On the Information Content of Decomposed Financial Return Series: A Wavelet Approach

Theo Berger*

January 15, 2014

Preliminary draft. Please do not quote without permission of the authors.

We decompose financial return series via wavelets into different time scales to analyse their information content regarding the volatility of the returns. Moreover, we investigate the information of each scale and discuss the decomposition of daily Value-at-Risk (VaR) forecasts.

By an extensive empirical analysis, we analyse financial assets in calm and turmoil market times and show that daily VaR forecasts are mainly driven by the volatility which is captured by the scales comprising the short-run information. Further, we apply Extreme-Value-Theory to each time scale and illustrate that the information which is stored by the short-run scales linked via copulas, outperforms classical parametric VaR approaches which incorporate all information available.

*thberger@uni-bremen.de

1 Introduction

Triggered by the financial crisis in 2008, the adequate prediction of future losses gained increasing attention in the field of risk measurement. Volatile market times, as observed in 2008 and 2009, stressed the assumptions of classical risk measurement techniques and shifted away the interest from return prediction to the adequate modeling of future volatility. For the practitioner as well as for the researcher, Value-at-Risk (VaR) became a popular standard to translate financial market risk into a single monetary risk figure. As it is mandatory for financial institutions to measure and communicate risk in terms of VaR, nowadays, nearly all listed companies present VaR figures in their annual reports to give an idea about their risk exposures.

As the modeling of VaR is widely discussed, the choice of marginal return distribution represents the balance point of VaR forecasting (see Jorion (2007) and Berger & Missong (2013)). According to Angelidis et al. (2004) amongst others, distribution functions which capture the fat tails of the marginal return distribution adequately (as i.e. the t distribution does) outperform models based on the classical assumption of normally distributed returns in volatile market times. Moreover, as the explicit modeling of the left tail of the distribution represents the scope of VaR modeling, the application of Extreme Value Theory became rather promising. Bekiros & Georgoutsos (2005), Brooks et. al (2005) and Berger (2013) confirm the adequate VaR performance of Extreme Value Theory (EVT) applied to to financial asset returns.

In this paper, we add to the exciting VaR literature and give innovative insight into the information content which is needed to forecast VaR. We decompose return series into different time scales and investigate the information content of each scale with respect to VaR. Ramsey (2002) and Renaud et al. (2002) are among the first ones who address the usage of wavelets in the field of finance. The advantage of wavelet analysis lies in the fact, that financial return series can be analysed apart of the choice of marginal distribution. As described by Murtagh et al. (2003), time-scale decomposition results in wavelet coefficients for each scale and a reconstruction of the original series is feasible.

Lo Cascio (2007) decomposes UK real GDP via wavelet to investigate the long-

run structure of the data apart form external shocks. Cifter (2011) makes use of the wavelet decomposition to determine the parameters for the Generalized Pareto Distribution. Reboredo and Rivera-Castro (2014) analyse dependency between oil prices and stock returns and Jammazi and Aloui (2012) combine wavelets with neural networks to investigate crude oil prices. In the context of forecasting, parsimonious wavelet decomposition is applied to filter the returns in order to measure denoised dependency between the investigated assets: Boubaker and Sghaier (2012) apply wavelet decomposition to show the stability of the applied dependency measures for all time scales, He et al. (2012) use a multivariate wavelet denoising in order to investigate the dynamics of correlations for international markets as well as Khalfaoui and Boutahar (2012) who investigate the dependency of indexes. Hence, wavelets are explicitly used to model dependency between different time scales. However, the reason for the main focus on dependency measurement is given as follows: As the classical wavelet decomposition summarizes information at each scale, the amount of data points decreases by each time-scale, and consequently, events of the original series can not be located in larger time-scales, so that point forecasts are difficult to achieve, since the data length varies by scale.

As independently explored by different authors using different notations, e.g. Shensa (1991), Percival and Walden (2000) Starck et al. (2007), the "undecimated wavelet decomposition" approach is introduced to overcome the so called shift variance of the wavelet decomposition. So to say, the data length does not decrease by time-scale, so that the localisation of an event is ensured for all time scales.

Based on the undecimated wavelet decomposition, in this paper, we explore innovative insight into the forecasting of VaR by decomposing the investigated return series via wavelet technique. We decompose each return series into different time-scales and analyse the information content of each scale by fitting generalized pareto distribution (GPD) to each series in order to adequately model the relevant quantiles. Moreover, we apply an extensive empirical study to investigate the necessary information for daily VaR forecasts. We investigate VaR forecasts of the decomposed return series based on GPD and link the information of different time scales via t copula. As a result, the relevant information for one day forecasts turns out to

be included essentially in the short-run scales. This can be seen, by adding successively more information to the VaR forecast in order to improve the out-of-sample forecasting quality of VaR forecasts.

The remainder of the paper is structured as follows: Section 2 presents the relevant methodology. In section 3, we present the analysis of decomposed time series and VaR forecasts and in section 4 we evaluate VaR forecasts via Out-of-Sample performance. Section 5 concludes.

2 Methodology

2.1 Wavelet Decomposition

In this paper we use the redundant, discrete wavelet transform also known as the Haar a trous wavelet decomposition as introduced by Murtagh et al. (2003). This algorithm is based on the Haar wavelet decomposition, but in contradiction to this approach, it is described by shift invariance.

As given in Murtagh et al. (2003), we need to perform successive convolutions with the discrete low-pass filter h:

$$c_{i+1}(k) = \sum_{l=-\infty}^{+\infty} h(l)c_i(k + 2^i l). \tag{1}$$

where the original time series $r(t)$ represents the lowest scale $c_0(t)$. The low-pass filter, h, is defined as $h = (1/2, 1/2)$.

In order to obtain the wavelet coefficients w_i we take the difference between two successive smoothed series in the following way:

$$w_i(k) = c_{i-1}(k) - c_i(k). \tag{2}$$

It is to note, that we decompose the time series up to eight scales.

2.2 GJR-GARCH

Based on the decomposed return series, we standardize the decomposed residuals via GARCH process. More concrete, we apply the GJR-Garch model (as introduced by

Glosten et al. (1993)) in order to take account for asymmetric shocks on volatility. The Model is given by:

$$\sigma_t^2 = \Omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma 1(\epsilon_{t-t} < 0)\epsilon_{t-1}^2. \tag{3}$$

In this context, γ takes account for the asymmetric reaturn behaviour which is caused by negative shocks.

2.3 Extreme Value Theory

In order to model VaR for the return series, we assume both normally and t distributed returns as a benchmark for the subsequent empirical analysis. However, in order to adequately model the marginal distribution of each individual scale, more flexible return models are needed. Moreover, in the context VaR forecasts, the tail behavior of scales matters. Therefore, using Extreme Value Theory (EVT) to model exceptional returns (and using the empirical distribution for the "interior" part of the marginals) seems to be a sensible alternative. Based on the GJR-GARCH filtered i.i.d. residuals, additionally to parametric distributions, the *peaks over threshold approach*, will be applied to each scale. Consequently, the data that exceeds a pre-defined threshold at level α, is modeled via the Generalized Pareto Distribution (GPD). In the succeeding empirical analysis, $\alpha = 10\%$ will be the threshold value. So to say, the extreme observations in the tails of the empirical distribution are modeled via EVT (see Longin and Solnik (2001)) and the cumulative distribution function F_ξ is given by:

$$F(x)_{\xi;\beta} = 1 - (1 + \xi\frac{x}{\beta})^{\frac{-1}{\xi}} \tag{4}$$

with $x \geq 0$, $\beta > 0$ and $\xi > -0,5$. In this context, x represent the exceedances, ξ the tail index (shape) parameter and β the scale parameter respectively. Let $g_{\xi,\beta}$ denote the density function, N_u the number over the threshold u, and X_j the observed values over the threshold u and $Y_j = X_j - u$. Then the log likelihood is given by:

$$l_t = \sum_{j=1}^{N_u} \ln g_{\xi,\beta}(Y_i) \tag{5}$$

$$= -N_u \ln \beta - \left(1 + \frac{1}{\xi} \sum_{j=1}^{N_u} \ln \left(1 + \xi \frac{Y_i}{\beta} \right) \right), \tag{6}$$

2.4 Copulas

In order to model joint distributions of the different time scales, we apply t-copulas to adequately model VaR. Copulas got initially introduced to financial time series by Embrechts et al. (2002) and represent a two step approach which separates the choice marginal distribution from dependency modeling. However, this idea goes back to Sklar's Theorem (1959), which ensures the separation between marginal distribution and dependency:

Let $r_1, ...r_n$ be random variables, $F_1, ..., F_n$ the corresponding marginal distributions and H the joint distribution, then there exists a copula C: $[0,1]^n \to [0,1]$ such that:

$$H(r_1, ..., r_n) = C(F_1(r_1), ...F_n(r_n)) \tag{7}$$

Conversely if C is a copula and $F_1, ..., F_n$ are distribution functions, then H (as defined above) is a joint distribution with margins $F_1, ...F_n$.

Due to the fact, that we need to model joint default probabilities of up to eight scales, we focus on the application of elliptical copulas. However, in order to account for tail dependency, we link the scales via t copula.

The t copula belongs to the family of elliptical copulas and is derived from the multivariate t distribution.

The setup of the t copula is given by:

$$C^t(r_1, ..., r_n) = t_{\rho,v}(\Phi^{-1}(r_1), ..., \Phi^{-1}(r_n)) \tag{8}$$

$$= \int_{-\infty}^{\Phi^{-1}(r_1)} ... \int_{-\infty}^{\Phi^{-1}(r_n)} \frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{n}{2}}|\rho|^{\frac{1}{2}}} \left(1 + \frac{1}{v} z^T \rho^{-1} z \right)^{-\frac{v+n}{2}} dz_1...dz_n \tag{9}$$

in this setup $t_{\rho,v}$ stands for the multivariate t distribution with correlation matrix ρ and v degrees of freedom (d.o.f.). For $v \to \infty$ the t distribution approximates a Gaussian. To make parameter estimation feasible even for large portfolios, all copula parameters are estimated in a two step maximum likelihood method as given

by Joe (1996). In the first step the GARCH parameters $\hat{\theta}_1$ related to the univariate margins are estimated by:

1.

$$\hat{\theta}_1 = ArgMax_{\theta_1} \sum_{t=1}^{T} \sum_{j=1}^{n} \ln f_j(r_{jt;\theta_1}) \tag{10}$$

Based on $\hat{\theta}_1$ the t copula parameters $\hat{\theta}_2$ are estimated in the second step using:

2.

$$\hat{\theta}_2 = ArgMax_{\theta_2} \sum_{t=1}^{T} \ln c(F_1(r_{1t}), F_2(r_{2t}), ...F_n(r_{nt}); \theta_2, \hat{\theta}_1). \tag{11}$$

2.5 Value-at-Risk Forecasts

Within a VaR framework, we are able to compare different approaches by the out-of-sample quality of VaR forecast.

Generally, VaR defines a maximum loss limit which will not be exceeded with a given probability and is defined as the quantile at level α of the distribution of portfolio returns:

$$VaR_\alpha = F^{-1}(\alpha). \tag{12}$$

In a parametric setup, if we model VaR for univariate time series, the respective quantiles are functions of the variances. Consequently, the respective quantiles of the modeled portfolio variances can be directly translated into VaR (see Jorion (2007)). Let α be the quantile, σ_t^2 the conditional variance, then VaR at time $t+1$ is given by: $VaR_{t+1} = -\alpha\sqrt{\sigma_t^2}$ for elliptical distributions. For instance the 95% VaR of PF return y_t represents the parametric 5% quantile.
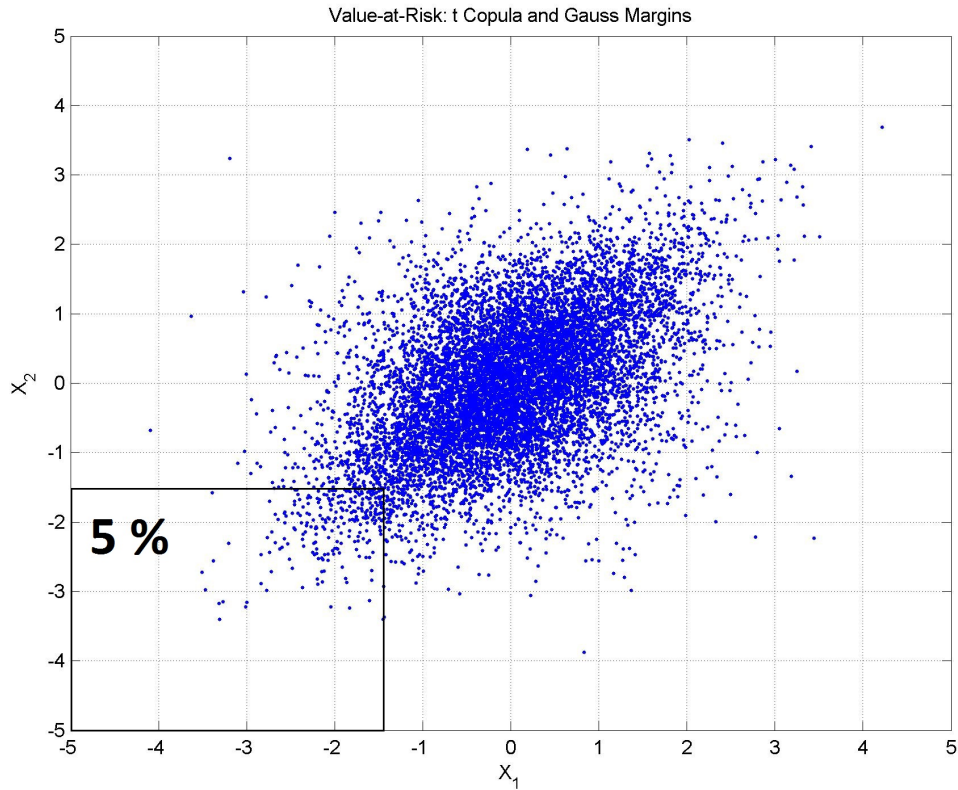
Figure 1: t copula with gaussian margins: Empirical 95% Value-at-Risk

However, in the context of copulas, due to the flexibility regarding the univariate marginal distribution functions, the estimated VaR at time $t + 1$ is not modeled parametrically. Based on the fitted copula, 10.000 returns are simulated and the VaR is simply the empirical quantile of the vector of simulated portfolio returns based on the information available at time t (see figure 1).

2.6 Backtesting

In order to evaluate the VaR performance, let a VaR misspecification be defined by the scenario in which the realized loss exceeds the VaR forecast. So that within the applied VaR framework, the applied models based on the information of different time scales can be evaluated by their empirical rate of VaR misspecifications via Out of sample analysis (always benchmarked against the realized returns of the underlying asset). Since for 95% VaR forecasts, we expect to observe 1% of VaR misspecifications. Empirically, if we analyse VaR forecast for one year (250 trading

days), we expect to observe 12 or 13 misspecifications. The next to subsections cover the applied quality criteria of our analysis, whereas we refer to Campbell (2006) for a thorough overview of VaR backtesting criteria.

Conditional Coverage

In order to investigate the absolute amount of misspecifications we apply the backtesting criteria of unconditional coverage by Kupiec (1995). More concrete, we test the null hypothesis, if the realized failure rate is in line with the expected failure rate. The relevant test statistic LR_{UC} is given by:

$$LR_{UC} = -2ln[(1-p)^{T-N}p^N] + 2ln[(1-(N/T))T - N(N/T)^N] \qquad (13)$$

p stands for the percent left tail level, T for the total days and N for the number of misspecifications. However, the null is given by $H_0 : p = 1/N$ whereas the test statistic is $\chi^2(1)$ distributed.

In addition to the amount of misspecifications, we focus on the clustering of VaR breaches. Due to the fact that the 1/N strategy is not dynamic in terms of portfolio weights, we analyse if the static strategy will be outperformed by the dynamic ones. Moreover, we investigate whether risk concentration on low-volatile assets bears the risk of VaR breaches which are not independently over time.

The test statistic, LR_{IND} is given by Jorion (1998):

$$LR_{IND} = -2ln[(1-\pi)^{(T_{00}+T_{01})}\pi^{(T_{01}+T_{11})}] + 2ln[(1-\pi_0)^{T_{00}}\pi_0^{T_{01}}(1-\pi_1)^{T_{10}}\pi_1^{T_{11}}]. \qquad (14)$$

Let T_{ij} be the number of observed values i followed by j. Whereas 1 represents a misspecification and 0 a correct estimation. π represents the probability of observing an exception and π_i the probability of observing an exception conditional on state i. Namely, we test the null hypothesis of independent VaR breaches over time, so that a model which creates too many clustered misspecifications gets rejected. Analogue to LR_{UC}, the test statistic is $\chi^2(1)$ distributed.
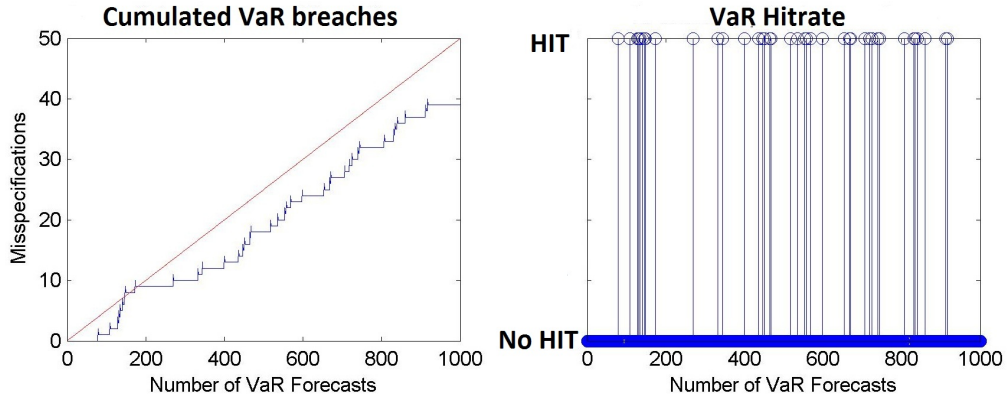
Figure 2: Example: backtesting criteria for 95% VaR forecasts (1000 observations): (left) Unconditional Coverage, (right) Independence

Figure 2 illustrates the backtesting criteria graphically. The graph on the left side describes the cumulated rate of VaR misspecifcations for an example comprising 1000 days and illustrates the idea of the unconditional coverage criterion. The realized amount of VaR misspecifications is benchmarked against the expected amount (straight line). Further, the independence criterion (right side) penalizes clustered VaR misspecifications if two misspecifications occur sequently. Based on both backtests, Christoffersen (1998) introduces the Conditional Coverage criterion as the sum of both UC and IND:

$$LR_{CC} = LR_{UC} + LR_{IND}. \tag{15}$$

This test statistic follows a $\chi^2(2)$ distribution and controls for both coverage rate and clustering.

3 Empirical Analysis

The data set comprises daily return series of stocks which are listed in the Dow Jones Industrial Average (DJIA). As we investigate the time from January 1st 2000 till September 30th 2013, due to its late listing in October 3rd in 2007, we excluded VISA from our analysis for sake of consistency. Consequently, we analyse 3600 daily log returns of 29 assets. The descriptive statistics for all assets are given in Table 1. All assets are described by an average rate of return close to zero,

10

whereas Kurtosis, Jarque Bera and the QQ statistics indicate non-normality of the return series. Interestingly, bank stocks as JP Morgan and Goldman Sachs show the highest standard deviations (0,027 and 0,025 respectively), whereas the highest losses are given by Home Depot (-0,339%), Procter & Gamble (-0,377%) and United Technologies (-0,332%).

	Mean	Std Dev	Max	Min	Skew	Kurt	JB-Stat	Q Stat	LM
3M	0,000	0,015	0,105	-0,094	0,057	8,094	2,1	115887,0	3576,2
AT&T	0,000	0,018	0,151	-0,135	0,096	9,318	398,6	118143,0	3568,2
AM EXP	0,000	0,025	0,188	-0,194	-0,007	11,733	11,6	117775,2	3571,5
BOEING	0,000	0,020	0,144	-0,194	-0,258	8,614	100,1	117970,2	3579,8
CATER	0,000	0,021	0,137	-0,157	-0,072	7,276	226,3	124074,9	3584,5
CHEVR	0,000	0,017	0,189	-0,133	0,066	15,026	259,1	123833,5	3584,0
CISCO	0,000	0,027	0,218	-0,177	0,177	10,623	12124,8	117941,9	3572,7
COCAC	0,000	0,014	0,130	-0,106	0,114	11,422	469,2	120343,7	3576,1
E I DU	0,000	0,019	0,109	-0,120	-0,180	8,248	622,8	101597,0	3550,8
EXXON	0,000	0,016	0,159	-0,150	0,063	13,894	311,5	124838,5	3581,3
GE	0,000	0,020	0,180	-0,137	0,017	10,984	102,7	123059,9	3580,5
GOLDS	0,000	0,025	0,235	-0,210	0,294	14,144	266,8	118995,0	3572,3
HOME	0,000	0,021	0,132	-0,339	-0,991	24,167	928,3	115497,2	3576,0
INTEL	0,000	0,025	0,183	-0,249	-0,475	11,206	15199,8	114745,9	3569,3
INTER	0,000	0,017	0,123	-0,169	-0,022	11,279	622,7	123842,8	3582,8
JP M	0,000	0,027	0,224	-0,232	0,262	14,945	68,2	102964,5	3532,8
JJ	0,000	0,013	0,115	-0,173	-0,567	20,063	693,2	112242,4	3570,1
MCD	0,000	0,016	0,090	-0,137	-0,196	9,113	342,4	126810,2	3588,4
MERCK	0,000	0,019	0,123	-0,312	-1,557	31,876	490,0	119511,8	3576,6
MS	0,000	0,020	0,179	-0,170	-0,146	12,399	12236,2	86496,6	3534,1
NIKE	0,001	0,020	0,134	-0,216	-0,546	14,885	454,8	121612,6	3586,7
PFIZER	0,000	0,017	0,097	-0,118	-0,297	8,264	200,2	124705,0	3579,1
P&G	0,000	0,015	0,097	-0,377	-4,870	125,870	119,8	122023,9	3579,3
TRAV	0,000	0,020	0,228	-0,201	0,329	17,854	998,9	115504,0	3574,0
UT	0,000	0,018	0,128	-0,332	-1,607	36,955	139,5	121136,5	3583,5
UNH	0,001	0,021	0,298	-0,206	0,272	22,407	177,1	121302,5	3584,7
VERI	0,000	0,017	0,137	-0,126	0,140	9,471	284,7	112518,0	3564,4
WAL	0,000	0,016	0,105	-0,093	0,196	8,607	1510,0	107502,8	3548,7
WD	0,000	0,020	0,148	-0,203	-0,059	11,901	1130,2	118845,7	3580,8

Table 1: Descriptive Statistics: All Assets

Further, to investigate the information content of the given return series, we decompose each series into 8 time scales via Haar a trous wavelet decomposition. Figure 3 exemplifies the wavelet decomposition for 3M and illustrates each scale[1]. As

---

[1] We illustrate the wavelet decomposition for 3M, since it is the first asset in DWJI, graphs for

can be seen, scale 1 represents the short-run noise and thus incorporates the short-run information. Generally we can see, that the larger the scale the less volatile the decomposed series gets. Technically speaking, scale 1 incorporates information and noise of 2 days per observation, whereas scale 8 incorporates the noise of 256 days per observation and thus the long run trend.
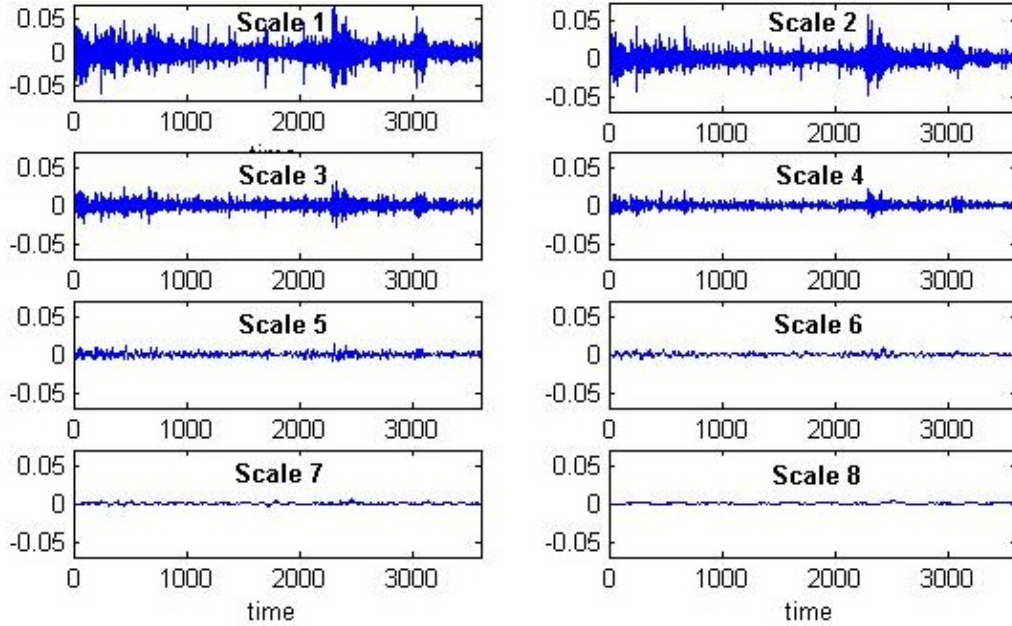

Figure 3: Decomposed Scales: 3M 2000-2013 !!DRAFT/BAD RESOLUTION!!

Note that the time scales are stochastically independent by construction, as illustrated in Table 2. In the empirical application, all correlations are close to zero so that there is no significant linear dependence between the scales.

---

all other assets are not stated due to page constraints.

Asset: 3M	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8
Scale 1	1,00	0,00	-0,02	0,03	0,00	0,02	0,00	0,00
Scale 2		1,00	-0,01	0,01	0,00	0,01	-0,01	0,00
Scale 3			1,00	-0,06	0,06	0,03	0,01	0,00
Scale 4				1,00	0,00	-0,03	0,02	0,01
Scale 5					1,00	-0,05	0,03	0,03
Scale 6						1,00	-0,03	-0,07
Scale 7							1,00	-0,06
Scale 8								1,00

Table 2: Correlation of Scales: 3M

As illustrated in Table 3, the volatility of the series is decomposed into eight scales, whereas the volatility declines for larger scales. As can be seen, due to the independence of the scales, the wavelet decomposition is "energy preserving" so that the variances of each scale sum up to the variance of the return series.

As well, the variance is mainly described by the first three scales, whereas the last scales do only account for a small part. Economically speaking, most of the risk is described by the short-run noise which is captured in the first three scales. Consequently, in the context of forecasting, the first scales will be in scope of interest and we will investigate how much information is needed to score adequate forecasts.

Asset: 3M								
Variances	$[\sigma^2 * 10^7]$							
Returns	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8
1636	892	391	199	83	38	18	8	7

Variances for each scale are presented in the format $\sigma^2 * 10^7$.

Table 3: Variance of Scales: 3M

Thus, turning the focus to the evaluation of information in the context of forecasting, we treat every series independently to identify the incorporated information content. So that, in order to filter the series at hand in terms of autocorrelation and heteroscedasticity, the GJR-GARCH model is fitted to each return series and each of its corresponding scale separately. For illustration purposes[2], the relevant parameters for the individual asset 3M are presented. Table 4 states the GJR-GARCH

---

[2]Tables for each individual asset are available upon request to the author.

parameters and the corresponding standard errors for each scale. Consequently, the GJR-GARCH model is fitted to each of the 29 assets and the corresponding 8 scales. As given in the example, we can see that the GARCH-parameters, α and β, are significant for all scales. However, as a result of the decomposition β (accounting for the persistency of the series) declines as the scale gets larger whereas α (accounting for the shocks) increases remarkably. In the illustrated example, β decreases from 0,91 (returns) down to 0,07 (scale 7) and increases to 0,41 again (scale 8). However, the increase in scale 8 can be explained by the long-run trend. As every observation incorporates the information of 256 return observations, no clear distinction between persistency and shock is possible. Interestingly, γ which accounts for the asymmetry in the process, is significant in the first scale and the last three scales. This can be explained by the fact that negative shocks are either relevant in the short run on an ad-hoc basis as in the return series and scale 1 or, as more days are comprised by the shock, if the shock can be interpreted as a negative market trend.

| **Asset: 3M** | | | | | | | | | |
Parameters	Returns	S1	S2	S3	S4	S5	S6	S7	S8
Ω	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
α	0,03	0,17	0,31	0,37	0,69	0,76	0,85	0,85	0,55
t-Value	452,89	306,38	391,72	482,31	491,36	558,04	580,50	808,87	576,68
β	0,91	0,77	0,65	0,58	0,16	0,17	0,11	0,07	0,41
t-Value	5141,06	1584,22	804,37	872,77	212,83	261,30	196,33	-963,75	1308,26
γ	0,07	0,03	0,00	0,00	0,00	0,00	0,01	0,03	0,08
t-Value	385,92	50,18	0,00	0,00	0,00	0,00	-16,37	42,08	79,64
LL	10055,60	11342,55	12650,55	14051,29	15596,95	17089,24	18520,67	20013,00	21709,81

Table 4: Estimated GJR-GARCH Parameters: 3M

Next, based on the filtered return series, to investigate the impact of the decomposed volatility on the forecasting of VaR, we translate the filtered return and scale series into 95% VaR forecast. As we am explicitly interested in the left tails of each time series, we do not assume any elliptical distribution and fit the Generalized Pareto distribution to each of the individual return series and its corresponding 8

scales to model the 5% quantile for every series. Unlike in the elliptical case, VaR is not a direct function of the volatility so that the individual VaRs for each scale do not sum up to the total VaR. However, the results allow us to investigate the tail of each series and thus the contribution to the overall VaR figure.

	Returns	Scale 1	Scale 2	Scale 3	Scale 4	Scale 5	Scale 6	Scale 7	Scale 8
3M	-3,00%	-1,22%	-0,67%	-0,47%	-0,23%	-0,11%	-0,05%	-0,03%	-0,01%
AT&T	-2,51%	-1,11%	-0,82%	-0,53%	-0,29%	-0,13%	-0,06%	-0,03%	-0,01%
AM EXP	-5,43%	-1,37%	-1,10%	-0,63%	-0,40%	-0,16%	-0,09%	-0,03%	-0,01%
BOEING	-3,48%	-1,39%	-0,97%	-0,66%	-0,32%	-0,16%	-0,07%	-0,03%	0,01%
CATER	-3,86%	-1,55%	-1,07%	-0,65%	-0,36%	-0,17%	-0,08%	-0,03%	0,02%
CHEVR	-2,91%	-1,07%	-0,77%	-0,50%	-0,28%	-0,13%	-0,06%	-0,03%	-0,01%
CISCO	-7,40%	-1,81%	-1,27%	-0,72%	-0,45%	-0,20%	-0,10%	-0,04%	0,01%
COCAC	-3,11%	-0,89%	-0,57%	-0,42%	-0,20%	-0,10%	-0,05%	-0,02%	0,01%
E I DU	-3,88%	-1,27%	-0,89%	-0,58%	-0,32%	-0,14%	-0,07%	-0,03%	-0,01%
EXXON	-3,10%	-1,01%	-0,75%	-0,46%	-0,26%	-0,13%	-0,06%	-0,03%	-0,01%
GE	-4,01%	-1,28%	-0,76%	-0,56%	-0,33%	-0,15%	-0,07%	-0,02%	-0,01%
GOLDS	-4,28%	-1,77%	-1,10%	-0,75%	-0,39%	-0,19%	-0,10%	-0,04%	0,01%
HOME	-5,97%	-1,39%	-0,87%	-0,62%	-0,34%	-0,16%	-0,07%	-0,04%	0,00%
INTEL	-4,79%	-1,51%	-1,09%	-0,72%	-0,43%	-0,19%	-0,10%	-0,04%	0,02%
INTER	-3,11%	-1,14%	-0,71%	-0,51%	-0,29%	-0,13%	-0,06%	-0,03%	-0,01%
JP M	-5,83%	-1,66%	-1,04%	-0,69%	-0,43%	-0,17%	-0,09%	-0,03%	-0,01%
JJ	-2,53%	-0,85%	-0,59%	-0,36%	-0,18%	-0,09%	-0,04%	-0,02%	0,00%
MCD	-3,58%	-1,10%	-0,77%	-0,48%	-0,26%	-0,11%	-0,06%	-0,02%	0,00%
MERCK	-3,66%	-1,11%	-0,81%	-0,57%	-0,28%	-0,13%	-0,07%	-0,02%	0,00%
MS	-5,48%	-1,40%	-0,89%	-0,55%	-0,34%	-0,16%	-0,07%	-0,03%	0,00%
NIKE	-3,43%	-1,28%	-0,83%	-0,60%	-0,30%	-0,13%	-0,07%	-0,03%	-0,01%
PFIZER	-7,50%	-1,20%	-0,78%	-0,53%	-0,28%	-0,13%	-0,06%	-0,03%	-0,01%
P&G	-3,42%	-0,85%	-0,60%	-0,37%	-0,19%	-0,10%	-0,05%	-0,02%	0,00%
TRAV	-4,41%	-1,30%	-0,86%	-0,55%	-0,31%	-0,14%	-0,07%	-0,03%	-0,01%
UT	-3,78%	-1,22%	-0,81%	-0,52%	-0,26%	-0,14%	-0,06%	-0,03%	-0,01%
UNH	-3,80%	-1,38%	-1,04%	-0,60%	-0,33%	-0,15%	-0,08%	-0,04%	-0,01%
VERI	-3,84%	-1,21%	-0,79%	-0,51%	-0,27%	-0,13%	-0,05%	-0,02%	-0,01%
WAL	-2,71%	-1,04%	-0,72%	-0,46%	-0,24%	-0,12%	-0,05%	-0,03%	-0,01%
WD	-5,82%	-1,34%	-0,89%	-0,63%	-0,30%	-0,14%	-0,07%	-0,03%	-0,01%

Based on every single return series and individual scales, the 95% VaR is modeled. VaR is modeled via GJR-GARCH approach and based on Generalized Pareto Distribution. Every forecast is based on 3600 observations.

Table 5: 95% Value-at-Risk Forecasts per scale: All Assets

Table 4, summarizes the decomposed VaR forecasts for each asset and its corresponding decomposed scales. As already indicated by the decomposition of the variance (illustrated in Table 3), the average VaR forecasts are declining for larger

scales. Interestingly, by focusing on the VaR forecasts for the different assets, VaR forecasts based on the return series vary stronger by each asset than the VaR figures for, lets say, scale 1. For instance, the highest overall VaR forecast is given for Pfizer and equals -7,50% and the corresponding VaR forecast based on the information exclusively incorporated in scale 1 equals -1,20%. In comparison to this, the lowest VaR forecast, given by AT&T, equals -2,51% and the corresponding VaR based on scale 1 is given by -1,11%. This illustration underlines the independence of each series, whereas it also shows, that the overall VaR forecasts seem to vary stronger by each individual asset than the volatility which is captured in the short-run scales.

Consequently, the overall risk seems to depend more on the investigated asset than the decomposed risk for each scale. More concrete, according to the stated VaR figures, forecasts based on scale 1 do not vary as the VaR forecasts based on the return series. Thus, all assets exhibit similar risk in the short-run, whereas the overall risk figures based on the return series seem to be more individual. However, analogue to the analysis of the variance, it is to state, that scale 1 - 4 seem to incorporate most of the information regarding the risk, whereas the larger scales appear to be of marginal impact regarding the overall VaR figure.

4 Value-at-Risk Analysis

To gain further insight into the information content of each time scale with respect to risk related questions, we go on by setting up an empirical VaR forecasting study, which enables us to judge the quality of the VaR forecasts based on different scales and thus its individual information content. To do so, we investigate the quality of VaR forecasts over time via a rolling window analysis. Namely, we model iterative VaR forecasts on past 1000 days to forecast the 1001st. Starting from January 1st 2000, the first day to forecast is December 20th 2003 and the last day is September 30th 2013. It is to note, that the rolling window stays constant and each forecast is based on the following steps:

i. Wavelet decomposition of 1000 observations

ii. Fitting GJR-Garch model to each series (returns and each of the 8 scales)

iii. Fitting Generalized Pareto Distribution to each series to model the quantile (Additionally we fit normal and t distribution as a benchmark)

iv. Compare VaR univariate forecast with the realized return, represented by the 1001st observation

v. As well, we link forecasts of different scales via t copula to analyse joint probability functions

Thus, this setup enables us to investigate 2600 VaR forecasts for each return series and each of its corresponding scale. As we also investigate elliptical distributions, to set the benchmark for our forecasts, we analyse more than 2 million daily VaR forecasts in total[3]. In the context of this analysis, the main criterion will be the failure rate of VaR misspecifications. As we forecast 95% VaR figures, we expect a failure rate of 5%. Consequently, the failure rate will be the main quality criteria of interest and thus VaR forecasts are preferred which result in an empirical rate close to the expected 5%.

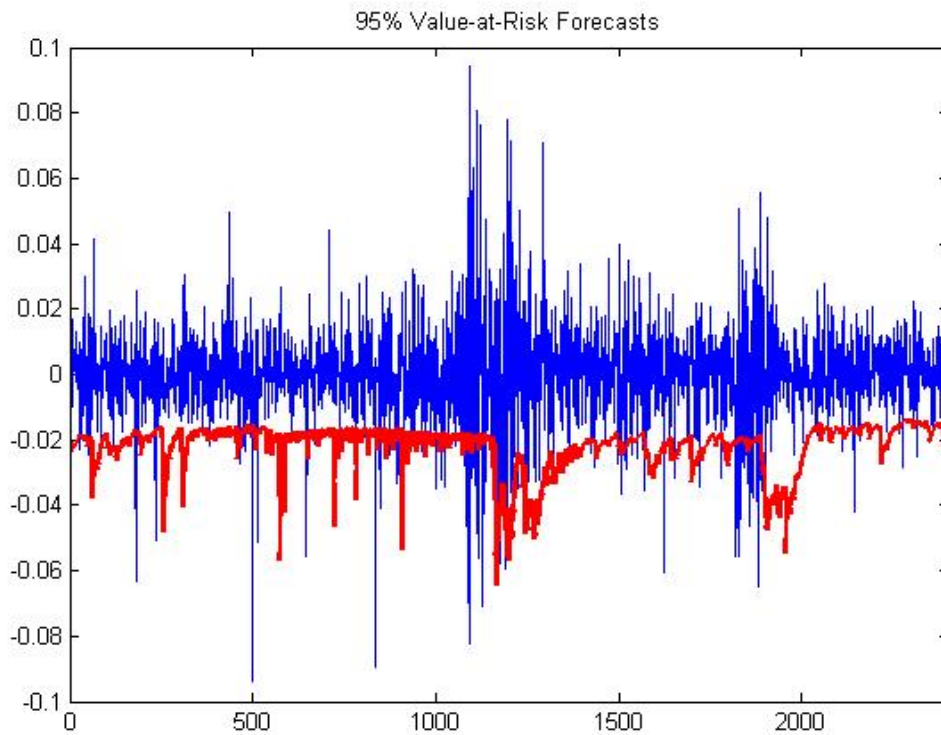[3]29 assets, 9 scales, 2600 days and 3 different distributional assumptions.

Figure 4: 95% Value-at-Risk Forecasts !!!DRAFT!!!

Table 4 illustrates the methodology of VaR forecasts. Every forecast is evaluated against the realized return, whereas a VaR breach is indicated if the negative return is larger than the VaR forecast. As a benchmark for our VaR performance analysis we also apply, lets say, classical parametric approaches by assuming normal as well as t distribution for the underlying series. In this context, the VaR is simply the square root of the GARCH-variance of the underlying time series multiplied by the quantile of the assumed elliptical return distribution (as presented in section 2.5). However, for sake of page constrains, in the remainder of this chapter we will focus on the most interesting results regarding VaR forecasts[4]. Namely, we present VaR forecasts based on the benchmark (normal and t distribution) and on the information of scale 1 modeled via GP distribution. As well, we present the effect of successively adding the information which are given in the larger scales by joining the scales via t copulas.

---

[4] All results are available upon request.

Moreover, in order to start our analysis we separate the investigated sample into a pre-crisis and post-crisis subsample, whereas we use the Lehman bankruptcy (September 15th 2008) as the starting point of the post-crisis subsample. Consequently, the pre-crisis sample covers 1026 days and the post-crisis sample 1574 days. We do so to take account for the changed market regimes and to avoid that conservative forecasts in calm market phases are compensated with breaches in volatile market phases as presented in Berger & Missong (2013).

Asset 3M	Gauss Returns	t Returns	GPD Scale 1	Scale 1&2	Scale 1&2&3
Pre-Crisis	$3,54\%^{**}$	$2,42\%^{**}$	$10,34\%^{**}$	$5,77\%$	$4,19\%$
Post-Crisis	$5,50\%$	$4,30\%$	$10,70\%^{**}$	$7,08\%^{**}$	$5,88\%$
Total	$4,63\%$	$3,46\%^{*}$	$10,54\%^{**}$	$6,50\%^{**}$	$5,13\%$

The percentages represent the relative amount of VaR breaches for the investigated period, return distribution and data (either returns or scales). */** indicate a rejection of the LR_{CC}-hypothesis at a 95%/99% significance level.

95% Value-at-Risk Empirical Failure Rate: 3M 2000-2013

Table 4 summarizes the relevant statistics, namely the empirical failure rate of the VaR forecasts, for 3M. Results regarding the VaR forecasts are given for the pre- and post-crisis as well as for the total sample[5]. We present the empirical failure rate for the two benchmark models, namely VaR forecasts based on the filtered return series under the assumption of normally and t-distributed returns. As well we show VaR forecasts based on scale 1, scale 1 and scale 2 linked via t copula and scale 1, scale 2 and scale 3 as well linked via t copula[6].

First and foremost it is to say, that the assumption of normally distributed returns leads to an acceptable performance if we solely focus on the empirical failure rate for the total sample (4,63%), whereas the assumption of t-distribution leads to too conservative VaR forecasts and thus less breaches (3,46%). However, if we investigate the predefined subsamples, we observe different performance, due to lower volatility regime in the pre-crisis period.

---

[5]Statistics for all investigated assets are available upon request

[6]It is to note, that forecasts i.e. based on scale 1, are benchmarked against the realized gains and losses stemming from the return series.

The second part of the Table illustrates the successive addition of information, starting from scale 1. We present the VaR forecasts, based on different time scales under the assumption of GP distributed returns. As can be seen, VaR forecasts solely based on scale 1 do lead to an inadequately high amount of VaR exceedamces. Having in mind that the VaR forecasts decline by larger scales, it gets clear that forecasts based on individual scales do not lead to adequate amount of mis-specifications. Thus there is no scale, which incorporates all information needed to forecast daily volatility. However, successively adding information to the first scale, by joining the VaR forecasts based on scale 2 via t copula, results in a decreased failure rate. If we also add the information incorporated in scale 3, we end up in a failure rate which is described by a higher precision than the applied benchmarks. If we continue by successively adding more scales (which is not presented here), we result in slightly too conservative VaR forecasts.

Next, in order to show the validity of the results, we go on by presenting the relevant empirical failure rates for the two benchmark approaches and the three joint scales, for each of the investigated assets in Table 6.

Asset	Gauss Returns	t Returns	GPD Scale 1&2&3	Asset	Gauss Returns	t Returns	GPD Scale 1&2&3
3M	4,63%	3,46*%	**5,13%**	JP M	4,79%	3,71%**	**5,04%**
AT&T	4,29%	3,46%	**4,96%**	JJ	3,58%**	2,88%**	**4,46%**
AM EXP	**5,00%**	3,50%**	5,50%	MCD	3,83%**	2,67%*	**4,00%**
BOEING	4,83%	3,79%*	**5,00%**	MERCK	3,25%**	2,92%**	**4,54%**
CATER	**5,04%**	3,50%*	5,04%	MS	4,42%	3,38%*	**5,21%**
CHEVR	6,21%**	**4,96%**	6,13%	NIKE	3,79%*	3,13%**	**4,58%**
CISCO	3,83%**	2,83%**	**4,67%**	PFIZER	4,04%*	3,13%**	**4,88%**
COCAC	4,38%	3,21%*	**5,08%**	P&G	5,08%	3,92%	**5,04%**
E I DU	4,58%	3,75%**	**5,33%**	TRAV	4,00%*	3,00%**	**4,96%**
EXXON	5,50%	4,58%*	6,04%*	UT	4,33%	3,58%**	**5,33%**
GE	**4,67%**	3,42%**	5,54%	UNH	4,46%	3,38%**	**5,25%**
GOLDS	**4,75%**	3,67%**	**4,75%**	VERI	4,75%	3,83%*	**5,04%**
HOME	4,29%	3,29%**	**4,88%**	WAL	**4,92%**	3,92%	4,79%
INTEL	4,50%	3,29%**	**4,92%**	WD	4,46%	3,54**%	**5,50%**
INTER	4,63%	3,67%*	**4,83%**				

The percentages represent the relative amount of VaR breaches for the investigated period, return distribution and data (either returns or scales linked via copula. The superior ampount of breaches for each stock is market in bold. */** indicate a rejection of the LR_{CC}-hypothesis at a 95%/99% significance level.

Table 6: Empirical Value-at-Risk Failure Rate: All Assets 2000-2013

By comparing the overall results, we see that VaR forecasts based on the assumption of normally distributed returns do only lead to more precise forecasts for 3 assets (Caterpillar, General Electrics and Goldman Sachs). Analogue, the assumption of t distribution leads for two assets to more precise VaR forecast (Exxon and Wal Mart). Hence, VaR forecasts based on the first three scales, modeled via GPD and linked via t copula do clearly outperform the elliptical benchmark in terms of precision. Consequently, the first three scales of each return series incorporate the necessary information to forecast daily VaR.

5 Conclusion

We decomposed financial return series into different time scales and investigated the information content of each time scale in a VaR framework. The empirical data set comprises daily returns of stocks listed in the Dow Jones Industrial Average from 2000 to 2013.

Based on the decomposed series, we fit GPD to each scale and demonstrate that the relevant information regarding daily VaR forecasts is captured in the scales comprising the short-run stochastic behavior of the series. Moreover, individual VaR forecasts based on filtered return series vary stronger by each individual asset than the VaR forecasts based on different time scales. The main information concerning VaR forecasts is inherent in the first three time scales, whereas the long-run information contributes only marginally to the overall VaR figure.

Further, we add to the analysis of decomposed return series by investigating the empirical quality of VaR forecasts for each scale. Moreover, we linked individual scales using elliptical copulas. By successively adding the information of larger scales to the VaR forecasts, we show that the relevant information for daily VaR forecasts seems to be captured by the first three scales.

We find that the information of the first scales linked via copulas outperformed the applied parametric benchmark approaches based on the assumption of normally and t distributed returns. Thus, combining the VaR forecasts, modeled via Extreme Value Theory, of the first three scales via t copula leads to VaR forecast which are described by a higher precision regarding the empirical failure rates in calm and turmoil market times for all investigated assets.

Based on our results, we illustrated that the relevant information for daily VaR forecasts is stored in the time scales comprising the short-run trend. Consequently, we strongly recommend to model daily VaR forecasts by capturing the VaR of the low time scales joint via copulas in order to achieve precise daily VaR forecasts.

References

[1] Angelidis T., A. Benos and S Degiannakis(2004), *The use of garch models in var estimation*, Statistical Methodology, 1, 105-128.

[2] Bekiros S. and D. Georgoutsos (2005), *Estimating value-at-risk by extreme value and conventional methods: a comparative evaluation of their predictive performance*, International Financial Markets, Institutions and Money, 15, 209-228.

[3] Berger T. (2013), *Forecasting Value-at-Risk using time varying copulas and EVT return distributions*, International Economics, 133, 93-106.

[4] Berger T. and M. Missong (2013), *Value-at-Risk Forecasts, and the Puzzle of Dependency Modeling*, International Review of Financial Analysis, forthcoming.

[5] Brooks C., A. Clare, J. Dalle Molle and G. Persand (2005), *A comparison of extreme value theory approaches for determining value at risk*, Journal of Empirical Finance, 12, 339-352.

[6] Boubaker H. and N. Sghaier (2012), *Portfolio optimization in the presence of dependent financial returns with long memory: a copula approach*, Journal of Banking and Finance, forthcoming.

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

[8] Cifter A. (2011), *value-at-risk estimation with wavelet-based extreme value theory: evidence from emerging markets*, Physica A, 390, 2356-2367.

[9] Embrechts, P., A. McNeil and D. Straumann (2002),*Correlation and Dependence in Risk Management: Properties and Pitfalls*, Risk Management: Value at Risk and Beyond, Cambridge University Press, 176-223.

[10] Glosten, L.,R. Jagannathan and D. Runkle (1993),*On the relatuion*, Risk Management: Value at Risk and Beyond, Cambridge University Press, 176-223.

[11] He K., K. Lai and G. Xiang (2012), *portfolio value at risk estimate for crude oil markets: a multivariate wavelet denoising approach*, Energies, 5, 1018-1043.

[12] Joe H. (1996), *Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters.*, In: Ruschendorf, L., Schweizer, B., Taylor, M.D. (Eds.), Distributions with Fixed Margins and Related Topics. In: IMS Lecture Notes Monograph Series, 28, 120-141.

[13] Jammazi R. and C. Aloui (2012), *Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modelling*, Energy Economics, 34(3), 828-841.

[14] Jorion, P. (2007), *Value at Risk. The new Benchmark for Controlling Derivatives Risk*, McGraw, New York.

[15] Khalfaoui R. and M. Boutahar (2012), *portfolio risk evaluation: an approach based on dynamic conditional correlations models and wavelet multiresolution analysis*, MPRA paper no 41624.

[16] Kupiec P. (1995), *Techniques for verifying the accuracy of risk measurement models*, Journal of Derivatives, 2, 73-84.

[17] Lo Cascio I. (2007), *wavelet analysis and denoising: new tools for economics*, Working paper no. 600, Quen Mary University London.

[18] Murtagh F., J.L. Starck and O. Renaud (2003), *On Neuro Wavelet Modeling*, working paper.

[19] Percival D. and A. Walden (2000), *Wavelet methods for time series analysis*, Cambridge University Press.

[20] Ramsey J. (2002), *Wavelets in economics and finance: Past and future*, Studies in Nonlinear Dynamics and Econometrics, 6(3), 1-27.

[21] Reboredo J. and M. Riveiro-Castro (2014), *Wavelet-based evidence of the impact of oil prices on stock returns*, International Review of Economics and Finance, 29, 145-176.

[22] Renaud O., J.L. Starck and F Murtagh (2002), *Prediction based on a multiscale decomposition*, report.

[23] Shensa M.J. (1991), *discrete wavelet transforms: wedding the a trous and mallat algorithms*, IEEE Transactions on signal processing, 40, 2464-2483.

[24] Sklar C. (1959), *Fonctions de repartition a n dimensions et leurs marges*, Publicationss de Institut Statistique de Universite de Paris, 8, 229-231.

[25] Starck J.L., J. Fadili and F. Murtagh (2007), *The undecimated wavelet decomposition and its reconstruction*, IEEE Transactions on Image Processing, 16, 297-309.