

Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies



MASTER THESIS

**Value-at-Risk based on Extreme Value
Theory method and copulas. Empirical
evidence from Central Europe.**

Author: **Krenar Avdulaj**

Supervisor: **PhDr. Jozef Baruník**

Academic Year: **2009/2010**

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, May 20, 2010

Signature

Abstract

Assessing the extreme events is crucial in financial risk management. All risk managers and financial institutions want to know the risk of their portfolio under rare events scenarios. We illustrate a multivariate Monte Carlo and semi-parametric method to estimate Value-at-Risk (VaR) for a portfolio of stock exchange indexes in Central Europe. It is a method that uses the non-parametric empirical distribution to capture the small risks and the parametric Extreme Value theory to capture large risks. We compare this method with historical simulation and variance-covariance method under low and high volatility samples of data. In general historical simulation method overestimates the VaR for extreme events, while variance-covariance underestimates it. The method that we illustrate gives a result in between because it considers historical performance of the stocks and also corrects for the heavy tails of the distribution. We conclude that the estimate method that we illustrate here is useful in estimating VaR for extreme events, especially for high volatility times.

JEL Classification C22, G17

Keywords Value-at-Risk, Extreme Value Theory, copula.

Author's e-mail krenar.avdulaj@gmail.com

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

Abstrakt

Odcenění extrémních událostí je pro finanční risk management rozhodující. Finanční instituce a jejich manažeři chtějí co nejlépe znát riziko svého portfolia při neočekávaných událostech. V práci používáme vícerozměrný simulační metodu odhadování Value-at-Risk (VaR) portfolia středoevropských tržních indexů. Tato metoda používá neparametrickou empirickou distribuci, která zachycuje malé riziko, a teorii extrémních hodnot k zachycení velkého rizika. V práci porovnáváme naši metodu s metodou historických simulací a varianční-kovarianční metodou. Ukazujeme, že historická simulace obecně nadhodnocuje VaR extrémních událostí na trhu, zatímco varianční-kovarianční metoda riziko podhodnocuje. Naše metoda uvažuje historický vývoj akcií ale zároveň počítá i s těžkými chvosty distribuce výnosů. Závěrem práce tedy je, že metoda použitá v této práci je pro odhadování rizika v časech s vysokou volatilitou užitečná.

Klasifikace JEL	C22, G17
Klíčová slova	Value-at-Risk, Extreme Value Theory, copula.
E-mail autora	krenar.avdulaj@gmail.com
E-mail vedoucího práce	barunik@utia.cas.cz

Contents

List of Tables	vi
List of Figures	vii
Acronyms	ix
Thesis Proposal	x
1 Introduction	1
2 ARMA and GARCH Models in Financial Time Series	8
2.1 ARMA Models	8
2.2 GARCH Models	10
3 Extreme Value Theory (EVT)	12
3.1 Generalized Extreme Value (GEV) Distribution	13
3.2 Threshold Exceedances	15
4 Copulas	20
4.1 Copulas	20
5 Application of the model	25
6 Conclusion	37
Bibliography	42
A Appendix One	I
B Appendix Two	III

List of Tables

5.1	Summary Statistics. 01-Jan-2000 to 31-Dec-2009.	26
5.2	The ARMA-GARCH models that best fitted our data are ARMA(1,0)-GARCH(1,1) for the case of Austria, the Czech Republic and Slovakia and constant conditional mean-GARCH(1,1) for the rest. Data in brackets represent standard errors.	27
5.3	Point estimates of the tail parameters.	29
5.4	Correlation matrix obtained from fitting t -copula to data. . . .	30
5.5	Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2000-2009.	33
5.6	Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2000-2005.	34
5.7	Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2005-2009.	35

List of Figures

1.1	Value at Risk	2
1.2	Value at Risk (returns distributed normally)	3
3.1	Block maxima and peaks over threshold method. The red dots are the observations used for modelling for each method.	13
3.2	Probability density functions and cumulative distribution functions. $\xi = 0$ (Gumbell), $\xi = -0.5$ (Weibull) and $\xi = 0.5$ (Frechét). In all cases $\mu = 0$ and $\beta = 1$	15
3.3	Probability density functions and cumulative distribution functions. $\xi = 0$ exponential, $\xi = -0.5$ Pareto type II and $\xi = 0.5$ Pareto distribution. In all cases $\beta = 1$ and $\nu = 0$	16
3.4	Conditional excess distribution over the threshold u . We use the Extreme Value Theory (EVT) to model the observations which excess the threshold u	17
3.5	Empirical mean excess plot for the left tail (a) and the right tail (b) for ATX (Austria).	19
4.1	The Gaussian and t copulas for a bivariate sample of 10,000 observations in uniform scale (a and c) and in their original scale (b and d). The correlations are 0.7 in both cases, while for t copula the degrees of freedom are 2.	23
4.2	For both probability distribution functions the correlation is 0.7, while for t copula the degree of freedom is $\nu = 2$	24
5.1	In (a) are prices of 5 indexes from Central Europe Financial Markets, Austria (ATX), Germany (DAX), Czech Republic (PX50), Slovakia (SAX) and Switzerland Swiss Mkt (SSMI). Prices are normalized to 100 to make comparable. In (b) are returns series for ATX.	27

5.2	In (a) are ACF of raw returns and squared raw returns for ATX before filtering. In (b) are ACF of standardized residuals and squared standardized residuals for ATX, while in (c) are Residuals and conditional standard deviations for ATX.	28
5.3	Semi-parametric cumulative distribution function (a) and fitted cumulative distribution function for upper tail (b) and lower tail (c).	29
5.4	Empirical cumulative distribution function of simulated portfolio returns. Circles (from left to right) represent VaR estimates at 99.995%, 99.99%, 99.95%, 99.9%, 99.5%, 99%, 95% and 90% confidence intervals.	32
B.1	Returns series.	III
B.2	Autocorrelation of return series.	IV
B.3	Autocorrelation after filtering.	IV
B.4	Residuals and conditional standard deviation.	V
B.5	Mean excess plot for the left tail.	V
B.6	Mean excess plot for the right tail.	VI
B.7	Fitted lower (left) tail.	VI
B.8	Fitted upper (right) tail.	VII

Abbreviations

ACF	Autocorrelation Function
ARCH	Autoregressive Conditional Heteroskedasticity
ARMA	Autoregressive Moving Average
CLT	Central Limit Theorem
cdf	cumulative distribution function
EVT	Extreme Value Theory
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GEV	Generalized Extreme Value
GPD	Generalized Pareto Distribution
i.i.d	independent and identically distributed
VaR	Value-at-Risk

Master Thesis Proposal

Author	Krenar Avdulaj
Supervisor	PhDr. Jozef Baruník
Proposed topic	Value-at-Risk based on Extreme Value Theory method and copulas. Empirical evidence from Central Europe.

Topic characteristics In the financial world, for many years several Value-at-Risk risk models were used, and almost all of them seems that failed to predict the severe financial crises that started in 2007. Furthermore, models easy to implement based on historical simulation, or more sophisticated ones based on fat tail GARCH provide inconsistent results [DANIELSSON 2008]. However, as we know, a model is as good to predict as good the assumptions under which it is created are. Most of them are based on assumptions like stability of statistical properties of financial data [DANIELSSON 2008], or/and on normal distribution. The empirical models, on the other hand, show that the portfolio returns are not normally distributed. Their distribution have heavy tails and are leptokurtic. Thus, the risk measuring models that are based on this non-realistic assumption may be misleading for the decision maker.

In the recent years new methods and theories are developed, such as *Extreme Value Theory* and *copulas*. *EVT* is a branch of statistics that measures the extreme deviations of probability distribution from the mean. This Theory is not new though because it has been used before by hydrologists. However, it is relatively new methodology in financial risk management models and it is still in its experimental phases.

In this research we will empirically apply a market risk measuring model which employs Extreme Value Theory and t-Copulas. A portfolio of 5-6 indexes belonging to Central Europe Stock Markets will be constructed, and analyses of data over the last 9 years will be used as input in order to make a VaR measurement for a certain period of time.

Hypotheses The main question that this paper will try to answer is finding the VaR of a portfolio based on new methods like *EVT* and copulas. In focus will be Central European market.

Methodology Firstly, we apply a GARCH model (the specific parameters of the model will be chosen in a later step) on returns and residuals of this model will then be standardized (in this way we obtain *i.i.d* errors).

Secondly, we graphically characterize the distribution function of errors (will help us to have a clear view on our data distribution) by the following approach: The standardised *i.i.d* errors will be used to create empirical Cumulative Distribution Function based on Gaussian kernel. Based on general characteristics of financial time series the kernel CDF estimation is expected to be well suited for the interior of the distribution and performing poorly in lower and upper tail (this will be checked whether it will be true or not for our dataset). For this reason we will apply EVT for all observations that fall in each tail. We choose thresholds levels i.e such that 10% of data belong to each tail, and then fit the data that satisfy our condition (fall below threshold). This is also known as *peaks over thresholds* or *distribution of exceedances* method (DAVISON AND SMITH 1990).

Thirdly, the *i.i.d* variables obtained at the beginning will be used to estimate the parameters of t-copula. After obtaining the parameters of the t-copula we then are able to simulate jointly dependent equity index returns for a period of time for which we are interested to make the prediction (i.e one month).

Finally, we create the portfolio by putting the average weight of each index and report maximum gain or loss, as well as VaR of the portfolio at different confidence intervals.

Outline

1. Introduction
2. Theoretical Background on EVT.
3. The Model
4. Empirical Verification of the Model
5. Conclusion

In the first half of the thesis, there will be a theoretical representation of *EVT*, its applications and discussion. The copulas will not be in the focus of

this work, but of course there will be short introduction on their function. The second half will be focused on empirical work and fitting the model.

Core bibliography

1. BELTRATTI A. & C. MORANA (1999): "Computing value at risk with high frequency data." *Journal of Empirical Finance* (6): pp. 431-455.
2. HARVEY A.C. & N. SHEPHARD (1993): "Structural Time Series Models." *Handbook of Statistics* (11): pp. 261-302.
3. DANIELSSON J. & C. G. DE VRIES (1997): "Value-at-Risk and Extreme Returns." *London School of Economics and Institute of Economic Studies at University of Iceland; Tinbergen Institute and Erasmus University*.
4. DIEBOLD F.X., SCHUERMANN T. & J.D. STROUGHAIR (1998): "Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management." *The Wharton School, University of Pennsylvania*.
5. MCNEIL A.J. & F. RUDIGER (2000): "Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach." *Journal of Empirical Finance* (7): pp. 271-300.
6. DANIELSSON J. (2008): "Blame the Models." *Journal of Financial Stability* (4): pp. 321-328.
7. NYSTROM K. & J. SKOGLUND (2002): "Univariate Extreme Value Theory, GARCH and Measures of Risk." *Swedbank, Group Financial Risk Control*.
8. NYSTROM K. & J. SKOGLUND (2002): "A Framework for Scenario-Based Risk Management." *Swedbank, Group Financial Risk Control*.
9. BOUYE E., DURRLEMAN V., NIKEGHBALI A., RIBOULET G., & RONCALLI T. (2000): "Copulas for Finance A Reading Guide and Some Applications." *Financial Econometrics Research Centre, City University Business School, London. & Groupe de Recherche Operationnelle, Credit Lyonnais, Paris*

Chapter 1

Introduction

Value-at-Risk (VaR) is a risk measure tool which is based on loss distributions. After the market crash of 1987 and the increase of the off-balance-sheet products there was need for a proper risk management among banks. In October 1994 J.P.Morgan released *RiskMetricsTM* system which included the VaR as a risk measure [Linsmeier & Pearson (1996)], since then the VaR is an industry-wide standard. For many years it has played an important role in: risk management, risk measurement, financial control, financial reporting and evaluation of capital requirement. There is a vast number of sound articles and books written for Value-at-Risk. The list may be very long, but we can mention a few of them Best (1998); Duffie & Pan (1997); Das (2006); Holton (2003); Dowd (2000); Smithson & Lyle (1996).

VaR is a quantile based method that gives a single number as output. This number represents the risk of the portfolio for a certain holding period of time t at a particular confidence interval (probability). Without these two specifications, the risk horizon and probability, the VaR is meaningless.

Below we give a formal definition for VaR and in figure 1.1 we make an illustration.

Definition 1.1 (McNeil *et al.* (2005),page 38). The VaR of our portfolio at the confidence level $\alpha \in (0, 1)$ is given by the smallest number l such that the probability that the loss L exceeds l is no larger than $(1 - \alpha)$. Formally,

$$VaR_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} \quad (1.1)$$

Financial institutions' reports are usually calculated for $\alpha = 95\%$ or/and $\alpha = 99\%$, that is for confidence intervals 95% or 99% respectively. As re-

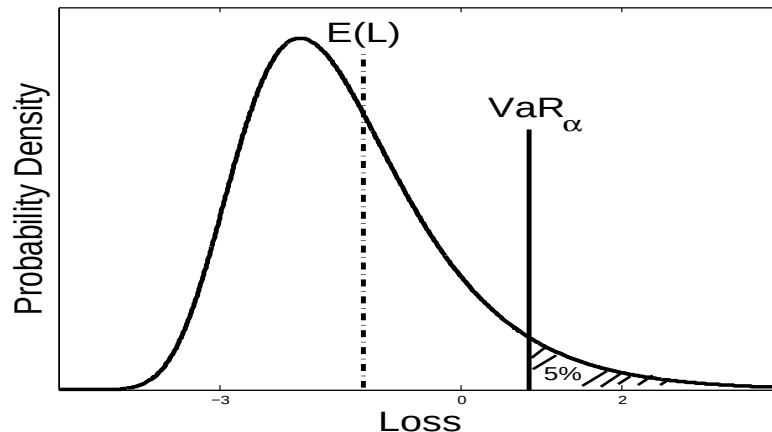


Figure 1.1: Value at Risk

gards the time horizon for which VaR is calculated it is usually 1 or 10 days (as required by Bank of International Settlements) in the case of market risk management and 1 year in the case of operational and credit risk management.

Methods used to estimate VaR

Historical Simulation

Historical simulation (HS) is a non-parametric and unconditional method. Its main advantage is that it is easy to implement and requires few assumptions. The key assumption is that the distribution of losses (or returns) in the portfolio is constant over the sample period and is a good predictor of future behaviour. For this method, i.e. if we invest in a simple stock or a portfolio, we take the historical returns (of the stock or of the portfolio) for the last (lets say) 100 days and sort from from the worst to the best return. If we want to make the valuation for a confidence interval i.e. 95% we state that the $VaR_{95\%}$ is equal to fifth lowest return. The interpretation of this value is "*with a probability 95% the loss for the next day will not be worse than the fifth lowest return*". However, we should be careful when applying this method because it is very sensitive to the length of data as the sample may not be a good representative of the proper distribution. In addition, with Historical Simulation we can not do out-of-sample predictions.

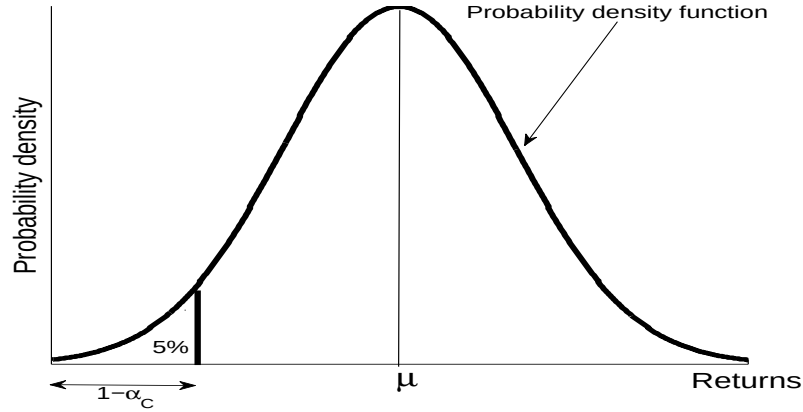


Figure 1.2: Value at Risk (returns distributed normally)

Variance-Covariance Method

This is a parametric method and can be either conditional or unconditional, based on the way we estimate the risk factor changes. More often, risk factors are logarithmic prices of financial assets, yields and logarithmic exchange rates. The method assumes that risk factor changes have multivariate normal distribution.

For a single asset the one day Value-at-Risk is calculated as: $VaR_\alpha = \alpha_c \sigma W_0$, where α_c is the value of the cut-off point (the inverse of standard normal cumulative distribution function) which is a function of confidence level c , σ is the daily standard deviation and W_0 the present value of the asset. If we want to calculate the VaR for T days, we multiply the one day VaR by $\sqrt{T}$, which is known as the time scaling component. For example, if we want to estimate the VaR for 5 days we multiply one day VaR by $\sqrt{5}$. If we want to calculate the VaR for a portfolio we just replace the standard deviation and the present value of the single asset with that of the portfolio $VaR_\alpha = \alpha_c \sigma_P W_P$ ($\sqrt{T}$ rule still applies). The advantage of this method is that the normal distribution assumption facilitates calculations. However the perfect symmetry of normal distribution is not in line with some type of risks (i.e. credit risk) and the thin tails of this distribution may not include extreme events, thus resulting in a underestimation of VaR.

Monte Carlo Simulation

The Monte Carlo method is based on simulations of an explicit parametric model for risk factor changes. The first step of this method is choice of the model and its calibration with historical data. Then we generate m realizations of the risk factor (m is the number of simulations). After we have the generated data we predict VaR on the simulated distribution at the required confidence interval. This can be done by empirical estimation, shortfall estimation or even by employing an extreme value model for the tails of the distribution. In general the EVT quantile estimators outperform the other methods [see Embrechts (2000a); Embrechts *et al.* (1999)]. The Monte Carlo simulation is sensitive to the model risk because if the stochastic process chosen in our model is not realistic so will be the VaR. Besides estimation of VaR, Monte Carlo simulation methods are used in option pricing, risk management and portfolio optimization. A good reference for Monte Carlo simulation methods in finance are Glasserman (2004) and Brandimarte (2006) (Chapter 4).

In the recent years academics have made extensive research on risk management. Artzner *et al.* (1999) introduce the concept of *coherent risk measures*. They propose that a risk measure is coherent if the following four properties are satisfied¹.

Consider an outcome $L \in \mathcal{L}$ (where $\mathcal{L}$ is a linear space of measurable functions defined on an appropriate probability space). The risk measure represented by function $\rho : \mathcal{L} \rightarrow \mathbb{R}$ is a coherent risk measure for $\mathcal{L}$ if the following properties hold:

- (*Translation Invariance*). $\forall L \in \mathcal{L}$ and every $l \in \mathbb{R}$, $\rho(L + l) = \rho(L) - l$. The risk of $\rho(L + l)$ is less than the risk of $\rho(L)$. The amount of cash l added to the position L acts like an insurance.
- (*Subadditivity*). $\forall L_1, L_2 \in \mathcal{L}$, $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$. The diversification principle. The risk of a merged portfolio cannot be greater than the sum of individual portfolio risks.
- (*Positive Homogeneity*). $\forall L \in \mathcal{L}$ and $\lambda \geq 0$ we have $\rho(\lambda L) = \lambda \rho(L)$. This property is similar to constant returns to scale. If you double your portfolio the risk will also double.

¹The authors in their work refer to as the *coherence axioms*. Because here we are not using rigorous formulation of these axioms we use the term *properties* instead of *axioms*.

- (*Monotonicity*). For $L_1, L_2 \in \mathcal{L}$ and $L_1 \leq L_2$ then $\rho(L_1) \leq \rho(L_2)$. Positions that lead to higher losses in every state of the world require more risk capital.

Artzner *et al.* (1997; 1999) criticize VaR because it exhibits poor aggregation properties. VaR suffers $q_\alpha(F_L) \leq q_\alpha(F_{L_A}) + q_\alpha(F_{L_B})$ (where $q_\alpha(\cdot)$ is the α quantile of F_L), the *Non-subadditivity* property, that is if we have two portfolios A and B with loss probability distributions F_{L_A} and F_{L_B} respectively and if we denote the total loss $L = L_A + L_B$ the inequality $q_\alpha(F_L) \leq q_\alpha(F_{L_A}) + q_\alpha(F_{L_B})$ does not hold all the times. The authors claim that there are cases where the VaR of total portfolio is higher than the sum of VaR of individual portfolios ($VaR_{A+B} \geq VaR_A + VaR_B$). In addition, they argue that VaR gives only an upper bound of losses that occur with a given frequency.

Another problem that one faces when implementing VaR analysis is the specification of probability distribution of losses $F_L(l)$. This problem sometimes is known as the *model risk*. The analyst may assume that the $F_L(l)$ follows normal distribution, while the real distribution may be heavy tailed (which is common in financial time series) and/or assymmetric.

In this work we should not worry about the satisfaction of the *subadditivity* property because it is usually not satisfied in portfolios that contain non linear derivatives. In our case the portfolio is composed of the same set of underlying elliptically distributed factors, and thus this property is satisfied [see McNeil *et al.* (2005), Theorem 6.8, p. 242].

The model that we estimate is mainly based on work of Nyström & Skoglund (2002a) and Nyström & Skoglund (2002b)². This is a two-step approach which enables the simulation of returns time series consistent with the historical performance and then computing the Value-at-Risk on the simulated returns series. We analyse the stock exchange indexes of 5 countries in the Central Europe for a period of ten years, from 01-Jan-2000 to 31-Dec-2009. The indexes that we use are Austria (ATX), Germany (DAX), the Czech Republic (PX50), Slovakia (SAX) and Switzerland (SSMI). The model that we estimate in this work can be seen as a univariate step and then the multivariate step. The univariate modelling involves ARMA-GARCH models and the Extreme Value Theory, while the multivariate one involves the copulas and their capability to conduct multivariate Monte Carlo simulations. Initially we filter the

²We have also consulted Matlab Digest article (July 2007) to implement the model in Matlab code. The article can be found on: <http://www.mathworks.com/company/newsletters/digest/2007/july/copulas.html>

log-returns by the asymmetric GARCH based on Glosten *et al.* (1993) model. The residuals of GARCH are standardized (by dividing with conditional standard deviations) and in this way we obtain approximately independent and identically distributed observations upon which the Extreme Value Theory estimation is based. This filtration procedure is in line with Diebold *et al.* (1998) suggestions how to implement the application of EVT in financial time series. At this moment we apply a semi-parametric cumulative distribution estimation, non-parametric for the internal part and parametric for the tails based on Extreme Value Theory. The parametric estimation for the tails makes possible the extrapolation, thus allowing to estimate quantiles outside the sample. Such a method of cumulative distribution estimation can be found in Daniélsson & de Vries (2000). The authors propose the semi-parametric method to estimate Value-at-Risk and compare it with historical simulation and RiskMetrics technique. They show that the semi-parametric method is more accurate than both other methods. The implementation of the Extreme Value Theory will be done through the *threshold exceedances* approach. Threshold exceedances and their distribution (distribution of exceedances), the Generalized Pareto Distribution, are treated thoroughly in Davison & Smith (1990). Other work in this direction, among the others, is done by McNeil & Frey (2000); McNeil *et al.* (2005) and Embrechts *et al.* (1997). After fitting the semi-parametric cumulative distribution function to the data we start the multivariate step by incorporating copulas.

The study of copulas and their role is a new growing field in statistics. Copulas are functions that join multivariate distribution functions to their one-dimensional margins [Nelsen (2006)]. This property, transforms the copulas into a bridge between univariate and multivariate distribution functions. We use the t copula for our analysis because this type is more appropriate due to their ability to better capture the phenomenon of extreme values [see Demarta, S. and McNeil, A. J. (2004)]. We transform standardized residuals to uniform variates and fit the copula to the transformed data. Having measured the dependence structure, we use copula capability to simulate uniform variates that preserve the measured dependence. The data are transformed back to standardized residuals form through the inverse cumulative distribution function that we measured for each index previously. We use the GARCH simulation on standardized residuals with the same parameters as the ones that we obtained earlier and thus obtain simulated returns. The Value-at-Risk can be calculated easily through the quantile estimation. We compare the estimate of

this approach with Historical Simulation and Variance Covariance method.

The paper is organized as follows: the first two chapters treat univariate models, in the fourth we introduce the multivariate model based on t copulas and in the fifth is the application of the model. In chapter 2 we introduce ARMA-GARCH models and the assymetric model that we use in our analysis. In chapter 3 we pay attention to the Extreme Value Theory and especially the *peaks-over-threshold* approach. Chapter 4 introduces copulas as a tool for multivariate modelling, and in chapter 5 we apply the model in practise and compare with Historical Simulation and Variance Covariance method.

Chapter 2

ARMA and GARCH Models in Financial Time Series

In this chapter we consider financial time series models, in particular logarithmic returns and the ARMA-GARCH model. In discrete time series the return for a time period T is calculated by $R_t = \sum_{t=1}^T R_t = \sum_{t=1}^T \frac{P_t - P_{t-1}}{P_t}$, where P_t is the price of the asset at time t . However if we decrease the time interval for which we calculate the return such that $\Delta t \rightarrow 0$ we get $R_T = \lim_{\Delta t \rightarrow 0} \sum_{t=1}^T \frac{\Delta P_t}{P_t}$. As $\Delta t \rightarrow 0$, $\Delta P_t \rightarrow dP$. Now we can extend the formula for the continuous time case and we can write $R_T = \int_{P_0}^{P_T} \frac{dP}{P} = \ln(P)|_{P_0}^{P_T} = \ln(P_T) - \ln(P_0)$.

We make a short review of ARMA and GARCH models which describe the dynamics of individual time series and then introduce the asymmetric ARMA-GARCH, a model that we will be used in our empirical work. We use the asymmetric ARMA-GARCH model to filter the time series in order to obtain independent and identically distributed (i.i.d) (or let better say approximately i.i.d) random variables which are necessary for the implementation of Extreme Value Theory which we introduce in Chapter 3.

2.1 ARMA Models

In financial time series modelling, the Autoregressive Moving Average (ARMA)(p, q) model is used extensively. This is a linear model and consists of two parts, the AR(p) and the MA(q) process. Before introducing these processes let us first denote the time series of logarithmic returns $(X_t)_{t \in \mathbb{Z}}$ which is defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We have the following linear processes:

- the AR(p) : $X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t$ where $\phi_1, \dots, \phi_p$ are the parameters

of the model, c is a constant and ϵ_t is white noise. The AR(p) process (without constant) can also be written as: $\phi(L)X_t = \epsilon_t$, where $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$. Based on the last equation and on *Wold* decomposition formula (Wold (1954)) we can write the AR(p) as an MA(∞) process: $X_t = \gamma(L)\epsilon_t$, where $\gamma(L) = (1 - \phi_1 L - \dots - \phi_p L^p)^{-1}$.

- the MA(q) : $X_t = d + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t$ where $\theta_1, \dots, \theta_q$ are the parameters of the model, d is a constant and ϵ_t is white noise. If we impose the condition $|\theta| < 1$ then we can express X_t as an AR(∞) process. This condition is known as *invertibility* condition. The MA models are weakly stationary by construction, they are finite linear combination of white noise.
- The ARMA(p, q) model is then: $X_t = c + \sum_{i=1}^p \phi_i X_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t$.

We say that ARMA(p, q) is a *causal* process if $X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}$ where ψ_i should satisfy $\sum_{i=0}^{\infty} |\psi_i| < \infty$. If we write the characteristic equation of the ARMA(p, q) in the complex plane:

$$\begin{cases} \tilde{\phi}(z) = 1 - \phi_1 z - \dots - \phi_p z^p \\ \tilde{\theta}(z) = 1 - \theta_1 z - \dots - \theta_q z^q \end{cases} \quad (2.1)$$

and provided that $\tilde{\phi}(z)$ and $\tilde{\theta}(z)$ have no common roots, the ARMA processes is causal if and only if $\tilde{\phi}(z)$ has no roots in the unit circle $|z| \leq 1$. So for the stationary ARMA process we should have $(\phi_1 + \dots + \phi_p) < 1$.

In time series analysis we may have to deal with a larger class of model known as ARIMA, where *I* stands for *integrated*. The ARIMA model has parameters (p, d, q) . Thus, the time series Y_t is said to follow an ARIMA(p, d, q) process if the differenced time series $X_t = \nabla^d Y_t$ follows an ARMA(p, q) process (the ∇^d is the difference operator of the order d). For example if the daily log-returns of a stock follows ARMA(p, q) process we can say that the stock log-price follows ARIMA($p, 1, q$)¹. The ARIMA processes where $d > 1$ are non-stationary. However, by differencing one time or more the time series may be transformed to stationary.

¹The log-returns time series is in fact a time series of log-prices differenced once ($d=1$). See the transformations at the beginning of the chapter.

2.2 GARCH Models

The Autoregressive Conditional Heteroskedasticity (ARCH) (Engle F. (1982)) is a non-linear model of modelling volatility. In this model the mean-corrected (returns) time series is serially uncorrelated, but dependent. This dependence is described by the ARCH(m):

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 \end{aligned} \quad (2.2)$$

where $a_t = X_t - \mu$ is the the mean corrected time series of log-returns and is strictly stationary, $\alpha_0 > 0$, $\alpha_i \geq 0$, ϵ_t is i.i.d random variable with mean 0 and variance 1 and m is the parameter of the model.

Looking at the ARCH model, the high volatility tends to be followed by high volatility and low volatility tends to be followed by low volatility, thus generating the volatility clusters. When we work with ARCH models we observe that positive and negative shocks have the same effect on volatility, however in empirical testing these effects are asymmetric. In addition ARCH models react slowly to large isolated shocks.

We apply the ARCH model to our data after removing any linear dependence via ARMA models and if the Langrange Multiplier test rejects the null hypothesis $H_0 : \alpha_i = 0$.

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (Bollerslev (1986)) is an extension of the ARCH model. The GARCH(m, s) is described by the following set of equations:

$$\begin{aligned} a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \end{aligned} \quad (2.3)$$

where $a_t = X_t - \mu$ is the the mean corrected time series and is strictly stationary, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, ϵ_t is i.i.d($0, 1$) random variable and m, s are the parameters of the model. In GARCH models the squared volatility depends on its squared past values and on the past squared value of the process. In the case when $s=0$ the GARCH is simply an ARCH model.

In times series modelling low order of GARCH, such as GARCH($1, 1$), GARCH($2, 1$) or GARCH($1, 2$) are very common. For GARCH($1, 1$) model the process is covariance-

stationary if and only if $\alpha_1 + \beta_1 < 1$, while for the GARCH(m, s) the condition is $\sum_{i=1}^m \alpha_i + \sum_{j=1}^s \beta_j < 1$.

In the same way as ARCH, the GARCH model introduced above (equation 2.3) is symmetric in a way that positive and negative shocks have the same impact on volatility. However, the empirical finding as Nyström & Skoglund (2002b) point out do not support this symmetry, and for this reason asymmetry is introduced in the model. The general case of asymmetric GARCH that we use (also known as *Threshold* GARCH) is obtained by introducing an additional parameter in the volatility equation [see Glosten *et al.* (1993) for details], and for ARMA(p, q)-GARCH(m, s) we have²:

$$\begin{aligned} X_t &= c + \sum_{k=1}^p \phi_k X_{t-k} + \sum_{l=1}^q \theta_l \epsilon_{t-l} + \epsilon_t \\ a_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i a_{t-i}^2 + \sum_{i=1}^m \psi_i \text{sgn}(a_{t-i}) a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \end{aligned} \quad (2.4)$$

where

$$\text{sgn}(a_{t-i}) = \begin{cases} 1 & \epsilon_{t-i} < 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$a_t = X_t - \mu, \quad \epsilon_t \text{ is i.i.d}(0, 1)$$

$$\left| \sum_{k=1}^p \phi_k \right| < 1 \quad k=1, 2, \dots, p$$

$$\left| \sum_{l=1}^q \theta_l \right| < 1 \quad l=1, 2, \dots, q$$

$$\sum_{i=1}^m \alpha_i + \frac{1}{2} \sum_{i=1}^m \psi_i + \sum_{j=1}^s \beta_j < 1$$

$$\alpha_0 > 0$$

$$\alpha_i \geq 0 \quad i=1, 2, \dots, m$$

$$\beta_j \geq 0 \quad j=1, 2, \dots, s$$

$$\alpha_i + \psi_i \geq 0 \quad i=1, 2, \dots, m$$

²In *Matlab*® GARCH toolbox this model is known as 'GJR' variance model, which comes from Glosten *et al.* (1993).

Chapter 3

Extreme Value Theory (EVT)

The EVT is a relatively new approach to risk management. This approach has been used for many years by hydrologists and lately is being applied in the insurance industries and in management of financial risks. EVT deals with events that are characterized by low frequency and high severity, that is the events that happen very rarely but their impact may be catastrophic.

There are two main approaches in EVT modelling, the *block maxima* and the modern approach of *threshold exceedances* (also known as *peaks over threshold*). The *block maxima*'s principle is dividing the time interval in equal chunks or blocks and modelling only the maximum loss for each of the blocks based on Generalized Extreme Value (GEV) distribution. This approach is considered wasteful of data because only one observation for each block is used. On the other hand the *threshold exceedances* approach models all the data that exceeds some predetermined "high level" (or *the threshold*). The data that exceed the threshold are fitted to the Generalized Pareto Distribution (GPD). Figure 3.1 illustrates both these methods. Modern risk management methods focus on the threshold exceedances approach. In our empirical work we will also use the *threshold exceedances* method.

In section 3.1 we give some short introduction on Generalized Extreme Value distribution, even if we do not use the block maxima approach in our work. We introduce the GEV distribution in order to make the reader familiar with Extreme Value Theory. In section 3.2 we treat the threshold exceedances approach, introduce the Generalized Pareto Distribution and analyse the threshold exceedances method.

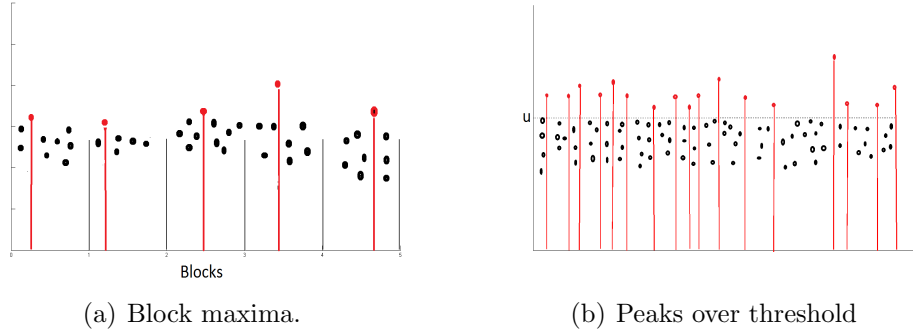


Figure 3.1: Block maxima and peaks over threshold method. The red dots are the observations used for modelling for each method.

3.1 Generalized Extreme Value (GEV) Distribution

Understanding of the GEV distribution is crucial in the extreme value theory. GEV distribution is of the same importance to the theory of extremes as the normal distribution is to the Central Limit Theorem (CLT) [McNeil *et al.* (2005) p.265]. Both, EVT and CLT tell us about the distribution in the limit. Before starting with GEV distribution let us first recall the CLT.

Let the sequence $X_1, \dots, X_n$ be n ($n \in \mathbb{N}$) observations of independent and identically distributed random variables. We denote the sum of observations $S_n = \sum_{i=1}^n X_i$. If the sum S_n is properly normalized it converges to the standard normal distribution.

Theorem 3.1 (Central Limit Theorem). *Suppose that $X_1, \dots, X_n$ are n independent and identically distributed random variables with mean μ and variance σ^2 . Then*

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = N(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-z^2/2) dz \quad (3.1)$$

On the other hand EVT gives similar results as CLT but for the maxima $M_n = \max\{X_1, \dots, X_n\}$. If we define $m_n = \min\{X_1, \dots, X_n\}$ we may write $m_n = -\max\{-X_1, \dots, -X_n\}$, that is the theory can be used even for describing m_n . Now we formally introduce the GEV distribution.

Definition 3.1 (Generalized Extreme Value pdf). The probability density function

of standard GEV distribution is given by function $h_{\xi,\beta,\mu}(x)$

$$h_{\xi,\beta,\mu}(x) = \begin{cases} \left(\frac{1}{\beta}\right) \left(1 + \xi \frac{x-\mu}{\beta}\right)^{(-1-\frac{1}{\xi})} \exp\left(-\left[1 + \xi \frac{x-\mu}{\beta}\right]^{-1/\xi}\right), & \xi \neq 0 \\ \left(\frac{1}{\beta}\right) \exp\left(-\exp\left[-\frac{(x-\mu)}{\beta}\right] - \frac{(x-\mu)}{\beta}\right), & \xi = 0 \end{cases} \quad (3.2)$$

where $1 + \xi \frac{x-\mu}{\beta} > 0$.

The parameter μ represents the location parameter, β represents the scale parameter whereas parameter ξ is the *shape* parameter ($1/\xi$ is sometimes known as the tail index). The *shape* parameter plays an important role because based on different values of this parameter the GEV distribution transforms in one of three classes of distributions. When $\xi = 0$ the distribution spreads out along all x axis and this type of distribution is known as *Gumbell* distribution. When $\xi < 0$ the distribution has upper bound and is known as *Weibull* distribution, and the last case where $\xi > 0$ the distribution has lower bound or belongs to *Fréchet* distribution. The probability density function of GEV is presented in figure 3.2(a)).

Definition 3.2 (The cumulative distribution function (CDF) of GEV). The distribution function of standard GEV distribution is given by

$$H_{\xi,\beta,\mu}(x) = \begin{cases} \exp\left(-\left[1 + \xi \frac{x-\mu}{\beta}\right]^{-1/\xi}\right), & \xi \neq 0 \\ \exp\left(-e^{-\frac{x-\mu}{\beta}}\right), & \xi = 0 \end{cases} \quad (3.3)$$

again, $1 + \xi \frac{x-\mu}{\beta} > 0$.

We can see the shape of cumulative distribution function of GEV for three different cases in figure 3.2(b).

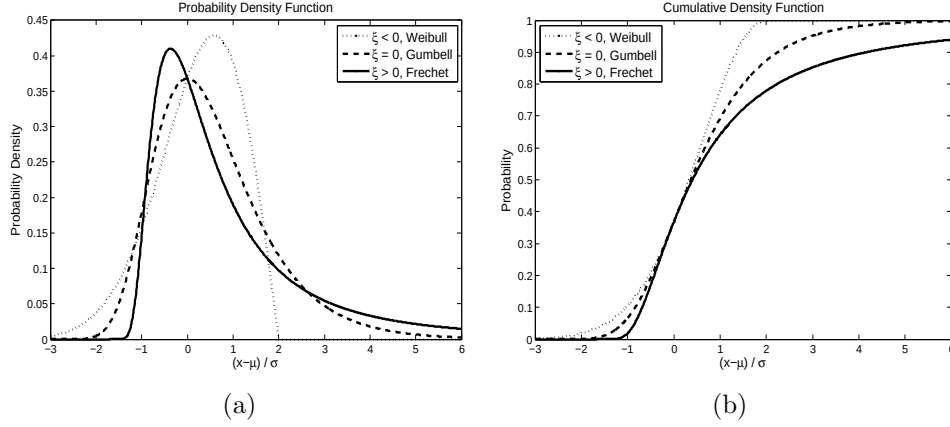


Figure 3.2: Probability density functions and cumulative distribution functions. $\xi = 0$ (Gumbell), $\xi = -0.5$ (Weibull) and $\xi = 0.5$ (Frechet). In all cases $\mu = 0$ and $\beta = 1$.

3.2 Threshold Exceedances

In the heart of threshold exceedances approach is the GPD. The data that exceed the threshold are modelled according to the GPD. Below is the formal definition of this distribution and its cdf.

Definition 3.3 (Generalized Pareto probability density function). Define the function $g_{\xi,\beta,\nu}(x)$,

$$g_{\xi,\beta,\nu}(x) = \begin{cases} \left(\frac{1}{\beta}\right) \left(1 + \xi \frac{x-\nu}{\beta}\right)^{-1-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ \left(\frac{1}{\beta}\right) e^{-\frac{(x-\nu)}{\beta}} & \text{if } \xi = 0 \end{cases} \quad (3.4)$$

$$\begin{aligned} \nu < x & \text{ when } \xi > 0, \text{ or } \nu < x < -\beta/\xi \text{ when } \xi < 0 \\ \nu < x, & \text{ when } \xi = 0 \end{aligned}$$

Definition 3.4 (Generalized Pareto cumulative distribution function (Embrechts *et al.* (1997), page 162)). Define the distribution function G_ξ by

$$G_\xi(x) = \begin{cases} \left(1 - [1 + \xi x]^{-1/\xi}\right), & \text{if } \xi \neq 0 \\ (1 - e^{-x}), & \text{if } \xi = 0 \end{cases} \quad (3.5)$$

where

$$\begin{aligned} x &\geq 0, & \text{if } \xi &\geq 0 \\ 0 \leq x &\leq -1/\xi, & \text{if } \xi &< 0 \end{aligned}$$

The location-scale family $G_{\xi;\beta,\nu}$ can be introduced by replacing x with $\frac{x-\nu}{\beta}$ for $\nu \in \mathbb{R}$, $\beta > 0$. In literature we can find the GPD in the form of

$$G_{\xi;\beta,0}(x) = \begin{cases} \left(1 - \left[1 + \xi \frac{x}{\beta}\right]^{-1/\xi}\right), & \text{if } \xi \neq 0 \\ \left(1 - e^{-\frac{x}{\beta}}\right), & \text{if } \xi = 0 \end{cases} \quad (3.6)$$

where $\beta > 0$ and for x the same conditions as in equation 3.5 apply (the location parameter $\nu = 0$). From equation 3.6 we see that the GPD is transformed in a distribution with two parameters, shape parameter ξ and scale parameter β ($G_{\xi,\beta}$).

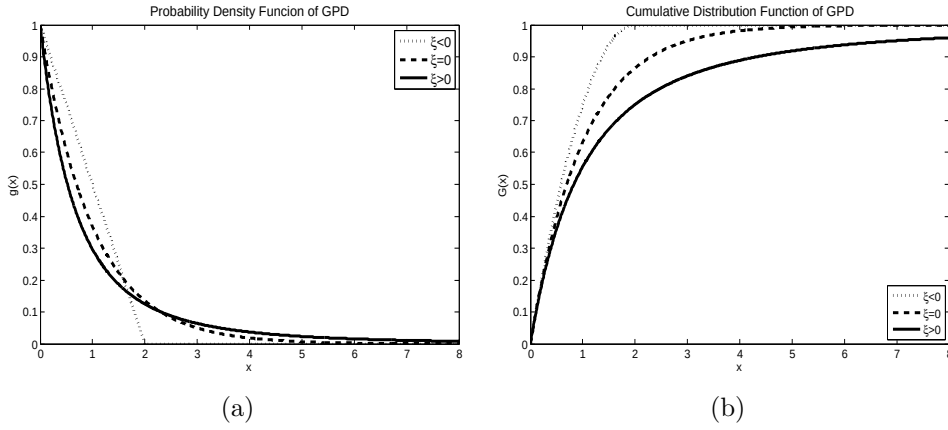


Figure 3.3: Probability density functions and cumulative distribution functions. $\xi = 0$ exponential, $\xi = -0.5$ Pareto type II and $\xi = 0.5$ Pareto distribution. In all cases $\beta = 1$ and $\nu = 0$.

The GPD includes three types of distribution as special cases. When $\xi > 0$ we have ordinary Pareto distribution, when $\xi = 0$ we have exponential distribution, and the case when $\xi < 0$ is known as Pareto-II type distribution. The ordinary Pareto distribution, where *shape parameter* $\xi > 0$, is most relevant in financial analysis since it is heavy tailed.

In threshold exceedances we model the excess distribution over the threshold u , like in figure 3.4. For a random variable X with distribution function F we formalize the conditional excess distribution over the threshold u :

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(y + u) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)} \quad (3.7)$$

for $0 \leq y < x_F - u$, where $x_F \leq \infty$ is the right endpoint of F and $y = x - u$.

The *mean excess function* plays an important role in EVT. For a random variable X with finite mean, the mean excess function for GPD is defined as:

$$e(u) = E(X - u | X > u) = \frac{\beta(u)}{1 - \xi} = \frac{\beta + \xi u}{1 - \xi} \quad (3.8)$$

where

$$\begin{cases} 0 \leq u < \infty & 0 \leq \xi < 1 \\ 0 \leq u < \frac{-\beta}{\xi} & \xi < 0 \end{cases}$$

From equation 3.8 we see that the mean excess is linear in u . This characteristic will be useful in determination of the threshold u when we estimate the shape and scale parameter for the GPD.

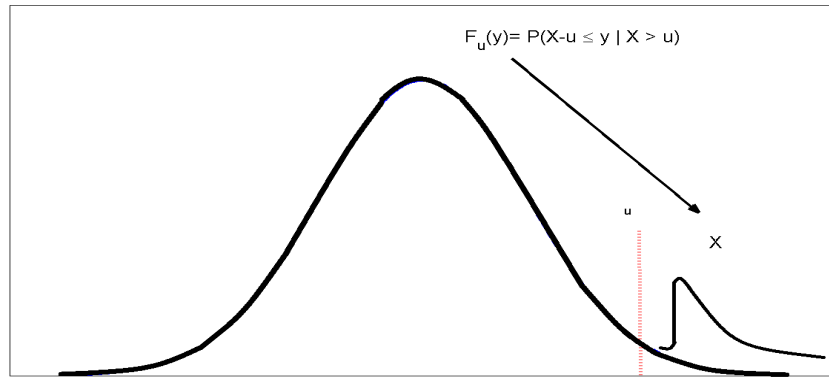


Figure 3.4: Conditional excess distribution over the threshold u . We use the EVT to model the observations which exceed the threshold u .

Balkema & de Haan (1974); Pickands (1975) discovered that the conditional excess distribution over the threshold has a nice property, the convergence to GPD.

Theorem 3.2 (Pickands-Balkema-de Haan). *For a large class of underlying distribution functions F the conditional excess distribution function $F_u(y)$, for u large, is well approximated to GPD*

$$F_u(y) \approx G_{\xi, \beta}(y) \quad \text{as } u \rightarrow \infty \quad (3.9)$$

Based on theorem 3.2 we make the following assumption:

Assumption 3.1. For some high threshold u from loss distribution F with right endpoint x_F we assume that $F_u(y) = G_{\xi, \beta}(y)$ for $0 \leq y < x_F - u$, $\xi \in \mathbb{R}$ and $\beta > 0$.

Using equation 3.7, assumption 3.1 and the empirical estimate $\frac{n - N_u}{n}$ for $F(u)$ [McNeil *et al.* (2005); Gilli & K llezi (2006)], where n is the total number of observations and N_u the number of observations over the threshold (consult appendix A for details) we can write:

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left[1 + \hat{\xi} \frac{x - u}{\hat{\beta}} \right]^{-\frac{1}{\hat{\beta}}} \quad (3.10)$$

Moreover, if we invert the equation 3.10 we get the high quantile or VaR of the underlying distribution. For $\alpha \geq F(u)$:

$$\widehat{VaR}_\alpha = q_\alpha(F) = u + \frac{\hat{\beta}}{\hat{\xi}} \left[\left(\frac{n(1 - \alpha)}{N_u} \right)^{-\hat{\xi}} - 1 \right] \quad (3.11)$$

Smith (1987) shows that estimates $\hat{\xi}$ and $\hat{\beta}$ are consistent and asymptotically normal as $N \rightarrow \infty$ and for $\xi > -1/2$. We recall the assumption that $X_1, \dots, X_n$ are realizations of independent random variables. We denote $K_1, \dots, K_{N_u}$ the data that exceed the threshold u and $Y_j = K_j - u$ for $j = 1, \dots, N_u$. We then calculate the log-likelihood by using the Generalized Pareto probability density function $g_{\xi, \beta}$ defined in equation 3.4 (with location parameter $\nu = 0$):

$$\begin{aligned} \ln L(\xi, \beta; Y_1, \dots, Y_{N_u}) &= \sum_{j=1}^{N_u} \ln g_{\xi, \beta}(Y_j) \\ &= \sum_{j=1}^{N_u} \ln \frac{1}{\beta} + \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{Y_j}{\beta} \right)^{-1 - \frac{1}{\xi}} \\ &= -N_u \ln \beta - \left(1 + \frac{1}{\xi} \right) \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{Y_j}{\beta} \right) \end{aligned} \quad (3.12)$$

To estimate the shape parameter $\hat{\xi}$ and scale parameter $\hat{\beta}$ we maximize the objective function in equation 3.12, of course subject to constraints $\beta > 0$ and $(1 + \xi \frac{Y_j}{\beta}) > 0, \forall j$.

As the EVT studies the tail distribution and the measurement of its thickness, one of the difficulties that rises is the determination of the amount of data

in the tail N_u or the start of the tail u . A very high u will include a few exceedances and will result in high variance estimators, while a small u will result in biased estimators. There is no such a rule for the choice of the threshold u , but a helpful tool is the mean excess plot $(u, e_n(u))$, where $x_1^n < u < x_n^n$. The empirical $e_n(u)$ is calculated like in Gilli & K llezi (2006):

$$e_n(u) = \frac{\sum_{i=k}^n (x_i^n - u)}{n - k + 1}, \quad k = \min\{i | x_i^n > u\} \quad (3.13)$$

where $n - k + 1$ is the number of observations exceeding u .

The rule on which we base the choice of threshold $u > 0$, is such that the mean excess is approximately linear for $x \geq u$ [see Embrechts *et al.* (1997), pp. 355]. The word *approximately* in this rule leaves place for interpretation, and for this reason we cannot expect a unique value of u . Figure 3.5 shows plots of empirical mean excess for ATX standardized residuals series¹. For the left tail this plot suggests a threshold level $u = 1.3814$ leaving 204 observations in the tail, while for the right tail $u = 1.3345$ leaving 204 observations in the tail.

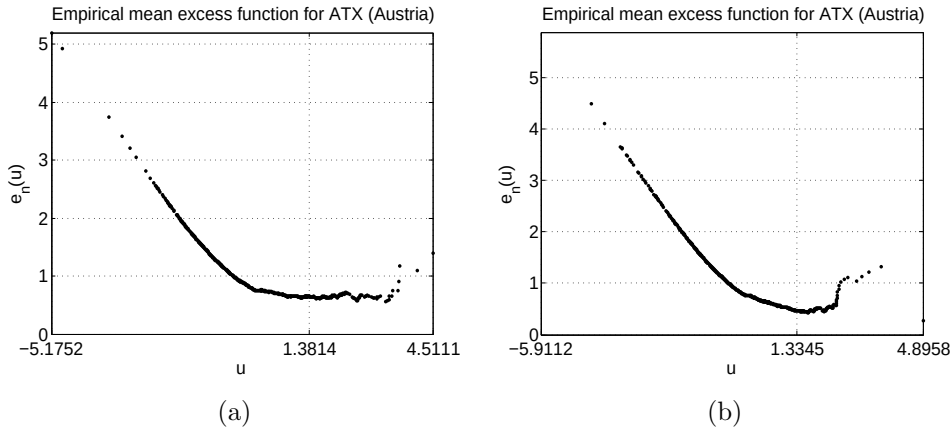


Figure 3.5: Empirical mean excess plot for the left tail (a) and the right tail (b) for ATX (Austria).

These values of threshold u suggest that we should reserve 8% of the data for each tail of the distribution. Thus only 16% of the total data are used to fit the Generalized Pareto distribution.

¹Standardized residuals are obtained by dividing residuals with conditional standard deviations of model 2.4, that is $\frac{a_t}{\sigma_t}$.

Chapter 4

Copulas

In this part we smoothly switch from univariate to multivariate modelling. We "combine" the univariate models that we obtained previously in Chapters 2 and 3, and create a multivariate model by the use of *copulas* (more specifically we will use the *t* copulas).

A copula is a multivariate probability distribution whose variables are uniformly distributed. Copulas are used to describe the dependence structure among elements of random vectors and are also used to build multivariate models, i.e. we can combine *univariate* distribution functions to construct *multivariate* distribution functions. Copulas are useful in financial risk management because by calibrating them on the dataset in which we are interested we measure the dependence structure between components of random vectors, i.e. vectors of financial risks factors. Furthermore, we can use copulas to simulate data with a specific dependence structure, thus helping to build multivariate risk models. Copulas application in finance have received much attention from academics in the recent years, and among the others we can mention Embrechts *et al.* (2003); Cherubini *et al.* (2004). This is because copulas have a wide use in financial industry, especially in credit scoring, risk management and derivative pricing.

4.1 Copulas

Definition 4.1 (Copula. McNeil *et al.* (2005), page 185.). A d -dimensional copula is a distribution function $[0, 1]^d$ with standard uniform marginal distributions. We reserve the notation $C(u) = C(u_1, \dots, u_d)$ for the multivariate distribution functions that are copulas. Hence C is a mapping of the form: $[0, 1]^d \rightarrow [0, 1]$,

i.e. a mapping of a unit hypercube into the unit interval. The following three properties must hold.

- $C(u) = C(u_1, \dots, u_d)$ is increasing in each component u_i .
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ for all $i \in \{1, \dots, d\}, u_i \in [0, 1]$.
- For all $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0, 1]^d$ with $a_i \leq b_i$ we have

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0 \quad (4.1)$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \dots, d\}$.

In our empirical work we will use some important properties of standard uniform distribution which are based on the following proposition.

Proposition 4.1. *Let G be a distribution function and let $G^{\leftarrow}$ denote its generalized inverse, that is $G^{\leftarrow}(y) = \inf\{x : G(x) \geq y\}$.*

- *Quantile transformation. If $U \sim U(0, 1)$ has a standard uniform distribution, then $P(G^{\leftarrow}(U) \leq x) = G(x)$.*
- *Probability transformation. If Y has a distribution function G , where G is a continuous univariate distribution function, then $G(Y) \sim U(0, 1)$.*

The relationship between the joint distribution and a copula is given by the Sklar's theorem.

Theorem 4.1 (Sklar 1959). *Let F be a joint distribution function with margins $F_1, \dots, F_d$. Then there exists a copula $C : [0, 1]^d \rightarrow [0, 1]$ such that, for all $x_1, \dots, x_d$ in $\mathbb{R} = [-\infty, \infty]$,*

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (4.2)$$

If the margins are continuous, then C is unique; otherwise C is uniquely determined on $\text{Ran}F_1 \times \text{Ran}F_2 \times \dots \times \text{Ran}F_d$, where $\text{Ran}F_i = F_i(\mathbb{R})$ denotes the range of F_i . Conversely, if C is a copula and $F_1, \dots, F_d$ are univariate distribution functions, then the function F defined in 4.2 is a joint distribution function with margins $F_1, \dots, F_d$.

This theorem is very important because it states that every multivariate distribution function has copulas and, the combination of copulas with univariate distribution functions can be used to obtain multivariate distribution functions.

Copulas can be divided in three categories: *fundamental*, *implicit* and the *explicit* copulas. For our purposes we will use implicit copulas, which are extracted from known multivariate distribution functions with the help of Sklar's theorem. For example, if we have a d -dimensional random Gaussian vector $X \sim N_d(\mu, P)$, then its copula is called *Gauss* copula. For the *Gauss* copula we have:

$$\begin{aligned} C_P^{Ga}(u) &= P(\Phi(X_1) \leq u_1, \dots, \Phi(X_d) \leq u_d) \\ &= \Phi_P(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)) \end{aligned}$$

where Φ is the standard normal univariate distribution function and Φ_P the respective joint distribution function. Meanwhile for the d -dimensional random vector from t distribution $X \sim t_d(\nu, 0, P)$, the t copula is defined as:

$$C_{\nu, P}^t(u) = t_{\nu, P}(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d))$$

where t_{ν} is the standard univariate t distribution function, P is the correlation matrix and $t_{\nu, P}$ is the joint distribution function of vector X with ν degrees of freedom. The t copula represents the dependence structure of a multivariate t distribution.

In figure 4.1 we plot a bivariate (2 dimensional) Gaussian copula and a bivariate t -copula in uniform scale (4.1(a), 4.1(c)), while in 4.1(b) and 4.1(d) we have transformed the simulated observations using the quantile function of the standard normal distribution.

We need to fit/calibrate copula to data in order to capture the dependence structure between components of random vectors, and to make the calibration we will employ log-likelihood method. However, before starting with the method we need to introduce the density of copula which we denote c

$$c(u_1, \dots, u_d) = \frac{\partial C(u_1, \dots, u_d)}{\partial u_1, \dots, u_d} \quad (4.3)$$

For the case of implicit copulas of an absolutely continuous joint distribution

function with strictly increasing and continuous marginal distribution functions [see McNeil *et al.* (2005), pp. 197] we can calculate the density like in equation 4.3:

$$\begin{aligned} c(u_1, \dots, u_d) &= \frac{\partial C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d} \\ &= \frac{\partial F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{\partial u_1 \dots \partial u_d} \\ &= \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))} \end{aligned} \quad (4.4)$$

where f is the joint density of F , $f_1, \dots, f_d$ are the marginal densities and $F_1^{-1}, \dots, F_d^{-1}$ are the inverses of marginal distribution functions. Figure 4.2 illustrates the probability density functions for Gaussian and t copula.

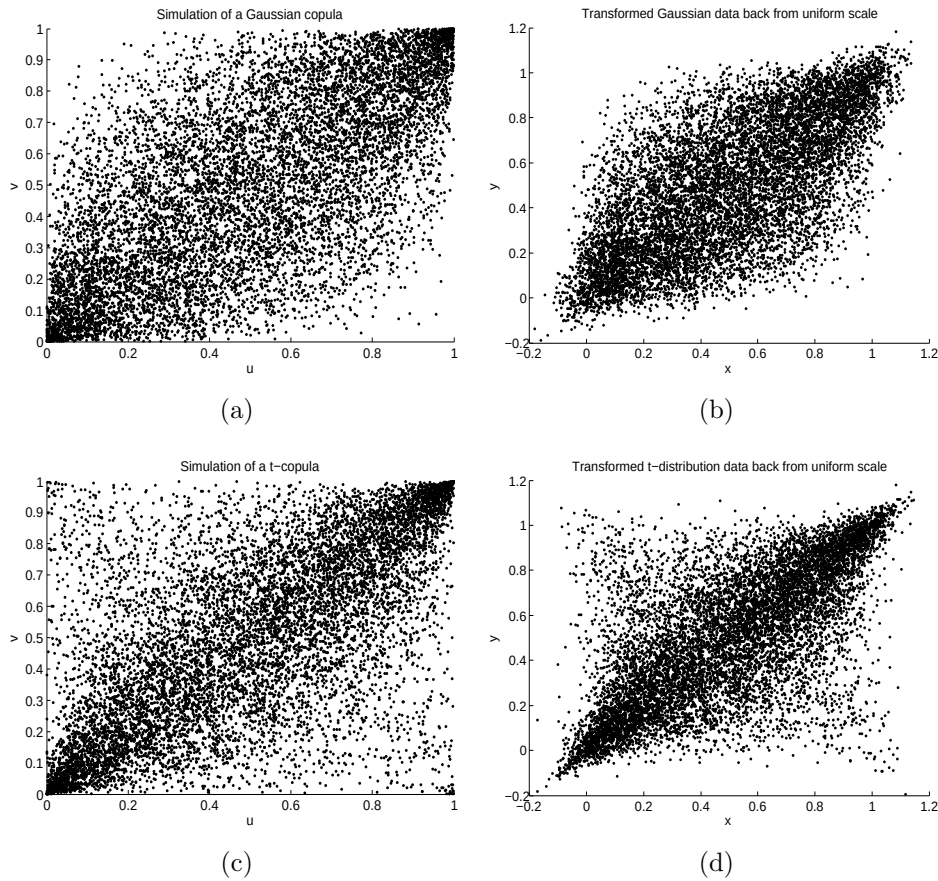


Figure 4.1: The Gaussian and t copulas for a bivariate sample of 10,000 observations in uniform scale (a and c) and in their original scale (b and d). The correlations are 0.7 in both cases, while for t copula the degrees of freedom are 2.

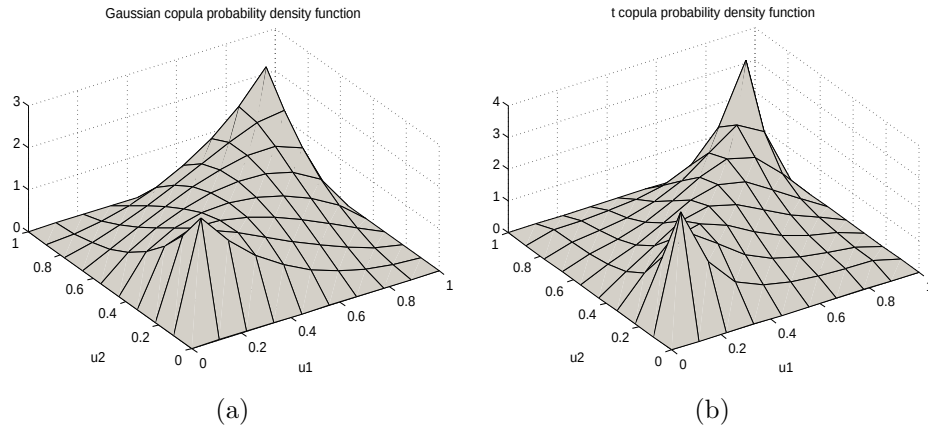


Figure 4.2: For both probability distribution functions the correlation is 0.7, while for t copula the degree of freedom is $\nu = 2$.

Suppose that $\{(X_{i1}, \dots, X_{id})^T, i = 1, \dots, n\}$ are n realizations from a multivariate distribution specified by d margins with cumulative probability distribution function F_i and probability density function $f_i, i = 1, \dots, n$, and a copula with density c . The parameter vector to be estimated is $\theta = (\beta^T, \alpha^T)^T$, where β is the vector of marginal parameters and α is the vector of copula parameters. The log-likelihood function then is given by:

$$l(\theta) = \sum_{i=1}^n \ln c\{F_1(X_{i1}; \beta), \dots, F_d(X_{id}; \beta); \alpha\} + \sum_{i=1}^n \sum_{j=1}^d \ln f_i(X_{ij}; \beta) \quad (4.5)$$

The maximum log-likelihood estimator $\hat{\theta}$ is

$$\hat{\theta} = \arg \max_{\theta \in \Theta} l(\theta),$$

where Θ is the parameter space.

In our model we will use maximum likelihood to calibrate the copula. We choose to use t copula, because for the case of financial returns data the t copula performs better than the Gaussian copula due to their ability to better capture the phenomenon of extreme values [see Demarta, S. and McNeil, A. J. (2004)].

Chapter 5

Application of the model

We create a multivariate market risk model based on stock exchange indexes of 5 countries in the Central Europe for a period of ten years, from 01-Jan-2000 to 31-Dec-2009. We choose this sample size because we are concerned to estimate the extreme outcomes, and in this situation it is suggested to use as large as possible sample with not lower than 1500 days [see Daníelsson & de Vries (2000)]. The construction of the model follows two steps: the first step is univariate modelling. The hypothetical equal weighted portfolio that we create includes indexes of: Austria (ATX), Germany (DAX), the Czech Republic (PX50), Slovakia (SAX) and Switzerland (SSMI)¹. Figure 5.1 is a representation of indexes price development for the last 10 years and of returns time series for Austria (ATX)² (the graphs for the other time series can be found in Appendix B). In figure 5.1(a) we see that there is high correlation between markets of Austria and the Czech Republic as well as between Germany and Switzerland. Other similarities, like mean, sign of skewness, minimum or maximum, for these markets can be obtained from summary statistics table 5.1. From that table we can see two groups of data with similar statistical properties, in one group are DAX and SSMI and in the other all the rest. In figure 5.1(b), we notice the high volatility levels after year 2008 which correspond to the start of credit crunch in the United States.

We filter the log-returns series using the asymmetric ARMA-GARCH model in equation 2.4 to remove the autocorrelation and to capture the conditional heteroskedasticity. The reason of using non-linear autoregressive model for

¹Data are downloaded from <http://www.stocktrading.cz> on 24 April 2010. We use *Matlab*® R2010a on *Windows*®7, 32-bit platform in this work. If different versions of either software's are used the output may be slightly different, but not changing the general result.

²We have chosen the Austria index ATX as an example to interpret the analysis. The comments and other interpretation for the other indexes are similar.

	Summary Statistics				
	ATX	DAX	PX50	SAX	SSMI
Mean	2.7801e-4	-7.2069e-5	3.2726e-4	4.8722e-4	-5.3415e-5
Std.Dev	0.01477	0.0164	0.015612	0.012101	0.012957
Skweness	-0.39699	0.042804	-0.49129	-0.13082	0.056544
Kurtosis	12.353	7.7345	16.111	15.153	9.4446
Minimum	-0.10253	-0.088747	-0.16185	-0.095775	-0.081081
Maximum	0.12021	0.10797	0.12364	0.1188	0.10788

Table 5.1: Summary Statistics. 01-Jan-2000 to 31-Dec-2009.

the returns series is motivated from figure 5.2(a), where the Autocorrelation Function (ACF) of squared returns signals persistence in variance. The non-linear model is supported from Q-test³ that we run on squared returns for each time series. This test confirms that there is autocorrelation in the second moment. Additionally, we run Engle's ARCH-test which confirms the presence of ARCH effects in our time series. We fitted the asymmetric model as in equation 2.4 to our dataset starting with low orders of GARCH and then selected the ones that fitted the data better. We tried models of different order for each time series and among the ones with significant coefficient we made the selection based on likelihood ratio test [see Hamilton (1994)]. Two were the competing models: ARMA(1,0)-GARCH(1,1) (which is the unrestricted model in the LR test) or the GARCH(1,1) (the restricted model). The LR test and the final parameters for each time series are summarized in table 5.2.

In the case of Austria, the Czech Republic and Slovakia the best fitting model is ARMA(1,0)-GARCH(1,1) while for the others it is GARCH(1,1). The residuals of the model are divided by conditional standard deviation, and thus obtaining standardized residuals which are approximately identically and independently distributed. If we compare figure 5.2(a) with 5.2(b) we see that all the autocorrelation is removed and data are approximately i.i.d. The Q-test on standardized squared residuals of each time series and the ARCH-test confirm that there is no autocorrelation and that the filtered residuals are i.i.d (there are no ARCH-effects). Now that we have the data in the required form for applying the Extreme Value Theory we can continue with the next step, the EVT application in the model.

³We run the test for 10, 15 and 20 lags. The same number of lags is used in Engle's ARCH-test.

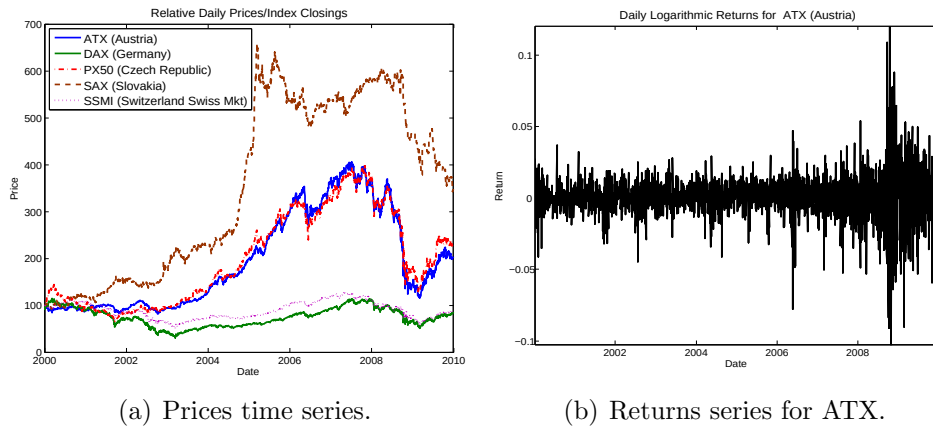


Figure 5.1: In (a) are prices of 5 indexes from Central Europe Financial Markets, Austria (ATX), Germany (DAX), Czech Republic (PX50), Slovakia (SAX) and Switzerland Swiss Mkt (SSMI). Prices are normalized to 100 to make comparable. In (b) are returns series for ATX.

	ATX	DAX	PX50	SAX	SSMI
c	7.3035e-4 (1.7838e-4)	2.4006e-4 (2.1798e-4)	7.9255e-4 (2.1384e-4)	3.8523e-4 (1.2898e-4)	2.4159e-4 (1.6579e-4)
ϕ_1	4.3378e-2 (2.0648e-2)	0 (0)	6.4525e-2 (2.068e-2)	-0.02908 (0.014651)	0 (0)
α_0	2.2946e-6 (5.6043e-7)	1.7961e-6 (4.1019e-7)	5.4863e-6 (1.2417e-6)	1.1164e-5 (4.0521e-6)	1.4483e-6 (3.1051e-7)
α_1	4.7939e-2 (1.5753e-2)	0 (1.2586e-2)	5.1687e-2 (1.5207e-2)	0.18157 (6.4762e-2)	0 (9.8912e-3)
ψ	8.3803e-2 (2.0355e-2)	0.13194 (0.01738)	0.10672 (2.3003e-2)	-0.076684 (4.1268e-2)	0.15309 (1.8373e-2)
β_1	0.89363 (1.3623e-2)	0.92288 (1.0653e-2)	0.86762 (1.6629e-2)	0.85678 (1.8809e-2)	0.90991 (1.0716e-2)
LR test (p-value)	0.0318	0.7022	0.0014	0.0299	0.8094

Table 5.2: The ARMA-GARCH models that best fitted our data are ARMA(1,0)-GARCH(1,1) for the case of Austria, the Czech Republic and Slovakia and constant conditional mean-GARCH(1,1) for the rest. Data in brackets represent standard errors.

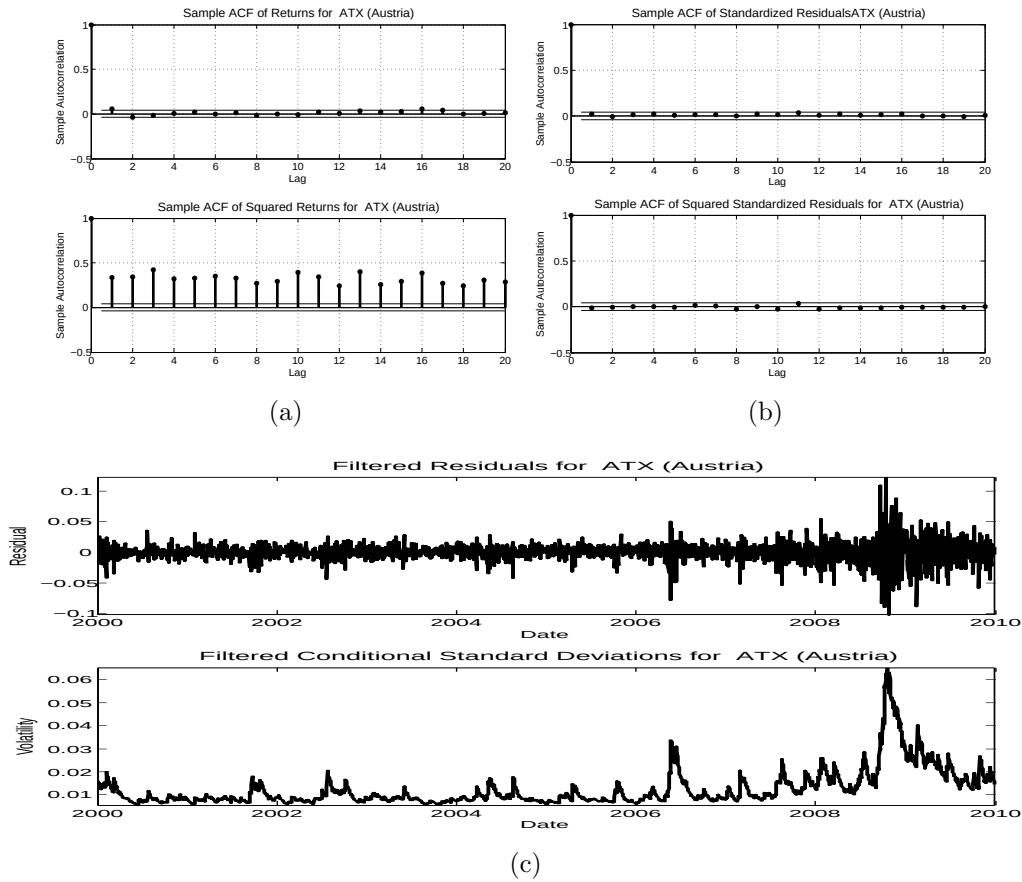


Figure 5.2: In (a) are ACF of raw returns and squared raw returns for ATX before filtering. In (b) are ACF of standardized residuals and squared standardized residuals for ATX, while in (c) are Residuals and conditional standard deviations for ATX.

We estimate the semi-parametric cumulative distribution function for the standardized residuals time series. The cumulative distribution function is semi-parametric because we fit a non-parametric kernel for the interior part of the distribution and a parametric GPD for the lower and upper tail of the distribution. In figure 5.3(b) we give the semi-parametric cumulative distribution function for the case of ATX, while in 5.3(a) and 5.3(c) we give a comparison between fitted and empirical cumulative distribution function for lower and upper tails (for the other indexes the graphs can be found in Appendix B). We see that the parametric Generalized Pareto distribution is a good fit for the tails, especially for the lower tail.

	ATX	DAX	PX50	SAX	SSMI
Left tail					
$\hat{\xi}$	0.0046	0.0307	0.0665	0.2072	-0.0079
$\hat{\beta}$	0.6356	0.5037	0.6048	0.6935	0.5756
$\widehat{VaR}_{99\%}$ (1 day)	2.7080	2.5842	2.7019	2.7313	2.6819
Right tail					
$\hat{\xi}$	0.0613	-0.2086	0.0753	0.1213	0.0104
$\hat{\beta}$	0.4361	0.4911	0.5410	0.5755	0.4032

Table 5.3: Point estimates of the tail parameters.

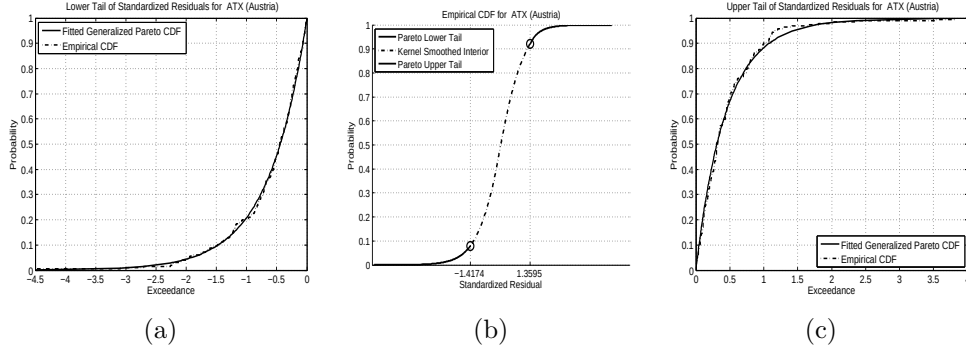


Figure 5.3: Semi-parametric cumulative distribution function (a) and fitted cumulative distribution function for upper tail (b) and lower tail (c).

We estimate the shape parameter $\hat{\xi}$ and scale parameter $\hat{\beta}$ for each index (time series) by maximizing the negative log-likelihood function in equation 3.12. The left tail and right tail estimates for shape and scale parameter are presented in table 5.3. Having the shape and scale parameter, as well as u , n and N_u that we obtained from empirical mean excess plot in section 3.2, it is possible to calculate the individual VaR for each index without involving any simulation⁴. For the left tail we can calculate the $\widehat{VaR}$ by using formula derived in equation 3.11. With probability 1% the next day loss will be greater than 2.7% for the ATX. (We calculated the VaR only for the left tail because we assume that we have a long position on the portfolio that we introduced earlier. However, it is easy to calculate the VaR for the right tail in the case we have also a short position.)

⁴We recall that our work is based on Monte-Carlo simulation, thus we will continue the development of the model further.

	ATX	DAX	PX50	SAX	SSMI
ATX	1.0000	0.5548	0.4661	0.0118	0.5144
DAX	0.5548	1.0000	0.4354	0.0307	0.7580
PX50	0.4661	0.4354	1.0000	0.0234	0.4138
SAX	0.0118	0.0307	0.0234	1.0000	0.0207
SSMI	0.5144	0.7580	0.4138	0.0207	1.0000

Table 5.4: Correlation matrix obtained from fitting t -copula to data.

Now we are done with univariate step and by the use of t copula we move to multivariate step. We transform the standardized residuals to uniform variates and calibrate the t copula to data by maximum log-likelihood method. In this way we capture the dependence structure between time series (the degrees of freedom and the linear correlation matrix). In table 5.3 we have the correlation matrix that we get from fitting the t copula while the degrees of freedom are 14.1805. After this step we use the copula capability to simulate random vectors with the same dependence structure as the one that we just calibrated. The uniform simulated variates are transformed to standardized residuals by inverting the semi-parametric cumulative distribution function that we fitted earlier. At this point we have obtained standardized residuals consistent with the ones that we obtained from ARMA-GARCH filtering process as in equation 2.4. We then introduce the autocorrelation and heteroskedasticity as in the original returns series via ARMA-GARCH simulation (consistent with parameters in table 5.2). In this way we simulated returns series that are consistent with historical performance.

Now we can continue with the last step, finding the Value-at-Risk of our portfolio. Unfortunately, there is one more thing that we should consider before calculating VaR. As we are working with log returns we have to be careful because log returns are time additive but not portfolio additive. On the other hand the simple returns are portfolio additive but not time additive. Thus when we construct the portfolio return series, we first convert the individual logarithmic returns to simple returns and multiply each return series with its weight in the portfolio. In this way we obtain the arithmetic return for the portfolio. Finally convert back portfolio return to logarithmic form. In order to see the reasoning clearly let us denote by $r_t = \log \frac{P_t}{P_{t-1}}$ the log returns , $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ the simple returns and w the weight of each index in the portfolio (w is a column vector). Let first convert from log return to simple return:

$$\begin{aligned}
r_t &= \log \frac{P_t}{P_{t-1}} \\
e^{r_t} &= \frac{P_t}{P_{t-1}} \\
e^{r_t} - 1 &= \frac{P_t}{P_{t-1}} - 1 \\
&= \frac{P_t - P_{t-1}}{P_{t-1}} \\
&= R_t
\end{aligned}$$

Here we weight the individual simple return for the portfolio at time t :

$$(e^{r_{i,t}} - 1) * w_i, \quad \text{where } i \in \{1, \dots, 5\} \quad \text{represents each index.}$$

Convert back to log returns and calculate the cumulative returns (the gain/loss during the risk horizon) which will be used to construct the empirical cumulative distribution function for the simulated returns:

$$\sum_{t=1}^T \sum_{j=1}^H \log(1 + (e^{r_{j,i,t}} - 1) * w_i)$$

where $i \in \{1, \dots, 5\}$ represents each index, H is the risk horizon in days and T is the length of the simulated time series (in our case $T = 8,000$).

In table 5.5 we have calculated the Value-at-Risk for our portfolio for a risk horizon of 1, 10 and 22 trading days, i.e. with a probability of 99% losses will not be higher than 10.14% in one month, while with probability 99.995% the losses will not be higher than 29.8%. For the same portfolio in table 5.5 we have summarized the results for two other methods of estimation of VaR, the simple Historical Simulation and the Variance-Covariance method. For these two methods we have calculated the VaR for one day and then we have used the scaling factor $\sqrt{T}$, where T is the number of days for which we want to make the estimation. This scale factor that we have used is the so called *square root of time method*, and requires that the portfolio value changes on successive days are identically and independently normally distributed with mean zero [see Hull (2006), Chapter 18 or Danielsson & de Vries (2000)]. If this requirement is violated the formula is an approximation. This requirement is in fact violated because portfolio returns are not i.i.d (we ran the ARCH test on mean corrected

portfolio returns to confirm it).

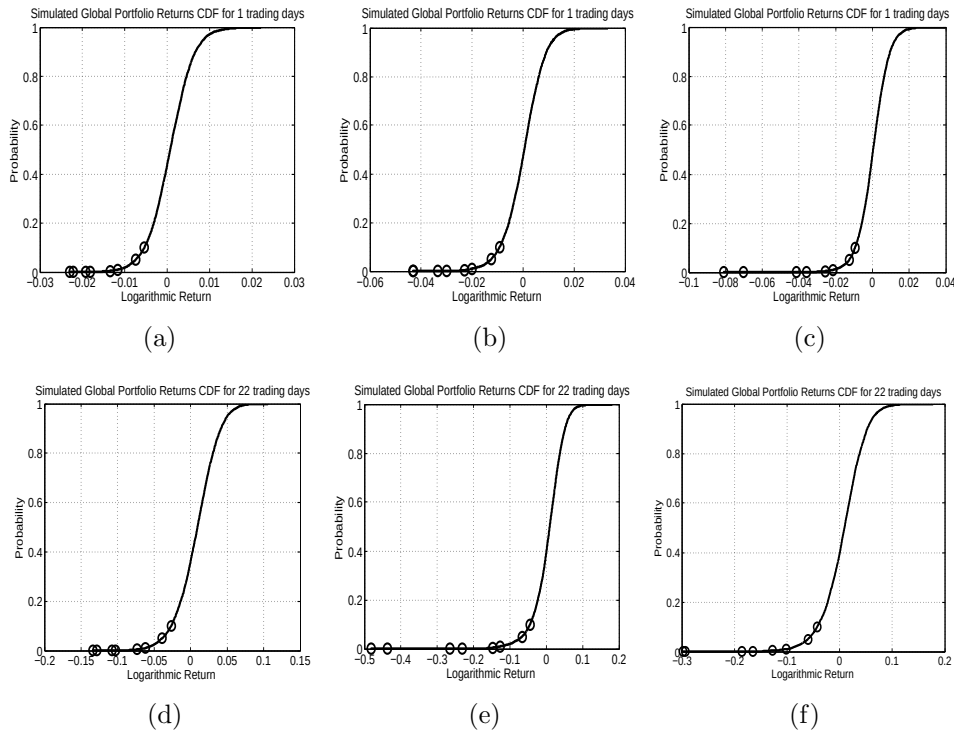


Figure 5.4: Empirical cumulative distribution function of simulated portfolio returns. Circles (from left to right) represent VaR estimates at 99.995%, 99.99%, 99.95%, 99.9%, 99.5%, 99%, 95% and 90% confidence intervals.

As expected, the historical simulation overpredicts the VaR for extreme quantiles in comparison with both other methods. This result is in line with Daniélsson & de Vries (2000) estimations. For extreme quantiles the historical simulation method gives the same estimates (i.e. at 99.99% or 99.995%) because this method can not give estimates outside the sample. The estimates for these quantiles correspond to the minimum value of the sample. Historical simulation method has another problem, it can not give estimates for probabilities lower than $\frac{1}{\text{sample size}}$.

VaR based on Monte-Carlo simulation and t -copula			
	1 trading day	10 trading days	22 trading days
90%	0.9525	2.9874	4.2784
95%	1.2672	4.1421	6.0302
99%	2.1446	6.6375	10.1483
99.5%	2.5625	7.8449	12.7315
99.9%	3.5893	10.4213	16.5097
99.95%	4.1429	12.3614	18.5877
99.99%	7.0327	15.2984	29.3324
99.995%	8.0879	16.1606	29.8068
VaR based on Historical simulation			
	1 trading day	10 trading days	22 trading days
90%	1.0411	3.2922	4.8832
95%	1.6165	5.1117	7.5818
99%	3.0863	9.7596	14.4758
99.5%	3.9033	12.3433	18.3080
99.9%	5.3589	16.9464	25.1356
99.95%	6.8103	21.5362	31.9433
99.99%	8.8696	28.0481	41.6021
99.995%	8.8696	28.0481	41.6021
VaR based on Variance-Covariance method			
	1 trading day	10 trading days	22 trading days
90%	1.2965	4.0999	6.0812
95%	1.6640	5.2622	7.8051
99%	2.3535	7.4424	11.0389
99.5%	2.6059	8.2405	12.2227
99.9%	3.1263	9.8862	14.6636
99.95%	3.3289	10.5270	15.6141
99.99%	3.7624	11.8978	17.6473
99.995%	3.9360	12.4467	18.4615

Table 5.5: Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2000-2009.

The variance-covariance method has the opposite problem of the historical simulation, it underpredicts VaR for extreme quantiles. In our tests up to 99.5% confidence interval this method gives higher estimates than Monte Carlo, but with small difference. When we test for quantiles higher than 99.5% the variance covariance seriously underestimates the VaR. At 99.995% confidence level this method gives an estimation such as 48% of EVT estimation, and even lower than the historical simulation. Our approach gives a result in between these two methods because it takes in consideration the historical performance

of the stocks but also employs a parametric method to correct for the fat tails. The EVT-Monte Carlo simulation that we have used has a drawback though, it assumes that the correlation among stocks (indexes in our case) is constant.

VaR based on Monte-Carlo simulation and t -copula			
	1 trading day	10 trading days	22 trading days
90%	0.5402	1.8373	2.6676
95%	0.7435	2.5385	3.8830
99%	1.1623	4.0472	6.2147
99.5%	1.3464	4.6864	7.3471
99.9%	1.8115	6.3608	10.3045
99.95%	1.9241	6.6832	10.7468
99.99%	2.2129	16.5106	12.8602
99.995%	2.3003	18.5950	13.3981

Historical Simulation			
	1 trading day	10 trading days	22 trading days
90%	0.9589	3.0324	4.4978
95%	1.3832	4.3741	6.4878
99%	2.3521	7.4379	11.0322
99.5%	2.6382	8.3429	12.3745
99.9%	3.4350	10.8625	16.1118
99.95%	3.8532	12.1850	18.0733
99.99%	3.9434	12.4702	18.4963
99.995%	3.9434	12.4702	18.4963

Variance-Covariance			
	1 trading day	10 trading days	22 trading days
90%	1.0683	3.3784	5.0109
95%	1.3712	4.3361	6.4315
99%	1.9393	6.1326	9.0962
99.5%	2.1473	6.7903	10.0717
99.9%	2.5761	8.1464	12.0830
99.95%	2.7431	8.6744	12.8662
99.99%	3.1003	9.8039	14.5416
99.995%	3.2433	10.2562	15.2125

Table 5.6: Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2000-2005.

In order to see that how these methods perform when we use samples of low volatility and high volatility we split the original data in two parts equally and run all the three methods on each part⁵. The first part includes data

⁵We are aware that the sub-samples size now is 1275, lower than the suggested minimum size of 1500 days.

from 2000-2005, a period in which the markets have been relatively stable. The second period from 2005-2009 is characterized from high volatility due the effects of financial crisis in the United States. The estimation of VaR for the period before crisis for all the three methods are summarized in table 5.6 while for the period after crisis in table 5.7.

VaR based on Monte-Carlo simulation and t -copula			
	1 trading day	10 trading days	22 trading days
90%	0.9146	2.9253	4.4578
95%	1.2403	4.1675	6.5669
99%	2.0031	7.2532	12.7604
99.5%	2.3006	8.5736	14.7838
99.9%	2.9889	11.4092	23.1007
99.95%	3.3438	11.6678	26.4669
99.99%	4.2977	14.2809	43.7527
99.995%	4.3082	14.4133	48.1122
Historical Simulation			
	1 trading day	10 trading days	22 trading days
90%	1.1765	3.7203	5.5181
95%	1.8238	5.7674	8.5545
99%	3.7811	11.9568	17.7347
99.5%	4.6995	14.8612	22.0427
99.9%	6.8077	21.5278	31.9308
99.95%	8.5029	26.8886	39.8822
99.99%	8.8696	28.0481	41.6021
99.995%	8.8696	28.0481	41.6021
Variance-Covariance			
	1 trading day	10 trading days	22 trading days
90%	1.4899	4.7115	6.9882
95%	1.9123	6.0471	8.9693
99%	2.7046	8.5525	12.6855
99.5%	2.9946	9.4697	14.0459
99.9%	3.5926	11.3609	16.8509
99.95%	3.8255	12.0972	17.9431
99.99%	4.3236	13.6725	20.2796
99.995%	4.5231	14.3033	21.2152

Table 5.7: Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon. Sample period 2005-2009.

During the quiet times the EVT underpredicts the VaR compared with the other methods in all but 10 days interval and quantiles $\geq 99.99\%$. We see from table 5.7, in the EVT section the strange result where $VaR_{10d,99.99\%} >$

$VaR_{22d,99.99\%}$ and $VaR_{10d,99.995\%} > VaR_{22d,99.995\%}$. However, for the EVT estimation for 10 days if we run the simulation again for several times (without resetting the random number generator) we get estimations around 10% for the 99.995% confidence interval. Thus, the high value that we have there is just because the generator randomly draws such numbers. For high volatility periods and for risk horizon more than 10 days the EVT estimates are higher. For 22 days risk horizon the VaR estimates have a value of 48% at 99.995% confidence level, which means that EVT is a good candidate to capture very rare events. Such events where the losses will be 48% in on month (22 trading days) are very rare. These results support the EVT as a risk measurement tool for the extreme quantiles (e.g. one tenth of one percent). We should note that in this estimation process the portfolio weights are held fixed throughout the risk horizon, and that any transaction costs required to rebalance the portfolio are ignored.

Chapter 6

Conclusion

We have illustrated the implementation of Extreme Value Theory as a tool in risk measurement in a multivariate distribution framework. There are different approaches to Value-at-Risk estimation, and most of them wrongly assume that stock returns follow normal distribution or multivariate normal distribution in the case of a portfolio of stocks. The two step approach that we illustrated is a semi-parametric method that uses the non-parametric empirical distribution to capture the small risks and the parametric method based on Extreme Value Theory to capture large risks. The use of Extreme Value Theory in the model improves the estimation of VaR for extreme quantiles because except modelling the fat tails it allows for extrapolation in the tails beyond the range of data. On the other hand the use of t copulas makes the connection between univariate and multivariate distribution and helps to conduct Monte Carlo simulations. It takes into consideration the historical dependence structure with the respective degrees of freedom as well as simulates multivariate t distributed observations. With the late developments in the computing capabilities of computers Monte Carlo simulations for larger portfolios are easy to implement.

Our conclusion is that the Extreme Value Theory is a good risk measurement tool for extreme events and especially for high volatility times. In our estimation process, for high volatility samples and for 22 trading days risk horizon, we got an VaR estimate of 48% at 99,995% confidence level which is reasonable if we compare with the losses suffered in the US markets in 2007-2008. For routine risk estimation i.e. 90% or 95% confidence intervals the simple methods of historical simulations and variance-covariance may provide good VaR estimations too.

We suggest that further work needs to be done to test the sensitivity of this

model based on the choice of threshold level u . An other point of interest may be the sensitivity analysis based on the choice of degrees of freedom of t copula when we make Monte Carlo simulations.

Bibliography

- ARTZNER, P., F. DELBAEN, J. EBER, & D. HEATH (1997): “Thinking coherently.” *Risk* **10(11)**: pp. 68–71.
- ARTZNER, P., F. DELBAEN, J. EBER, & D. HEATH (1999): “Coherent measures of risk.” *Mathematical Finance* **9(3)**: pp. 203–228.
- BALKEMA, A. A. & L. DE HAAN (1974): “Residual life time at great age.” *The Annals of Probability* **2(5)**: pp. 792–804.
- BEST, P. (1998): *Implementing Value at Risk*. Baffins Lane Chichester, West Sussex PO19 1UD, England: John Willey & Sons.
- BOLLERSLEV, T. (1986): “Generalized autoregressive conditional heteroskedasticity.” *Journal of Econometrics* **31(3)**: pp. 307–327.
- BRANDIMARTE, P. (2006): *Numerical methods in finance and economics: a MATLAB-based introduction*. John Wiley & Sons, New York.
- BRIGO, D., A. DALESSANDRO, M. NEUGEBAUER, & F. TRIKI (2009): “A stochastic processes toolkit for risk management: Mean reverting processes and jumps.” *Journal of Risk Management in Financial Institutions* **3(1)**: pp. 65 – 83.
- CHERUBINI, U., E. LUCIANO, & W. VECCHIATO (2004): *Copula Methods in Finance*. John Wiley and Son Ltd.
- DANÍELSSON, J. (2008): “Blame the models.” *Journal of Financial Stability* **4(4)**: pp. 321–328.
- DANÍELSSON, J. & C. G. DE VRIES (2000): “Value-at-Risk and Extreme Returns.” *Annales d’Economie et de Statistique* **60**: pp. 239–270.
- DAS, S. (2006): *Risk Management*. 2 Clementi Loop, #02-01, Singapore 129809: John Willey & Sons (Asia).

- DAVISON, A. C. & R. L. SMITH (1990): "Models for exceedances over high thresholds." *Journal of the Royal Statistical Society. Series B (Methodological)* **52(3)**: pp. 393–442.
- DEMARTA, S. AND MCNEIL, A. J. (2004): "The t copula and related copulas." *Technical report*, Department of Mathematics, Federal Institute of Technology, ETH Zentrum CH-8092 Zurich.
- DIEBOLD, F. X., T. SCHUERMANN, & J. D. 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.
- DOWD, K. (2000): "Estimating value at risk: A subjective approach." *The Journal of Risk Finance* **1(4)**: pp. 43–46.
- DUFFIE, D. & J. PAN (1997): "An overview of value at risk." *The Journal of Derivatives* **4(3)**: pp. 7–49.
- DURRLEMAN, V., A. NIKEGHBALI, & T. RONCALLI (2000): "Which copula is the right one?" *Technical report*, Groupe de Recherche Opérationnelle Crédit Lyonnais, France.
- EMBRECHTS, P. (2000a): "Extreme value theory: Potentials and limitations as an integrated risk management tool." *Technical report*, Department of Mathematics, ETH, Swiss Federal Technical University, Zurich, Switzerland.
- EMBRECHTS, P., C. KLÜPPELBERG, & T. MIKOSCH (1997): *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer.
- EMBRECHTS, P., F. LINDSKOG, & A. MCNEIL (2003): "Modelling Dependence with Copulas and Applications to Risk Management." In S. RACHEV (editor), "Handbook of Heavy Tailed Distributions in Finance," chapter 8, pp. 329–384. Elsevier.
- EMBRECHTS, P., S. RESNICK, & G. SAMORODNITSKY (1999): "Extreme value theory as a risk management tool." *North American Actuarial Journal Society of Actuaries* pp. 30–41.
- ENGLE F., R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* **50(4)**: pp. 987–1007.

- GENÇAY, R. & F. SELÇUK (2004): “Extreme value theory and value-at-risk: Relative performance in emerging markets.” *International Journal of Forecasting* **20(2)**: pp. 287 – 303. Forecasting Economic and Financial Time Series Using Nonlinear Methods.
- GILLI, M. & E. KËLLEZI (2006): “An application of extreme value theory for measuring financial risk.” *Computational Economics* **27(2)**: pp. 207–228.
- GLASSERMAN, P. (2004): *Monte Carlo Methods in Financial Engineering*, volume 53 of *Stochastic Modelling and Applied Probability*. Springer-Verlag.
- GLOSTEN, L. R., R. JAGANNATHAN, & D. E. RUNKLE (1993): “On the relation between the expected value and the volatility of the nominal excess return on stocks.” *The Journal of Finance* **48(5)**: pp. 1779–1801.
- HAMILTON, J. D. (1994): *Time Series Analysis*. Princeton University Press.
- HARVEY, A. C. & N. SHEPHARD (1993): “Structural time series models.” In G. S. MADDALA, C. R. RAO, & H. D. VINOD (editors), “Handbook of Statistics,” volume 11. Elsevier.
- HOLTON, G. A. (2003): *Value-at-Risk, Theory and Practice*. Academic Press.
- HULL, J. C. (2006): *Options, Futures and Other Derivatives*. Upper Saddle River, New Jersey, 07458: Prentice Hall.
- LINSMEIER, T. L. & N. D. PEARSON (1996): “Risk measurement: An introduction to value at risk.” *Technical report*, Department of Accountancy and Department of Finance, University of Illinois at Urbana-Champaign.
- MASHAL, R. & A. ZEEVI (2002): “Beyond correlation: Extreme co-movements between financial assets.” *Technical report*, Columbia University.
- 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, A. J., R. FREY, & P. EMBRECHTS (2005): *Quantitative Risk Management: Concepts, Techniques and Tools*. 3 Market Place, Woodstock, Oxfordshire OX20 1SY, United Kingdom: Princeton University Press.
- NELSEN, R. B. (2006): *An Introduction to Copulas*. Springer Series in Statistics. Springer.

- NYSTRÖM, K. & J. SKOGLUND (2002a): “A Framework for Scenario-Based Risk Management.” *Technical report*, Swedbank, Group Financial Risk Control, S-105 34 Stockholm, Sweden.
- NYSTRÖM, K. & J. SKOGLUND (2002b): “Univariate Extreme Value Theory, GARCH and Measures of Risk.” *Technical report*, Swedbank, Group Financial Risk Control, S-105 34 Stockholm, Sweden.
- PICKANDS, J. (1975): “Statistical inference using extreme value order statistics.” *The Annals of Statistics* **3**(1): pp. 119–131.
- SMITH, R. L. (1987): “Estimating tails of probability distributions.” *The Annals of Statistics* **15**(3): pp. 1174–1207.
- SMITHSON, C. & M. LYLE (1996): “Value at risk.” *Risk* pp. 25–27.
- WOLD, H. O. A. (1954): *A Study in the Analysis of Stationary Time Series*. Uppsala: Almqvist and Wiksell Book Co.

Appendix A

Appendix One

Derivation of VaR.

First we define the tail of the distribution function $F : \bar{F} = 1 - F$. If we denote the number of observations n and the number of observations that exceed the threshold u by N_u , we can write the empirical estimate of $F(u)$ equal to $\frac{n-N_u}{n}$. Embrechts *et al.* (1997) p. 354 suggests using the empirical distribution function to estimate $\bar{F}(u) = \frac{N_u}{n}$ or if we calculate for $F(u)$: $\bar{F}(u) = 1 - F(u) = \frac{N_u}{n} \rightarrow F(u) = \frac{n-N_u}{n}$. We also use theorem of *Pickands-Balkema-de Haan* based on which we replace $F_u(y) = G_{\xi,\beta}(y)$. Thus, from equation 3.7 we have:

$$\begin{aligned}\hat{F}(x) &= F(u) + F_u(y)\bar{F}(u) \\ &= 1 - \bar{F}(u) + F_u(y)\bar{F}(u) \\ &= 1 + \bar{F}(u)[F_u(y) - 1] \\ &= 1 + \frac{N_u}{n} \left[1 - \left(1 + \hat{\xi} \frac{x - u}{\hat{\beta}} \right)^{\frac{-1}{\hat{\xi}}} - 1 \right] \\ &= 1 - \frac{N_u}{n} \left[1 + \hat{\xi} \frac{x - u}{\hat{\beta}} \right]^{\frac{-1}{\hat{\xi}}}\end{aligned}\tag{A.1}$$

Now we invert the result in A.1 to obtain the high quantile estimator or the VaR for $\alpha \geq \frac{n-N_u}{n}$. So,

$$\begin{aligned}\alpha &= 1 - \frac{N_u}{n} \left[1 + \hat{\xi} \frac{q_\alpha(F) - u}{\hat{\beta}} \right]^{\frac{-1}{\hat{\xi}}} \\ \left[1 + \hat{\xi} \frac{q_\alpha(F) - u}{\hat{\beta}} \right]^{\frac{-1}{\hat{\xi}}} &= \frac{n}{N_u}(1 - \alpha)\end{aligned}$$

$$\begin{aligned}
\hat{\xi} \frac{q_\alpha(F) - u}{\hat{\beta}} &= \left[\frac{n}{N_u} (1 - \alpha) \right]^{-\hat{\xi}} - 1 \\
q_\alpha(F) - u &= \frac{\hat{\beta}}{\hat{\xi}} \left[\left[\frac{n}{N_u} (1 - \alpha) \right]^{-\hat{\xi}} - 1 \right] \\
\widehat{VaR}_\alpha = q_\alpha(F) &= u + \frac{\hat{\beta}}{\hat{\xi}} \left[\left[\frac{n}{N_u} (1 - \alpha) \right]^{-\hat{\xi}} - 1 \right] \tag{A.2}
\end{aligned}$$

Appendix B

Appendix Two

In this appendix we show figures and tables for time series other than Austria (ATX).

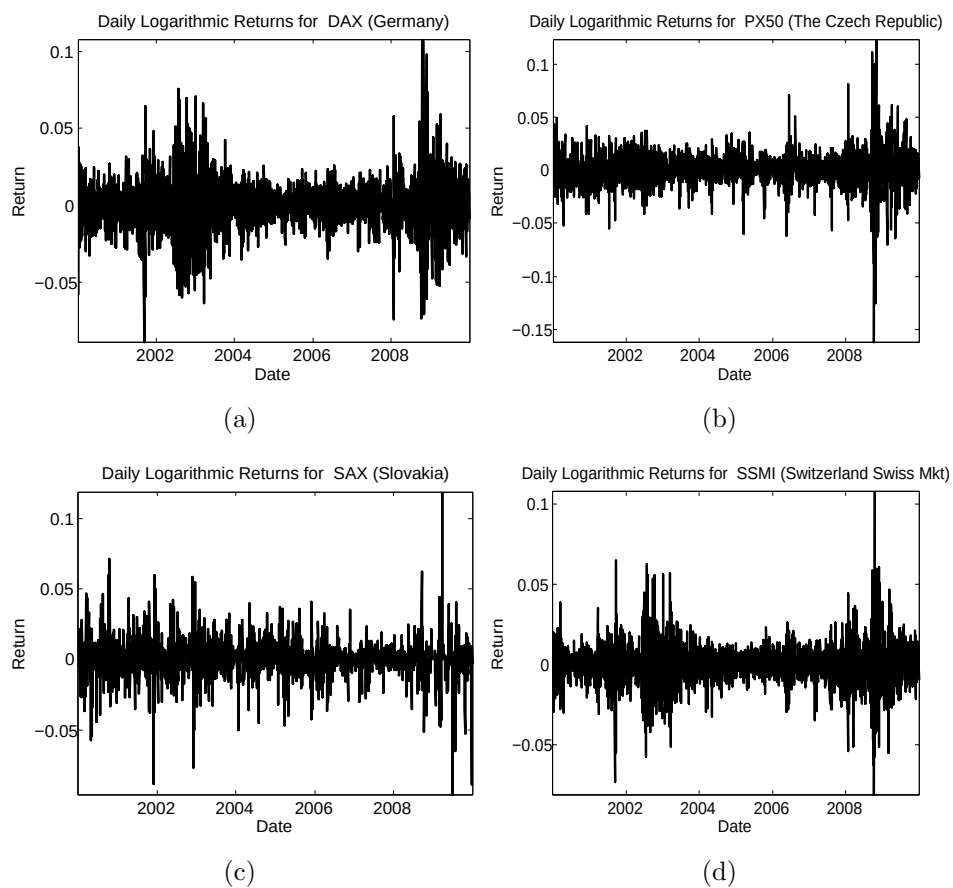


Figure B.1: Returns series.

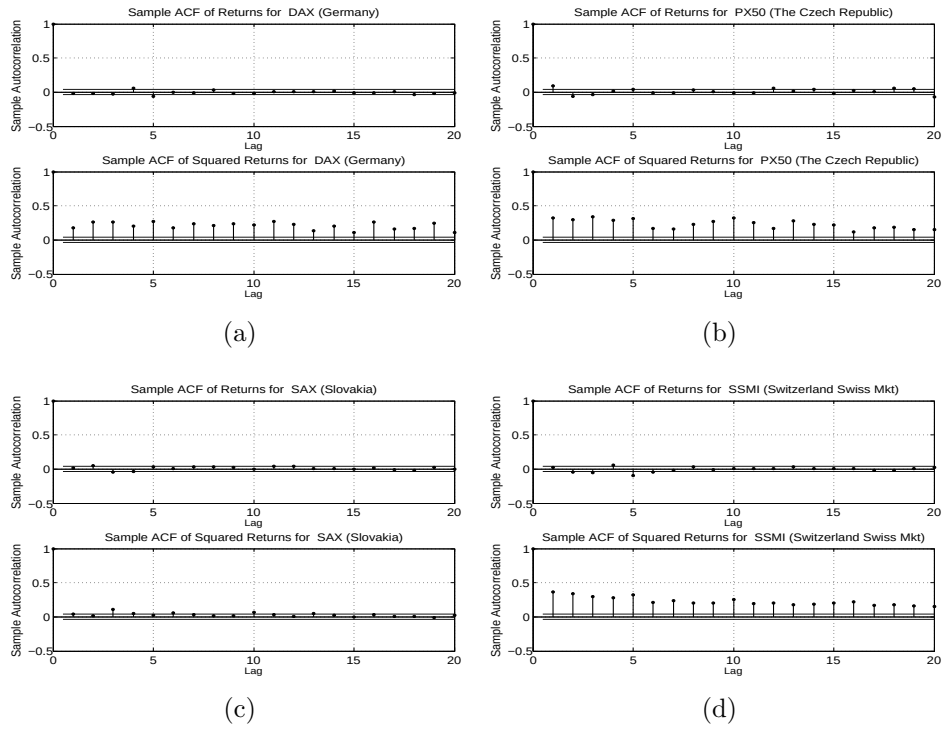


Figure B.2: Autocorrelation of return series.

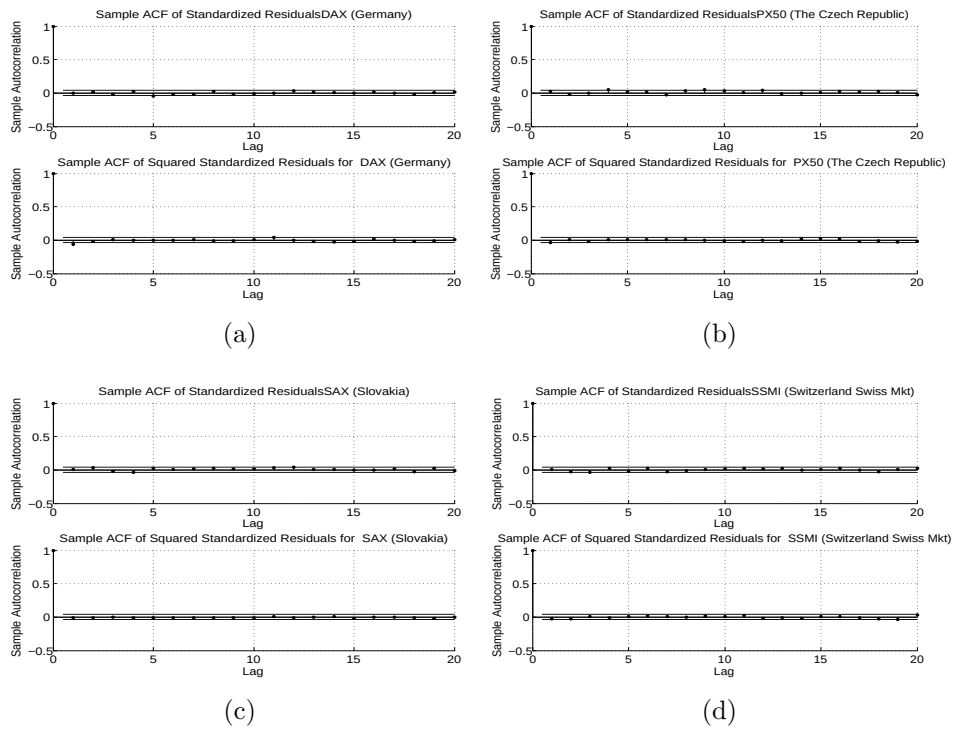


Figure B.3: Autocorrelation after filtering.

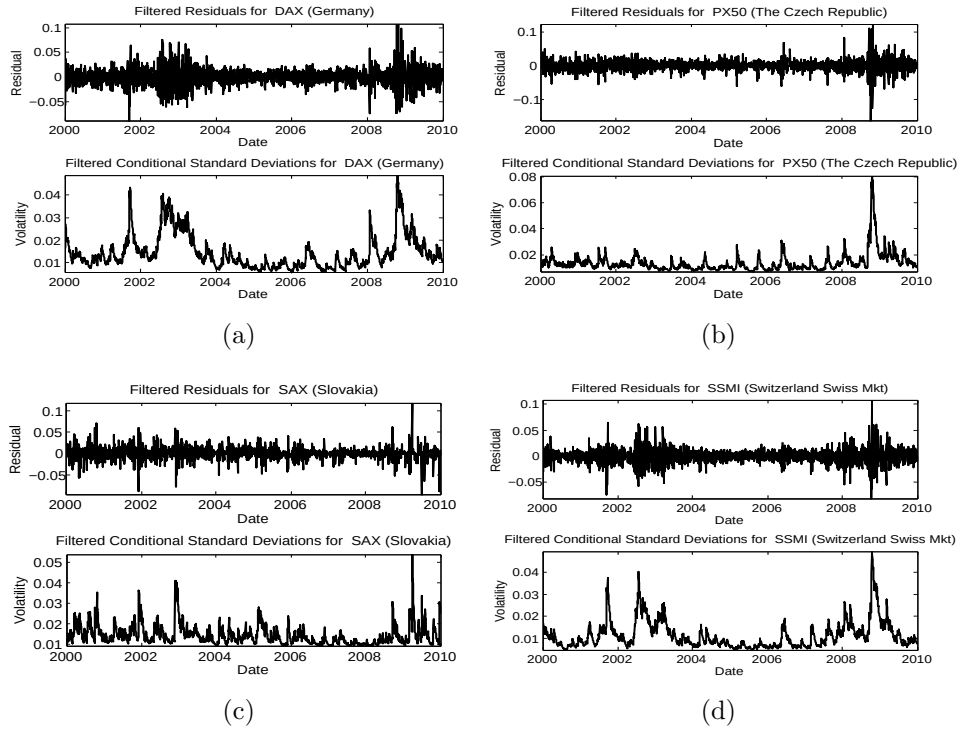


Figure B.4: Residuals and conditional standard deviation.

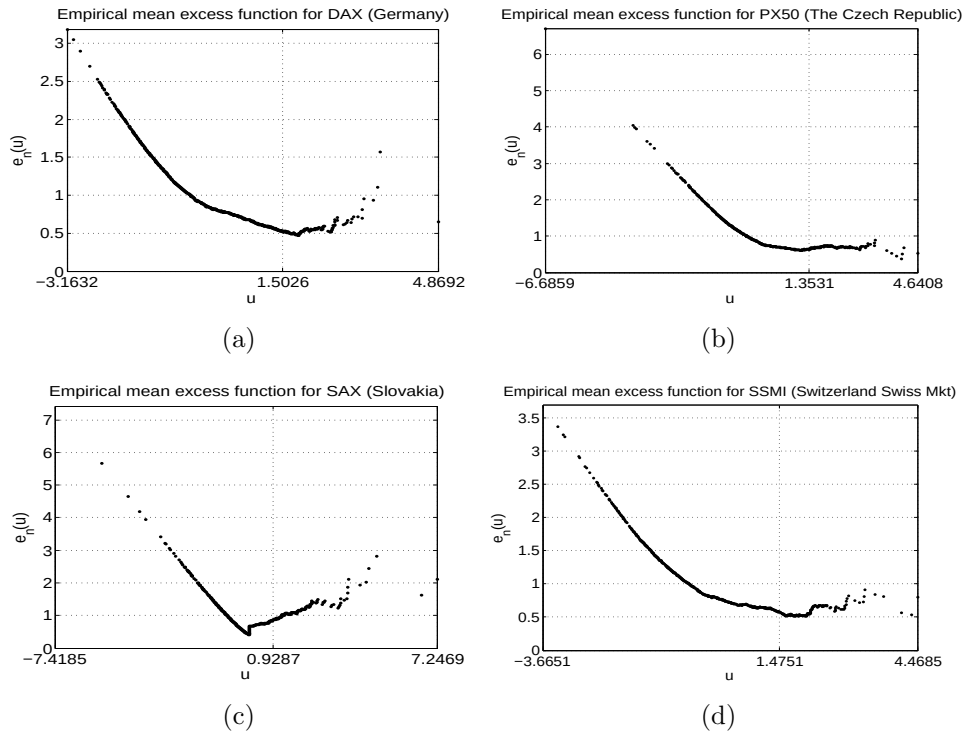


Figure B.5: Mean excess plot for the left tail.

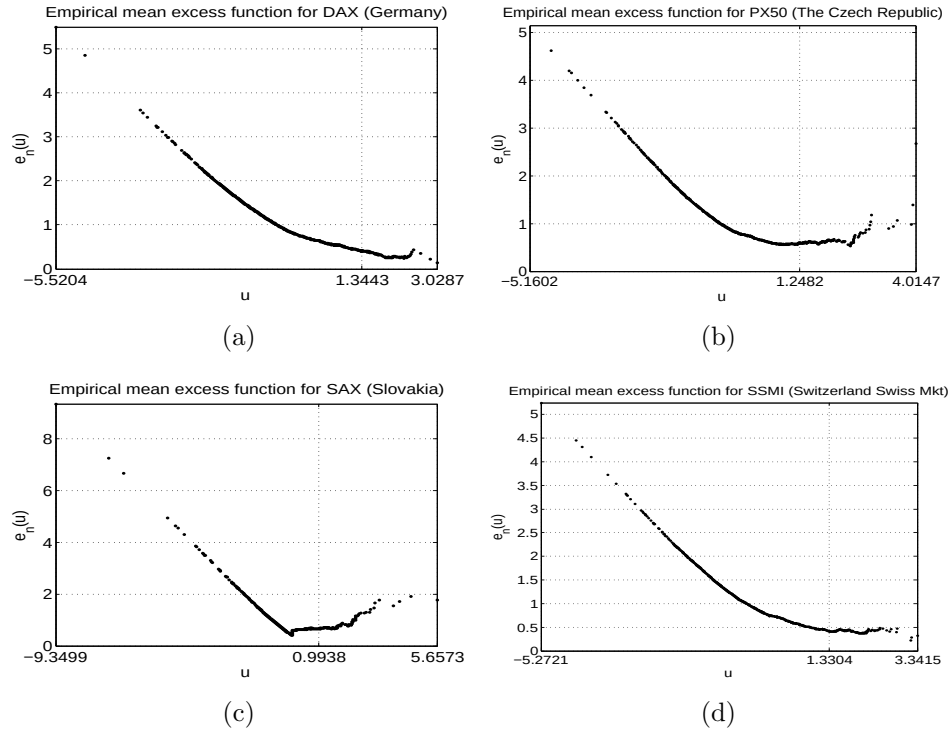


Figure B.6: Mean excess plot for the right tail.

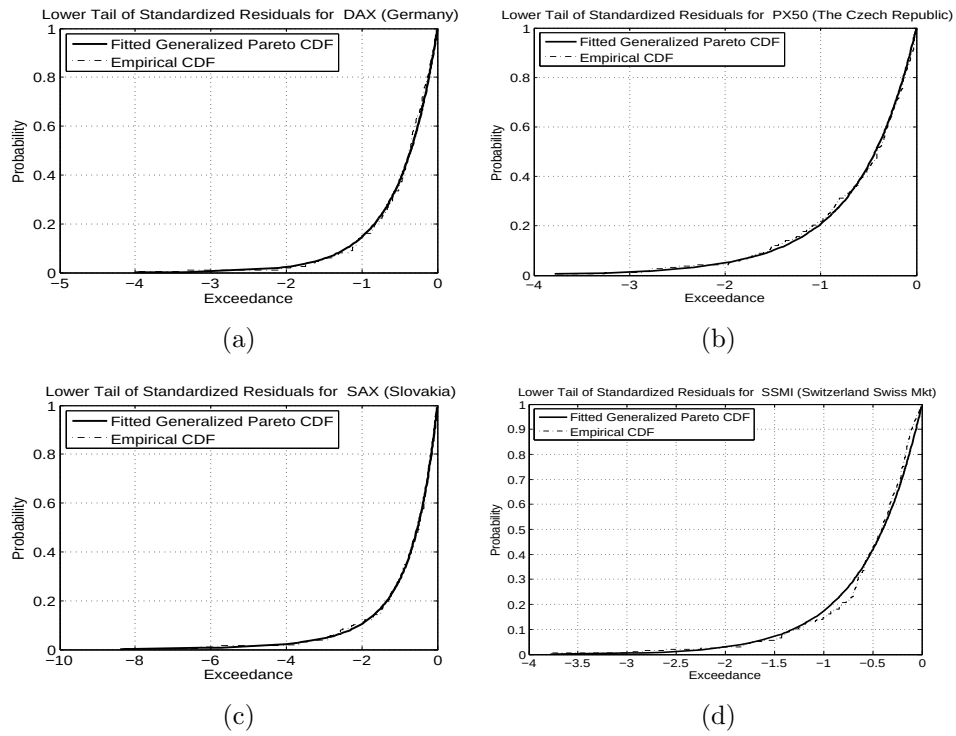


Figure B.7: Fitted lower (left) tail.

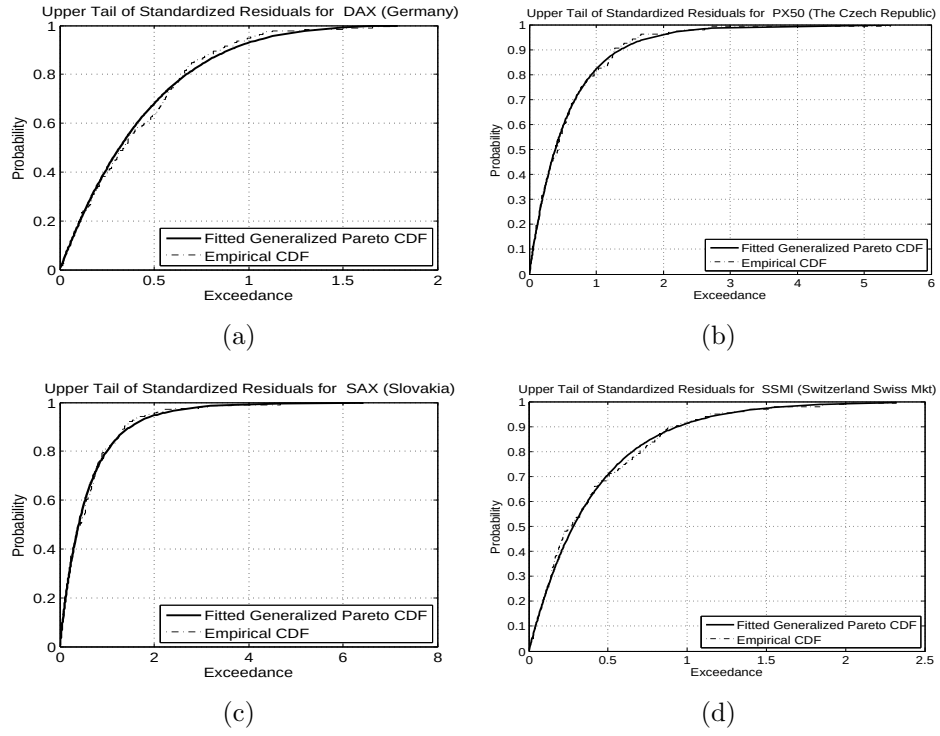


Figure B.8: Fitted upper (right) tail.