

Forecasting one-day-ahead Value-at-Risk with a Duration based POT method

P. Araújo Santos^{a,1}, M.I. Fraga Alves^b

^a*Escola Superior de Gestão e Tecnologia de Santarém and CEAUL, Portugal*

^b*Faculdade de Ciências da Universidade de Lisboa and CEAUL, Portugal*

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, two models for forecasting one-day-ahead Value-at-Risk were suggested and applied to real data. Comparative studies provides evidence that they can perform better than state-of-the art risk models, both in terms of out-of-sample accuracy and minimization of capital requirements under the Basel II Accord.

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

¹Corresponding address: Departamento de Informática e Métodos Quantitativos, Escola Superior de Gestão e Tecnologia, Instituto Politécnico de Santarém, Complexo Andaluz, Apartado 295, 2001-904 Santarém, Portugal. Tel.: +351 243 303 200; fax: +351 243 332 152. E-mail addresses: paulo.santos@esg.ipsantarem.pt (P.Araújo Santos), isabel.alves@fc.ul.pt (M.I. Fraga Alves).

1. Introduction

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 (AIDs) communicate their daily VaR forecasts to monetary authorities (typically, a central bank) at the beginning of each trading day and define 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 correspondent reported VaR.

The rest of the paper is organized as follows. In Section 2, we review the peaks over threshold (POT) method with one example that illustrates the *tendency to clustering of violations* problem. 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 10.

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

The Generalized Pareto Distribution (GPD) has the form

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

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

$$E[Y] = \frac{\sigma}{1-\gamma} \quad (\gamma < 1),$$

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

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*

$$F_u(y) = P[X - u \leq y | X > u] = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad (2)$$

for $0 \leq y < x_F - u$, where x_F is the (finite or infinite) right endpoint of F . The Extreme Value Theory (EVT), with the following theorem, suggests the GPD (1) as an approximation for the excess distribution (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*

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

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) over a high threshold u can be approximated by the GPD (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 exceedances above u in a sample X_1, \dots, 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

$$\hat{\bar{F}}(x) = \frac{n}{n_x} \left(1 + \hat{\gamma} \frac{x-u}{\hat{\sigma}} \right)^{-1/\hat{\gamma}}, \quad \text{valid for } x > u. \quad (3)$$

For $p < \bar{F}(u)$ and inverting the tail estimator formula (3), we get the VaR POT estimator

$$\hat{\text{VaR}}_{t+1}^{POT}(p) = u + \frac{\hat{\sigma}}{\hat{\gamma}} \left(\left(\frac{n}{n_x p} \right)^{\hat{\gamma}} - 1 \right). \quad (4)$$

Turning theory into practice, one example is presented to illustrate the application of the VaR POT estimator (4) and the problem of tendency to clustering of violations, using the 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. In Figure 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 using the data, adjust very well to the excesses. In

this example, we obtain a $\text{VaR}(0.05)$ equal to 1.42 and a $\text{VaR}(0.01)$ equal to 2.67. In Figure 2, instead of considering 15190 daily returns to estimate the $\text{VaR}(0.01)$, 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%.

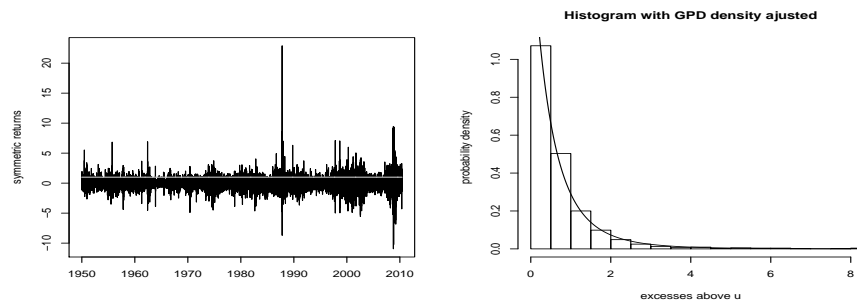


Figure 1: Symmetric returns (left) and Histogram of 1519 excesses above the threshold $u = x_{13671:15190} = 0.9897$ (right) for the S&P 500 Index from January 4, 1950 through May 18, 2010.

However, the serious problem of POT method and other unconditional models, is tendency to clustering of violations associated with the volatility clustering phenomenon. Figure 3 illustrates this problem during the 2008 financial crisis period. Between January 2, 2008 and February 12, 2009, we have a cluster of violations, which correspond to 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%.

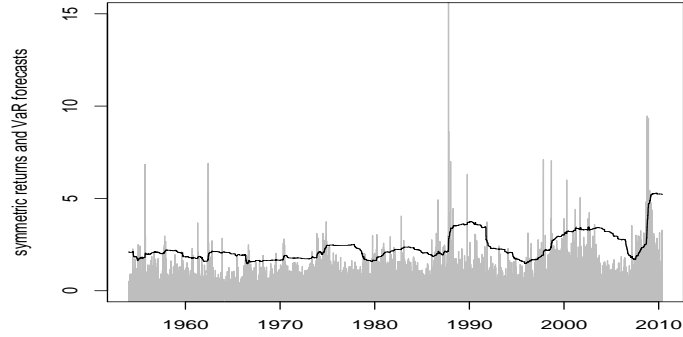


Figure 2: Symmetric returns (grey) of S&P 500 Index from January 4, 1950 through May 18, 2010, and one-day-ahead VaR(0.01) forecasts with POT method (black line) and a rolling window of size 1000.

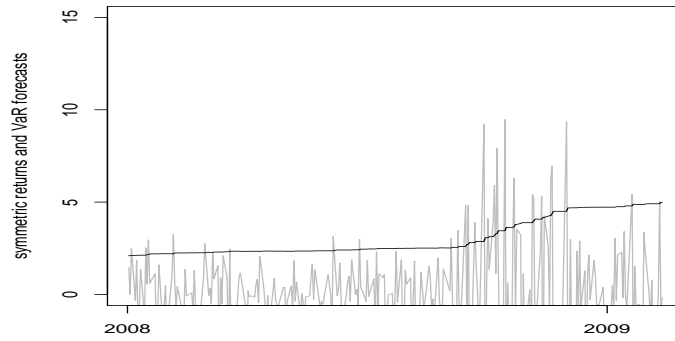


Figure 3: Symmetric returns of S&P 500 Index from January 4, 1950 through May 18, 2010 (282 trading days), and one-day-ahead VaR(0.01) forecasts with POT method (black line) and a rolling window of size 1000.

3. A duration based POT method (DPOT)

Our main goal is 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), develop 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, \dots, y_n be excesses above a high threshold u , d_1 the duration until the first excess and d_2, \dots, d_n , defined by

$$d_i = t_i - t_{i-1}, \quad (5)$$

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, $d_n, d_{n-1}, \dots, d_{n-v+1}$ and the duration since the excess n which we define by d_{n+1}^t .

With the durations d_i, \dots, 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

$$d_{i,v} = d_i + \dots + d_{i-v+1} = t_i - t_{i-v}. \quad (6)$$

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

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

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 usually holds for in various financial time series. Figure 4 (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 4 (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. Adding durations we get the duration since the preceding v excesses defined in (6) and the correlation increase a little more when we compute the correlation between excesses and the inverse of these durations. The empirical results show some nonlinear association between excesses and durations.

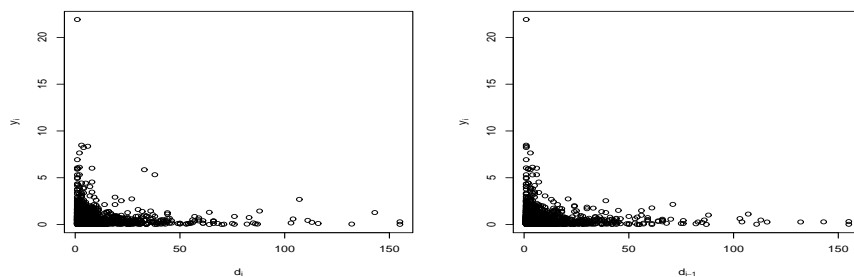


Figure 4: S&P Index from January 4, 1950 through May 18,2010. Scatter plot of excess above a high threshold ($u = 0.9897$) and duration since the preceding excess (left) and scatter plot of excess above a high threshold ($u = 0.9897$) and duration between the 2 preceding excesses (right).

Table 1

S&P 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 (5) and the duration since the excess n , d_{n+1}^t , we assume the GPD for the excesses Y_i above u , such that

$$Y_t \sim \text{GPD}\left(\gamma, \sigma_t = g(\alpha_1, \dots, \alpha_k, \dots, d_{n+1}^t, d_n, d_{n-1}, \dots, d_{n-v+2})\right),$$

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

$$\widehat{\text{VaR}}_{t+1|t}^{DPOT}(p) = u + \frac{\hat{\sigma}_t}{\hat{\gamma}} \left(\left(\frac{n}{n_x p} \right)^{\hat{\gamma}} - 1 \right), \quad (7)$$

$$\hat{\sigma}_t = g(\hat{\alpha}_1, \dots, \hat{\alpha}_k, \dots, d_{n+1}^t, d_n, d_{n-1}, \dots, d_{n-v+2}).$$

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_{n+1}^t and the last v durations between excesses,

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

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

The empirical results of Subsection 3.1 suggest an 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

$$\widehat{\text{VaR}}_{t+1|t}^{DPOT(v,c)}(p) = u + \frac{\hat{\alpha}}{\hat{\gamma}(d_{t,v})^c} \left(\left(\frac{n}{n_x p} \right)^{\hat{\gamma}} - 1 \right), \quad (8)$$

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

$$\begin{aligned} \log L(\gamma, \alpha) &= \log \prod_{i=v}^n f_{Y_i}(y_i) \\ &= \log \prod_{i=v}^n \left(\frac{\alpha}{(d_{i,v})^c} \right)^{-1} \left(1 + \frac{\gamma}{\alpha} y_i (d_{i,v})^c \right)^{-(1/\gamma+1)} \\ &= - \sum_{i=v}^n \log \left(\frac{\alpha}{(d_{i,v})^c} \right) - \left(\frac{1}{\gamma} + 1 \right) \sum_{i=v}^n \log \left(1 + \frac{\gamma}{\alpha} y_i (d_{i,v})^c \right). \quad (9) \end{aligned}$$

We present results for $v = 3$, $c \in (2/3, 3/4)$ and apply an implementation of Nelder and Mead algorithm, using the stats package of R (R Development Core Team, 2008), to maximize (9).

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, 134 (0.9443%) and 140 (0.9866%) violations, respectively with $c = 2/3$ and $c = 3/4$. These percentages are much closer to the expected 1% than the 1.367% obtained with the unconditional POT model. In Figure 5, the grey line correspond to the S&P 500 returns, the slashed black, the dotted black and the black lines correspond to one-day-ahead VaR forecasts calculated respectively with the DPOT($c = 3/4$), DPOT($c = 2/3$) and the POT models. For the 2008 global financial crises period, Figure 5, 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($c = 3/4$) was 8 and with DPOT($c = 2/3$) 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($c = 3/4$) and DPOT($c = 2/3$) is very small, suggesting that the method

is robust for different values of c between this two values, in spite of the empirical findings in Section 4 will suggest that a choice of $c = 3/4$ is preferable.

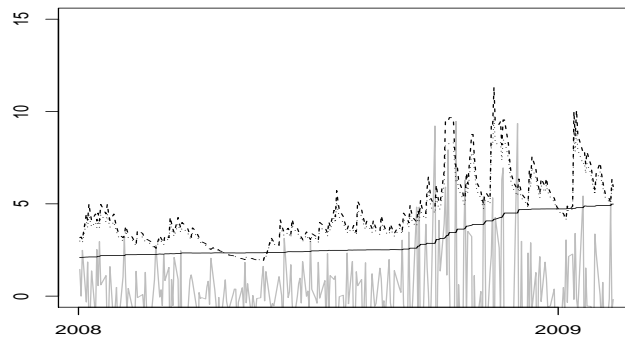


Figure 5: Symmetric returns of S&P 500 Index from January 4, 1950 through May 18, 2010 (grey), and one-day-ahead VaR(0.01) forecasts with POT method (black), DPOT($c = 3/4$) (slashed black line), DPOT($c = 2/3$) (dotted black line) and a rolling window of size 1000.

4. Comparative studies

Using the returns from S&P 500 Index, 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. Previous comparative studies show that this method performs better. We employ the R language in order to develop the programs. In Section 4.1 we briefly review the Conditional EVT method, in Section 4.2 we evaluate the accuracy of out-of-sample interval forecasts produced with the different methods and in Section 4.3 we compare the performance in terms of minimization of capital requirements under the Basle II Accord.

4.1. Conditional EVT

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 for

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 s th order autoregressive process, AR(s),

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

The unpredictable component is expressed by $\varepsilon_t = z_t \sigma_t$, where the innovations, z_t , is 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 $i = 0, 2, \dots, p$ and $j = 1, 2, \dots, q$. This time-varying volatility model for the unpredictable component, is a Generalised Autoregressive Conditional Heteroscedasticity (GARCH) process, proposed by Bollerslev (1986). The GARCH process 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 asymmetric skewed t distribution. We will denote this model as CEVT-sst. Applying Conditional EVT, VaR is calculated using the following equation:

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

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. \hat{z}_p is a quantile p estimate, obtained with the POT method and the standardized residuals calculated as

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

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), Bystrm (2004), Bekiros and Georgoutsos (2005), Kuester et al. (2006), Ghorbel and Trabelsi (2008), Ozun et al. (2010)), and this is the reason why we choose CEVT-n and CEVT-sst models for the comparative studies.

4.2. Interval forecasts evaluation

In this section we compare the CEVT-sst, CEVT-n and DPOT models with $v = 3$, $c \in (2/3, 3/4)$, denoted respectively by DPOT(2/3) and DPOT(3/4). We examine the one-day-ahead VaR(0.01) forecasts performance for a portfolio that is long in the S&P 500 Index, considering 15190 returns produced by all the historical data from January 4, 1950 until May 18, 2010. Using a rolling window of size 1000, 1490 one-day-ahead VaR(0.01) forecasts were obtained for each model. 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 cover-

age (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

$$I_t = \alpha + \beta_1 I_{t-1} + \beta_2 \text{VaR}_{t|t-1}(p) + \varepsilon_t, \quad (10)$$

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 and does not rely on an asymptotic 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 are summarized in Table 2. 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. In terms of the UC hypothesis, both DPOT and CEVT models performs very well taking into account that in no case the hypothesis is reject since all p values are very high. Is interesting to note the impressive performance of CEVT models in terms of UC, with 142

violations in 14190 out-of-sample forecasts was impossible to obtain a better result (violation frequency equal to 0.010007). In terms of IND hypothesis, the DPOT models performs clearly better than the CEVT models, especially the DPOT(3/4). For this model, both independence tests do not reject the IND hypothesis with the usual significance levels. This empirical evidence shows that the DPOT model can produce iid violations and be successful in removing the tendency to clustering of violations, which was our main objective.

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($c = 3/4$)	0.009443	0.5011	0.1018	0.1048
DPOT($c = 2/3$)	0.009866	0.8724	0.4066	0.0351
EVTC-sst	0.010007	0.9933	0.0236	0.0314
EVTC-n	0.010007	0.9933	0.0145	0.0166

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 $E_{t+1} \geq CR_{t+1}A_{t+1}$. The Basel II Accord stipulates CR_{t+1} as

$$CR_{t+1} = \sup\left\{(3+k)\overline{\text{VaR}}_{60}, \text{VaR}_t\right\}, \quad (11)$$

where $\overline{\text{VaR}}_{60}$ is the average VaR over the previous 60 trading days 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 manner as McAleer et al. (2009), we can write the ADI profit for day $t + 1$ as follows

$$\Pi_{t+1} = r_{At+1}A_{t+1} - r_{Dt+1}D_{t+1} - r_{Et+1}E_{t+1},$$

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 returns for the period 3 January to 12 February 2009, which includes the global financial crisis, taking into account the comparability with this previous study. Using equation (11) and the models CEVT-sst, CEVT-n, DPOT(3/4) and DPOT(2/3), we calculated CR_t for each model over this period. The solid and dotted black lines in Figure 4 corresponds to daily

capital requirements obtained respectively with the CEVT-st and CEVT-n risk models. The grey lines in Figure 4, solid and dotted, corresponds respectively to the DPOT(3/4) and DPOT(2/3) risk models. From Figure 4 is evident that, for this Index and this period, the DPOT models performs much better than the CEVT models. Only in few days and with very small differences, the CEVT models produced lower daily capital requirements. Table 3 gives the maximum number of violations in the previous 250 trading days and the average capital requirements. None of the DPOT and CEVT models exceeds 10 violations, the limit allowed by the Basel II accord. In the case of the DPOT models, the average CR_t is almost the same with $c = 3/4$ and $c = 2/3$, however, the choice of $c = 3/4$ leads to a maximum N_v equal to 8, while with $c = 2/3$ we obtain 10. The DPOT models lead to substantially lower average capital requirements (0.1496 and 0.1495) than the CEVT models (0.1781 and 0.1825).

In the study of McAleer et al. (2009c), the time-varying volatility models that do not exceeded the limit of 10 violations and with lower average capital requirements were the GJR model of Glosten, Jagannathan and Runkle (1992) and the Exponential GARCH (EGARCH) model of Nelson (1991). These two models, lead respectively to 0.157 and 0.153 averages CR_t , higher averages than those obtained with DPOT. However, now the differences are smaller suggesting that a risk management strategy which combines a DPOT model with these models may produce better results than using only single models.

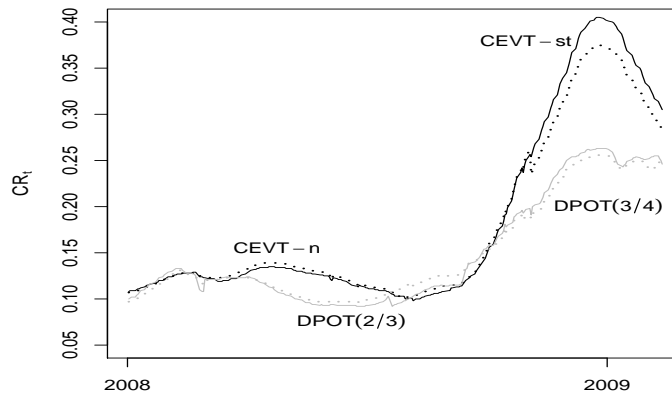


Figure 6: Daily capital requirements (CR_t) for S&P 500 Index returns between 3 January and 12 February 2009, under the Basel II Accord, applying the CEVT-st (solid black line), CEVT-n (dotted black line), DPOT(3/4) (solid grey line) and DPOT(2/3) (grey and dotted line) models.

Table 3

Maximum N_v and average CR_t for S&P 500 Index from January 3, 2008 until February 12, 2009.

Model	Maximum N_v	Average capital requirements (CR_t)
POT	29	–
DPOT($c = 3/4$)	8	0.1495
DPOT($c = 2/3$)	10	0.1496
EVTC-sst	8	0.1781
EVTC-n	10	0.1825

5. Conclusions

In this work we propose a POT method that uses the durations between excesses as covariates. Based on this method, two DPOT models for forecasting one-day-ahead VaR were proposed and compared with other models. Empirical findings presented in Section 4.1 show that they perform well in terms of unconditional coverage and performs better than state-of-the art models in terms of producing iid violations and in removing the tendency to clustering of vio-

lations. Moreover, the empirical findings presented in Section 4.1, suggest that the DPOT models can have an important role in the minimization of capital requirements under the Basel II Accord. Considering the empirical evidence presented in previous studies (McAleer et al., 2009b and McAleer et al., 2009c), it is possible that better results can be achieved 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

We thank Juan-Ángel Jiménez Martín and other participants at the 30th International Symposium on Forecasting, at San Diego CA, USA, June 2010, for useful suggestions. Research partially supported by Fundação para a Ciência e Tecnologia (FCT/PROTEC and FCT/POCI 2010 project) and Center of Statistics and Applications of the University of Lisbon (CEAUL).

References

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

- Basel Committee on Banking Supervision, (1995), An Internal Model-Based Approach to Market Risk Capital Requirements, BIS, Basel, Switzerland.
- Basel Committee on Banking Supervision, (1996), Amendment to the Capital Accord to incorporate market risks, BIS, Basel, Switzerland.
- Basel Committee 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.
- Basel Committee on Banking Supervision, (2006), International Convergence of Capital Measurement and Capital Standards, a Revised Framework Comprehensive Version, BIS, Basel, Switzerland.
- 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, 209-228.
- 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*.
- 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.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**,307-327.
- Christoffersen P. (1998). Evaluating Intervals Forecasts. *International Economic Review*, **39**, 841-862.

- Diebold, F.X., Schuermann, T. and Strouhair, 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.
- Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). Modeling Extremal Events for Insurance and Finance (Berlin: Springer).
- 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.
- 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.
- 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.
- Jorion, P. (2000). *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.
- Kuester, K., Mittik, S. and Paoletta, M.S., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics*, 4(1), 53-89.
- Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives*, 3, 73-84.
- 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*.

- 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.
- 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.
- Nelson, D.B. (1991). Conditional heteroscedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370.
- 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.
- 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.
- Pickands III, J. (1975). Statistical Inference using Extreme value Order Statistics, *Ann. Statist.*, 3, 119-131.
- 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>.
- Smith, R. (1987). Estimating tails of probability distributions. *Ann. Statist.*, (15)1174-1207.
- Smith, R. (1990). Models for Exceedances over High Thresholds. *J. R. Statist. Soc. B*, (52), no.3, 393-442.
- Tsay, R.(2010). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics.

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