Forecasting Value-at-Risk with a Duration based POT method

P. Araújo Santos
Escola Superior de Gestão e Tecnologia de Santarém and CEAUL
paulo.santos@esg.ipsantarem.pt

M.I. Fraga Alves
Faculdade de Ciências da Universidade de Lisboa and CEAUL
isabel.alves@fc.ul.pt

March 15, 2011

Abstract: Threshold methods, based on fitting a stochastic model to the excesses over a threshold, were developed under the acronym POT (peaks over threshold). In order to eliminate the tendency to clustering of violations, a model based approach within the POT framework, that uses the durations between excesses as covariates, is proposed. Based on this approach, models for forecasting one-day-ahead Value-at-Risk were suggested and applied to real data. Comparative studies provide evidence that they can perform better than state-of-the art risk models and much better than the widely used RiskMetrics model, both in terms of out-of-sample accuracy and under the Basel II Accord.

Keywords: quantitative risk management; extreme value theory; financial time series; clustering of violations; Basel II Accord.

1 Introduction

Investors and traders must pay attention not only to the expected return from their activities but also to the risks that they incur. It is widely accepted that risk-adjusted performance measures can guide institutions toward a better risk/return profile and can play a relevant role to achieve a more secure financial system. This justify the interest of developing more accurate risk models. Value-at-Risk (VaR) aggregates several components of risk into a single number and has emerged as the standard measure in quantitative risk management. In terms of regulation, the Basel II Accord requires that banks and other Authorized Deposit-taking Institutions (ADIs) to report their daily VaR forecasts to the monetary authorities (typically, a central bank) at the beginning of each trading day and defines daily capital requirements based on these forecasts (for a detailed discussion of VaR, see Jorion, 2000). We will consider the symmetric of daily log returns, $R_{t+1} = -\log(V_{t+1}/V_t) \times 100$, where V_t is the value of the portfolio at time t. The one-day-ahead VaR forecast made at time t for time $t+1$, $VaR_{t+1|t}(p)$, is defined by

$$P[R_{t+1} > VaR_{t+1|t}(p)|\Omega_t] = p,$$

where Ω_t is the information set up to time-t and p is the *coverage rate*. A *violation* occurs when the symmetric daily return exceeds the reported VaR, i.e., when $R_{t+1} > VaR_{t+1|t}(p)$. The rest of the paper is organized as follows. In Section 2, we review the peaks over threshold (POT) method with an example that illustrates the problem of *tendency to clustering of violations*. In Section 3, in order to solve this problem, we propose risk models based on durations and within the POT framework. Comparisons between the proposed risk models and other models are made in Section 4. Finally, conclusions and directions for future research are given in Section 5.

2 The POT method and the tendency to clustering of violations problem

The Generalized Pareto Distribution (GPD) has the form

$$G_{\gamma,\sigma}(y) = \begin{cases} 1-(1+\gamma y/\sigma)^{-1/\gamma}, & \gamma \neq 0 \\ 1-\exp\left(-y/\sigma\right), & \gamma = 0, \end{cases} \tag{2.1}$$

where $\sigma > 0$, and the support is $y \geq 0$ when $\gamma \geq 0$ and $0 \leq y \leq -\sigma/\gamma$ when $\gamma < 0$. The expected value and variance are given by

$$E[Y] = \frac{\sigma}{1-\gamma} \quad (\gamma < 1), \qquad VAR[Y] = \frac{\sigma^2}{(1-\gamma)^2(1-2\gamma)} \quad (\gamma < 1/2).$$

Generally, with $\gamma > 0$, $E[Y^c]$ does not exist for $\gamma \geq 1/c$. The probability that the random variable (r.v.) X assumes a value that exceeds a threshold u by at most y, given that it does exceed the threshold, is given by the *excess distribution*

$$F_u(y) = P[X-u \leq y | X > u] = \frac{F(y+u)-F(u)}{1-F(u)}, \tag{2.2}$$

for $0 \leq y < x^F - u$, where x^F is the (finite or infinite) right endpoint of F, defined by $x^F := \sup\{x : F(x) < 1\}$. The Extreme Value Theory (EVT), with the following theorem, suggests the GPD (2.1) as an approximation for the excess distribution (2.2), for a sufficiently high threshold u.

Theorem 2.1. *(Balkema and de Haan (1974) and Pickands (1975)) It is possible to find a function $\beta(u)$ such that*

$$\lim_{u \to x_F} \sup_{0 \leq y < x_F - u} |F_u(y) - G_{\gamma,\beta(u)}(y)| = 0,$$

if and only if F is in the maximum domain of attraction of an extreme value distribution.

For a wide class of distributions, the excess distribution (2.2) over a high threshold u can be approximated by the GPD (2.1) and this result holds for essentially all common continuous distributions; more precisely, Theorem 2.1 holds for all distributions in some max-domain of attraction of an extreme value distribution, i.e., distributions for which the sequence of maxima linearly normalized converges to a non degenerate limit law. To estimate the parameters γ and σ we fit the GPD to the excesses over the conveniently chosen threshold u. For $\gamma > -1/2$, the standard properties of the maximum likelihood (ML) estimators have been proved by Smith (1987) and extended for $\gamma > -1$ by Zhou (2010). Furthermore, it is possible to show, using simulations, that inference is often robust to choice of the threshold u, when u is big enough. Smith (1987) proposed a tail estimator based on a GPD approximation to the excess distribution. We denote n the number of excesses above u in a sample $X_1, ..., X_{n_x}$. Using n/n_x as estimator of $\bar{F}(u)$ the relation $\bar{F}_u(x-u) = \bar{F}(x)/\bar{F}(u)$ and $\bar{F}_u(x-u)$ estimated by a GPD approximation, we obtain the tail estimator

$$\hat{\bar{F}}(x) = \frac{n}{n_x}\Big(1 + \hat{\gamma}\frac{x-u}{\hat{\sigma}}\Big)^{-1/\hat{\gamma}}, \quad \text{valid for } x > u. \tag{2.3}$$

For $p < \bar{F}(u)$ and inverting the tail estimator formula (2.3), we get the VaR POT estimator

$$\hat{\text{VaR}}^{POT}_{t+1|t}(p) = u + \frac{\hat{\sigma}}{\hat{\gamma}}\Big(\Big(\frac{n}{n_x p}\Big)^{\hat{\gamma}} - 1\Big). \tag{2.4}$$

Now, turning theory into practice, one example is presented to illustrate the problem of tendency to clustering of violations which occurs when we apply the VaR POT estimator (2.4) to financial time series. The data consist of 15190 daily returns of Standard & Poor's Index (S&P 500), from January 4, 1950 through May 18, 2010. We choose the threshold, $u = x_{13671:15190} = 0.9897$, such that 10% of the values are larger than the threshold; see McNeil and Frey (2000) for a simulation study that support a similar choice. In Figure 2.1 we present the returns with the threshold (grey line) and a histogram where we can observe how the GPD, with the parameters estimated by ML estimation, adjust very well to the excesses. In this example, we obtain a VaR(0.05) equal to 1.42 and a VaR(0.01) equal to 2.67. In Figure 2.2(a), instead of considering 15190 daily returns to obtain one VaR(0.01) estimate, we present one-day-ahead VaR forecasts with a rolling window of size 1000 ($n_w = 1000$). The percentage of days where the symmetric returns exceeds the correspondent VaR forecast, i.e., the percentage of violations, equals 1.367% of the 14190 days used for the out-of-sample forecasts, when the expected is 1%.

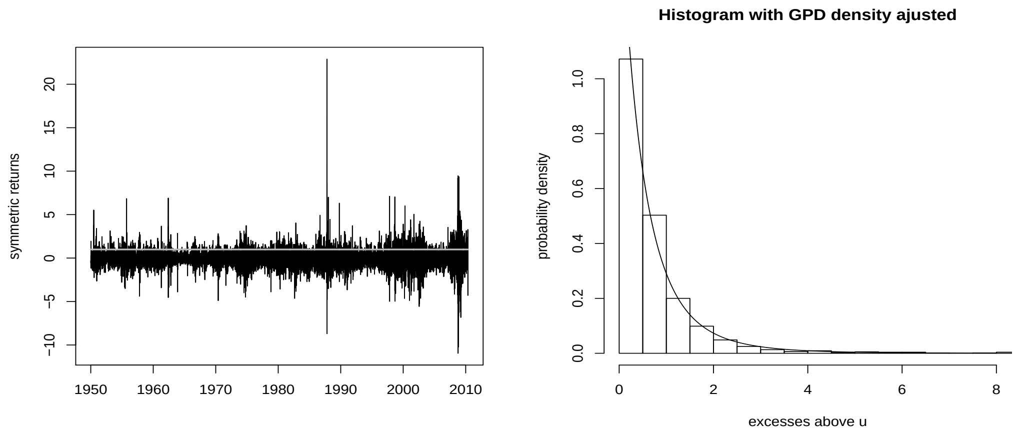
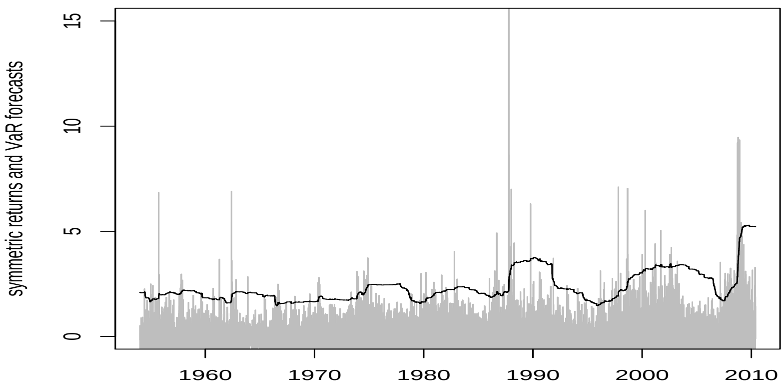
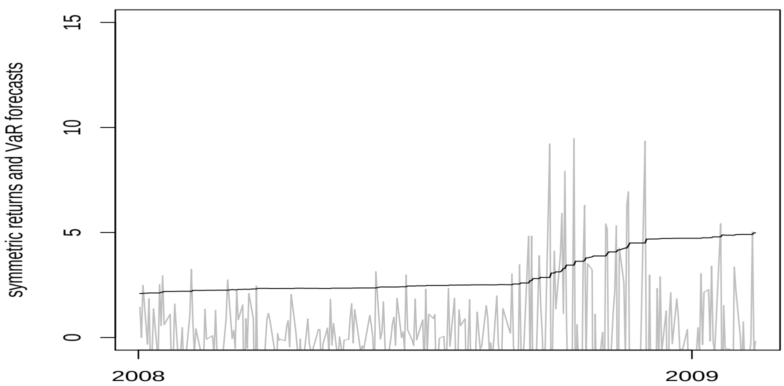


Figure 2.1: Symmetric returns (left) and Histogram of 1519 excesses above the threshold $u = x_{13671:15190} = 0.9897$ (right) for the S&P 500 Index from January 4, 1950 through May 18, 2010.

However, the serious problem of POT method and other unconditional models, is tendency to clustering of violations associated with the volatility clustering phenomenon. Figure 2.2(b) illustrates this problem during the 2008 financial crisis period. Between January 2, 2008 and February 12, 2009, we have a large number of violations in a short period of time. Over this period, the number of violations was 29, representing 10.28% of the 282 trading days, when the expected value for the percentage of violations is 1%.



(a) January 4, 1950 through May 18, 2010 (14190 trading days)



(b) January 2, 2008 through February 12, 2009 (282 trading days)

Figure 2.2: Symmetric returns of S&P 500 Index (grey line) and one-day-ahead VaR(0.01) forecasts with POT method (black line) and a rolling window of size 1000.

3 A duration based POT method (DPOT)

Our main goal is to eliminate the tendency to clustering of violations that occurs with the POT method. To achieve this goal, within the POT framework we propose the presence of durations between excesses as covariates. Smith (1990), developed ML and Least Squares estimation procedures under the POT framework with the shape and scale parameters dependent on covariates. For a general overview of EVT and its application to VaR, including the use of explanatory variables, see, for instance, Tsay (2010). For details about the mathematical theory of EVT and its applications to risk management, see Embrechts et al. (1997).

Let $y_1, ..., y_n$ be the excesses above a high threshold u, d_1 the duration until the first excess and $d_2, ..., d_n$, defined by

$$d_i = t_i - t_{i-1}, \tag{3.1}$$

where t_i denotes the day of excess i. We propose to use from the information set up to time t (Ω_t), the last v durations between excesses, $d_n, d_{n-1}, ..., d_{n-v+1}$ and the duration since the excess n which we define by d^t. With the durations $d_i, ..., d_{i-v+1}$, it is possible to consider at the time of excess number i, the duration since the preceding v excesses, defined by

$$d_{i,v} = d_i + ... + d_{i-v+1} = t_i - t_{i-v}. \tag{3.2}$$

At day t, after the excess n, we define $d_{t,1} = d^t$, $d_{t,2} = d^t + d_n$ and for $v = 3, 4, \cdots$,

$$d_{t,v} = d^t + d_{n,v-1} = d^t + d_n + \cdots + d_{n-v+2},$$

which represents the duration until t since the preceding v excesses.

3.1 Empirical Motivation

The motivation for the presence of durations between excesses as covariates has mainly been based on the relation between the amount of the excess and durations which we observe in various financial time series. Figure 3.1 (left) presents for the S&P 500 Index example of Section 1, the scatterplot of excesses (y_i) and durations since the preceding excess (d_i). Clearly, large excesses tend to be associated with short durations and small excesses tend to be associated with long durations. In Figure 3.1 (right) we observe a similar pattern for excesses and durations between the 2 preceding excesses (d_{i-1}). Table 1 gives Pearson correlations between excesses, durations and the inverse of durations. The linear association between excesses and durations is weak, but increases when we take the inverse of durations, as expected. Adding durations we get the duration since the preceding v excesses defined in (3.2) and the correlation increases a little more when we compute the correlation between excesses and the inverse of these durations. In short, the empirical results show some nonlinear association between excesses and durations. We also observe that the excesses have higher mean and higher variance with short durations, and lower mean and lower variance with long durations. Based on these empirical results, we propose to define the expected value and variance of the excesses dependent on the durations.

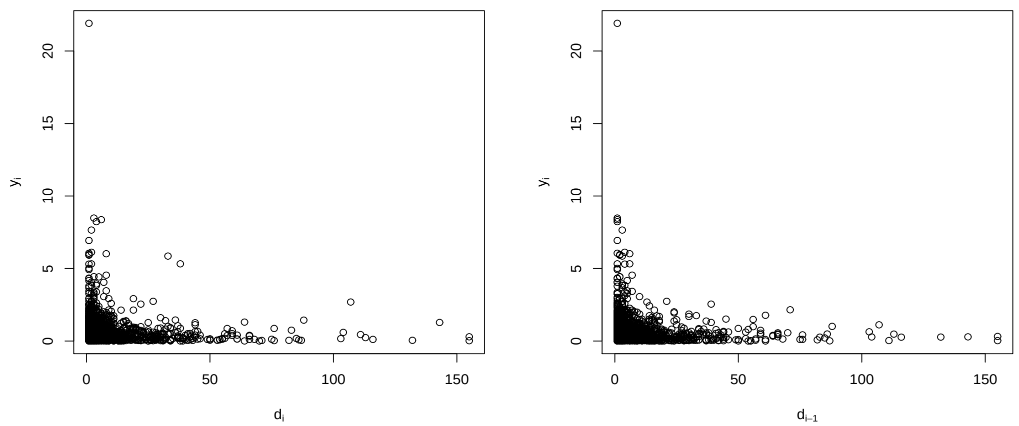


Figure 3.1: S&P 500 Index from January 4, 1950 through May 18,2010. Scatter plot of excesses above a high threshold ($u = 0.9897$) and durations since the preceding excess (left) and scatter plot of excesses and durations between the 2 preceding excesses (right).

Table 1

S&P 500 Index. Pearson correlation between y_i, d_{i-j}, $\frac{1}{d_{i-j}}$ and $\frac{1}{d_{i,v}}$.

j	$\text{Corr}(y_i, d_{i-j})$	$\text{Corr}(y_i, \frac{1}{d_{i-j}})$	v	$\text{Corr}(y_i, \frac{1}{d_{i,v}})$
0	-0.123	0.193	2	0.284
1	-0.127	0.174	3	0.325
2	-0.096	0.149	4	0.335
3	-0.126	0.148	5	0.346

3.2 DPOT Model

With the durations (3.1) and the duration since the excess n, d^t, we assume the GPD for the excesses Y_i above u, such that

$$Y_t \sim GPD\Big(\gamma, \sigma_t = g(\alpha_1, ..., \alpha_k, ..., d^t, d_n, d_{n-1}, ..., d_{n-v+2})\Big),$$

where $\gamma, \alpha_1, ..., \alpha_k$, are parameters to be estimated. And we propose the following class of estimators

$$\widehat{\text{VaR}}_{t+1|t}^{DPOT}(p) = u + \frac{\hat{\sigma}_t}{\hat{\gamma}}\Big(\Big(\frac{n}{n_x p}\Big)^{\hat{\gamma}} - 1\Big), \tag{3.3}$$

with $\hat{\sigma}_t = g(\hat{\alpha_1}, ..., \hat{\alpha_k}, ..., d^t, d_n, d_{n-1}, ..., d_{n-v+2})$.

The proposed $DPOT$ method implies, for $\gamma < 1$, a conditional expected value for excesses, and for $\gamma < 1/2$, a conditional variance, both dependent on d^t and the last v durations between excesses,

$$E[Y_t|\Omega_t] = \frac{\sigma_t}{1-\gamma} \quad (\gamma < 1), \quad VAR[Y_t|\Omega_t] = \frac{(\sigma_t)^2}{(1-2\gamma)} \quad (\gamma < 1/2).$$

The empirical results of Section 3.1 suggest a inverse relation between excesses and the durations since the preceding v excesses, with $1/(d_{i,v})^c$, $c > 0$, which leads to the specification $\sigma_t = \alpha \frac{1}{(d_{t,v})^c}$ and the VaR estimator

$$\widehat{\text{VaR}}_{t+1|t}^{DPOT(v,c)}(p) = u + \frac{\hat{\alpha}}{\hat{\gamma}(d_{t,v})^c}\Big(\Big(\frac{n}{n_x p}\Big)^{\hat{\gamma}} - 1\Big), \tag{3.4}$$

where $\hat{\gamma}$ and $\hat{\alpha}$ are estimators of the parameters γ and α. Applying the maximum likelihood theory to estimate the parameters, the log likelihood obtained is

$$\begin{aligned}
\log \mathrm{L}(\gamma,\alpha) &= \log \prod_{i=v}^{n} f_{Y_i}(y_i) \\
&= \log \prod_{i=v}^{n} \Big(\frac{\alpha}{(d_{i,v})^c}\Big)^{-1} \Big(1 + \frac{\gamma}{\alpha} y_i (d_{i,v})^c\Big)^{-(1/\gamma+1)} \\
&= -\sum_{i=v}^{n} \log \Big(\frac{\alpha}{(d_{i,v})^c}\Big) - \Big(\frac{1}{\gamma}+1\Big) \sum_{i=v}^{n} \log\Big(1 + \frac{\gamma}{\alpha} y_i (d_{i,v})^c\Big). \qquad (3.5)
\end{aligned}$$

We present results for $v = 3$, $c \in \{0.8, 0.75, 0.7\}$ and apply an implementation of Nelder and Mead algorithm, using the stats package of R (R Development Core Team, 2008), to maximize (3.5).

Using the proposed models with the S&P 500 Index returns presented in the Section 1 example, we obtain for 14190 one-day-ahead VaR forecasts, 138 (0.9725%), 134 and 134 (0.9443%) violations, respectively with $c = 0.8$, $c = 0.75$ and $c = 0.7$. These percentages are much closer to the expected 1% than the 1.367% obtained with the unconditional POT model. In Figure 3.2, the grey line corresponds to the S&P 500 returns, the dotted, the longslash, the solid and the longslash grey lines, correspond to one-day-ahead VaR forecasts calculated respectively with the DPOT($c = 0.8$), DPOT($c = 0.75$), DPOT($c = 0.7$) and the POT models. For the 2008 global financial crises period, Figure 3.2, shows how the DPOT models solve the problem of tendency to clustering of violations, producing much better risk forecasts that adjust quickly to the high volatility in the returns during September and October. Within this period of 282 days, the number of violations with DPOT($c = 0.8$) was 8, with DPOT($c = 0.75$) was 8 and with DPOT($c = 0.7$) was 11, much less than the 29 violations obtained with the unconditional POT method. Moreover, notice that with some exceptions, in the majority of the days the difference between DPOT($c = 0.8$), DPOT($c = 0.75$) and DPOT($c = 0.7$) forecasts, is very small, suggesting that the method is robust for different values of c in the interval between

0.7 and 0.8. Empirical findings in Section 4 will suggest that a choice of $c = 0.75$ is preferable. We also study the model with c estimated, but we achieve poor results.

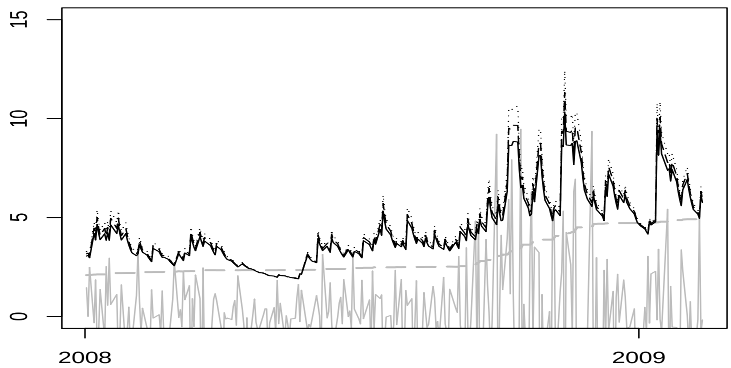


Figure 3.2: Symmetric returns of S&P 500 Index from January 2, 2008 through February 12,2009 (solid grey), and one-day-ahead VaR(0.01) forecasts with DPOT($c = 0.8$) (dotted), DPOT($c = 0.75$) (longdash), DPOT($c = 0.7$) (solid), POT method (longdash grey) and a rolling window of size 1000.

4 Comparative studies

Using the returns from S&P 500 Index, German stock market Index (DAX) and Financial Times London Stock Exchange Index (FTSE), we compare the proposed DPOT method with a two-stage hybrid method which combines a time-varying volatility model with the EVT approach, known as Conditional EVT, and with two conditional parametric models. We employ the R language in order to develop the programs. The data were obtained from the Web site http://finance.yahoo.com/. In Section 4.1 we briefly review the Conditional EVT method, the Asymmetric Power Autoregressive Conditional Heteroscedasticity (APARCH) model and the widely used RiskMetrics model. In Section 4.2 we evaluate the accuracy of out-of-sample interval forecasts produced with the risk models and in Section 4.3 we compare the performance under the Basel II Accord.

4.1 Conditional EVT, APARCH and RiskMetrics

The EVT procedure described in Section 2 is unconditional, however, to solve or reduce the problem of clustering, we can apply EVT to returns adjusted by some dynamic structure. It is usual to assume for the returns, $r_t = \mu_t + \varepsilon_t$, where ε_t is the unpredictable component and μ_t the conditional mean expressed as a sth order autoregressive process, AR(s),

$$\mu_t = \phi_0 + \sum_{i=1}^{s} \phi_i r_{t-i}.$$

The unpredictable component can be expressed by $\varepsilon_t = z_t \sigma_t$, where the innovations, z_t, are a sequence of independently and identically distributed random variables with zero mean and unit variance, and the conditional variance is

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

where $\alpha_i > 0$ and $\beta_j > 0$, for $i = 0, 1, ..., p$ and $j = 1, 2, ..., q$. This time-varying volatility model for the unpredictable component, is a Generalised Autoregressive Conditional Heteroscedasticity (GARCH) process, proposed by Bollerslev (1986). The GARCH model with $p = 1$ and $q = 1$, usually captures with success several stylized facts of financial time series. Diebold et al. (1998) proposed in a first step the standardization of the returns through the conditional means and variances estimated with a time-varying volatility model, and in a second step, estimation of a p quantile using EVT and the standardized returns. McNeil and Frey (2000) combine a AR(1)-GARCH(1,1) process assuming normal innovations with the POT method from EVT. We will denote this model as CEVT-n. The filter with normal innovations, while capable of removing the majority of clustering, will frequently be a misspecified model for returns. For accommodate this misspecification, Kuester et al. (2006) suggested a filter with the skewed t distribution. We will denote this model as CEVT-sst. Applying Conditional EVT, the VaR estimator is

$$VaR_{t+1|t}^{CEVT}(p) = \hat{\mu}_{t+1|t} + \hat{\sigma}_{t+1|t} \hat{z}_p,$$

where $\hat{\mu}_{t+1|t}$ and $\hat{\sigma}_{t+1|t}$ are the estimated conditional mean and conditional standard deviation for $t+1$, obtained with a AR(1)-GARCH(1,1) process. Moreover, $\hat{z}_p$ is a quantile p estimate, obtained with the POT method and the standardized residuals calculated as

$$(z_{t-n+1}, ..., z_t) = \left(\frac{r_{t-n+1} - \hat{\mu}_{t-n+1}}{\hat{\sigma}_{t-n+1}}, ..., \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t} \right).$$

Several studies conclude that conditional EVT is the method with better out-of-sample performance to forecast one-day-ahead VaR (e.g. McNeil and Frey (2000), Byström (2004), Bekiros and Georgoutsos (2005), Kuester et al. (2006), Ghorbel and Trabelsi (2008), Ozun et al. (2010)), and this is the reason wy we choose CEVT-n and CEVT-sst models for the comparative studies.

Empirical evidence shows that the increase in volatility is larger when the returns are negative than when they are positive. This asymmetric evolution of the conditional variance is known as leverage effect (Black, 1976). We also choose for the comparative study one asymmetric GARCH-type model, the APARCH model introduced by Ding, Granger and Engle (1993). The conditional variance of the APARCH(p, q) model can be written as

$$\sigma_t^\delta = w + \sum_{i=1}^{p} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta,$$

where $\delta > 0$ and $-1 < \gamma < 1$. The asymmetric coefficient γ, takes the leverage effect into account. We consider this model as it is a very general GARCH-type model, including as special cases several GARCH-type models and asymmetric GARCH-type models: ARCH Model of Engle ($\delta = 2$, $\gamma_i = 0$ and $\beta_j = 0$), GARCH Model of Bollerslev ($\delta = 2$, $\gamma_i = 0$), TS-GARCH Model of Taylor and Schwert ($\delta = 2$, $\gamma_i = 0$), GJR-GARCH Model of Glosten, Jagannathan and Runkle ($\delta = 2$), T-ARCH Model of Zakoian ($\delta = 1$), N-ARCH Model of Higgens and Bera ($\gamma_i = 0$ and $\beta_j = 0$) and log-ARCH Model of Geweke and Pentula ($\delta \to 0$). The model chosen for the comparative studies was the AR-APARCH(1,1) with skewed t innovations, which we denote by APARCH-sst.

$$\widehat{\text{VaR}}_{t+1|t}^{APARCH}(p) = \hat{\phi}_0 + \hat{\phi}_1 r_t + s_p \times \hat{\sigma}_{t+1|t},$$

with s_p a quantile p of the skewed t distribution with parameters estimated using the data. In a comparative study for the Asian markets, Tu, Wong and Chang (2008) found that the APARCH model with the skewed t distribution performs better than with the normal or with the student distribution. GARCH-type models with skewed t innovations have been frequently found to provide excellent forecast results; see, for example, Mittnik and Paolella (2000), Giot and Laurent (2004).

Finally, for the comparative study, we also choose the widely used RiskMetrics model developed by J.P. Morgan (J.P. Morgan´s Riskmetrics Technical Document, 1996). This model assumes that the return follows a conditional normal distribution $\text{N}(0, \sigma_t^2)$, with the dynamic of volatility modeled using a exponential weighted moving average (EWMA) method

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda)\varepsilon_{t-1}^2.$$

RiskMetrics (1996) suggests $\lambda = 0.94$ for daily data. The recursion can be initialized by the sample variance ($\sigma_1^2 = \hat{\sigma}^2$) or the square of the first return ($\sigma_1^2 = r_1^2$).

$$\widehat{\text{VaR}}_{t+1|t}^{RM}(p) = z_p \times \hat{\sigma}_{t+1|t},$$

with z_p a quantile p of the standard normal distribution. The empirical results of the following section will clearly suggest that with the normality assumption we obtain underestimated VaR forecasts and more violations than the expected.

4.2 Interval forecasts evaluation

In this section we compare the CEVT-sst, CEVT-n, APARCH-sst, RiskMetrics and DPOT models with $v = 3$, $c \in \{0.8, 0.75, 0.7\}$, denoted respectively by DPOT(0.8), DPOT(0.75) and DPOT(0.7). We examine the one-day-ahead VaR(0.01) forecasts performance with the S&P 500 Index, DAX Index and FTSE Index, considering returns produced by all the historical data until May 18, 2010. Using a rolling window of size 1000 we obtain 14190, 3917 and 5599 one-day-ahead VaR(0.01) forecasts for each model, respectively with the S&P 500, DAX and FTSE. As usual, the threshold u was chosen such that 10% of the values are larger than the threshold. The primary tool for assessing the accuracy of the interval forecasts is to monitor the binary sequence generated by observing if the return on day $t+1$ is in the tail region specified by the VaR at time-t, or not. This is referred to as the hit sequence

$$I_{t+1}(p) = \begin{cases} 1 & \text{if } R_{t+1} > VaR_{t+1|t}(p) \\ 0 & \text{if } R_{t+1} \leq VaR_{t+1|t}(p). \end{cases}$$

Christoffersen (1998) showed that evaluating interval forecasts can be reduced to examining whether the hit sequence satisfies the unconditional coverage (UC) and independence (IND) properties. To test the UC hypothesis we apply the Kupiec test (Kupiec, 1995). To test the IND hypothesis we apply two tests. In the same line as Engle and Manganelli (2004), Berkowitz *et al.* (2009) consider the autoregression

$$I_t = \alpha + \beta_1 I_{t-1} + \beta_2 \text{VaR}_{t|t-1}(p) + \varepsilon_t, \tag{4.1}$$

and propose the logit model. We can test the IND hypothesis with a likelihood ratio test considering for the null $\beta_1 = \beta_2 = 0$ and in this case the asymptotic distribution is chi-square with 2 degrees of freedom. We refer to this test as the CAViaR independence test of Engle and Manganelli (CAViaR). The other independence test applied was recently introduced in the literature (Araújo Santos and Fraga Alves, 2010) and is based on the ratio $(D_{N:N} - 1)/D_{[N/2]:N}$, where $D_{N:N}$ and $D_{[N/2]:N}$, are the maximum and the median of durations between consecutive violations and until the first violation. This new test is suitable for detect models with a tendency to generate clusters of violations, is based on an exact distribution, is pivotal in the sense that is based on a distribution that does not depend on an unknown parameter and outperforms, in terms of power, existing procedures in realistic settings. We refer to this test as MM ratio test.

The empirical findings, with the p values of the tests, are presented in Tables 2, 3 and 4. Table 5, summarize the results in terms of number of times that the hypotheses are rejected. As the unconditional POT model do not account for volatility clustering, is unable to produce iid violations and both independence tests reject the IND hypothesis with very small p values. With a violation frequency equal to 0.01367, the UC hypothesis is also clearly reject in the case of the POT model. We only considered the POT method in the Table 2. The performance of RiskMetrics is very poor. The violation frequency is even much worse than with the unconditional POT model. With RiskMetrics the violation frequencies equals 0.018675, 0.016845 and 0.018222, respectively for the S&P 500, DAX and FTSE indexes, much higher than the expected 0.01. For this model and with all indexes the UC hypothesis is rejected with very small p-values. Both DPOT, CEVT and APARCH-sst models performs very well in terms of the UC hypothesis, taking into account that in no case the hypothesis is rejected since all p-values are very high. It is interesting to note the impressive performance of CEVT models in terms of UC in Table 2, with 142 violations in 14190 out-of-sample forecasts it was impossible to obtain a better result (the violation frequency is equal to 0.01000705). The same impressive performance occurs with the DPOT(0.7) in Table 4, with 56 violations in 5599 out-of-sample forecasts was impossible to obtain a better result (the violation frequency is equal to 0.01000179). In terms of IND hypothesis, the DPOT models performs clearly better than the CEVT models and than APARCH-sst. Considering the eighteen cases with three DPOT models, three indexes and two independence tests, with DPOT models the IND hypothesis is rejected only in one case. For the CEVT-n, CEVT-sst and APARCH-sst models, the IND hypothesis is rejected, respectively, 3, 3 and 4 times. Table 5, summarize these results. This empirical evidence shows that the DPOT models can be successful in removing the tendency to clustering of violations, which was our main objective, can perform better than state of the art risk models and much better than the widely used RiskMetrics model.

Table 2

Out-of-sample accuracy for VaR(0.01) applied to S&P 500 Index returns from January 4, 1950 until May 18, 2010, with a rolling window of size 1000. Unconditional coverage and independence tests.

Model	Violation frequencies	Kupiec p-value	CAViaR p-value	MM Ratio p-value
POT	0.013672	0.0000	0.0000	0.0000
DPOT(0.8)	0.009725	0.7410	0.0189	0.7902
DPOT(0.75)	0.009443	0.5011	0.1018	0.1048
DPOT(0.7)	0.009443	0.5011	0.8659	0.0566
CEVT-n	0.010007	0.9933	0.0145	0.0166
CEVT-sst	0.010007	0.9933	0.0236	0.0314
APARCH-sst	0.009015	0.2305	0.0064	0.0717
RiskMetrics	0.018675	0.0000	0.0000	0.4391

Table 3

Out-of-sample accuracy for VaR(0.01) applied to DAX Index returns from November 27, 1990 until May 18, 2010, with a rolling window of size 1000. Unconditional coverage and independence tests.

Model	Violation frequencies	Kupiec p-value	CAViaR p-value	MM Ratio p-value
DPOT(0.8)	0.08425	0.3085	0.6918	0.8821
DPOT(0.75)	0.008935	0.4953	0.7175	0.8597
DPOT(0.7)	0.010722	0.6533	0.2786	0.1886
CEVT-n	0.010467	0.7706	0.0156	0.6180
CEVT-sst	0.009446	0.7250	0.0030	0.7227
APARCH-sst	0.009191	0.6058	0.0079	0.0358
RiskMetrics	0.016845	0.0001	0.1187	0.3046

Table 4

Out-of-sample accuracy for VaR(0.01) applied to FTSE Index returns from April 3, 1984 until May 18, 2010, with a rolling window of size 1000. Unconditional coverage and independence tests.

Model	Violation frequencies	Kupiec p-value	CAViaR p-value	MM Ratio p-value
DPOT(0.8)	0.009109	0.4962	0.1033	0.6359
DPOT(0.75)	0.009109	0.4962	0.3405	0.6373
DPOT(0.7)	0.010002	0.9989	0.8646	0.5687
CEVT-n	0.011073	0.4275	0.4037	0.7410
CEVT-sst	0.011073	0.4275	0.4143	0.7423
APARCH-sst	0.008573	0.2143	0.0047	0.2338
RiskMetrics	0.018222	0.0000	0.2644	0.6569

Table 5

Number of rejections of the UC and IND hypotheses with significance level equal to 0.05.

	Number of rejections	
	UC hypothesis	IND hypothesis
DPOT(0.8)	0	1
DPOT(0.75)	0	0
DPOT(0.7)	0	0
CEVT-n	0	3
CEVT-sst	0	3
APARCH-sst	0	4
RiskMetrics	3	1

4.3 Minimization of capital requirements under the Basel II Accord

Under the Basel II Accord, ADIs have to communicate their daily risk forecasts to the monetary authority (typically a central bank) at the beginning of the trading day, using a VaR model. Too high forecasts will lead to large capital requirements. On the other hand, too low forecasts will lead to excessive violations and consequently to a penalty that increases capital requirements. The penalty can be an increase in a multiplicative factor to calculate capital requirements or the imposition of a standard model when the number of violations exceeds 10. Let us consider an ADI that invest at day $t+1$ an amount A_{t+1} in a portfolio of risky assets. The portfolio is financed by deposits (D_{t+1}) and equity (E_{t+1}). At day $t+1$ the ADI must satisfy capital requirements for market risk (CR_{t+1}) such that $\mathrm{E}_{t+1} \geq \mathrm{CR}_{t+1}\mathrm{A}_{t+1}$. Note that for a given CR_{t+1}, to satisfy this inequality the ADI can increase the equity or reduce the amount invested. Of course, even without this rule, risk averse investors will reduce this amount during periods of high risk. The Basel II Accord stipulates CR_{t+1} as

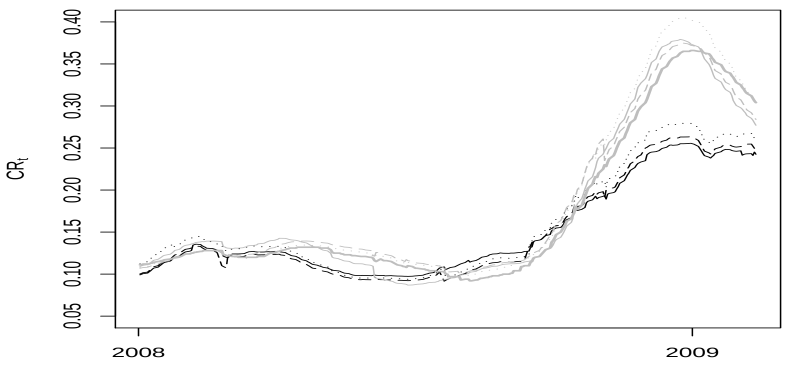
$$\mathrm{CR}_{t+1} = \sup\Big\{(3+k)\overline{\mathrm{VaR}}_{60}, \mathrm{VaR}_t\Big\}, \tag{4.2}$$

where $\overline{\mathrm{VaR}}_{60}$ is the average VaR over the previous 60 trading day´s and k is a multiplicative factor that depends on the number of violations in the previous 250 trading days (N_v), according to the following function,

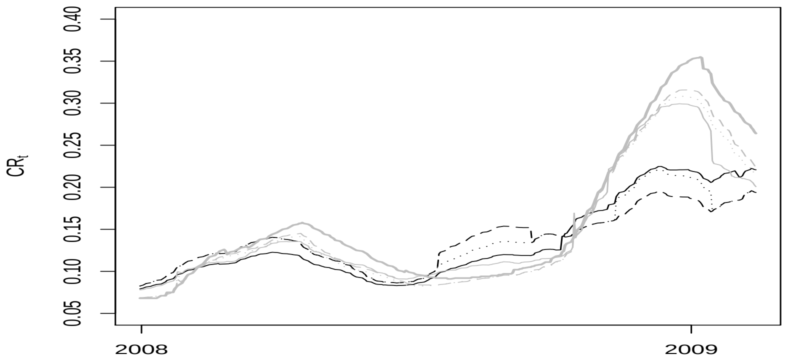
$$k = \begin{cases} 0 & \text{if } N_v \leq 4 \\ 0.3 + 0.1(N_v - 4) & \text{if } 5 \leq N_v \leq 6 \\ 0.65 & \text{if } N_v = 7 \\ 0.65 + 0.1(N_v - 7) & \text{if } 8 \leq N_v \leq 9 \\ 1 & \text{if } N_v = 10. \end{cases}$$

In the same way as McAleer et al. (2009), we can write the ADI profit for day $t+1$ as $\Pi_{t+1} = r_{At+1}A_{t+1} - r_{Dt+1}D_{t+1} - r_{Et+1}E_{t+1}$, where r_{At+1} denotes the return on the ADI portfolio on day $t+1$, r_{Dt+1} the rate for deposits on day $t+1$

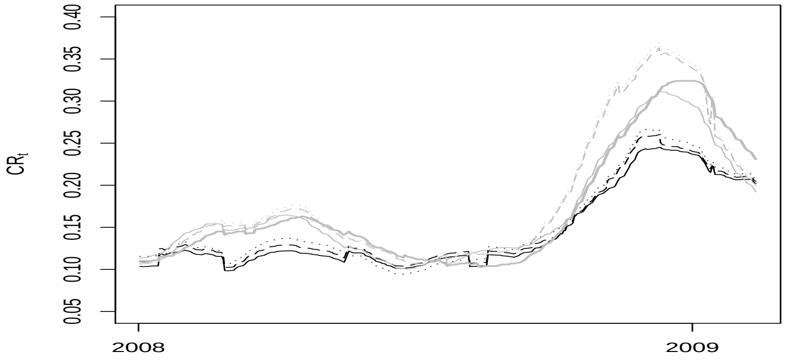
and r_{Et+1} the cost of holding equity. An increase in E_{t+1} will reduce expected profits and for that reason an ADI is interested in the minimization of CR_{t+1}. In a recent work, McAleer et al. (2009c) compare, in terms of minimization of capital requirements, well known and widely used time-varying volatility models applied in one-day-ahead VaR forecast. These authors advanced the idea and conclude that optimal risk management within the Basel II Accord requires to use combinations of models. In this Section we choose the S&P 500 index, DAX index and FTSE index returns for the period, January 2, 2008, to February 12, 2009, which includes the global financial crisis, taking into account the comparability with this previous study. Using equation (4.2) and the DPOT, CEVT, APARCH-sst and RiskMetrics models, we calculated CR_t for each model over this period and for each index. In Figure 4.1, the dotted, longdash, solid, dotted grey, longdash grey, solid grey and bold grey lines, correspond to daily capital requirements obtained respectively with the DPOT(0.8), DPOT(0.75), DPOT(0.7), CEVT-n, CEVT-sst, APARCH-sst and RiskMetrics models. From Figure 4.1, it is evident that, for these indexes and this period, the DPOT models perform much better than the other models under study. Only in few days and with very small differences, the DPOT models produced higher capital requirements and this is mainly before the high volatile period, suggesting that DPOT models anticipate better these periods than the other models. Tables 6, 7 and 8 gives the maximum number of violations in the previous 250 trading days and the average capital requirements. In terms of number of violations, the DPOT models with $c = 0.8$ and $c = 0.75$ perform better than with $c = 0.7$ with which we had 11 violations in the previous 250 trading days. Although this occurs during a very severe crisis, exceeds 10 violations and falls in the red zone defined by the Basel II Accord. In terms of capital requirements, Tables 6,7 and 8 show that, in the period under study, the DPOT models lead to substantially lower average capital requirements than the other models under study. The differences are in the majority of cases higher than 200 basis points and in some cases higher than 300 basis points.



(a) S&P 500 index



(b) DAX index



(c) FTSE index

Figure 4.1: Daily capital requirements (CR_t)between 2 January 2008 and 12 February 2009, under the Basel II Accord, applying the DPOT(0.8) (dotted), DPOT(0.75) (longdash), DPOT(0.7) (solid), CEVT-n (dotted grey), CEVT-sst (longdash grey), APARCH-sst (solid grey) and RiskMetrics (bold grey) models.

Table 6

Maximum N_v and average CR_t for S&P 500 index from January 2, 2008 until February 12, 2009.

Model	Maximum N_v	Average capital requirements (CR_t)
DPOT(0.8)	8	0.1583
DPOT(0.75)	8	0.1495
DPOT(0.7)	9	0.1505
CEVT-n	10	0.1825
CEVT-sst	8	0.1781
APARCH-sst	7	0.1739
RiskMetrics	11	0.1715

Table 7

Maximum N_v and average CR_t for DAX index from January 2, 2008 through February 12, 2009.

Model	Maximum N_v	Average capital requirements (CR_t)
DPOT(0.8)	5	0.1385
DPOT(0.75)	5	0.1373
DPOT(0.7)	11	0.1351
CEVT-n	4	0.1457
CEVT-sst	4	0.1484
APARCH-sst	5	0.1474
RiskMetrics	11	0.1601

Table 8

Maximum N_v and average CR_t for FTSE index from January 2, 2008 until February 12, 2009.

Model	Maximum N_v	Average capital requirements (CR_t)
DPOT(0.8)	9	0.1522
DPOT(0.75)	10	0.1500
DPOT(0.7)	11	0.1446
CEVT-n	9	0.1850
CEVT-sst	9	0.1821
APARCH-sst	7	0.1688
RiskMetrics	11	0.1705

5 Conclusions

In this work we propose a POT method that uses the durations between excesses as covariates. Based on this method, three DPOT models for forecasting one-day-ahead VaR were compared with other models. Empirical findings presented in Section 4.2 show that they perform very well in terms of unconditional coverage and better than state-of-the art models in terms of removing the tendency to clustering of violations. In terms of out-of-sample accuracy, DPOT models perform much better than the widely used RiskMetrics model. Moreover, the empirical findings presented in Section 4.3, suggest that the DPOT models can have an important role in the minimization of capital requirements under the Basel II Accord. In the period under study, the DPOT models lead to substantially lower average capital requirements. It is possible that we can achieve lower average capital requirements by integrating DPOT in a combination of models strategy or, for example, in a dynamic learning strategy such as the one proposed by McAleer et al. (2009). The study of these issues remains for future research. Finally, we notice that in order to deal with the volatility clustering, the proposed models do not assume a parametric distribution for the entire distribution of the returns, as the CEVT or GARCH-type models, but assumes a parametric model only on the tail and based on solid asymptotic theory.

Acknowledgements

A preliminary version of this paper was presented at the 30th International Symposium on Forecasting, at San Diego CA, USA, June 2010, and at the 4rd International Conference on Computational and Financial Econometrics and 3rd Workshop of the ERCIM Working Group on Computing and Statistics, at University of London, UK, December 2010. We thank Juan-Ángel Jiménez-Martin, Pilar Abad, Sonia Benito and other participants, for useful suggestions and comments. Research partially supported by Fundação para a Ciência e Tecnologia (FCT/PROTEC and FCT/OE) and Center of Statistics and Applications of the University of Lisbon (CEAUL).

References

[1] Araújo Santos, P. and Fraga Alves, M.I. (2010). A new class of independence tests for interval forecasts evaluation. Computational Statistics and Data Analysis. In press. doi:10.1016/j.csda2010.10.002.

[2] Balkema, A.A. and de Haan, L. (1974). Residual Life Time at Great Age. Ann. Probab., 2,792-804.

[3] Basel Commitee on Banking Supervision, (1995), An Internal Model-Based Approach to Market Risk Capital Requirements, BIS, Basel, Switzerland.

[4] Basel Commitee on Banking Supervision, (1996), Amendment to the Capital Accord to incorporate market risks, BIS, Basel, Switzerland.

[5] Basel Commitee on Banking Supervision, (1996), Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Model-Based Approach to Market Risk Capital Requirements, BIS, Basel, Switzerland.

[6] Basel Commitee on Banking Supervision, (2006), International Convergence of Capital Measurement and Capital Standards, a Revised Framework Comprehensive Version, BIS, Basel, Switzerland.

[7] Bekiros, S.D. and Georgoutsos, D.A. (2005). Estimation of Value-at-Risk by extreme value and conventional methods: a comparative evaluation of their predictive performance. Journal of International Financial Markets, Institutions and Money, 15, Issue 3, 2009-228.

[8] Berkowitz, J., Christoffersen P. and Pelletier D., 2009. Evaluating Value-at-Risk models with desk-level data, Management Science, Published online in Articles in Advance.

[9] Black R., 1976. Studies in stock price volatility changes. Proceedings of the 1976 Business Meeting of the Business and Economics Statistics Section. American Statistical Association, 177-181.

[10] Bystrom, H. (2004). Managing Extreme Risks in Tranquil and Volatile Markets Using Conditional Extreme Value Theory. *International Review of Financial Analysis*, 13, no.2, 133-152.

[11] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, **31**,307-327.

[12] Christoffersen P. (1998). Evaluating Intervals Forecasts. International Economic Review, **39**, 841-862.

[13] Diebold, F.X., Schuermann, T. and Stroughair, J.D. (1998). Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management. Working Paper, 98-10, Wharton School, University of Pennsylvania.

[14] Ding, Z., Engle, R.F and Granger, C.W.J., 1993. A long memory property of stock market return and a new model, Journal of Empirical Finance, 1, 83-106.

[15] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). Modeling Extremal Events for Insurance and Finance (Berlin: Springer).

[16] Engel, R.F. and Manganelli, S., (2004). CAViaR: Conditional Autoregressive Value-at-Risk by Regression Quantiles, Journal of Business and Economics Statistics, 22, 367-381.

[17] Giot, P. and Laurent, S. (2004). Modelling Daily Value-at-Risk Using Realized Volatility and ARCH type Models. Journal of Empirical Finance, 11, 379-398.

[18] Glosten, L., Jagannathan and R. and Runkle, D. (1992). On the relation between the expected value and volatility of nominal excess return on stocks, Journal of Finance, 46, 1779-1801.

[19] Ghorbel, A. and Trabelsi, A. (2008). Predictive performance of conditional Extreme Value Theory in Value-at-Risk estimation. Int. J. Monetary Economics and Finance, 1, No.2, 121-147.

[20] Jorian, P. (2000). Value at Risk: The New Benchmark for Managing Financial Risk, McGraw-Hill, New York.

[21] Kuester, K., Mittik, S. and Paolella, M.S., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. Journal of Financial Econometrics, 4(1), 53-89.

[22] Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models, Journal of Derivatives, 3, 73-84.

[23] McAleer, M., Jiménez-Martin, J.A. and Pérez-Amaral, T. (2009). A decision rule to minimize daily capital charges in forecasting value-at-risk. To appear in Journal of Forecasting.

[24] McAleer, M., Jiménez-Martin, J.A. and Pérez-Amaral, T. (2009b). Optimal Risk Management Before, During and After the 2008-09 Financial Crisis. Medium for Econometric Applications, 17(3), 20-27.

[25] McAleer, M., Jiménez-Martin, J.A. and Pérez-Amaral, T. (2009c). Has the Basel II Accord Encouraged Risk Management During the 2008-09 Financial Crisis? Tinbergen Institute Discussion Paper. TI 2009-039/4.

[26] Mittnik, S. and Paolella, M.S. (2000). Consitional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates. Journal of Forecasting, 19, 313-333.

[27] Nelson, D.B. (1991). Conditional heteroscedasticity in asset returns: a new approach, Econometrica, 59, 347-370.

[28] McNeil, A.J. and Frey, R. (2000). Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: An extreme value approach. Journal of Empirical Finance, 7, 271-300.

[29] Ozun, A., Cifter, A. and Yilmazer, S. (2010). Filtered Extreme Value Theory for Value-At-Risk Estimation: Evidence from Turkey. The Journal of Risk Finance Incorporating Balance Sheet, 11, no.2, 164-179.

[30] Pickands III, J. (1975). Statistical Inference using Extreme value Order Statistics, Ann. Statist., 3, 119-131.

[31] Riskmetrics (1996), J.P. Morgan Technical Document, 4^{th} Edition, New York, J.P. Morgan.

[32] R Development Core Team, (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[33] Smith, R. (1987). Estimating tails of probability distributions. Ann. Statist., (15)1174-1207.

[34] Smith, R. (1990). Models for Exceedances over High Thresholds. J. R. Statist. Soc. B, (52), no.3, 393-442.

[35] Tsay, R.(2010). Analysis of Financial Time Series. Wiley Series in Probability and Statistics.

[36] Tu, A.H., Wong, W.K. and Chang, M.C. (2008). Value-at-Risk for Long and Short Positions of Asian Stock Markets. International Research Journal of Finance and Economics, 22, 135-143.

[37] Zhou, C.(2010). The extent of the maximum likelihood estimator for the extreme value index, J. Multivariate Anal. 101-4, 971-983.