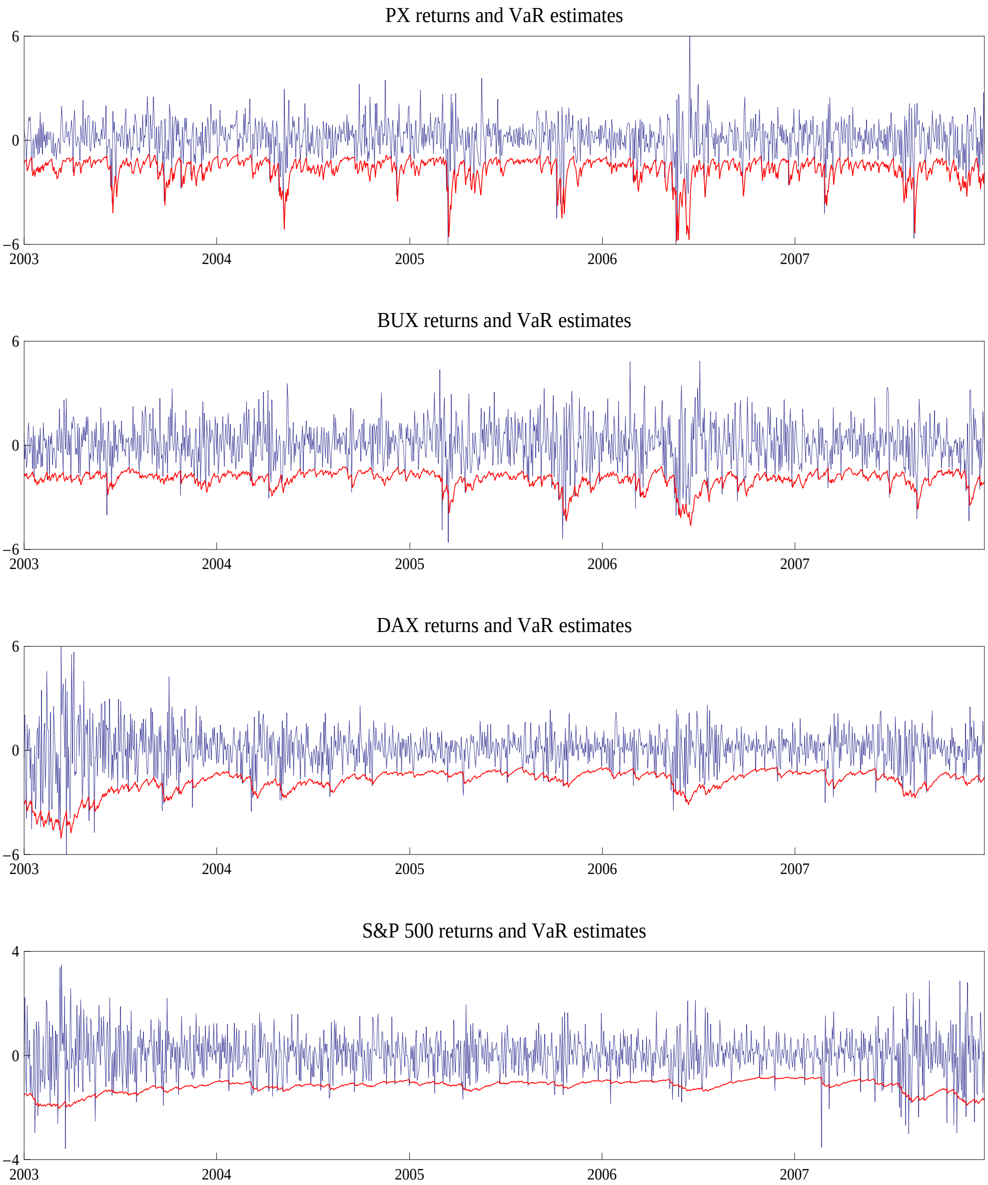
	Symmetric Absolute Value				**Asymmetric Slope**			
	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
99% VaR								
β_1	**0.4648**	**0.4337**	**0.1901**	**0.2556**	0.0965	0.1583	**0.0410**	**0.1351**
s.e.	0.0846	0.1680	0.0642	0.0596	0.1585	0.1590	0.0201	0.0496
p-value	0.0000	0.0049	0.0015	0.0000	0.2714	0.1596	0.0205	0.0032
β_2	**0.7997**	**0.8370**	**0.8988**	**0.8801**	**0.8535**	**0.8702**	**0.9699**	**0.8981**
s.e.	0.0643	0.0422	0.0270	0.0300	0.0662	0.0617	0.0156	0.0506
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_3	**0.5182**	**0.4155**	**0.2638**	**0.2740**	**0.2465**	0.1689	-0.0443	0.0683
s.e.	0.2449	0.1113	0.0410	0.0604	0.1455	0.1519	0.0536	0.1623
p-value	0.0172	0.0001	0.0000	0.0000	0.0451	0.1332	0.2039	0.3368
β_4					**0.6914**	**0.4998**	**0.1705**	**0.3941**
s.e.					0.3288	0.1486	0.0248	0.2283
p-value					0.0177	0.0004	0.0000	0.0422
RQ	63.3490	61.7005	49.9293	50.7525	58.7019	58.2166	47.0342	49.2272
95% VaR								
β_1	0.1544	**0.1046**	**0.1860**	**0.1607**	0.0871	0.0202	0.0831	**0.0882**
s.e.	0.0976	0.0252	0.0542	0.0654	0.1727	0.0413	0.0706	0.0274
p-value	0.0569	0.0000	0.0003	0.0070	0.3071	0.3123	0.1198	0.0006
β_2	**0.8606**	**0.8967**	**0.8695**	**0.8856**	**0.8373**	**0.8899**	**0.8913**	**0.9178**
s.e.	0.1045	0.0306	0.0499	0.0381	0.1308	0.0232	0.0620	0.0221
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_3	0.2657	**0.2199**	**0.2500**	**0.2138**	0.1605	**0.1348**	0.0745	-0.0018
s.e.	0.2001	0.0760	0.1106	0.0651	0.1558	0.0808	0.1175	0.0591
p-value	0.0922	0.0019	0.0119	0.0005	0.1515	0.0476	0.2631	0.4878
β_4					**0.3859**	**0.2796**	**0.2982**	**0.2269**
s.e.					0.1950	0.0598	0.0867	0.0534
p-value					0.0239	0.0000	0.0003	0.0000
RQ	205.1004	217.5601	193.9097	188.4962	201.1709	215.7554	186.7689	184.1124

Source: Author's computations.

terms in CAViaR regression equations and hence impacts the future quantile most substantially. Another similarity to the previous empirical research refers to the parameter of lagged absolute return, β_3, in SAV specification and the parameter corresponding to the lagged negative return, β_4, in AS specification. The parameter of lagged absolute return comes out relatively smaller compared to β_2, but it also is highly significant in most cases. However, when focusing on the coefficient β_3 in AS specification, where this only refers to the positive part of the lagged returns, the results present a different picture. We find that the parameter of the lagged negative return, β_4, is always strongly statistically significant (with only one exception), while the coefficient associated with positive returns (β_3) in many cases comes out insignificant. This implies that the impact of the lagged returns on VaR forecasts is often asymmetric. Figure 5.1 depicts one-step-ahead VaR forecasts of AS specification for the first (pre-crisis) period and Figure 5.2 for the second (crisis) period. Figures for SAV specification can be found in Appendix B.

Figure 5.1: AS CAViaR at 95% confidence level, first period

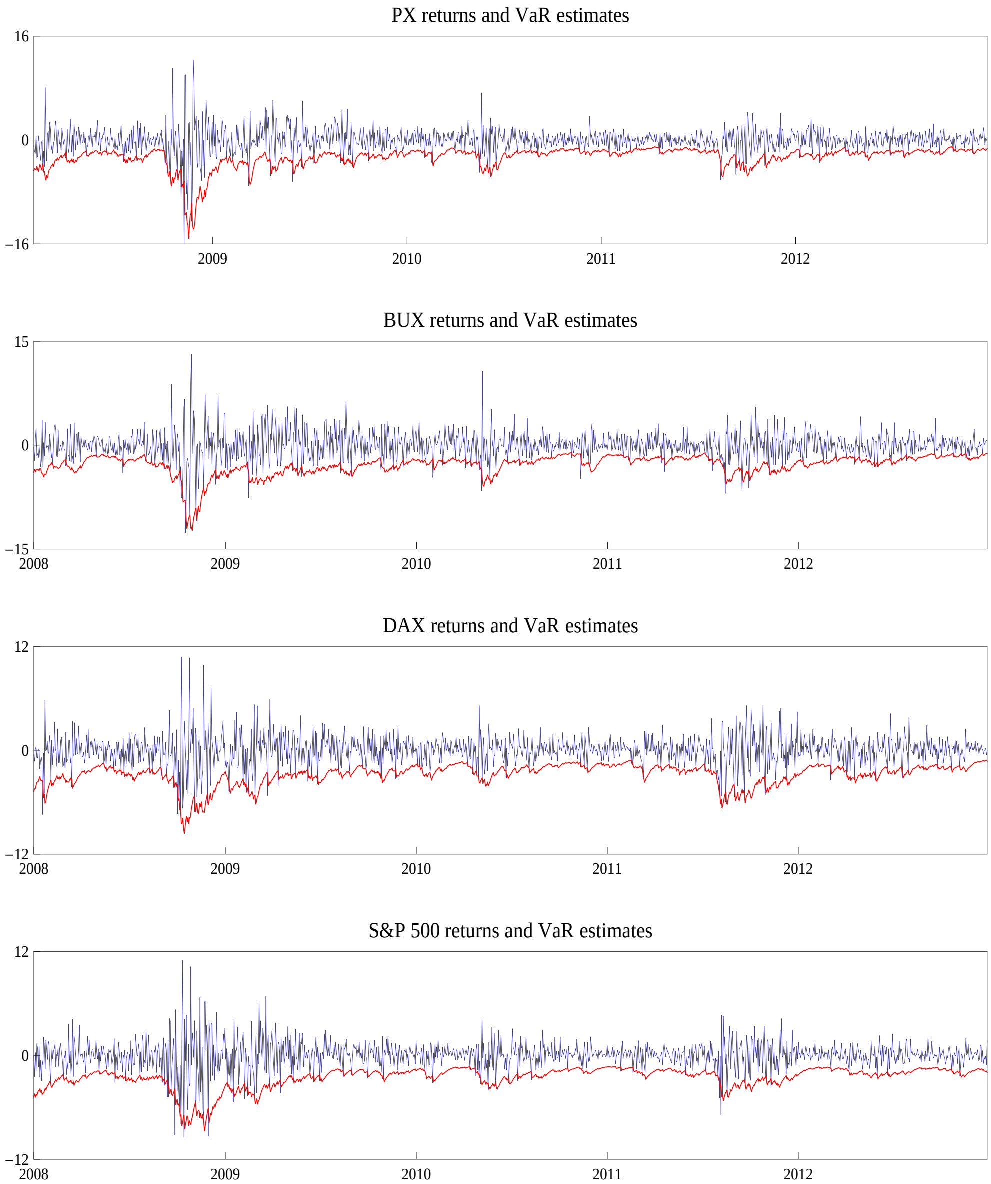
Returns and one-step-ahead VaR forecasts of Asymmetric Slope CAViaR at 95% confidence level. In-sample covers the period of 01/2003 - 12/2006, out-of-sample the year of 2007.



Source: Author's computations.

Figure 5.2: AS CAViaR at 95% confidence level, second period

Returns and one-step-ahead VaR forecasts of Asymmetric Slope CAViaR at 95% confidence level. In-sample covers the period of 01/2008 - 12/2011, out-of-sample the year of 2012.



Source: Author's computations.

5.3 Evaluation of performance

Having discussed the results from both parametric and semiparametric estimations, we proceed with the evaluation of in-sample and out-of-sample performance of all four models for VaR calculation under study. The results are reported separately for the first (pre-crisis) period and the second (crisis) period. To save space, in the text we only present the tables with backtesting results for PX as a representative of Central European market, which is the main focus of this thesis. Tables with the results for BUX and other indices used as benchmarks can be found in Appendix A.

For each VaR and model, we evaluate the absolute performance using the methods characterized in Section 4.1 and report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$) and p-value of the dynamic quantile (DQ) test of Engle & Manganelli (2004) for correct dynamic specification ($p - DQ$). Similarly to Kuester *et al.* (2006) or Huang *et al.* (2010), we use such a specification of DQ test, where the regression matrix $\mathbf{X}$ contains constant and four lagged hits, $Hit_{t-1}(\theta), .., Hit_{t-4}(\theta)$. Furthermore, we evaluate the relative performance of the models as described in Section 4.2 and report the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normally distributed errors serving as the benchmark. For Diebold-Mariano test, we consider two-sided test and 5% significance level. We therefore reject the null hypothesis of equal predictive ability of models when the absolute value of DM test statistic is higher than 1.96. To choose the best performing model, we first consider the absolute performance indicators and after that we decide possibly between more models based on the results of the relative performance indicators.

5.3.1 First (pre-crisis) period

In-sample fit

Starting with PX index, for which we report the backtesting results in the left-hand side panel of Table 5.7, we find that both CAViaR specifications perform very well, having the proportion of violations very close to the nominal levels in case of 99% and 95% VaR as well. On the other hand, in terms of unconditional coverage our benchmark GARCH models do not deliver such accurate results. Compared to GARCH models, the unconditional coverage of CAViaR is especially better for 95% VaR. For 99% confidence level, GARCH seems to underestime VaR, while GARCH-t appears to do just the opposite, as could be expected since we move further in the left tail of the distribution.

Table 5.7: Backtesting for PX, first period

Absolute and relative performance of GARCH and CAViaR models for daily PX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	2.0854	0.5958	1.0924	1.1917	0.8032	0.8032	0.8032	1.2048
$p - LR_{uc}$	0.0025	0.1633	0.7716	0.5530	0.7466	0.7466	0.7466	0.7530
$p - LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.0104	0.3785	0.9587	0.8386	0.9491	0.9491	0.9491	0.9517
$p - DQ$	0.0192	0.0025	0.0009	0.0001	0.7602	0.7602	0.7602	0.2170
$\mathcal{T}_\alpha$	0.0436	0.0404	0.0402	0.0345	0.0437	0.0449	0.0432	0.0436
DM		-2.3602	-2.5496	-2.0368		-1.3462	-1.3611	-1.1170
95% VaR								
$PoF(\%)$	3.7736	2.6812	4.9652	4.8659	4.8193	2.0080	8.4337	5.2209
$p - LR_{uc}$	0.0624	0.0002	0.9596	0.8446	0.8953	0.0142	0.0230	0.8738
$p - LR_{ind}$	0.6425	1.0000	0.7385	0.2905	0.5964	0.0761	0.4928	1.0000
$p - LR_{cc}$	0.1583	0.0011	0.9446	0.5611	0.8617	0.0103	0.0596	0.9874
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0003	0.0000	0.0881	0.0185
$\mathcal{T}_\alpha$	0.1296	0.1352	0.1287	0.1231	0.1233	0.1293	0.1294	0.1239
DM		-3.2434	3.6771	-1.6672		-1.3570	2.4243	-0.7798

Source: Author's computations.

In case of 95% VaR, for normal GARCH the proportion of violations is less than 4 %. GARCH-t produced around 2.7 % violations, not passing the unconditional coverage test. Based on the p-values of independence test, all of the models considered manage to produce cluster-free violations. Evaluating the models more generally with conditional coverage test, we find that all models for both levels pass the test at 1% siginificance level, except for GARCH-t for 95% VaR, whereas the normal GARCH for 99% VaR passes the test only marginally. However, the DQ test clearly rejects the null hypothesis of correct dynamic specification for all the models at 1% significance level, excluding 99% GARCH VaR. This could be attributed to the fact that unlike conditional coverage test, this one tests the sequence of violations for much higher-order dependencies. Having assessed the absolute performance, we conclude that both SAV and AS are the most precise by the definition and turn to the evaluation of the relative performance. While CAViaR models manage to be correct on average, as the results show they also produce the smallest loss for both levels and provide statistically more accurate forecasts for 99% VaR at 5% significance level, as indicated by DM test statistic. In terms of unconditional coverage and loss, AS appears to be the best model for PX.

The results for BUX are summarized in Table A.7. In this case, we observe very similar performance of all models for 95 % VaR as for PX. CAViaR models again outperform GARCH and GARCH-t, delivering the number of violations very close to 5 % and comfortably passing the unconditional coverage test. As for normal GARCH, this passes the test only marginally and GARCH-t is strongly rejected. Focusing on the 99% VaR, all the models seem to be correct on average according to the p-value of the unconditional coverage test. However, the results come out best for GARCH and SAV, having the proportion of failures almost equal to 1 %. While the independence test suggests independent sequence of violations for all models, the DQ test indicates dynamic misspecification. To decide between GARCH and SAV for 99% VaR, we look at the value of the tick-loss function, which is slightly lower for GARCH. Also DM test statistic suggests superior performance of GARCH, therefore we consider this one to be the best fit. As for 95% VaR, the smallest average loss is produced by AS, yet statistically indistinguishable from benchmark GARCH model. We consider AS to be the best model as it delivers the proportion of violations closest to 5 % and lowest tick-loss at the same time.

Turning to DAX index, we report the corresponding results in Table A.8. All the models considered perform adequately at all levels with regard to unconditional coverage, comfortably passing the unconditional coverage, independence and conditional coverage test. Only GARCH seems to underestimate VaR for 99% confidence level and GARCH-t somewhat overestimate it for 95 % confidence level. DQ test again signals significant misspecification almost for all cases, except for 99% GARCH VaR. In terms of relative performance, we find that compared to GARCH models, CAViaR specifications deliver the lower tick-loss for both confidence levels, with AS being the absolute winner again for 95% VaR and 99% as well, where its superior performance is even supported with the DM test statistic.

Finally, we summarize the backtesting results for S&P 500 in Table A.9. Generally, both GARCH and CAViaR models do a good job describing the evolution of the left tail for both levels. Nevertheless, CAViaR specifications seem to be more accurate as their proportion of failures is closer to either 1 % or 5%, depending on the corresponding VaR confidence level. Turning to the information in the sequence of violations, all the models for both levels manage to produce cluster-free violations, as indicated by the p-value of the independence test. However, the DQ test suggests dynamic misspecifiaction as the null hypothesis is rejected at 1% significance level in most cases. Only 99% GARCH, GARCH-t and AS do not seem to suffer from any misspecification. Having discussed the absolute performance, we now proceed with the evaluation of the relative performance. Here we find that CAViaR specifications appear to perform on par with GARCH in a statistical sense (DM test statistic), though they deliver lower values of the tick-loss function. Both SAV and AS come out similarly for 99% and 95% VaR, however, in case of S&P 500 index SAV specification seems to be more precise for a change.

Out-of-sample performance

Now we assess the accuracy of models out-of-sample. The backtesting results are reported in the right-hand side panel of Table 5.7 for PX, Table A.7 for BUX, Table A.8 for DAX and Table A.9 for S&P 500.

Turnig to PX first, we find that all the models under study deliver similar results for 99% VaR in terms of the absolute performance, producing the proportion of violations close to 1 % and all being accepted by unconditional coverage, independence, conditional coverage and even DQ test. As for the relative performance, DM test statistic does not indicate any statistical difference from the benchmark GARCH. However, judging by the value of the tick-loss function, CAViaR specification SAV appears to be slightly more precise compared to AS and GARCH. Turning to the 5% quantile, we observe different performance. Two models come out most accurate by the definition

- GARCH with normally distributed erros and CAViaR specification AS with PoF very close to the nominal level. GARCH-t clearly overestimates the risk, producing around 2 % of violations, whereas SAV tends to underestimate it as the proportion of violations reaches almost 8.5 %. Despite that, even these two models pass the unconditional coverage test. The sequences of violations appear to be cluster-free, all passing the independence test. In case of GARCH type models, DQ test indicates some misspecification, CAViaR specifications seem to be correctly specified at 1% significance level. In terms of loss, GARCH model appears to be best choice, though the difference in comparison with AS is statistically insignificant (DM test statistic).

In case of BUX index, GARCH-t seems to perform best for the 99% VaR, having the proportion of failures closest to the nominal level, though GARCH and SAV perform quite good as well. AS does not seem to do a good job and tends to underestimate VaR as it delivers nearly 3 % of violations, almost being rejected by the unconditional coverage test at 1% significance level. Moving to the 95% VaR, the most precise unconditional coverage is observed for SAV, delivering around 4.5 % of violations and the least precise for GARCH-t with 3 % of violations, but still passing the unconditional coverage test. Turning to the independence property of the sequence of violations, this is not rejected by the indepence test for all the models at both levels. According to the conditional coverage test, all the models are correct, though AS for 99% VaR passes only marginally. However, the DQ test rejects the correct specification in all cases. Assessing the relative performance, we find that for 99% VaR GARCH-t produces the lowest value of the tick-loss function, therefore we consider it the best choice. As for 95% VaR, CAViaR models perform better in terms of loss, though not statistically different from the benchmark. By the definition, SAV appears to be the best fit as it delivers the most precise unconditional coverage and yet produces quite low tick-loss.

As for DAX index, all the models perform very well for 99% VaR with regard to the unconditional coverage, independence, conditional coverage and dynamic specification. For 95% VaR, all perform adequately, with GARCH-t slightly overestimating and SAV underestimating VaR, but passing the unconditional coverage test and satisfying the independence property according to the independence test. Nevertheless, the most accurate proportion of failures (approximately 5.2 %) is delivered by GARCH, followed by AS with almost 4.4 % of violations. To choose the best fit for both levels, we look at the relative performance statistics. For 99% VaR, forecasts are statistically indistinguishable (DM test statistic) compared to the benchmark GARCH model, though AS delivers the lowest value of the tick-loss. For 95% VaR, our conclusions are very similar. We find GARCH and AS to be the best performing models in terms of average loss, providing similar forecasts in a statistival sense.

Results for S&P 500 present quite a different picture. All the models provide very poor performance out-of-sample, resulting in too many VaR violations and noticeably underestimating VaR. The results are especially inferior further in the left tail (99% VaR), where the proportion of failures is approximately four times higher than the nominal level. Here all the models are strongly rejected by all the tests. In case of 95% VaR, GARCH, GARCH-t and SAV pass at 1% significance level, only AS is formally rejected by the unconditional coverage test (producing 10.4 % of violations). Possibly the best fit here is GARCH-t with PoF of 6.8 %, also producing the lowest value of the tick-loss function and performing superiorly compared to the benchmark. Considerably worse out-of-sample performance of both GARCH type and CAViaR models for S&P 500 in comparison with other indices can be probably explained by the earlier kick-off of the GFC in the U.S., resulting in more extremely negative returns and hence causing more violations.

5.3.2 Second (crisis) period

In-sample fit

Now we evaluate the performance of our investigated models in the crisis period and proceed in the same manner as for the first (pre-crisis) period. Starting with PX index again, the backtesting results are summarized in Table 5.8. Similarly to the first period, we find that both CAViaR specifications do a very good job describing the evolution of the left tail, having the proportion of violations very close to either 1 % or 5 %, depending on the corresponding confidence level. For 99% VaR, GARCH-t seems to slightly overestimate the risk, while GARCH with normally distributed errors appears to do just the opposite, but this is again not surprising given the fact that we focus on the more extreme quantile of the distribution. Still, all the models pass the conditional coverage test, hence satisfying jointly the unconditional coverage and independence properties. Though, the DQ test suggests dynamic misspecification for all of them except for GARCH. In case of the 95% VaR, we observe the same unconditional coverage for GARCH, SAV and AS (approximately 5 % of violations). GARCH-t delivers around 3.7 % of violations, thus overstating the risk, but passing the unconditional coverage test. However, turning to the indepence property of the sequence of violations, only AS manages to produce cluster-free violations, whereas the rest of the models are rejected at 5% significance level.

Table 5.8: Backtesting for PX, second period

Absolute and relative performance of GARCH and CAViaR models for daily PX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark. NA is reported in case of no violation.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	1.5889	0.7944	0.9930	0.9930	2.8000	0.0000	0.0000	0.8000
$p - LR_{uc}$	0.0837	0.4965	0.9823	0.9823	0.0190	NA	NA	0.7419
$p - LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	NA	NA	1.0000
$p - LR_{cc}$	0.2240	0.7935	0.9997	0.9997	0.0640	NA	NA	0.9474
$p - DQ$	0.5392	0.0002	0.0010	0.0010	0.0796	NA	NA	0.5523
$\mathcal{T}_\alpha$	0.0634	0.0608	0.0629	0.0583	0.0322	0.0328	0.0339	0.0333
DM		-1.6536	0.5133	-1.5168		-1.3360	-1.2283	-1.2693
95% VaR								
$PoF(\%)$	5.0645	3.6743	5.0645	5.0645	4.4000	4.0000	4.4000	4.4000
$p - LR_{uc}$	0.9253	0.0432	0.9253	0.9253	0.6571	0.4529	0.6571	0.6571
$p - LR_{ind}$	0.0146	0.0107	0.0007	0.1555	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.0504	0.0050	0.0033	0.3631	0.9061	0.7545	0.9061	0.9061
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0025	0.0008	0.0025	0.0025
$\mathcal{T}_\alpha$	0.2015	0.2039	0.2037	0.1998	0.1255	0.1272	0.1230	0.1264
DM		-2.5252	1.4902	0.2991		-2.3007	-1.3575	0.4972

Source: Author's computations.

Evaluating the models more generally with conditional coverage test, apart from AS, GARCH model passes as well, but only marginally. Not surprisingly, the DQ test for correct specification rejects all the models. Moving on to the relative performance, we find AS to be the best performing model as it produces the lowest value of the tick-loss function at both levels, though DM test statistic does not indicate statistically significant difference in forecast accuracy compared to the benchmark. In spite of that, we consider it the best fit given the very precise unconditional coverage and independent violations (as the only model for 95% confidence level).

Now we turn to BUX index, for which we present the results in Table A.10. Here we observe very similar pattern as for PX, with CAViaR models being more precise in terms of unconditional coverage at both levels. For 99% VaR, GARCH and GARCH-t exhibit the same tendency, the former slightly understating and the latter overstating the risk, but along with CAViaR comfortably passing unconditional coverage, independence, conditional coverage and even DQ test for correct specification (except for GARCH-t). Focusing on the 95% VaR, the proportions of failures are very accurate for GARCH, SAV and AS, with GARCH-t slightly overestimating VaR with almost 3.5 % of violations. All the models satisfy the correct conditional coverage as suggested by the p-values of the test. However, the DQ test rejects the correct dynamic specification for all of them. To decide between the models with the best absolute performance, we look at the the value of the tick-loss function and DM test statistic. As for 99% CAViaR, the average loss is more in favour of AS, whereas SAV manages to deliver quite low loss as well while having the most precise unconditional coverage at the same time. As suggested by DM test, performance of both CAViaR specifications is statistically superior to the benchmark, therefore we consider both of them to be a very good fit for BUX at this level. Even though the CAViaR forecasts do not seem to be different from the benchmark in a statistical sense at 95% confidence level, the results referring to the average loss come out similarly, with AS appearing to the best choice here.

Moving on to DAX index, we report the corresponding results in Table A.11. CAViaR models do a good job at both levels with regard to the unconditional coverage, passing also the independence, conditional coverage test and DQ test at 1% significance level. The forecasting tendencies of GARCH type models observed in case of PX and BUX are apparent also here, especially for GARCH-t (resulting in 0.3 % of violations). As for 95% VaR, all the models deliver similar proportion of failures, coming close to 5 %, except for GARCH. This produces 6.5 % of violations, resulting in the underestimatoin of the underlying risk and not passing the conditional coverage test. DQ test again signals significant misspecification for all the models at this level. In terms of the relative performance, we find that CAViaR specifications deliver the lower tick-loss for both confidence levels, also providing statistically more accurate forecasts compared to the benchmark for 5% quantile (DM test statistic), with AS to be producing absolutely the lowest value of the tick-loss function.

Last but not least, we summarize the results for S&P 500 in Table A.12. In this case, we come to the same conclusions as before. CAViaR models perform very well at both levels with regard to the unconditional coverage. GARCH type models follow similar pattern as for PX, BUX and DAX. Turning to the information in the sequence of violations, all the models for both confidence levels manage to produce cluster-free violations, as indicated by the p-value of the independence test. Apart from 99% GARCH VaR, all the models also have the correct conditional coverage. However, the DQ test suggests dynamic misspecifiaction as the null hypothesis is rejected at 5% significance level in all cases. Having discussed the absolute performance, we now proceed with the evalution of the relative performance, where we try to decide for a possible winner. Again, we observe the best performance of CAViaR specifications in terms of the average loss at both levels, especially for 1% quantile, with AS proving to be the best choice for both 99% and 95% VaR, even though in the second case there is no significant difference from the benchmark.

Out-of-sample performance

As before, we assess the accuracy of the models also out-of-sample. The back-testing results are reported in the right-hand side panel of Table 5.8 for PX, Table A.10 for BUX, Table A.11 for DAX and Table A.12 for S&P 500.

Turnig to PX first, we find quite different results for the 1% quantile compared to the pre-crisis period. Two of four models, namely GARCH-t and SAV, fail to produce any violation and hence are rejected by all the tests referring to the absolute performance. GARCH clearly understates the risk, resulting in too many VaR violations (2.8 %), but still marginally passing the unconditional coverage test at 1% significance level. The only model producing proportion of failures closest to the nominal level is AS, with 0.8 % of violations, having also the correct condtitional coverage and dynamic specification. For 95% VaR, all the models perform adequately, delivering cluster-free volations and PoF around 4 %, all unambiguously passing the more general conditional coverage test. Though, according to the DQ test, the models do not seem to be correctly specified at this level. To choose between three possibilies with the most precise conditional coverage, we turn to the evaluation of the relative performance. Even though the DM test does not suggest statistically more accurate forecasts of either SAV or AS compared to the benchmark GARCH, in this case we consider SAV to be the best fit, as it manages to produce the most precise PoF and the lowest value of the tick-loss function at the same time.

Continuing with BUX index, we observe similar results at 99% confidence level as for PX, but here only GARCH results in some violations, not being rejected by the unconditional coverage, independence, conditional coverage and DQ test. The rest of the models clearly overestimate the risk as they fail to produce any violation. Turning to the 5% quantile, we find CAViaR specifications to be more aacurate, delivering around 4 % and 6.1 % of violations (SAV and AS respectively), whereas GARCH type models produce PoF below 3 %, but still passing the unconditional coverage test. All of the models are not rejected by the independence test. DQ test however again indicates misspecification, only except for AS, which eventually favours this CAViaR specification. Looking at the value of the tick-loss function, the previous conclusion about the AS accuracy is supported as this succeeds in producing the lowest average loss.

Similarly to BUX, also in case of S&P 500 we observe very poor out-of-sample performance of three out of four models at 99% confidence level, delivering no violations and hence being rejected by all the tests regarding the absolute performance. Again, the only good fit is GARCH, with 1.2 % of violations, also producing independent violations according to the p-value of the independence test, comfortably passing both conditional coverage and DQ test for correct dynamic specification and having the lowest value of the tick-loss function at the same time. As for 95% VaR, we find that all the models considered tend to overestimate the risk, resulting in less violations. Here again GARCH seems to be the best choice by the definition and also in relative terms, having the unconditional coverage of 3.6 % and passing also independence and conditional coverage test. However, it is strongly rejected by the DQ test at this level. Despite that, we consider it the best choice for S&P 500.

For DAX index, the results are generally better than for the previous indices, especially for 1% quantile. All the models do quite a good job describing the evolution of the left tail, where GARCH and SAV deliver the same and the most precise results with respect to all the absolute performance tests. The remaining two models pass as well, though they are less precise by the definition. In case of the 95% VaR, the most precise model in terms of unconditional coverage is GARCH with normally distributed errors (4.3 % of violations), followed by SAV (3.5% of violations). The models again pass the tests, except for the DQ test that suggest misspecification for all of them. To choose the best fit, we turn to the evaluation of the relative performance. According to the DM test statistic, the forecsts of all the models are not statistically different from the benchmark at both levels. Though deciding between GARCH and SAV, the value of the tick-loss function is slightly lower for GARCH. Considering that it delivers more precise unconditional coverage for 95% VaR, we can conclude that this model comes out best at both confidence levels.

5.4 Summary and discussion of the results

Having described the backtesting results for all the indices in detail, we now proceed with the summary and discussion of the overall performance. First we focus on each period separately, since we want to examine also how the performance of the investigated models is influenced by the inclusion of the GFC in the sample and then we try to draw some general conclusion.

In the first (pre-crisis) period, CAViaR models perform very well in-sample for all indices at both 99% and 95% confidence levels in terms of the absolute performance, having the unconditional coverage generally closest to the nominal level (either 1 % or 5 %). At the same time, these models also succeed in producing the lowest value of the tick-loss function, which means that their adaptation to the returns is more flexible in time, unlike GARCH models that generally exhibit relatively long memory. The precision of both Symmetric Absolute Value (SAV) and Asymmetric Slope (AS) CAViaR specifications is not suprising, due to the design of the RQ objective function. GARCH type models often perform less adequately in-sample, delivering very poor results in some cases. As the results suggest, AS appears to be the best choice in most cases, followed by SAV that works somewhat better for S&P 500 at both confidence levels. The only exception is presented for 99% VaR in case of BUX, where GARCH could possibly come out as a winner in comparison with SAV, based on the value of the tick-loss function and DM test statistic for equal predictive ability. However, out-of-sample does not present such definite conclusions in favour of CAViaR models as in the in-sample. Having evaluated the performance also out-of-sample, we find that at least one GARCH specification is always able to do comparably good job as some CAViaR specification, if not better, depending on the confidence level. Only in case of S&P 500, the results show poor performance of both GARCH and CAViaR models, especially for the lower left tail. For 5% quantile, GARCH-t performs best. This could be possibly attributed to the fact that the model considers heavier tails and unlike the remaining models it is able to deal with the sudden spur of volatility and occurrence of more extreme observations caused by the forthcoming financial crisis, which is already covered in the out-of-sample for this index.

Turning to the second period covering the 08/09 GFC, as a group, CAViaR models again perform uniformly better in-sample than GARCH models at both confidence levels for all the indices, delivering the most precise unconditional and conditional coverage and being able to produce the lowest average tick-loss at the same time. Even though DM test does not always indicate significantly more accurate forecats, AS specification again seems to be generally the best choice across indices and confidence levels due to the lowest value of the tick-loss function. However, the performance of all the models out-of-sample is again commonly worse than in-sample. As could be expected, compared to the pre-crisis period, we observe generally poorer out-of-sample performance, with the unconditional coverage mostly below the nominal levels. The results come out inferior especially further in the left tail. For 1% quantile, in case of BUX and S&P 500, GARCH-t, SAV and AS produce no violations at all, hence highly overstate the risk. Only GARCH with normally distributed errors results in some VaR violations, being the only possible fit by the definition for these indices at this level. GARCH and also SAV seem to perform adequately for DAX, while AS manages to do quite a good job describing the evolution of the left tail for PX. At the 95% confidence level, CAViaR models provide relatively accurate forecasts in terms of the unconditional coverage and the average loss for PX (SAV) and BUX (AS), while GARCH appears to be more precise for S&P 500 and possibly DAX. Not surprisingly, worse performance of the models might be explained by the fact that their parameters were estimated on the sample that comprises highly volatile years, which consequently leads to the overestimation of the risk out-of-sample. In particular, for GARCH-t this resulted in quite low values of the degrees of freedom (ν) across indices, which did not necessarily coincide with the observations in the out-of-sample within this period.

Generally, as a group, CAViaR models manage to deliver very accurate, often superior, in-sample performance with respect to the unconditional and conditional coverage, and the average loss, though they are often rejected by the DQ test for correct dynamic specification. However, as this tests for much higher-order dependencies in the data, it is much stricter compared to the others.[3] Very accurate in-sample performance of CAViaR in terms of conditional coverage and loss, even in the crisis period, leads us to believe that this approach may be generally a suitable tool for VaR estimation, avoiding the restrictive assumptions about the distribution of return innovations being its greatest advantage. We suppose that even better overall performance would be achieved having estimated the one-step-ahead VaR as the rolling sample, as commonly done in practice, and possibly on the larger sample. Furthermore, we suppose that the performance of this approach on Central European stock market data could be likely improved by extending the quantile regressions with variables that incorporate more information about the financial returns, such as implied volatility etc., as has been mentioned in Chapter 3. However, we leave this for further research.

[3] Our results referring to the DQ test are consistent with other works, for example Kuester *et al.* (2006).

Chapter 6

Conclusion

The thesis concentrates on the direct conditional Value at Risk modeling, applying the flexible quantile regression and hence avoiding making strict assumptions about the distribution of return innovations. We investigate how the introduction of dynamics to VaR estimation using semiparametric Conditional Autoregressive Value at Risk (CAViaR) models of Engle & Manganelli (2004) impacts VaR performance and predictive ability. In our empirical application, we compare this semiparametric method for VaR calculation to the more common parametric approach that serves as the benchmark. More precisely, we compare our key CAViaR models to the GARCH models with two different error distributions: normal and also Student t-distribution as this is more in accord with the characteristics of financial data. Among four different CAViaR specifications, we estimate two, namely Symmetric Absolute Value (SAV) and Asymmetric Slope (AS). Our first main contribution stems from the fact that we examine the performance of these models on Central European stock price indices. Using ten years of data from January 2003 to December 2012 and dividing the dataset into two time periods (01/2003 - 12/2007 and 01/2008 - 12/2012) also enables us to study how the performance of these models is affected by the 08/09 GFC, which is another contribution.

In the first part of the thesis, we present the theoretical background for Value at Risk concept, where we describe the common methods for its calculation, with their stronger and weaker points. Further, we focus of the fundamentals of GARCH and CAViaR models and their estimation techniques. We conclude the theoretical part with the methods used for forecast evaluation of the absolute as well as the relative performance.

In the empirical part, we start with the analysis of the corresponding market data and proceed with the estimation of considered VaR models. Then we turn to the evaluation of their performance, where we focus on each period separately (as well as the in-sample fit and the out-of-sample performance) and describe the results for all the indices in detail. Finally, we provide the summary and discussion of the results.

The results indicate that the CAViaR models perform better in-sample, delivering very accurate unconditional and conditional coverage and lower loss at the same time, compared to the parametric GARCH models. Very precise estimates are obtained in the first (pre-crisis) period and also the second (crisis) period. The results are generally much less the same for all the indices. GARCH with normally distributed errors often tends to underestimate the risk, while the assumption of t-distributed errors leads to its overestimation, especially further in the left tail. From all the models, AS CAViaR seems to be the best fit in most cases, with regard to the unconditional and conditional coverage and the average loss. As for the predictive ability, CAViaR specifications provide also accurate out-of-sample forecasts. As expected, better results are obtained in the first period. The forecasts are accurate at both 99% and 95% confidence levels, especially for PX and DAX. In case of S&P 500 we obtain uniformly inferior results, which can be explained by the earlier kick-off of the GFC in the U.S. Out-of-sample performance in the second period generally comes out poorer. This is not surprising since the in-sample within this period comprises two highly volatile years and exhibits quite different characteristics compared to the data in the out-of-sample. However, CAViaR models still manage to do quite a good job, especially in case of PX at both levels, BUX at 95% level or DAX at 99% level.

In conclusion, we showed that the direct modeling of conditional quantiles is a reasonable approach to VaR estimation, which improves VaR accuracy. Precise performance of the CAViaR models as a group, even in the crisis period, leads us to believe that this approach may be very useful as a risk management tool. As we considered the basic CAViaR in our application, it might be of interest to examine the performance of its extended versions that could possibly deliver even better results. We see this as a base for future work.

Bibliography

ANGELIDIS, T., BENOS, A. & S. DEGIANNAKIS (2004): "The Use of GARCH Models in VaR Estimation." *Statistical Methodology* **1**: pp. 105–128.

ARTZNER, P., DELBAEN, F., EBER, J.M. & D. HEATH (1999): "Coherent Measures of Risk." *Mathematical Finance* **9(3)**: pp. 203–228.

BAILLIE, R.T., BOLLERSLEV, T. & H.O. MIKKELSEN (1996): "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* **74(1)**: pp. 3–30.

BAMS, D., LEHNERT, T. & C.C.P. WOLFF (2005): "An Evaluation Framework for Alternative VaR-models." *Journal of International Money and Finance* **24(6)**: pp. 944–958.

BARUNIK, J. & F. ZIKES (2013): "Semiparametric Conditional Quantile Models forFinancial Returns and Realized Volatility." *Working Papers Series* Available at SSRN: http://ssrn.com/abstract=2313047.

BASEL COMMITTEE ON BANKING SUPERVISION (1996): "Amendment to the Capital Accord to Incorporate Market Risks." Available at http://www.bis.org.

BERKOWITZ, J., CHRISTOFFERSEN, P. & D. PELLETIER (2011): "Evaluating Value-at-Risk Models with Desk-Level Data." *Management Science* **57(12)**: pp. 2213–2227.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* **31**: pp. 307–327.

BOLLERSLEV, T. & J.M. WOOLRIDGE (1992): "Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances." *Econometric Reviews* **11(2)**: pp. 143–172.

Boudoukh, J., Richardson, M. & R. Whitelaw (1998): "The Best of Both Worlds: A Hybrid Approach to Calculating Value at Risk." *RISK* **11(5)**: pp. 64–67.

Campbell, S.D. (2007): "A Review of Backtesting and Backtesting Procedures." *Journal of Risk* **9(2)**: pp. 1–17.

Chen, C.W.S., Gerlach, R., Hwang, B.B.K. & M. McAleer (2012): "Forecasting Value-at-Risk Using Nonlinear Regression Quantiles and the Intra-Day Range." *International Journal of Forecasting* **28(3)**: pp. 557–574.

Chernozhukov, V. & L. Umantsev (2001): "Conditional Value-at-Risk: Aspects of Modeling and Estimation." *Empirical Economics* **26(1)**: pp. 271–292.

Christoffersen, P.F., (1998): "Evaluating Interval Forecasts." *Econometrica* **39(4)**: pp. 841–862.

Danielsson, J., & C.G. de Vries(2000): "Value-at-Risk and Extreme Returns." *Annales d'Economie et de Statistique* **60**: pp. 239–270.

Diebold, F.X., & R.S. Mariano(1995): "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* **13(3)**: pp. 253–263.

Diebold, F.X., Schuermann, T. & J. Stroughair (1998): "Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management." *Center for Financial Institutions Working Papers 98-10* Wharton School Center for Financial Institutions, University of Pennsylvania.

Ding, Z., Granger, C.W.J. & R.F. 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.

Engle, R.F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. inflation." *Econometrica* **50(4)**: pp. 987–1007.

Engle, R.F. & S. Manganelli (2001): "Value at Risk Models in Finance." *Working Paper Series No. 75.* ECB.

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.

ENGLE, R.F. & V.K. NG (1993): "Measuring and Testing the Impact of News on Volatility." *The Journal of Finance* **48(5)**: pp. 1749–1778.

FAMA, E.F. (1965): "The Behavior of Stock Market Prices." *Journal of Business* **38(1)**: pp. 34–105.

GERLACH, R., CHEN, C.W.S., & N.Y.C.. CHAN (2011): "Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets." *Journal of Business and Economic Statistics* **29(4)**: pp. 481–492.

GIACOMINI, R. & I. KOMUNJER (2005): "Evaluation and Combination of Conditional Quantile Forecasts." *Journal of Business and Economic Statistics* **23(4)**: pp. 416–431.

HUANG, D., YU, B., LU, Z., FABOZZI, F.J., FOCARDI, S., & M. FUKUSHIMA (2010): "Index-Exciting CAViaR: A New Empirical Time-Varying Risk Model." *Studies in Nonlinear Dynamics and Econometrics* **14(2)**: pp. 1–26.

JEON, J. & J.W. TAYLOR (2013): "Using CAViaR Models with Implied Volatility for Value-at-Risk Estimation." *Journal of Forecasting* **32(1)**: pp. 62–74.

JORION, P. (1996): "Risk:Measuring the Risk in Value at Risk." *Financial Analysts Journal* **52(6)**: pp. 47–56.

KOENKER, R. & G. BASSETT (1978): "Regression Quantiles." *Econometrica* **46**: pp. 33–50.

KOENKER, R. & G. BASSETT (1982): "Robust Tests for Heteroscedasticity Based on Regression Quantiles." *Econometrica* **50(1)**: pp. 43–61.

KOURETAS, G.P. & L. ZARANGAS (2005): "Conditional Autoregresive Value-at-Risk by Regression Quantiles: Estimation Market Risk for Major Stock Markets." *Working Papers Series 0521*. Available at: http://ideas.repec.org/p/crt/wpaper/0521.html.

KUESTER, K., MITTNIK, S., & M.S. PAOLELLA(2006): "Value-at-Risk Prediction: A Comparison of Alternative Strategies." *Journal of Financial Econometrics* **4(1)**: pp. 53–89.

KUPIEC, P.H. (1995): "Techniques for Verifying the Accuracy of Risk Measurement Models." *Journal of Derivatives* **3(2)**: pp. 73–84.

LJUNG, G.M. & G.E.P. BOX (1978): "On a Measure of Lack of Fit in Time Series Models." *Biometrika* **65(2)**: pp. 297–303.

MANDELBROT, B. (1963): "The Variation of Certain Speculative Prices." *Journal of Business* **36(4)**: pp. 394–419.

MCNEIL, A.J. & R. FREY (2000): "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach." *Journal of Empirical Finance* **7(3-4)**: pp. 271–300.

MCNEIL, E., FREY, R. & P. EMBRECHTS (2005): *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press.

MITTNIK, S., PAOLELLA, M.S., & S.T. RACHEV (2000): "Diagnosing and Treating the Fat Tails in Financial Returns Data." *Journal of Empirical Finance* **7(3-4)**: pp. 389–416.

MORGAN, J.P. & REUTERS (1996): *RiskMetricsTM - Technical Document.* 4th ed. New York: J.P.Morgan.

NELSON, D.B. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* **59(2)**: pp. 347–370.

PORTNOY, S. (1991): "Asymptotic Behaviour of Regression Quantiles in Non-Stationary, Dependent Cases." *Journal of Multivariate Analysis* **38(1)**: pp. 100–113.

TAYLOR, S.J. (1986): *Modelling Financial Time Series.* John Wiley & Sons, Ltd.

WHITE, S. (1994): *Estimation, Inference and Specification Analysis.* Cambridge University Press. Cambridge, U.K.

YU, P. L. H., LI, W. K. & S. JIN (2010): "On Some Models for Value-at-Risk." *Econometric Reviews.* **29(5-6)**: pp. 622–641.

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, first period

Statistics	PX	BUX	DAX	S&P 500
Observations	1007	1009	1022	1007
Mean	0.1221	0.1135	0.0738	0.0442
Median	0.1756	0.1385	0.1199	0.0807
Std. dev.	1.0811	1.3200	1.2437	0.7774
Minimum	-6.1250	-5.6027	-6.3360	-3.5867
Maximum	7.0482	4.8660	6.6446	3.4814
Skewness	-0.5519	-0.2436	-0.1753	0.0198
Ex. kurtosis	4.9091	1.0641	3.3232	1.3918
Jarque-Bera test				
p-value	3.57639e-231	3.21794e-013	6.97703e-104	2.25864e-018
test statistic	1061.2500	57.5297	475.0520	81.2635
ADF test				
p-value	4.904e-009	8.931e-011	7.931e-009	4.404e-015
test statistic	-6.8967	-7.4819	-6.8236	-8.7650

Source: Author's computations.

Table A.2: Descriptive Statistics: out-of-sample, first period

Statistics	PX	BUX	DAX	S&P 500
Observations	249	244	251	250
Mean	0.0461	0.0229	0.0751	0.0144
Median	0.1091	0.0993	0.1203	0.0813
Std. dev.	1.0722	1.1800	0.9766	1.0112
Minimum	-5.6732	-4.3649	-3.0036	-3.5343
Maximum	2.7433	3.3440	2.5214	2.8790
Skewness	-0.8876	-0.1324	-0.3134	-0.4947
Ex. kurtosis	3.1820	1.1716	0.3611	1.4321
Jarque-Bera test				
p-value	1.22753e-030	0.000653091	0.0648523	1.40127e-007
test statistic	137.7450	14.6676	5.47129	31.5614
ADF test				
p-value	8.921e-012	0.0324	4.496e-028	0.0228
test statistic	-7.7979	-3.5703	-15.6394	-3.6923

Source: Author's computations.

Table A.3: Descriptive Statistics: in-sample, second period

Statistics	PX	BUX	DAX	S&P 500
Observations	1007	1009	1026	1009
Mean	-0.0682	-0.0419	-0.0291	-0.0139
Median	-0.0516	-0.0334	0.0292	0.0903
Std. dev.	2.0585	2.1599	1.8320	1.8104
Minimum	-16.1850	-12.6490	-7.4335	-9.4695
Maximum	12.3640	13.1780	10.7970	10.9570
Skewness	-0.3945	-0.0030	0.1583	-0.2193
Ex. kurtosis	9.7730	4.9324	4.6612	5.7526
Jarque-Bera test				
p-value	0	1.30899e-222	3.75849e-203	2.74216e-304
test statistic	4029.5800	1021.8100	932.2020	1397.9500
ADF test				
p-value	0.0008392	8.823e-005	2.279e-006	5.849e-006
test statistic	-4.6394	-5.1571	-5.8782	-5.7030

Source: Author's computations.

Table A.4: Descriptive Statistics: out-of-sample, second period

Statistics	PX	BUX	DAX	S&P 500
Observations	250	244	255	249
Mean	0.0469	0.0239	0.0826	0.0444
Median	0.0814	-0.0797	0.1126	0.0195
Std. dev.	1.0496	1.2367	1.1669	0.7997
Minimum	-3.4715	-3.1675	-3.4773	-2.4951
Maximum	3.3581	4.1324	4.2401	2.4615
Skewness	-0.1065	0.3000	-0.1513	0.0350
Ex. kurtosis	0.4053	0.5600	1.2879	0.8694
Jarque-Bera test				
p-value	0.335571	0.0326188	9.16231e-005	0.0193327
test statistic	2.1838	6.8457	18.5957	7.8919
ADF test				
p-value	0.01203	1.363e-025	2.148e-016	2.101e-016
test statistic	-3.9005	-14.3260	-9.1228	-9.1254

Source: Author's computations.

Table A.5: Akaike Information Criteria for GARCH(p,q), first period

	GARCH				GARCH-t			
(p,q)	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
(1,1)	2874.0	2869.4	2870.9	2871.3	2795.1	2795.9	2870.9	2871.3
(1,2)	3350.2	3347.6	3343.8	3342.4	3345.2	3342.6	3340.5	3339.1
(2,1)	3003.7	3003.0	3004.1	3004.8	2994.2	2992.6	2994.3	2994.5
(2,2)	2228.3	2223.3	2226.2	2224.9	2229.2	2223.8	2227.1	2225.4

Source: Author's computations.

Table A.6: Akaike Information Criteria for GARCH(p,q), second period

	GARCH				GARCH-t			
(p,q)	PX	BUX	DAX	S&P 500	PX	BUX	DAX	S&P 500
(1,1)	3760.1	3761.4	3761.0	3763.0	3737.9	3739.6	3739.6	3735.5
(1,2)	4122.5	4124.5	4124.5	4126.4	4100.7	4102.7	4102.7	4104.6
(2,1)	3827.1	3821.5	3827.3	3823.0	3807.1	3794.9	3805.9	3796.6
(2,2)	3559.7	3535.7	3544.9	3536.9	3533.0	3513.5	3522.0	3514.0

Source: Author's computations.

Table A.7: Backtesting for BUX, first period

Absolute and relative performance of GARCH and CAViaR models for daily BUX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	0.9911	0.5946	0.9911	0.8920	1.6393	1.2295	1.6393	2.8689
$p - LR_{uc}$	0.9772	0.1616	0.9772	0.7253	0.3581	0.7280	0.3581	0.0168
$p - LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.9995	0.3754	0.9995	0.9401	0.6555	0.9412	0.6555	0.0572
$p - DQ$	0.0038	0.0024	0.0009	0.0001	0.0007	0.0000	0.0007	0.0030
$\mathcal{T}_\alpha$	0.0392	0.0397	0.0406	0.0397	0.0406	0.0391	0.0402	0.0474
DM		-2.1214	2.1777	-1.2051		-1.4883	-0.8979	1.0489
95% VaR								
$PoF(\%)$	3.3697	2.6759	4.8563	4.9554	3.6885	2.8689	4.5082	3.6885
$p - LR_{uc}$	0.0118	0.0002	0.8333	0.9481	0.3251	0.0980	0.7202	0.3251
$p - LR_{ind}$	1.0000	1.0000	0.6846	0.2660	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.0418	0.0010	0.5375	0.9007	0.6162	0.2543	0.9378	0.6162
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0013	0.0001	0.0107	0.0013
$\mathcal{T}_\alpha$	0.1375	0.1402	0.1382	0.1366	0.1291	0.1315	0.1280	0.1273
DM		-2.7605	3.0272	1.6776		-1.5805	0.9793	-0.8144

Source: Author's computations.

Table A.8: Backtesting for DAX, first period

Absolute and relative performance of GARCH and CAViaR models for daily DAX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	1.8591	0.9785	0.9785	1.0763	1.1952	0.7968	1.1952	1.1952
$p - LR_{uc}$	0.0137	0.9447	0.9447	0.8086	0.7630	0.7373	0.7630	0.7630
$p - LR_{ind}$	0.3616	0.0826	0.0826	0.1035	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.0316	0.2211	0.2211	0.2581	0.9555	0.9453	0.9555	0.9555
$p - DQ$	0.2988	0.0012	0.0000	0.0007	0.2291	0.7802	0.2291	0.2290
$\mathcal{T}_\alpha$	0.0366	0.0352	0.0336	0.0333	0.0325	0.0328	0.0323	0.0321
DM		-2.8414	-2.6966	-2.7720		-1.3090	-1.1522	-1.2346
95% VaR								
$PoF(\%)$	5.4795	3.9139	5.1859	5.0881	5.1793	3.5857	6.3745	4.3825
$p - LR_{uc}$	0.4883	0.0982	0.7863	0.8975	0.8969	0.2799	0.3370	0.6468
$p - LR_{ind}$	0.1509	0.6152	0.2056	0.2268	1.0000	1.0000	0.9797	1.0000
$p - LR_{cc}$	0.2803	0.2246	0.4326	0.4778	0.9916	0.5577	0.6305	0.9003
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0085	0.0010	0.0020	0.0023
$\mathcal{T}_\alpha$	0.1316	0.1323	0.1298	0.1271	0.1112	0.1126	0.1121	0.1083
DM		-3.7377	-1.0983	-0.8611		-2.0003	2.0213	-0.5282

Source: Author's computations.

Table A.9: Backtesting for S&P 500, first period

Absolute and relative performance of GARCH and CAViaR models for daily S&P 500 returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	1.2910	0.8937	1.0924	1.0924	4.4000	3.6000	3.6000	4.4000
$p - LR_{uc}$	0.3745	0.7300	0.7716	0.7716	0.0001	0.0014	0.0014	0.0001
$p - LR_{ind}$	1.0000	1.0000	0.1053	1.0000	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.6741	0.9421	0.2582	0.9587	0.0003	0.0060	0.0060	0.0003
$p - DQ$	0.6222	0.0544	0.0009	0.0219	0.0000	0.0000	0.0052	0.0000
$\mathcal{T}_\alpha$	0.0205	0.0203	0.0198	0.0203	0.0451	0.0420	0.0440	0.0466
DM		-2.1048	-0.9976	-1.1677		-2.5399	-0.6327	1.1980
95% VaR								
$PoF(\%)$	4.4687	4.0715	4.9652	5.2632	8.0000	6.8000	7.6000	10.4000
$p - LR_{uc}$	0.4312	0.1630	0.9596	0.7039	0.0444	0.2146	0.0787	0.0006
$p - LR_{ind}$	0.1909	0.1060	0.7385	0.2041	0.5801	1.0000	1.0000	0.1931
$p - LR_{cc}$	0.3119	0.1024	0.9446	0.4153	0.1139	0.4629	0.2132	0.0011
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.5284	0.1900	0.1450	0.5500
$\mathcal{T}_\alpha$	0.0798	0.0802	0.0787	0.0788	0.1305	0.1296	0.1310	0.1367
DM		-3.2046	0.5253	0.8015		-3.2635	0.9634	2.7513

Source: Author's computations.

Table A.10: Backtesting for BUX, second period

Absolute and relative performance of GARCH and CAViaR models for daily BUX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark. NA is reported in case of no violation.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	1.3875	0.6938	0.9911	1.1893	0.4098	0.0000	0.0000	0.0000
$p - LR_{uc}$	0.2426	0.3008	0.9772	0.5573	0.2933	NA	NA	NA
$p - LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	NA	NA	NA
$p - LR_{cc}$	0.5051	0.5854	0.9995	0.8418	0.5756	NA	NA	NA
$p - DQ$	0.1887	0.0076	0.0214	0.0214	0.5962	NA	NA	NA
$\mathcal{T}_\alpha$	0.0645	0.0644	0.0612	0.0577	0.0337	0.0391	0.0405	0.0373
DM		-2.6540	-2.6471	-2.4832		1.1356	1.6598	0.5423
95% VaR								
$PoF(\%)$	4.7572	3.4688	4.9554	4.9554	2.8689	2.0492	4.0984	6.1475
$p - LR_{uc}$	0.7213	0.0184	0.9481	0.9481	0.0980	0.0170	0.5053	0.4265
$p - LR_{ind}$	0.6345	0.8350	0.0426	0.7407	1.0000	1.0000	0.4116	0.9353
$p - LR_{cc}$	0.8381	0.0609	0.1278	0.9447	0.2543	0.0578	0.5718	0.7265
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0009	0.1025
$\mathcal{T}_\alpha$	0.2168	0.2201	0.2156	0.2138	0.1299	0.1356	0.1307	0.1265
DM		-2.9143	-1.1900	-0.1426		-0.6226	1.7444	1.1552

Source: Author's computations.

Table A.11: Backtesting for DAX, second period

Absolute and relative performance of GARCH and CAViaR models for daily DAX returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p - LR_{uc}$), p-value of independence test ($p - LR_{ind}$), p-value of conditional coverage test ($p - LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p - DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	1.4620	0.2924	0.9747	1.0721	0.7843	0.3922	0.7843	0.3922
$p - LR_{uc}$	0.1641	0.0073	0.9347	0.8185	0.7190	0.2660	0.7190	0.2660
$p - LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.3798	0.0274	0.9966	0.9739	0.9373	0.5386	0.9373	0.5386
$p - DQ$	0.9891	0.0000	0.0960	0.0427	0.8202	0.5194	0.8202	0.5194
$\mathcal{T}_\alpha$	0.0498	0.0537	0.0487	0.0458	0.0371	0.0405	0.0377	0.0356
DM		-1.5293	-1.0778	-1.1426		-1.0489	-1.2052	-1.1946
95% VaR								
$PoF(\%)$	6.5302	4.7758	4.7758	4.8733	4.3137	3.1373	3.5294	3.1373
$p - LR_{uc}$	0.0314	0.7400	0.7400	0.8517	0.2646	0.1440	0.2564	0.1440
$p - LR_{ind}$	0.1787	1.0000	1.0000	0.2767	1.0000	1.0000	1.0000	1.0000
$p - LR_{cc}$	0.0399	0.9464	0.9464	0.5438	0.8760	0.3438	0.5251	0.3438
$p - DQ$	0.0000	0.0000	0.0000	0.0000	0.0002	0.0045	0.0007	0.0002
$\mathcal{T}_\alpha$	0.1891	0.1886	0.1890	0.1820	0.1390	0.1434	0.1392	0.1340
DM		-3.7892	-2.8192	-2.1476		-1.7261	-1.8808	-1.8778

Source: Author's computations.

Table A.12: Backtesting for S&P 500, second period

Absolute and relative performance of GARCH and CAViaR models for daily S&P 500 returns. The left-hand side panel reports results for in-sample performance and the right-hand side panel reports results for out-of-sample performance. For each VaR we report the actual proportion of failures (PoF), p-value of unconditional coverage test ($p-LR_{uc}$), p-value of independence test ($p-LR_{ind}$), p-value of conditional coverage test ($p-LR_{cc}$), p-value of DQ test of Engle & Manganelli (2004) ($p-DQ$), the value of tick-loss function ($\mathcal{T}_\alpha$) and Diebold-Mariano (DM) test statistic for equal predictive accuracy with GARCH model with normal errors serving as the benchmark. NA is reported in case of no violation.

	in-sample				out-of-sample			
	GARCH	GARCH-t	SAV	AS	GARCH	GARCH-t	SAV	AS
99% VaR								
$PoF(\%)$	2.1804	0.4955	1.1893	1.0902	1.2048	0.0000	0.0000	0.0000
$p-LR_{uc}$	0.0011	0.0743	0.5573	0.7766	0.7530	NA	NA	NA
$p-LR_{ind}$	1.0000	1.0000	1.0000	1.0000	1.0000	NA	NA	NA
$p-LR_{cc}$	0.0049	0.2034	0.8418	0.9605	0.9517	NA	NA	NA
$p-DQ$	0.0356	0.0000	0.0009	0.0009	0.2170	NA	NA	NA
$\mathcal{T}_\alpha$	0.0520	0.0523	0.0503	0.0488	0.0245	0.0277	0.0270	0.0265
DM		-1.7970	0.5309	-2.0883		-0.8956	-0.9724	-0.9912
95% VaR								
$PoF(\%)$	6.1447	4.2616	4.8563	5.0545	3.6145	2.8112	2.0080	2.0080
$p-LR_{uc}$	0.1067	0.2701	0.8333	0.9368	0.2923	0.0852	0.0142	0.0142
$p-LR_{ind}$	0.2813	0.4841	0.2918	1.0000	1.0000	1.0000	1.0000	1.0000
$p-LR_{cc}$	0.1523	0.4261	0.5611	0.9968	0.5743	0.2272	0.0495	0.0495
$p-DQ$	0.0000	0.0000	0.0000	0.0000	0.0008	0.0001	0.0000	0.0000
$\mathcal{T}_\alpha$	0.1862	0.1864	0.1868	0.1825	0.0934	0.0984	0.0991	0.0988
DM		-3.2593	-2.0466	-0.6972		-1.3101	-1.3301	-1.2630

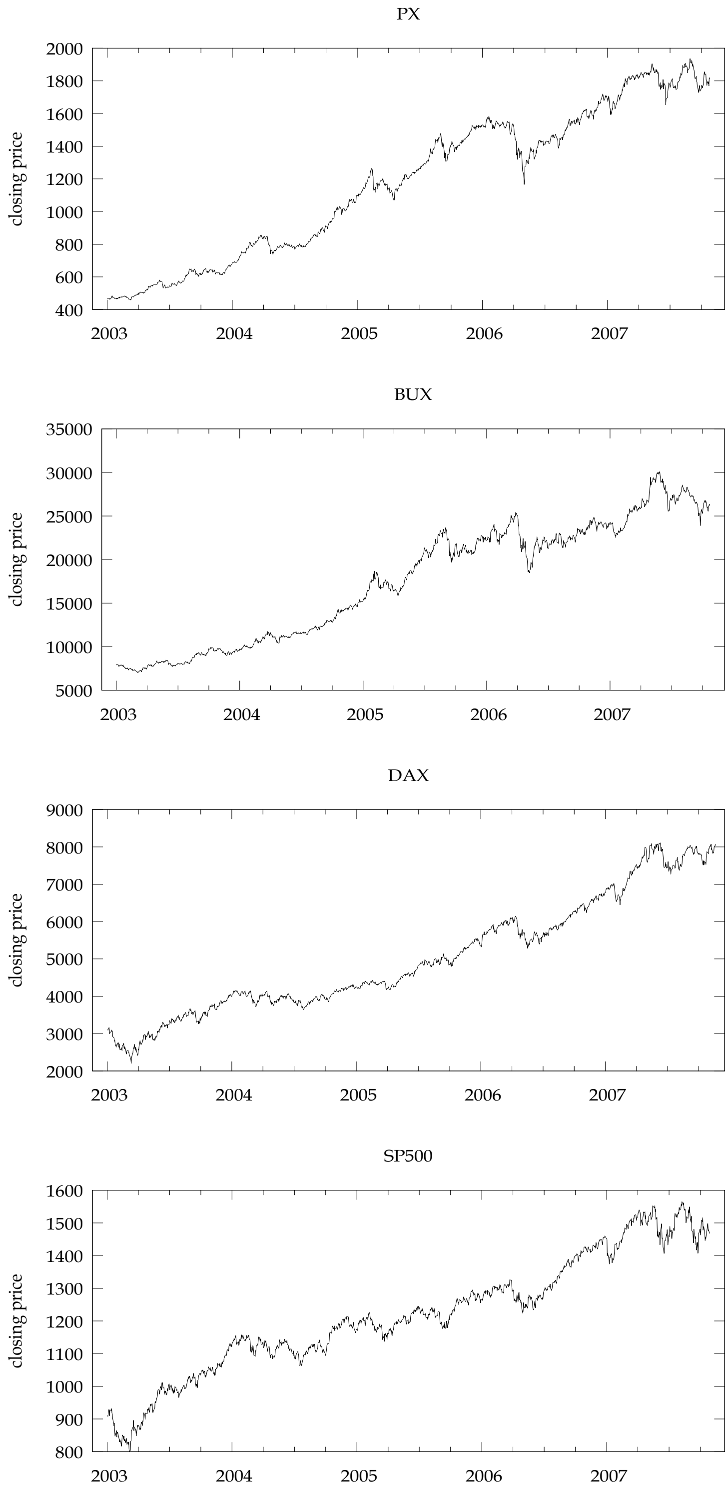
Source: Author's computations.

Appendix B

Figures

Figure B.1: Closing prices of all indices, first period

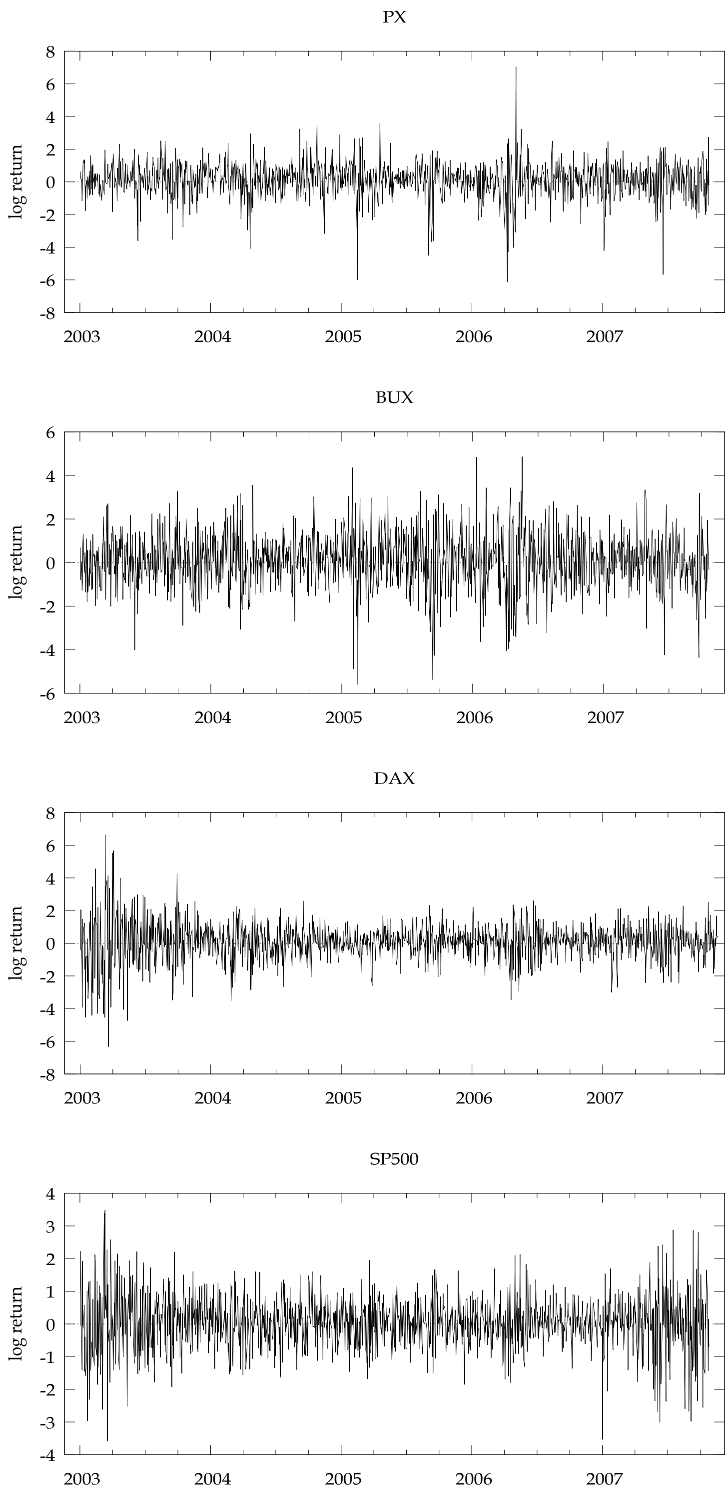
In-sample covers the period of 01/2003 - 12/2006, out-of-sample the year of 2007.



Source: Author's computations.

Figure B.2: Log returns of all indices, first period

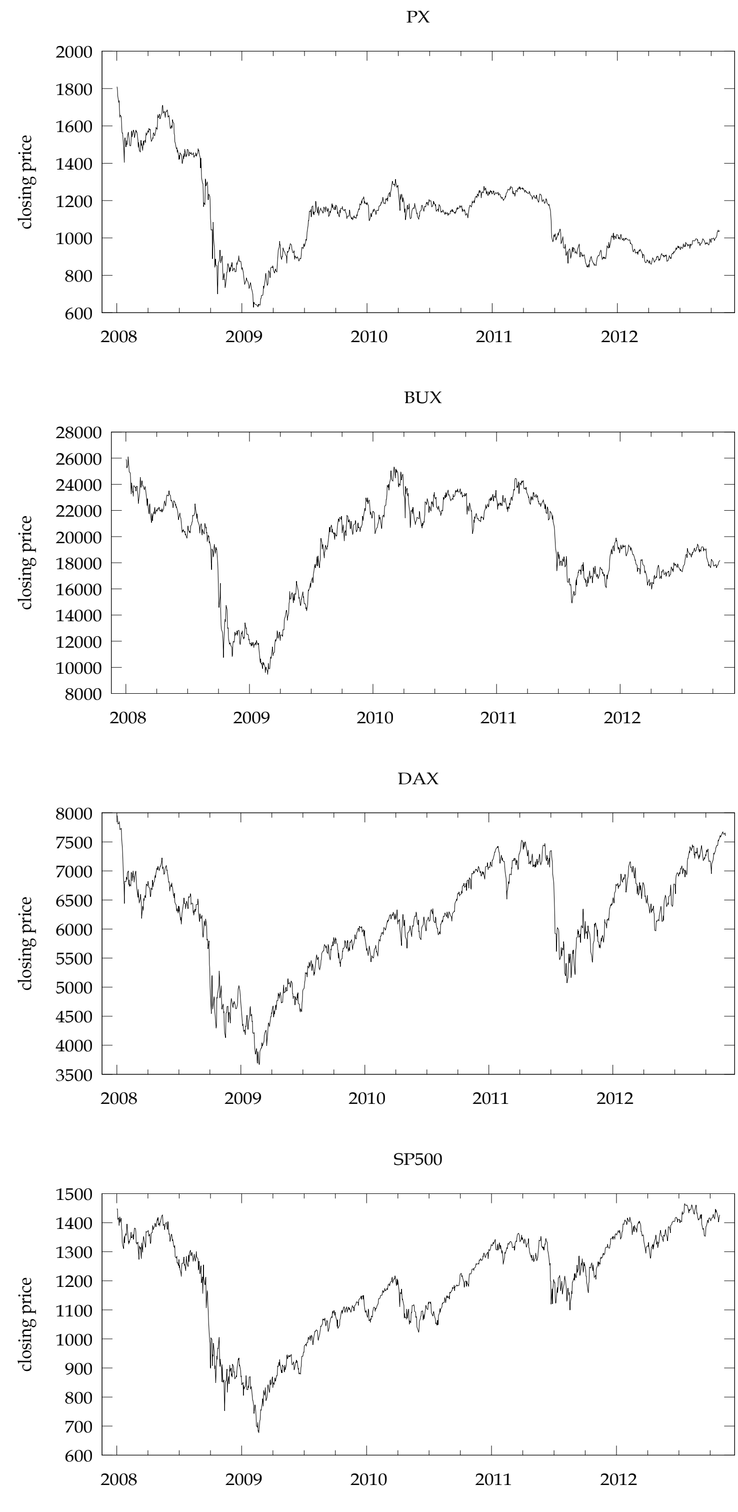
In-sample covers the period of 01/2003 - 12/2006, out-of-sample the year of 2007.



Source: Author's computations.

Figure B.3: Closing prices of all indices, second period

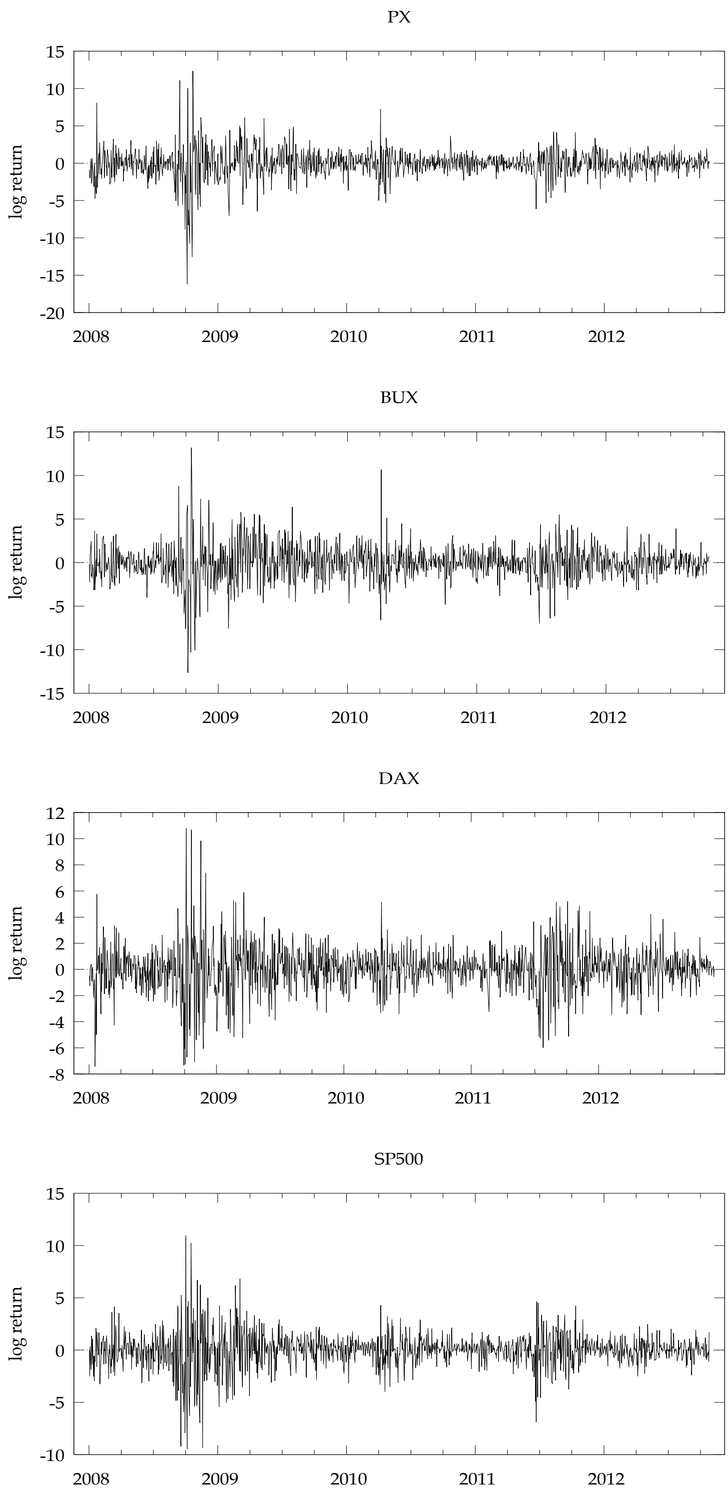
In-sample covers the period of 01/2008 - 12/2011, out-of-sample the year of 2012.



Source: Author's computations.

Figure B.4: Log returns of all indices, second period

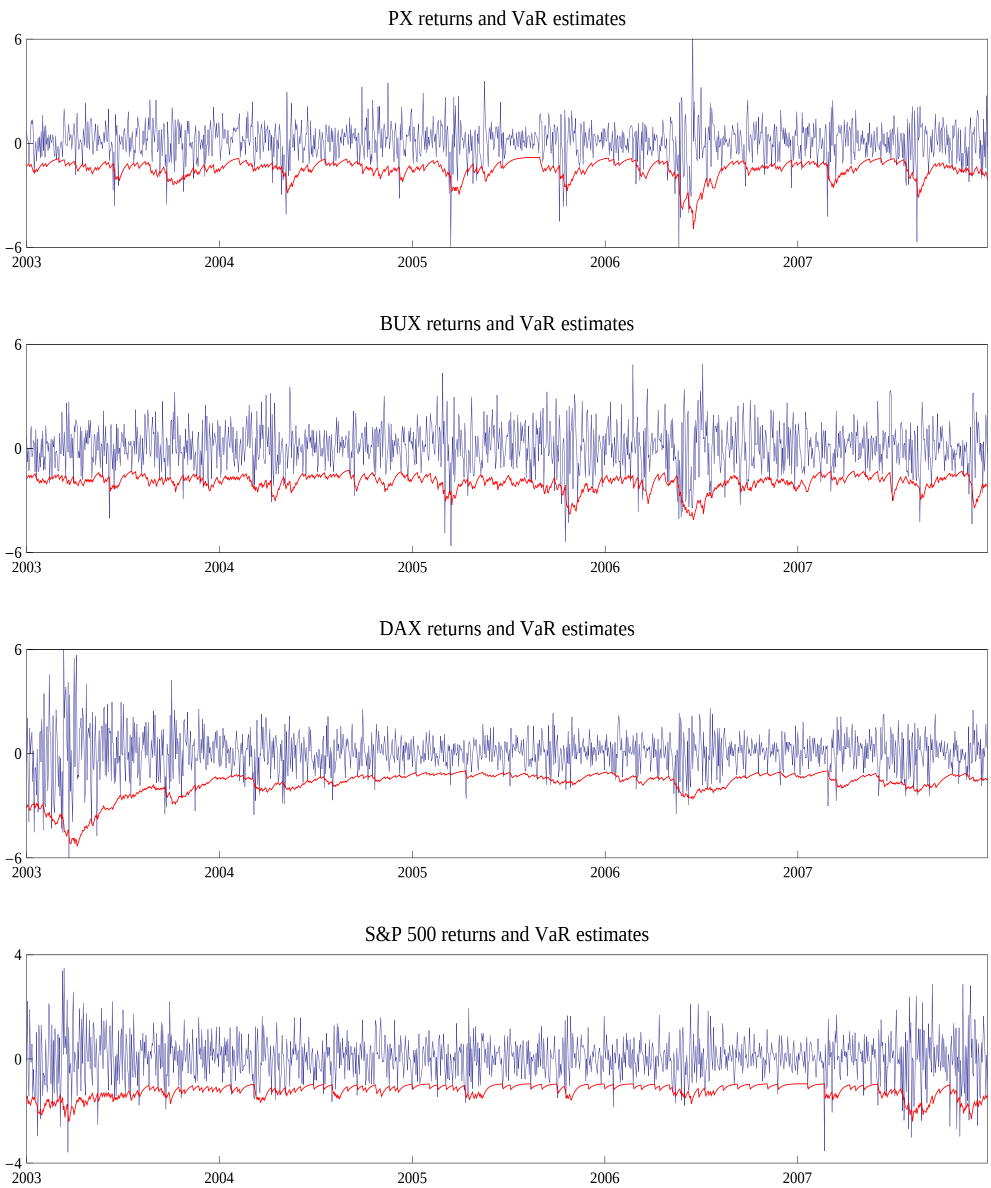
In-sample covers the period of 01/2008 - 12/2011, out-of-sample the year of 2012.



Source: Author's computations.

Figure B.5: SAV CAViaR at 95% confidence level, first period

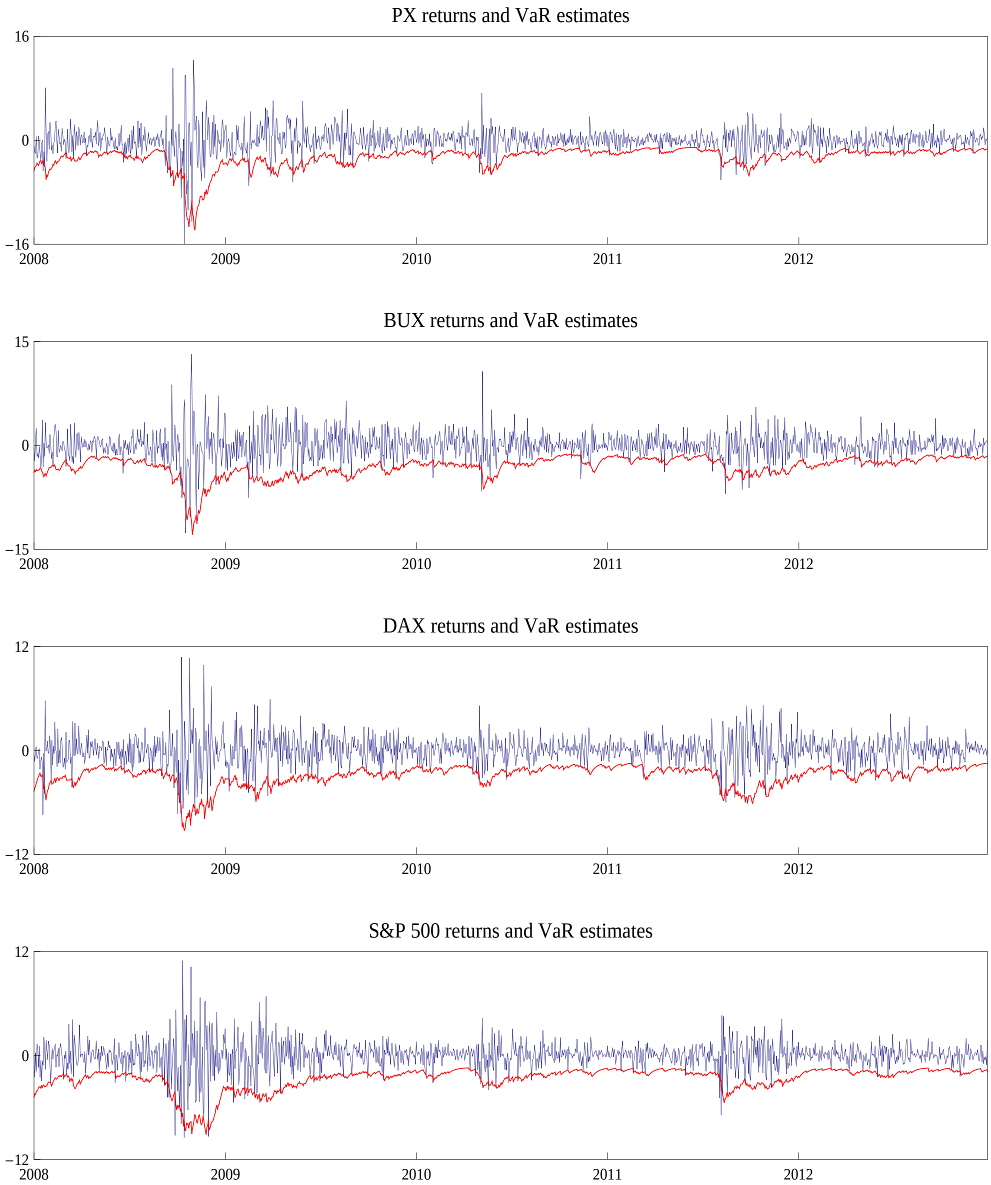
Returns and one-step-ahead VaR forecasts of Symmetric Absolute Value CAViaR at 95% confidence level. In-sample covers the period of 01/2003 - 12/2006, out-of-sample the year of 2007.



Source: Author's computations.

Figure B.6: SAV CAViaR at 95% confidence level, second period

Returns and one-step-ahead VaR forecasts of Symmetric Absolute Value CAViaR at 95% confidence level. In-sample covers the period of 01/2008 - 12/2011, out-of-sample the year of 2012.



Source: Author's computations.