The Extreme Value Theory as a Tool to Measure Market Risk

Krenar Avdulaj
Disclaimer: The IES Working Papers is an online paper series for works by the faculty and students of the Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Czech Republic. The papers are peer reviewed, but they are not edited or formatted by the editors. The views expressed in documents served by this site do not reflect the views of the IES or any other Charles University Department. They are the sole property of the respective authors. Additional info at: ies@fsv.cuni.cz

Copyright Notice: Although all documents published by the IES are provided without charge, they are licensed for personal, academic or educational use. All rights are reserved by the authors.

Citations: All references to documents served by this site must be appropriately cited.

Bibliographic information:

This paper can be downloaded at: http://ies.fsv.cuni.cz
The Extreme Value Theory as a Tool to Measure Market Risk

Krenar Avdulaj#

# IES, Charles University Prague
E-mail: krenar.avdulaj@gmail.com

July 2011

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 market risk estimating method which employs Monte Carlo simulations to estimate Value-at-Risk (VaR) for a portfolio of 4 stock exchange indexes from Central Europe. The method uses the non-parametric empirical distribution to capture small risks and the parametric Extreme Value theory to capture large and rare risks. We compare estimates of 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.

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

JEL: C22, G17

Acknowledgements:
The author wishes to thank Jozef Baruník for insightful reading of the manuscript and for many helpful comments. The suggestions and comments of an anonymous reviewer are highly acknowledged. All errors and omissions remain entirely the fault of the author.
1. Introduction

Value-at-Risk (VaR) is a risk measure tool 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 RiskMetrics™ system which included the VaR as a risk measure [see Linsmeier and 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, e.g. Best (1998), Duffie and Pan (1997), Das (2006), Holton (2003), Dowd (2000), Smithson and 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 the probability, the VaR is meaningless.

In recent years academics have made extensive research on risk management. Artzner et al. (1999) introduce the concept of coherent risk measures. In work of Artzner et al. (1997) as well as in Artzner et al. (1999) the authors criticize VaR because it exhibits poor aggregation properties. They argue that VaR suffers the Non-subadditivity property, \( q_\alpha(F_L) \leq q_\alpha(F_{LA}) + q_\alpha(F_{LB}) \) (where \( q_\alpha(\cdot) \) is the \( \alpha \) quantile of \( F_L \)), that is if we have two portfolios \( A \) and \( B \) with loss probability distributions \( F_{LA} \) and \( F_{LB} \) respectively and if we denote the total loss \( L = L_A + L_B \) the inequality \( q_\alpha(F_L) \leq q_\alpha(F_{LA}) + q_\alpha(F_{LB}) \) 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 \( \text{VaR}_{A+B} \geq \text{VaR}_A + \text{VaR}_B \). In addition, they argue that VaR gives only an upper bound of losses that occur with a given frequency.

In this work we should not worry about the satisfaction of the subadditivity property because it may be a problem only in portfolios that contain non linear derivatives. In our case the portfolio is composed of the same set of underlying elliptically distributed risk 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 the work of Nyström and Skoglund (2002a) and Nyström and Skoglund (2002b)\(^1\). It is a two-

\(^1\) We 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/
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 4 countries in the Central Europe for a period of ten years, from January 1st, 2000 to December 31st, 2009. The indexes are Austria (ATX), Germany (DAX), the Czech Republic (PX50) and Switzerland (SSMI). The model we estimate is composed of two phases: first univariate modelling and then multivariate model. The univariate modelling involves AR-GARCH models and the Extreme Value Theory, while the multivariate involves $t$-copulas and their capability to conduct multivariate Monte Carlo simulations.

Initially we filter the log-returns by constant mean-GARCH or the AR-GARCH based on Glosten et al. (1993) model. The standardized residuals are then used by EVT. This filtration procedure is in line with Diebold et al. (1998) suggestions on implementing the application of EVT in financial time series. At this moment we make a semi-parametric cumulative distribution estimation, non-parametric for the internal part and parametric for the tails based on Extreme Value Theory (EVT). 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 Danielsson and 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 (the Generalized Pareto Distribution), are treated thoroughly in Davison and Smith (1990). Other work in this direction, among the others, is done by McNeil and 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 the ability to capture better 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 simulated data are transformed back to standardized residuals form through the inverse of semi-parametric 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 estimate during filtration and thus obtain simulated returns. The Value-at-Risk then is calculated easily through the quantile estimation. Finally, we compare the estimate of this approach with Historical Simulation and Variance Covariance method.

The paper is organized as follows: section two formalizes univariate modelling, we introduce AR-GARCH models and the Extreme Value Theory following the peaks-over-threshold approach. In the third section we introduce the multivariate model based on t-copulas and in the fourth, the application of the model and interpretation of the results follows.

2. Model specification

2.1. Univariate modelling

We start filtering the data using two AR-GARCH models: the standard conditional constant mean-GARCH and a model allowing for asymmetry. The latter is in line with work of Nyström and Skoglund (2002b). The asymmetric GARCH 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]. Let us first denote the time series of logarithmic returns \((X_t)_{t \in \mathbb{Z}}\), defined on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\). The asymmetric AR(1)-GARCH(1,1) model is:

\[
\begin{align*}
X_t &= c + \phi_1 X_{t-1} + a_t \\
\epsilon_t &= \sigma_t \epsilon_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \psi \text{sgn}(a_{t-1}) a_{t-1}^2 + \beta \sigma_{t-1}^2
\end{align*}
\]

where

\[
\text{sgn}(a_{t-1}) = \begin{cases} 
1 & \epsilon_{t-1} < 0 \\
0 & \text{otherwise}
\end{cases}
\]
and
\[ a_t = X_t - \mu, \quad \epsilon_t \text{ is i.i.d } (0,1) \]
\[ |\phi| < 1, \quad \alpha_1 + \frac{1}{2} \psi + \beta < 1, \quad \alpha_0 > 0 \]
\[ \alpha_1 \geq 0, \quad \beta \geq 0, \quad \alpha_1 + \psi \geq 0 \]

**Extreme value theory**

Extreme Value Theory 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 into 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 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). In this work we use the latter approach.

Let us introduce formally the GPD.

**Definition 2.1 (Generalized Pareto probability density function).** Define the function \( g_{\xi,\beta,\nu}(x) \),

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

\( \nu < x \text{ when } \xi > 0, \text{ or } \nu < x < -\beta/\xi \text{ when } \xi < 0 \)

\( \nu < x, \text{ when } \xi = 0 \)

**Definition 2.2 (Generalized Pareto cumulative distribution function).** Define the distribution function \( G_{\xi}(x) \) by

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

\[ x \geq 0, \quad \text{if } \xi \geq 0 \]
\[ 0 \leq x \leq -1/\xi, \quad \text{if } \xi < 0 \]

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

\[ G_{\xi,\beta,0}(x) = \begin{cases} 
\left( 1 - \left[ 1 + \frac{\xi}{\beta} \right]^{-1/\xi} \right), & \text{if } \xi \neq 0 \\
\left( 1 - e^{-\frac{x}{\beta}} \right), & \text{if } \xi = 0
\end{cases} \quad (4) \]

where \( \beta > 0 \) and for \( x \) the same conditions as in Equation 3 apply (the location parameter \( \nu = 0 \)). From Equation 4 we see that the GPD is transformed into a distribution with two parameters, shape parameter \( \xi \) and scale parameter \( \beta \) (\( G_{\xi,\beta} \)).

![Probability Density Function of GPD](a)

![Cumulative Distribution Function of GPD](b)

Figure 1: 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 the most relevant in financial analysis since it is heavy tailed.
In threshold exceedances we model the excess distribution over the threshold $u$, like in the Figure 2. For a random variable $X$ with distribution function $F$ we formalize the conditional excess distribution over the threshold $u$:

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

(5)

for $0 \leq y < x_F - u$, where $x_F \leq \infty$ is the right end point 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 the GPD is defined as:

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

(6)

where

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

From the Equation 6 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.

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

Figure 2: Conditional excess distribution over the threshold $u$. We use the EVT to model the observations which exceed the threshold $u$.
Theorem 2.1 (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

\[
F_u(y) \approx G_{\xi,\beta}(y) \quad \text{as} \quad u \to \infty
\]  

(7)

Based on Theorem 2.1 we make the following assumption:

**Assumption 2.1.** 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 \( 0 < y < x_F - u \), \( \xi \in \mathbb{R} \) and \( \beta > 0 \).

Using the Equation 5, assumption 2.1 and the empirical estimate \( \frac{n - N_u}{n} \) for \( F(u) \) [see McNeil et al. (2005), Gilli and Kellezi (2006)], where \( n \) is the total number of observations and \( N_u \) the number of observations over the threshold we can write (consult Appendix A.1 for details):

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

(8)

Moreover, if we invert the Equation 8 we get the high quantile or VaR of the underlying distribution. For \( \alpha \geq F(u) \):

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

(9)

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

---

\(^2\)Note the change in notation in the EVT modelling, the \( X_t \) used here are in fact \( \epsilon_t \) from the filtering process in the first step.
parameter $\nu = 0$):

$$
\ln L(\xi, \beta; Y_1, \ldots, Y_{N_u}) = \sum_{j=1}^{N_u} \ln g_{\xi,\beta}(Y_j)
$$

$$
= \sum_{j=1}^{N_u} \ln \frac{1}{\beta} + \sum_{j=1}^{N_u} \ln \left( 1 + \frac{\xi Y_j}{\beta} \right)^{-1 - \frac{1}{\xi}}
$$

$$
= -N_u \ln \beta - \left( 1 + \frac{1}{\xi} \right) \sum_{j=1}^{N_u} \ln \left( 1 + \frac{\xi Y_j}{\beta} \right) (10)
$$

To estimate the shape parameter $\xi$ and scale parameter $\beta$ we maximize the objective function in the Equation 10, 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 $x^n_1 < u < x^n_n$. The empirical $e_n(u)$ is calculated like in Gilli and Kellezi (2006):

$$
e_n(u) = \sum_{i=k}^{n} \frac{(x^n_i - u)}{n - k + 1}, \quad k = \min \{ i | x^n_i > u \} (11)
$$

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 $u$. For example, Figure 3 shows plots of empirical mean excess for the ATX standardized residuals series$^3$. 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 again 204 observations in the tail. 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.

---

$^3$Standardized residuals are obtained by dividing residuals $a_t$ with conditional standard deviations of Equation 1, i.e. $\frac{a_t}{\sigma_t}$. 

8
2.2. Multivariate modelling: t-copulas

Copulas application in finance have received much attention from academics in the recent years. We mention here the work of Embrechts et al. (2003), Cherubini et al. (2004), Dias and Embrechts (2010) and Patton (2006) among many others. This is because copulas have a wide use in financial industry, especially in credit scoring, risk management and derivative pricing.

Let us define the concept of copula.

Definition 2.3. A $d$-dimensional copula is a distribution function $[0,1]^d$ with standard uniform marginal distributions. We reserve the notation $C(u) = C(u_1, \ldots, u_d)$ 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 relationship between the joint distribution and a copula is given by the Theorem of Sklar.

Theorem 2.2 (Sklar 1959). Let $F$ be a joint distribution function with margins $F_1, \ldots, F_d$. Then there exists a copula $C : [0,1]^d \rightarrow [0,1]$ such that, for all $x_1, \ldots, x_d$ in $\mathbb{R} = [-\infty, \infty],

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \quad (12)$$
If the margins are continuous, then $C$ is unique; otherwise $C$ is uniquely determined on $\text{Ran}F_1 \times \text{Ran}F_2 \times \ldots \times \text{Ran}F_d$, where $\text{Ran}F_i = F_i(\mathbb{R})$ denotes the range of $F_i$. Conversely, if $C$ is a copula and $F_1, \ldots, F_d$ are univariate distribution functions, then the function $F$ defined in 12 is a joint distribution function with margins $F_1, \ldots, F_d$.

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.

We use $t$ copula in multivariate model because for the case of financial returns data the $t$-copula performs better than the Gaussian copula [see Demarta, S. and McNeil, A. J. (2004)]. If we have a $d$-dimensional random vector from $t$ distribution $X \sim t_d(\nu, 0, P)$, the $t$ copula is defined as:

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

where $t_{\nu}$ 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 $\nu$ degrees of freedom. The $t$ copula represents the dependence structure of a multivariate $t$-distribution.

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

$$l(\theta) = \sum_{i=1}^{n} \ln c\{F_1(X_{i1}; \beta), \ldots, F_d(X_{id}; \beta); \alpha\} + \sum_{i=1}^{n} \sum_{j=1}^{d} \ln f_i(X_{ij}; \beta) \quad (13)$$

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

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

where $\Theta$ is the parameter space.
3. Application of the model

We remind the reader that our dataset includes daily returns of stock exchange indexes of Austria (ATX), Germany (DAX), the Czech Republic (PX50) and Switzerland (SSMI). The analysis covers the period from January 1st, 2000 to December 31st, 2009. Table 1 shows a statistical summary of the dataset. We note that the data are highly non-normal and there is presence of autocorrelations. We filter raw returns series through AR-GARCH models and summarize the results in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>ATX</th>
<th>DAX</th>
<th>PX50</th>
<th>SSMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.7801e-4</td>
<td>-7.2069e-5</td>
<td>3.2726e-4</td>
<td>-5.3415e-5</td>
</tr>
<tr>
<td>Std.Dev</td>
<td>0.0147</td>
<td>0.0164</td>
<td>0.0156</td>
<td>0.0129</td>
</tr>
<tr>
<td>Skweness</td>
<td>-0.3969</td>
<td>0.0428</td>
<td>-0.4912</td>
<td>0.0565</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>12.35</td>
<td>7.73</td>
<td>16.11</td>
<td>9.44</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.1025</td>
<td>-0.0887</td>
<td>-0.1618</td>
<td>-0.0810</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.1202</td>
<td>0.1079</td>
<td>0.1236</td>
<td>0.1078</td>
</tr>
<tr>
<td>Jarque-Bera stat.</td>
<td>9.360*</td>
<td>2.382*</td>
<td>18.368*</td>
<td>4.412*</td>
</tr>
<tr>
<td>Arch test stat.</td>
<td>788.8*</td>
<td>467.4*</td>
<td>705.3*</td>
<td>605.5*</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics. 01-Jan-2000 to 31-Dec-2009. (*) shows the rejection of $H_0$ that data are normal for the J-B test and $H_0$ of no autocorrelation for Engle’s Arch test. Confidence level is 5% in both cases, while number of lags is 10.

We then estimate the semi-parametric cumulative distribution function for the standardized residuals time series. In Figure 4(b) we give the semi-parametric cumulative distribution function for the case of ATX, while in 4(a) and 4(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.7). We see that the parametric Generalized Pareto distribution is a good fit for the tails, especially for the lower tail.

We estimate the shape parameter $\xi$ and scale parameter $\beta$ for each index by maximizing log-likelihood defined in Equation 10. The left tail and right tail estimates for shape and scale parameter are presented in table 3. We note that having the shape and scale parameter, as well as $u$, $n$ and $N_u$ obtained from empirical mean excess plot, it is possible to calculate the individual
Table 2: The models that best fitted our data were asymmetric AR(1)-GARCH(1,1) for the case of Austria and the Czech Republic and the constant conditional mean-GARCH(1,1) for the rest. In all cases the innovations are assumed to have $t$ distribution. Data in brackets represent standard errors.

<table>
<thead>
<tr>
<th></th>
<th>ATX</th>
<th>DAX</th>
<th>PX50</th>
<th>SSMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>7.30e-4</td>
<td>6.47e-4</td>
<td>7.92e-4</td>
<td>5.26e-4</td>
</tr>
<tr>
<td></td>
<td>(1.78e-4)</td>
<td>(2.19e-4)</td>
<td>(2.13e-4)</td>
<td>(1.67e-4)</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>4.33e-2</td>
<td>-</td>
<td>6.45e-2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.06e-2)</td>
<td>( - )</td>
<td>(2.06e-2)</td>
<td>( - )</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>2.29e-6</td>
<td>1.44e-6</td>
<td>5.48e-6</td>
<td>1.36e-6</td>
</tr>
<tr>
<td></td>
<td>(5.60e-7)</td>
<td>(5.03e-7)</td>
<td>(1.24e-6)</td>
<td>(4.1e-7)</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>4.79e-2</td>
<td>8e-2</td>
<td>5.16e-2</td>
<td>10.1e-2</td>
</tr>
<tr>
<td></td>
<td>(1.57e-2)</td>
<td>(0.01)</td>
<td>(1.52e-2)</td>
<td>(1.27e-2)</td>
</tr>
<tr>
<td>$\psi$</td>
<td>8.38e-2</td>
<td>-</td>
<td>0.10</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(2.03e-2)</td>
<td>( - )</td>
<td>(2.29e-2)</td>
<td>( - )</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.893</td>
<td>0.914</td>
<td>0.867</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>(1.36e-2)</td>
<td>(0.01)</td>
<td>(1.66e-2)</td>
<td>(1.27e-2)</td>
</tr>
<tr>
<td>DoF</td>
<td>7.24</td>
<td>11.02</td>
<td>6.12</td>
<td>9.01</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(1.9)</td>
<td>(0.72)</td>
<td>(1.39)</td>
</tr>
<tr>
<td>$\mathcal{L}L$</td>
<td>7,954</td>
<td>7,427</td>
<td>7,590</td>
<td>8,095</td>
</tr>
<tr>
<td>AIC</td>
<td>-15,895</td>
<td>-14,845</td>
<td>-15,167</td>
<td>-16,181</td>
</tr>
</tbody>
</table>

Figure 4: Semi-parametric cumulative distribution function (a) and fitted cumulative distribution function for upper tail (b) and lower tail (c).

VaR for each index without involving any simulation\(^4\). For illustration, in

\(^4\)We recall that our work is based on Monte-Carlo simulation, thus we will continue the
Table 3: Point estimates of the tail parameters.

<table>
<thead>
<tr>
<th></th>
<th>ATX</th>
<th>DAX</th>
<th>PX50</th>
<th>SSMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left tail</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.0046</td>
<td>0.0307</td>
<td>0.0665</td>
<td>-0.0079</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.6356</td>
<td>0.5037</td>
<td>0.6048</td>
<td>0.5756</td>
</tr>
<tr>
<td>$\hat{\text{VaR}}_{99%}$ (1 day)</td>
<td>2.7080</td>
<td>2.5842</td>
<td>2.7019</td>
<td>2.6819</td>
</tr>
<tr>
<td>Right tail</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\xi}$</td>
<td>0.0613</td>
<td>-0.2086</td>
<td>0.0753</td>
<td>0.0104</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.4361</td>
<td>0.4911</td>
<td>0.5410</td>
<td>0.4032</td>
</tr>
</tbody>
</table>

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 using the semi-parametric $c.d.f$ and then 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 4 we have the correlation matrix that we get from fitting the $t$ copula. After this step we use the copula capability to simulate random vectors with the same dependence structure as the one that we just calibrated\(^6\). 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 AR-GARCH filtering process as in Equation 1. We then introduce the autocorrelation and heteroskedasticity as in the original returns series via AR-GARCH simulation (consistent with parameters in table 2). In this way we simulated returns series that are consistent with historical performance.

Now we continue with the last step, finding the Value-at-Risk of our portfolio. We calculate the cumulative returns of the portfolio based on the development of the model further.

\(^5\)We calculate 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 want to know also for a short position.

\(^6\)The number of simulations is 10,000 for every forecasting day.
Table 4: Correlation matrix obtained from fitting t-copula to data with degrees of freedom 9.58.

<table>
<thead>
<tr>
<th></th>
<th>ATX</th>
<th>DAX</th>
<th>PX50</th>
<th>SSMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATX</td>
<td>1</td>
<td>0.5560</td>
<td>0.4646</td>
<td>0.5152</td>
</tr>
<tr>
<td>DAX</td>
<td>0.5560</td>
<td>1</td>
<td>0.4368</td>
<td>0.7657</td>
</tr>
<tr>
<td>PX50</td>
<td>0.4646</td>
<td>0.4368</td>
<td>1</td>
<td>0.4141</td>
</tr>
<tr>
<td>SSMI</td>
<td>0.5152</td>
<td>0.7657</td>
<td>0.4141</td>
<td>1</td>
</tr>
</tbody>
</table>

following Equation

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

where $i \in \{1, \ldots, 4\}$ represents each index, $H$ is the risk horizon in days and $T$ is the length of the simulated time series [consult Appendix A.2 for details].

We calculate the Value-at-Risk for our portfolio for a risk horizon of 1, 10 and 22 trading days. With a probability of 99% losses will not be higher than 11.74% in one month, while with probability 99.995% the losses will not be higher than 23.69%. For the same portfolio in table C.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 time scale factor 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) or Danielsson and de Vries (2000)]. If this requirement is violated the formula is an approximation. Indeed, this requirement is violated because portfolio returns are not i.i.d [see the data summary in Table 1].

We find that the historical simulation overpredicts the VaR for extreme quantiles in comparison with both other methods. This result is in line with Danielsson and 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

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 and 1 day market risk prediction this method gives an estimation such as 90% 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 assumes that the correlation among stocks (indexes in our case) is constant.

In order to see how these methods perform when we use samples of low sample. Historical simulation method has another problem, it can not give estimates for probabilities lower than 1/sample size.
volatility and high volatility we split the original data in two parts equally and run all the three methods on each part\(^7\). The first part includes data from 2000-2004, a period in which the markets have been relatively stable. The second period from 2005-2009 is mostly characterised by high volatility due the effects of financial crisis in the United States (2007-2009).

During the quiet times the EVT underpredicts the VaR compared with the other methods. For high volatility periods the EVT estimates are again in between for 1 and 10 days of forecast. For 22 days risk horizon the VaR estimates have a value of 48.04\% at 99.995\% confidence level, which means that EVT is a good candidate to capture very rare events. In this particular case we observe that our approach gave an out-of-sample prediction, the extrapolation claimed earlier in the text. Such events where losses will be 48.04\% in one month (22 trading days) are very rare. These results support the EVT as a risk measurement tool for the extreme quantiles. 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.

4. Conclusions

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 the Extreme Value Theory to capture large risks. The use of the 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 a smooth connection between univariate and multivariate distribution and helps to conduct Monte Carlo simulations. It takes into consideration

\(^7\)We are aware that the sub-samples size now is 1275, lower than the suggested minimum size of 1500 days for an accurate estimate of tail index [see Danielsson and de Vries (2000)]. For each of these datasets the filtering was made with similar AR-GARCH models as for the whole sample, but for the sake of brevity we are not including those results here.
the historical dependence structure with the respective degrees of freedom as well as simulates multivariate $t$ distributed observations.

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 48.04% at 99.995% confidence level which is reasonable if we compare with the losses suffered in the US markets in 2007-2009. 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 the Monte Carlo simulations.
Appendix A.

Appendix A.1. 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 the Equation 5 we have:

$$\hat{F}(x) = F(u) + F_u(y)\bar{F}(u)$$
$$= 1 - \bar{F}(u) + F_u(y)\bar{F}(u)$$
$$= 1 + \bar{F}(u) [F_u(y) - 1]$$
$$= 1 + \frac{N_u}{n} \left[ 1 - \left( 1 + \xi \frac{x-u}{\beta} \right) \frac{1}{\tau} - 1 \right]$$
$$= 1 - \frac{N_u}{n} \left[ 1 + \xi \frac{x-u}{\beta} \right] \frac{1}{\tau} \quad (A.1)$$

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,

$$\alpha = 1 - \frac{N_u}{n} \left[ 1 + \xi \frac{q_\alpha(F) - u}{\beta} \right] \frac{1}{\tau}$$

$$\left[ 1 + \xi \frac{q_\alpha(F) - u}{\beta} \right] \frac{1}{\tau} = \frac{n}{N_u} (1 - \alpha)$$

$$\xi \frac{q_\alpha(F) - u}{\beta} = \left[ \frac{n}{N_u} (1 - \alpha) \right]^{-\xi} - 1$$

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

$$VaR_\alpha = q_\alpha(F) = u + \frac{\hat{\beta}}{\xi} \left[ \left[ \frac{n}{N_u} (1 - \alpha) \right]^{-\xi} - 1 \right] \quad (A.2)$$

18
Appendix A.2. Calculation of portfolio returns.

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:

\[
r_t = \log \frac{P_t}{P_{t-1}} \Rightarrow e^{r_t} = \frac{P_t}{P_{t-1}}
\]

\[
e^{r_t} - 1 = \frac{P_t}{P_{t-1}} - 1 = \frac{P_t - P_{t-1}}{P_{t-1}} = R_t
\]

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

\[(e^{r_{i,t}} - 1) \ast w_i, \quad \text{where} \quad i \in \{1, \ldots, 4\} \quad \text{represents each index.}\]

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:

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

where \( i \in \{1, \ldots, 4\} \) represents each index, \( H \) is the risk horizon in days and \( T \) is the length of the simulated time series (in our case \( T = 10,000 \)).
Appendix B. Figures

Figure B.6: Mean excess plot for the lower (left) tail [a-c-e] and upper (right) tail [b-d-f].
Figure B.7: Fitted lower (left) tail [a-c-e] and upper (right) tail [b-d-f].
### Appendix C. Tables

<table>
<thead>
<tr>
<th></th>
<th>VaR based on Monte-Carlo simulation and t-copula</th>
<th>VaR based on Historical simulation</th>
<th>VaR based on Variance-Covariance method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>90%</strong></td>
<td>1.0814</td>
<td>3.3494</td>
<td>5.0343</td>
</tr>
<tr>
<td><strong>95%</strong></td>
<td>1.4390</td>
<td>4.5653</td>
<td>7.1088</td>
</tr>
<tr>
<td><strong>99.5%</strong></td>
<td>2.6364</td>
<td>8.6744</td>
<td>13.9580</td>
</tr>
</tbody>
</table>

| **90%**        | 1.2449   | 3.9366   | 5.8389    | 1.1066   | 3.4993   | 5.1903    | 1.3916    | 4.4007   | 6.5273    |

| **90%**        | 1.5690   | 4.9617   | 7.3594    | 1.2514   | 3.9365   | 5.8715    | 1.8321    | 5.7936   | 8.5933    |
| **95%**        | 2.0138   | 6.3683   | 9.4456    | 1.6067   | 5.0807   | 7.5359    | 2.3515    | 7.4360   | 11.0284   |

**Table C.5:** Value-at-Risk (in %) calculated for a portfolio with equal weights at different risk horizon.

### References


IES Working Paper Series

2011
1. Roman Horváth, Jakub Matějů : How Are Inflation Targets Set?
2. Jana Procházková, Lenka Šťastná : Efficiency of Hospitals in the Czech Republic
3. Terezie Výprachtická : The Golden Rule of Public Finance and the Productivity of Public Capital
4. Martina Mysíková : Income Inequalities within Couples in the Czech Republic and European Countries
5. Veronika Holá, Petr Jakubík : Dopady změn parametrů pojištění vkladů v roce 2008
7. Aleš Maršál : The Term Structure of Interest Rates in Small Open Economy DSGE Model
8. Robert Flasza, Milan Rippel, Jan Šolc : Modelling Long-Term Electricity Contracts at EEX
10. Tomáš Havránek, Zuzana Iršová, Karel Janda : Demand for Gasoline Is More Price-Inelastic than Commonly Thought
11. Martina Mysíková : Personal Earnings Inequality in the Czech Republic
12. Ondřej Lopušník : Reflections on the reconciliation problem
13. Martin Gregor, Lenka Šťastná : The Decentralization Tradeoff for Complementary Spillovers
15. Andrea Klimešová, Tomáš Václavek : Pricing of Gas Swing Options using Monte Carlo Methods
17. Karel Báta : Equity Home Bias Among Czech Investors: Experimental Approach
18. Karel Janda : Credit Guarantees and Subsidies when Lender has a Market Power
19. Roman Horváth : Research & Development and Long-Term Economic Growth: A Bayesian Model Averaging Analysis
20. Petr Jakubík : Household Balance Sheets and Economic Crisis
21. Josef Brechler, Adam Geršl: Political Legislation Cycle in the Czech Republic
23. Michal Skořepa: A convergence-sensitive optimum-currency-area index
24. Marek Rusnák, Tomáš Havránek, Roman Horváth: How to Solve the Price Puzzle? A Meta-Analysis
26. Krenar Avdulaj: The Extreme Value Theory as a Tool to Measure Market Risk

All papers can be downloaded at: http://ies.fsv.cuni.cz