



UNIVERSITY  
OF TRENTO - Italy  
Department of Economics

# Dynamic VaR models and the Peaks over Threshold method for market risk measurement: an empirical investigation during a financial crisis

Marco Bee, Fabrizio Miorelli

n.  
9/2010



# Dynamic VaR models and the Peaks over Threshold method for market risk measurement: an empirical investigation during a financial crisis

Marco Bee

Fabrizio Miorelli

June 14, 2010

## Abstract

This paper presents a backtesting exercise involving several VaR models for measuring market risk in a dynamic context. The focus is on the comparison of standard dynamic VaR models, *ad hoc* fat-tailed models and the dynamic Peaks over Threshold (POT) procedure for VaR estimation with different volatility specifications. We introduce three different stochastic processes for the losses: two of them are of the GARCH-type and one is of the EWMA-type. In order to assess the performance of the models, we implement a backtesting procedure using the log-losses of a diversified sample of 15 financial assets. The backtesting analysis covers the period March 2004 - May 2009, thus including the turmoil period corresponding to the subprime crisis. The results show that the POT approach and a Dynamic Historical Simulation method, both combined with the EWMA volatility specification, are particularly effective at high VaR coverage probabilities and outperform the other models under consideration. Moreover, VaR measures estimated with these models react quickly to the turmoil of the last part of the backtesting period, so that they seem to be efficient in high-risk periods as well.

**Keywords:** Market risk, Extreme Value Theory, Peaks over Threshold, Value at Risk, Fat tails.



# 1 Introduction

Market risk was the earliest type of risk systematically tackled both by practitioners and academics. Massive developments in the methodology have been triggered by the release of the Market Risk Amendment by the Bank of International Settlement (BIS, 1996). In order to keep a certain capital buffer against adverse market movements, banks and financial institutions with relevant trading activity have to measure regularly the exposure on the financial assets held in their trading portfolios. This exposure has to be converted in a monetary amount, the so-called regulatory capital required against market risk. Standard approaches to market risk measurement are mostly based on the normality assumption, which is often inadequate from the empirical point of view.

A well-known measure of market risk, used both for internal risk management and regulatory purposes, is the so-called *Value-at-Risk* (VaR), which estimates, given a certain time horizon, the maximum loss that a bank is going to suffer with a certain probability level. The VaR, originally proposed by J.P. Morgan (RiskMetrics, 1996), has become a standard tool in market risk management.

Employing the VaR parametric setup requires various statistical assumptions. A commonly accepted starting point is that, when dealing with the production of short-term VaR estimates, a dynamic approach is preferable, because it allows to capture the empirical properties of the loss time series. In this case, the primary concern is the choice of an accurate econometric model that takes into account some stylized facts of financial time series. Among them (see, for example, Cont, 2001), the main features considered in this article are heteroskedasticity, persistence and fat tails. The first two are strictly connected with the empirical autocorrelation function of squared losses, which is typically significantly positive for a long time, while the third one is observed both in the filtered conditional loss distribution and in the unconditional loss distribution.

Many VaR models try to incorporate one or more of these stylized facts. The earliest dynamic VaR approaches assemble conditional volatility models like ARCH (Engle, 1982) or GARCH (Bollerslev, 1986), which directly model the persistence in squared returns/losses, with the standard normal distribution. Although they usually fit the data reasonably well, the normality assumption is generally a cause of VaR underestimation, because the filtered conditional loss distribution, i.e. the distribution of standardized losses, is often heavier-tailed than the normal. Thus, especially at high coverage levels, these models are quite inadequate.

An obvious solution (see, for example, Angelidis *et al.*, 2004; Hung *et al.*, 2008) consists in fitting a distribution with tails fatter than the normal: the data generating processes range from the simple standardized Student- $t$  or Generalized Error Distribution (GED), to the more complex Edgeworth-Sargan distribution (see Baixauli and Alvarez, 2006). From a theoretical point of view, each of these models may suit the filtered loss data better than the normal distribution. However, they are based on *a priori* hypotheses on the data-generating process. This fact, which entails some lack of generality, is considered their major drawback, although the performance may be good in specific cases.

Recently, there have been a lot of efforts towards the implementation of techniques and methods belonging to the field of Extreme Value Theory (EVT). In particular, the Peaks Over Threshold (POT)

method seems to be the most effective way of applying EVT to VaR estimation: here, we refer mainly to the dynamic POT approach developed in McNeil and Frey (2000), Fernandez (2003), and Battacharya and Ritolia (2008). This procedure allows to simultaneously take into account persistence in squared returns/losses, heteroskedasticity and fat tails. Roughly speaking, it applies the traditional POT framework (Embrechts *et al.*, 1997; McNeil and Saladin, 1997) to the GARCH residuals or, more generally, to any volatility-filtered loss time series. The main advantages of the approach are generality, strong theoretical basis and acceptable computational tractability, as well as accuracy when dealing with fat-tailed data.

The analysis presented in this paper provides the results of a backtesting exercise aimed at assessing several dynamic VaR estimation models applied to the log-losses of 15 financial instruments - equities, currencies and indexes. We employ three different dynamic POT models, some thin-tailed and several *ad hoc* fat-tailed alternative VaR models. The goal is to compare the dynamic POT models with alternative VaR models and, in particular, assess their accuracy with respect to the fat-tailed *ad hoc* models. Moreover, we assess the impact on the results of two different (EWMA and GARCH) volatility filters for the losses.

The paper is organized as follows: Section 2 briefly summarizes several VaR-based techniques for market risk measurement. Section 3 outlines the core EVT theoretical results, focusing on the dynamic POT procedure for VaR estimation. Section 4 presents the procedure employed to perform the comparison of the models. Section 5 reports the main findings of the analysis. Finally, Section 6 contains concluding remarks and discusses some issues open to future research.

## 2 Measuring market risk: methods and approaches

The exponential increase in the amount of traded assets, the broader and growing financial integration through countries, and the uncertainty that accompanied the last two decades of the '90s have led to an augmented concern about systemic risk. In particular, the need to monitor the risk of big losses caused by large drops in prices became more and more relevant to both regulators and financial institutions. One of the most successful market risk measurement models, namely the VaR system based on the RiskMetrics<sup>TM</sup> framework (RiskMetrics, 1996), was developed by J.P. Morgan Chase in the late '80s. At present, VaR is considered a standard tool, largely employed both for internal risk management purposes and for regulatory capital charging. The success of VaR can be explained by its theoretical and computational tractability, its flexibility and its adoption as a risk measure for internal models in the BIS Market Risk Amendment.

The VaR measures the maximum loss that a bank or a financial institution will suffer on his trading book over a certain time horizon with a predetermined probability level  $q = 1 - \alpha$ , called coverage level. VaR is a large quantile of the loss distribution (unconditional or conditional), and aims at summarizing the downside risk of a portfolio or a single asset. From a mathematical point of view, the VaR at level  $q$  is the real number  $k$  such that:

$$\mathbb{P}(X \leq k) = \int_{-\infty}^k f(x) dx = 1 - \alpha, \quad k = VaR(q),$$

where  $f(x)$  is the density function of the random variable  $X$  used for modelling the losses. The probability level  $\alpha$  indicates the expected relative number of violations (losses bigger than VaR), while  $q$  is the expected coverage level or, in other words, the probability that losses will be smaller or equal than the VaR value.

One of the major drawbacks of VaR is its inability to give an indication about the entire tail of the loss distribution: VaR gives the maximum loss that will not be exceeded with a certain coverage level  $q$ , but does not tell anything about the potential losses above VaR, which are indeed of primary importance. This problem can be taken into account by means of the so-called *Expected Shortfall* (ES), i.e. the conditional mean of the losses that exceed VaR:

$$ES(\alpha) = \mathbb{E}\left(X \mid X > VaR(q)\right) = \frac{1}{\alpha} \int_{VaR(q)}^{\infty} x f(x) dx .$$

Given a certain coverage level  $q$ , VaR estimates the potential maximum loss that will not be exceeded with probability  $q$ , and ES supplements this information by giving a measure of the average loss above VaR. Using both risk measures is sometimes very useful, particularly when the risk manager is concerned about big but rare loss events in the tail of the distribution.

## 2.1 The VaR approach in an econometric framework

For market risk measurement over short time intervals, a dynamic approach is strongly recommended, as it allows to extrapolate sensible information from the loss time series and include it in the VaR estimates, leading to a forward-looking VaR measure that adapts daily to current market conditions. The most common framework for dynamic VaR modelling assumes that losses follow a stochastic process with certain peculiar features, modelled in an econometric framework.

Let  $x_t = -\ln(P_t/P_{t-1})$  be the log-loss, or the negative log-return, from day  $t-1$  to day  $t$ , where  $P_t$  is the asset price on day  $t$ . The series  $x_t$  represents the path of log-losses over time, and is modelled as the product of two components: a stochastic part  $z_t$  and a deterministic part  $\sigma_{t|t-1}$ . The deterministic component allows to bind losses through time. The random variable  $z_t$  is a zero mean and unitary variance random variable with distribution  $\mathcal{D}(\Theta)$  and parameter space  $\Theta$ , while  $\sigma_{t|t-1}^2$  is the conditional variance of  $x_t$ :

$$\begin{aligned} x_t &= \sigma_{t|t-1} \cdot z_t, & z_t &\sim \mathcal{D}(\Theta), \quad \mathbb{E}(z_t) = 0, \quad \text{Var}(z_t) = 1, \quad \forall t, \\ \sigma_{t|t-1}^2 &= f(x_{t-1}, \dots, x_1; \zeta), \end{aligned} \tag{1}$$

where  $f(x_{t-1}, \dots, x_1; \zeta)$  is the equation for the conditional variance with parameters  $\zeta$ . It is possible to include a component for the conditional mean: in this case,  $x_t = \mu_{t|t-1} + \sigma_{t|t-1} \cdot z_t$ .

The VaR can be scaled virtually at any time horizon: a short-time horizon (typically, 1 or 10 days) seems to be more useful than very long time horizons. Generally, for internal market risk assessment, the one-day-ahead VaR is the most sensible choice. According to (1), the daily VaR is a quantile of  $z_t$  with

coverage probability  $q$  times the standard deviation forecast:

$$\text{VaR}(q)_{t+1} = \sigma_{t+1|t} \cdot z_q, \quad \sigma_{t+1|t}^2 = f(x_{t-1}, \dots, x_1; \zeta), \quad (2)$$

where  $z_q$  is a quantile of  $z$  at coverage level  $q$  and  $\sigma_{t+1|t}$  is the one-step-ahead standard deviation forecast.

## 2.2 Empirical properties of financial time series: modelling issues in a VaR framework

As noted by many contributions in the literature (see, for example, Mandelbrot, 1963; Fama, 1965; Ding and Granger, 1996; Cont, 2001) some common statistical features affect financial return/loss time series. Identifying and modelling these characteristics is crucial for risk management purposes, because it is necessary to extrapolate predictable movements from the loss time series in order to improve the accuracy of VaR estimates. We now detail some stylized facts we will focus on in this paper.

\* *Strong serial correlation and persistence in squared returns.*

Returns and losses are generally not significantly correlated. However, this does not imply that they are independent over time. The persistence effect of some positive transformations of losses, such as squared or absolute values, is indeed well-known; see, for example, Ding and Granger (1996), who also show that the empirical autocorrelation tends to decay rapidly in the first lags and more slowly at higher lags. Thus, a shock on losses takes a long time to be reabsorbed, leading to consecutively large returns/losses.

\* *Fat-tailed empirical distributions.*

A key and commonly investigated issue in the VaR literature is the presence of fat tails in the return/loss distribution. Even filtering the loss data with a volatility model, in order to gather the persistence in squared losses described above, the distribution of filtered losses still tends to show heavy tails. This feature is neglected in the earliest VaR models, which fit the filtered losses with a standard normal distribution, leading to a normal conditional loss distribution  $N(0, \sigma_{t+1|t}^2)$ . This assumption is consistent with the hypothesis of log-normal prices and allows to simplify considerably the VaR estimation procedure, but seems to be unrealistic and bias-leading.

Strong serial autocorrelation and persistence in squared returns are explicitly modelled by Engle (1982) ARCH and Bollerslev (1986) GARCH approaches. The GARCH(1,1) volatility framework is the most common GARCH model in the literature and can be formalized as follows:

$$x_t = \sigma_t \cdot z_t, \quad z_t \sim \mathcal{D}(\Theta), \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \forall t.$$

It takes into account serial correlation and volatility clustering, but does not fully consider volatility persistence and long memory. As  $\beta$  departs from 1, the GARCH(1,1) process gradually gives less weight to the past squared losses, as can be seen by substituting recursively  $\sigma_{t-j}^2$  in (3):

$$\sigma_t^2 = \omega(1 - \beta)^{-1} + \alpha \sum_{i=1}^{\infty} \beta^i x_{t-i}^2.$$



The process estimated for the analysis presented in this paper produces an average  $\beta$  equal to 0.89 for equities and currencies and 0.875 for indexes; moreover,  $\alpha + \beta$  results mostly very close to 1. This seems to be a typical behavior of GARCH processes on returns/losses (see, for example, Angelidis *et al.*, 2004). For this reason, we call a GARCH model with  $\beta < 0.90$  a ‘typical GARCH model’. Such a low  $\beta$  keeps the weight structure from accounting for long memory in squared losses; furthermore, the autocorrelation function is not in accord with its empirical counterpart, which often remains significant for many lags. Roughly speaking, the weight structure starts too high, decays too rapidly and approaches zero too soon.

The presence of long memory in squared returns and the need to model this feature appropriately seems to be an emergent topic in the financial literature for returns/losses time series (Ding and Granger, 1996). An elegant approach is the so-called Fractionally Integrated GARCH (FIGARCH), proposed by Baillie *et al.* (1996), which models the conditional variance allowing for an hyperbolic rate of decay (see also Christodoulou-Volos and Siokis, 2006; Gil-Alana, 2006). This process, however, is considerably more complicated than the standard GARCH from the statistical and computational point of view.

A simple model for the conditional volatility, which partially accounts for high persistence and long memory, is the Exponentially Weighted Moving Average (EWMA) framework developed by J.P. Morgan in its RiskMetrics<sup>TM</sup> environment (RiskMetrics, 1996). The EWMA volatility model is a non-parametric truncated exponential weighted average of past squared losses. Its expression, for the univariate case, is the following:

$$\sigma_{t+1}^2 = \frac{1}{T} \frac{\sum_{\tau=1}^T \lambda^{\tau-1} x_{t-\tau+1}^2}{\sum_{\tau=1}^T \lambda^{\tau-1}}, \quad T = 75, \lambda = 0.94. \quad (3)$$

For  $T \mapsto \infty$ , (3) converges to the recursive equation  $\sigma_t^2 = (1 - \lambda) x_{t-1}^2 + \lambda \sigma_{t-1}^2$ . Thus, the EWMA model is a restricted IGARCH specification, with  $\alpha = 1 - \lambda$  and  $\beta = \lambda$ , so that  $\alpha + \beta = 1$ .

Compared to the GARCH(1,1) specification, the RiskMetrics<sup>TM</sup> EWMA(75,0.94) variance partially captures the volatility persistence effect when  $\lambda > \beta$ . In the standard case,  $\lambda = 0.94$  is sufficiently large to improve the typical GARCH process. Therefore, if one is worried about persistence effects and correlation in squared losses and tries to account for these phenomena with a simple but reasonably accurate variance model, the EWMA specification is certainly of interest.

Both the GARCH(1,1) process and the EWMA model require a distributional assumption for the stochastic component  $z_t$ . If the filtered return distribution is fat-tailed,  $z_t$  should reflect this feature. The original RiskMetrics<sup>TM</sup> framework is based on the normal distribution, as well as most of the basic first-generation GARCH models employed for dynamic VaR estimation. This choice may be reasonable at low VaR coverage level (say, up to 95%) where estimating the density in the tail of the empirical filtered loss distribution is not as important as in the case of high coverage levels (99% or higher). In the latter case, neglecting the presence of fat tails and excess kurtosis may lead to a systematic underestimation of VaR. Several models aim at solving this issue: the most common ones use *ad hoc* hypotheses on the law governing the stochastic component, such as the standardized Student-*t* distribution and the GED distribution. Both these distributions have tails fatter than those of the normal and are well suited

for modelling filtered losses. A more convincing approach is based on Extreme Value Theory.

### 3 Extreme Value Theory for VaR estimation: a dynamic approach

Extreme Value Theory (EVT) is a comprehensive set of statistical procedures for the analysis of extreme data. Originally, EVT concepts were applied mainly to the study of natural extreme and rare events, such as floods and earthquakes. However, EVT quickly became popular in the financial and actuarial literature, in particular for modelling very large insurance losses. In the following, we mostly follow the setup proposed by Embrechts *et al.* (1997), adding some innovative features regarding the volatility model.

Within the EVT framework, there are essentially two kinds of methodologies. Even though they are related, each of them treats extreme data in a different manner.

\* *Block Maxima Method (BMM).*

The BMM method focuses on the largest values (maxima) taken from samples of independent and identically distributed (*iid*) observations. It is moderately expensive in terms of data, because it only uses periodical maxima and, therefore, requires wide datasets.

\* *Peaks Over Threshold (POT) method.*

The POT method focuses on observations that exceed a high threshold. It is defined on the excesses (i.e., the observations larger than some threshold  $u$ ) and is generally considered more efficient than BMM, because it uses all the excesses, not only periodical maxima.

Recently, there has been a lot of interest about POT models in the financial literature: in particular, McNeil and Frey (2000) propose an innovative application of the POT procedure to dynamic VaR and ES estimation; see also Fernandez (2003) and Batthacharyya and Ritolia (2008).

#### 3.1 The Peaks Over Threshold method

Let  $(x_1, \dots, x_n)$  be a sequence of *iid* observations from a random variable  $X$  with unknown cumulative distribution function  $F(x)$ . Let  $u \in \mathbb{R}^+$  be some predefined large value in the support of  $F$  and let  $x_0 \leq \infty$  be the right endpoint of  $F$ . The POT procedure focuses on the excess distribution over a high threshold  $u$ , that is the distribution of  $Y = X - u$ . It can be described by the conditional distribution function of  $Y$  given  $u$ , i.e. the probability that the losses exceed the threshold by no more than an amount  $y \geq 0$ , given that the threshold has been exceeded:

$$\mathbb{P}(X - u \leq y \mid X > u) = F_u(y) = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad 0 \leq y < x_0 - u. \quad (4)$$

The procedure relies on an important EVT result, the Balkema, de Haan and Pickands (BHP) theorem (see, for example, McNeil *et al.*, 2005, p. 277). It says that, under some conditions and for a certain

class of underlying distributions, the excess distribution converges to the Generalized Pareto Distribution (GPD):

$$\exists \beta(u) > 0 : \lim_{u \rightarrow x_F} \sup_{0 \leq y < x_F - u} |F_u(y) - G_{\xi, \beta(u)}(y)| = 0 \iff F \in \text{MDA}(H_\xi), \quad \xi \in \mathbb{R}, \quad (5)$$

where

$$G_{\xi, \beta}(y) = \begin{cases} 1 - \left(1 + \xi \frac{y}{\beta}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0, \\ 1 - \exp\left(-\frac{y}{\beta}\right) & \text{if } \xi = 0, \end{cases} \quad D(G_{\xi, \beta}) = \begin{cases} y \geq 0 & \text{if } \xi \geq 0, \\ 0 \leq y \leq \frac{\beta}{\xi} & \text{if } \xi < 0, \end{cases}$$

are respectively the cumulative distribution function of the GPD and its support. The BHP theorem (5) ensures that, if  $F$  belongs to the maximum domain of attraction of the *Generalized Extreme Value distribution* (GEV), for some real  $\xi$ , and if the threshold  $u$  tends to the right endpoint of the support, there exist a positive measurable function  $\beta(u)$  such that the GPD is the limiting distribution of the excesses  $Y$  above the threshold. In practice, this means that the distribution of the excess losses  $Y = X - u$  is well approximated by a GPD, given that  $u$  is sufficiently large. For our purposes, it is worth adding some comment about the two main assumptions of the theorem.

\*  $F \in \text{MDA}(H_\xi)$ .

Let  $x_1, \dots, x_n$  be a sequence of *iid* observations from  $F$  and let  $M_n = \max\{x_1, \dots, x_n\}$ . The Fisher-Tippet theorem states that, if there exist appropriate norming constants  $c_n > 0$  and  $d_n \in \mathbb{R}$  such that the normalised maxima converge to a non-degenerate distribution  $H$ , this limit distribution is a GEV:

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H, \quad H = H_{\xi, \beta}.$$

All distributions for which this condition holds are referred to as distributions in the maximum domain of attraction of the GEV, for some  $\xi \in \mathbb{R}$ . In other terms:  $F \in \text{MDA}(H_{\xi \in \mathbb{R}})$ .

\*  $u \rightarrow x_0$ .

The choice of  $u$  is the critical issue in the POT procedure. When fitting the GPD to data, a high threshold can lead to a small sample (too few excesses) and a low threshold causes a departure from the limiting result of the BHP theorem. This suggests that the choice of the threshold  $u$  is essentially related to the trade-off between variance and bias of the estimators.

The standard versions of the Fisher-Tippet and BHP theorems are based on *iid* data, but returns and losses are generally dependent. Fortunately, the convergence law for normalized maxima and for the excess distribution also holds for processes with extremal index  $\theta = 1$  (such as, for example, ARMA processes; see McNeil *et al.*, 2005, p. 270, for a definition of extremal index). For processes with

extremal index  $\theta < 1$  (this class includes, among others, ARCH and GARCH processes) the limit result is not completely justified because of the presence of extremal clusters and, therefore, non-*iid* excesses. In the latter case, the application of the POT procedure is somewhat problematic. It seems therefore more convenient to work with approximately *iid* data, applying an appropriate filter to the losses.

The core of the POT procedure is the use of the GPD  $G_{\xi,\beta}(y)$  as an approximation of the excess distribution  $F_u(y)$ . As said above, the underlying distribution  $F$  must belong to the  $MDA(H_{\xi \in \mathbb{R}})$  for the limit law to work. For the purposes of this paper, it is enough to note that (McNeil *et al.*, 2005, p. 278) ‘... essentially all the common continuous distributions of statistics or actuarial science are in  $MDA(H_{\xi})$  for some value of  $\xi$ ’. This means that the GPD can be thought of as the general model for excesses over a high threshold, without imposing any *ad hoc* assumption on  $F$ .

Using the limit result in the BHP theorem, (4) implies that the tail of the underlying distribution  $F$  has the following representation:

$$F(x) = [1 - F(u)] G_{\xi,\beta(u)}(y) + F(u), \quad x \geq u. \quad (6)$$

Substituting in (6) the GPD density and inverting, we obtain the expression for the quantile at coverage level  $q$ :

$$z_q = u + \frac{\beta}{\xi} \left\{ \left[ \frac{\alpha}{\bar{F}(u)} \right]^{-\xi} - 1 \right\}, \quad \bar{F}(u) \geq \alpha, \quad (7)$$

where  $\alpha = \mathbb{P}(X > z_q)$  and  $\bar{F}(u) = 1 - F(u)$ .

### 3.2 The POT procedure in a dynamic market risk measurement framework

The POT procedure for dynamic market risk measurement is mostly based on McNeil and Frey (2000). They suggest to perform the following steps.

- \* Fit an AR-GARCH-type process to the loss data, in order to capture the persistence in squared losses, using the Quasi-Maximum Likelihood Estimation (QMLE) procedure.
- \* Apply the POT procedure to AR-GARCH residuals by fitting with Maximum Likelihood estimation (MLE) the GPD on the excess filtered loss distributions  $F_u(z - u)$ . This allows to estimate the tail of the filtered loss distribution.
- \* Use the POT quantile estimator (7), where the estimator of  $\bar{F}_z(u)$  is  $\widehat{\bar{F}}_z(u) = n^{-1} \sum_{i=1}^n \mathbf{1}_{\{z_i > u\}}$ , and the volatility forecasts to compute daily VaR estimates.

The model allows to take into consideration both volatility clustering and fat tails in the filtered loss distribution, bypassing the problems related to the application of the POT procedure directly to losses.

The use of the GPD in the tail allows an accurate estimation of the quantile  $z_q$ , with no *ad hoc* assumption on the innovation distribution  $F_z(z)$ . Although not generally true in the case of loss time series, the

*iid* hypothesis for the residuals is usually acceptable when working with standardized losses  $z$ , given that the volatility model is sufficiently accurate to capture the stylized facts cited above.

According to McNeil and Frey (2000), VaR estimates obtained with this procedure are preferable to both dynamic methods that neglect the presence of fat tails in the residuals and unconditional methods. Moreover, the dynamic POT applied to daily loss data gives slightly better results than the GARCH Student- $t$  VaR model. As described in section 4.2, we slightly modify the procedure outlined above by including different volatility models (two GARCH-type and the EWMA set-up).

## 4 Backtesting: datasets, models and procedures

### 4.1 Datasets and methodology

The backtesting analysis is based on 15 assets (including three indexes) diversified by country (mainly UE and USA), currency, market and business activity. Such a relatively wide list should improve the robustness of the results. We use the log-losses  $x_t$  instead of the negative returns for convenience, because the GPD is defined on a non-negative support. For each asset, the time horizon goes from January 2, 2001 to May 9, 2009. The entire period of the loss time series can be split into two sub-periods.

- \* From January 2, 2001 to May 8, 2004. This contains the loss samples used for VaR estimation. The samples are constructed as follows: the first period begins on January 2, 2001, ends on March 1, 2004 and is used to obtain the first VaR forecast (for March 2, 2004). Every VaR forecast is obtained using a sample determined by shifting the previous sample one day ahead. We call these sub-periods ‘*sample periods*’.
- \* From March 2, 2004 to May 9, 2009. We use this period to evaluate the VaR estimates obtained with the models employed in the analysis. The choice of the time horizon is motivated by the presence of both relatively quiet and uprising market conditions (until February 2007) and the turmoil caused by the sub-prime crisis that began (in terms of increased volatility) in March 2007 and came to its peak around mid March 2009. We refer to this period as to the ‘*backtesting period*’.

Thus, the estimation scheme uses the ‘temporal moving window’ procedure: each VaR estimate is obtained on a different loss sample (the previous sample translated one day forward) with constant size. For practical purposes, we introduce the following notation.

1.  $\mathbb{T}$  is the entire period considered.
2.  $\mathbb{T}_j$  is the  $j$ -th sample period and  $\mathbb{X}_j = \{x_{j,1}, \dots, x_{j,n}\}$  is the corresponding estimation sample. Each sample  $\mathbb{X}_j$  contains the  $n$  realized losses of the sample period. The estimation sample moves forward every day, dropping the oldest observation and adding the newest, so the number of losses  $n$  is constant for every sample  $\mathbb{X}_j$ ,  $\forall j$  and equals to the number of trading days between January 2, 2001 and March 1, 2004.

3.  $\mathbb{T}_{bp}$  is the backtesting period and  $\mathbb{X}_{bp} = \{x_{bp,1}, \dots, x_{bp,m}\}$  is the set of losses used for the backtesting analysis. The length  $m$  of the backtesting period is equal to the number of trading days between March 2, 2004 and May 9, 2009.

The backtesting exercise is performed in an univariate framework: from a financial point of view, this means that each asset is treated as a single position in the trading portfolio. As for the indexes, we neglect the dependence structure among their components; we consider losses/returns of the indexes to be a proxy of losses/returns of index-tracking funds or ETFs (thus, single positions), whose prices are not available for such a long time backwards. The complete list of assets is as follows:

- \* Luxottica, Italcementi, Unicredit (Milan);
- \* Credit Suisse, Vodafone Group Plc, Citigroup Inc., Ford Motors, Microsoft Corp, General Motors Corp. (New York);
- \* MIBTEL (now, FTSE Italia All-Share), DAX 30, NIKKEI 225;
- \* GB Pound, BRA Real, ZAR Rand (all expressed in Euro);

Financial and banking stocks are included because they suffered the largest impact during and after the financial crisis. Moreover, we chose a couple of ‘less extreme’ assets (Luxottica, Vodafone), automotive firms and some indexes, mainly composed by financial, insurance and industrial companies. For currencies, two of them are more volatile (Real and Rand) and the remaining one is more stable (British Pound). All currencies are expressed in Euro.

Summarizing, we are interested in measuring the VaR models performance in the period  $\mathbb{T}_{bp}$ . The models are estimated with the ‘temporal moving window’ method that creates  $m$  equally-sized ( $n$ ) samples and, therefore,  $m$  VaR estimates, each of them associated to the trading dates in  $\mathbb{T}_{bp}$ .

## 4.2 The models

The backtesting analysis involves dynamic as well as unconditional VaR models. In particular, we use two different volatility models (GARCH(1,1) and EWMA(75,0.94)), some ad hoc fat tails assumptions and the dynamic POT procedure. Dynamic VaR models are grouped into three classes: dynamic POT, dynamic fat-tailed and dynamic thin-tailed models.

### 4.2.1 Dynamic POT models

Our implementation of the POT approach estimates daily dynamic VaR measures by applying the POT procedure to the volatility-filtered losses. With respect to McNeil and Frey (2000) we introduce some modifications.

1. McNeil and Frey (2000) consider an AR(1) component; rather, we neglect the conditional mean specification  $\mu_{t+1|t}$  in the GARCH(1,1) process. The reason is that, as pointed out by Angelidis *et al.* (2004), it does not seem to influence significantly the accuracy of the VaR estimates. In any

case, to motivate empirically this remark, we tried to include an AR(1), MA(1) and ARMA(1,1) component for some assets (MIBTEL, Credit Suisse, Microsoft and Rand): in almost all cases (with a few exceptions within the estimation windows  $\mathbb{T}_j$ ), the AR, MA and ARMA parameters are not significant, whereas the total estimation time increases significantly.

2. As anticipated in Sect. 3.2, the POT procedure is applied both to GARCH residuals and EWMA standardized losses. Moreover, we employ two different GARCH processes. The first one, proposed by McNeil and Frey (2000), is defined without assuming any specific distribution for the innovations  $z$  (QMLE estimation procedure); the second one assumes standardized Student- $t$  innovations. The latter assumption is justified because, if we constrain the GDP shape parameter  $\xi$  to be positive (which is reasonable if one thinks that residuals are fat-tailed),  $F_Z(z)$  belongs to the  $\text{MDA}(H_{\xi>0})$ , that is the MDA of a Fréchet, which includes the Student- $t$  distribution. Therefore, we assume implicitly that the distribution of the GARCH innovation  $z$  is Student- $t$  standardized, but we estimate the quantile  $z_q$  by modelling its tail with the accurate POT approximation instead of the standard Student- $t$  quantile. To sum up, we test three dynamic POT models: QMLE-GARCH POT,  $t$ -GARCH POT and EWMA POT.
3. We choose the residual that gives a number of excesses larger than 100. This results in a threshold equal to the 87-th empirical quantile in each sample period  $\mathbb{T}_j$ . We solve the threshold trade-off using, for all assets, the same criterion, thus ensuring that the POT estimation procedure makes use of at least 100 residuals: we believe that this number is a reasonable compromise between variance and bias. The graphical technique often proposed in the literature (Embrechts *et al.*, 1997, Sect. 6.2.2), consisting in setting the threshold at the point where the empirical mean excess function becomes approximately linear, has revealed to be a useless tool when working with large samples of data and assets.

The decision of applying the POT procedure to EWMA-filtered losses is due not only to the computational tractability of the EWMA volatility model but also to its ability in capturing squared residuals persistence more deeply than the ‘typical GARCH’ model. The decay factor  $\lambda$  of the EWMA volatility model is set to 0.94 (as in the RiskMetrics<sup>TM</sup> framework), thus leading to a decay of the weight structure of the squared residuals slower than the ‘typical GARCH’ model.

The dynamic POT models employed in the analysis can be formalized as follows:

$$\text{VaR}(q)_{t+1} = \sigma_{t+1|t} \cdot \left\{ u + \frac{\beta}{\xi} \left\{ \left[ \frac{\alpha}{\overline{F}_z(u)} \right]^{-\xi} - 1 \right\} \right\},$$

$$[a] \quad \sigma_{t+1|t}^2 = \omega + \alpha_1 x_t^2 + \beta_1 \sigma_t^2,$$

$$[b] \quad \sigma_{t+1|t}^2 = \frac{1}{\sum_{\tau=1}^{75} \lambda^{\tau-1}} \sum_{\tau=1}^{75} \lambda^{\tau-1} x_{t-\tau+1}^2, \quad \lambda = 0.94.$$

The two volatility specifications [a] and [b] remain valid for all the models in the next two subsections.

### 4.2.2 Dynamic fat-tailed models

It is possible to take into account heavy-tailed residuals by an *ad hoc* assumption on the distribution of  $z$ . The use of a standardized Student- $t$  random variable as a model for the stochastic component of the GARCH process can be considered acceptable, since the Student- $t$  distribution with few degrees of freedom is fat-tailed. The VaR is given by:

$$\text{VaR}(q)_{t+1} = \sigma_{t+1|t} \cdot z_q \sqrt{\frac{\nu-2}{\nu}}, \quad z_q = T_\nu^{-1}(q), \quad \nu > 2, \quad (8)$$

where  $z_q$  is the quantile of the Student- $t$  distribution with  $\nu$  degrees of freedom at the coverage level  $q$ ,  $\sqrt{\nu-2/\nu}$  is the reciprocal of the standard deviation for a Student- $t$  r.v. with  $\nu > 2$  and  $T_\nu(z)$  denotes the cumulative distribution function of the Student- $t$  r.v. The GARCH parameters are obtained optimizing the standardized Student- $t$  likelihood function (see next section). We will refer to 8 with volatility specifications [a] and [b] respectively as ‘ $t$ -GARCH Student- $t$ ’ and ‘EWMA Student- $t$ ’.

Another possibility (Angelidis *et al.*, 2004) consists in modelling the stochastic component  $z$  by means of a GED. The GED is a symmetric distribution characterized by a shape parameter  $\nu$ . For  $\nu < 2$ , it is leptokurtic. A dynamic VaR model with GED innovations can be written as follows:

$$\text{VaR}(q)_{t+1} = \sigma_{t+1|t} \cdot z_q, \quad z_q = G_{\nu,0,1}^{-1}(q), \quad (9)$$

where  $z_q$  is the quantile of a GED with shape parameter  $\nu$ , mean zero and unitary variance and  $G_{\nu,0,1}(z)$  is the cumulative distribution function of the standardized GED. The estimators of the GARCH parameters are obtained maximizing the GED likelihood function. 9 will be called ‘GED-GARCH GED’ and ‘EWMA GED’ models according to the volatility specifications.

A simple semi-parametric approach, not requiring any distributional assumption for the residuals, combines historical simulation and volatility filtering: we refer to this model as ‘Dynamic Historical Simulation’ (DHS). Two similar techniques have been proposed by Boudoukh *et al.* (1998), who apply the EWMA weights directly to losses, and by Fernandez (2003), who works with GARCH residuals. This method estimates the quantile  $z_q$  by means of the empirical quantile of the residual/filtered loss distribution rather than employing a predefined distributional assumption, is computationally light and quite accurate, according to the results in Fernandez (2003). Given the sample of filtered losses  $Z_j$  (with both EWMA and GARCH conditional volatility) we compute, in each estimation period  $\mathbb{T}_j$ ,  $\forall j = 1, \dots, m$ , the ordered residuals  $z_{j,[1]}, \dots, z_{j,[k]}, \dots, z_{j,[n]}$ . The empirical quantile at coverage level  $q$  is just the  $k$ -th ordered residual, where  $k = q \cdot n$ . The DHS model for dynamic VaR estimation can be written as follows:

$$\text{VaR}(q)_{t+1} = \sigma_{t+1|t} \cdot z_q, \quad z_q = z_{[q \cdot n]}.$$

We apply this method to EWMA-filtered losses and to  $t$ -GARCH, N(0,1)-GARCH and GED-GARCH residuals, leading to 4 dynamic VaR models called ‘EWMA DHS’, ‘ $t$ -GARCH DHS’, ‘N(0,1)-GARCH DHS’ and ‘GED-GARCH DHS’.



### 4.2.3 Dynamic thin-tailed models and unconditional models

Finally, we consider the distributional hypothesis used in the standard VaR estimation procedure, that is the normal distribution. As often pointed out in the literature (see, for example, McNeil and Frey, 2000 and Angelidis *et al.*, 2004), the two normal-driven models corresponding to [a] and [b] usually underestimate the quantile  $z_q$ , in particular at high coverage levels. Nevertheless, they are commonly used in practical applications: in particular, the latter is the univariate EWMA RiskMetrics<sup>TM</sup> approach, namely the standard solution for market risk measurement:

$$VaR(q)_{t+1} = \sigma_{t+1|t} \cdot z_q, \quad z_q = \Phi^{-1}(q),$$

where  $\Phi^{-1}(q)$  is the inverse of the standard normal distribution function at coverage level  $q$ . In the following, these models will be called respectively ‘N(0,1)-GARCH N(0,1)’ and ‘EWMA N(0,1)’ or ‘RiskMetrics VaR’.

Although conditional approaches for measuring market risk on a short-term basis are certainly preferable, we also consider two unconditional approaches: standard historical simulation (‘Unconditional HS’) and unconditional POT. The first one estimates VaR as the empirical quantile of the loss distribution, while unconditional POT estimates the tail of the loss distribution and high quantiles applying the POT procedure directly to losses (in each  $\mathbb{T}_j$ ), even though the *iid* assumption for the excesses (see Sect. 3) may not be satisfied. A summary of the models employed in the backtesting exercise is given in Table 1.

### 4.3 The estimation procedure

Let  $\theta' = (\omega, \alpha, \beta)$  be the vector of parameters of the GARCH(1,1) process with conditional variance  $\sigma_t^2 = \omega + \alpha_1 x_{t-1}^2 + \beta_1 \sigma_t^2$ . The estimators of the parameters are obtained maximizing the log-likelihood function of the GARCH model in use.

- \* For the GARCH POT model, estimation is based on QMLE. This allows to avoid any explicit distributional assumption for the innovations  $z$ , leading to a log-likelihood function derived from the normal assumption, even though it may not hold true. It can be shown (see McNeil *et al.*, 2005, p. 152, and the references therein) that, even when the GARCH innovations are not Gaussian, the estimators are consistent and asymptotically normal. The same procedure is used for the N(0,1) GARCH model estimation.
- \* For the  $t$ -GARCH (three models), the vector of parameters is  $\theta' = (v, \omega, \alpha_1, \beta_1)$  and the log-likelihood function used to obtain the estimates is based on the standardized Student- $t$  hypothesis for GARCH innovation.
- \* For the GED-GARCH process (two models), where  $\theta' = (v, \omega, \alpha_1, \beta_1)$ , the GED log-likelihood function is used.

**Table 1:** The table summarizes the VaR models tested in the backtesting period  $\mathbb{T}_{bp}$ .

Name	Conditional variance	Distribution: filtered losses	Estimation procedure: conditional variance
QMLE–GARCH POT	GARCH(1,1)	$F_Z \in \text{MDA}(H_{\xi \in \mathbb{R}})$	QMLE – N(0,1) log–L
$t$ –GARCH POT	GARCH(1,1)	$F_Z \in \text{MDA}(H_{\xi > 0})$	std Student– $t$ log–L
EWMA POT	EWMA(75,0.94)	$F_Z \in \text{MDA}(H_{\xi \in \mathbb{R}})$	–
N(0,1)–GARCH N(0,1)	GARCH(1,1)	N(0,1)	N(0,1) log–L
$t$ –GARCH Student– $t$	GARCH(1,1)	std $t(\nu)$	std Student– $t$ log–L
GED–GARCH GED	GARCH(1,1)	GED( $\nu, 0, 1$ )	GED log–L
EWMA N(0,1)	EWMA(75,0.94)	N(0,1)	–
EWMA Student– $t$	EWMA(75,0.94)	std $t(\nu)$	–
EWMA GED	EWMA(75,0.94)	GED( $\nu, 0, 1$ )	–
N(0,1)–GARCH DHS	GARCH(1,1)	N(0,1)	N(0,1) log–L
$t$ –GARCH DHS	GARCH(1,1)	std $t(\nu)$	std Student– $t$ log–L
GED–GARCH DHS	GARCH(1,1)	GED( $\nu, 0, 1$ )	GED log–L
EWMA DHS	EWMA(75,0.94)	–	–
Unconditional POT	–	–	–
Unconditional HS	–	–	–

For a detailed review of the log-likelihood functions see Angelidis *et al.* (2004). The whole procedure can be subdivided in the following five steps.

1. Compute, in each  $\mathbb{T}_j$ , the GARCH parameter estimates  $\hat{\theta}$ , the vector of past variance estimates (both EWMA and GARCH), a vector of  $n$  GARCH residuals  $\hat{Z}_j = (\hat{z}_1, \dots, \hat{z}_n)^T$ , the EWMA standardized losses  $\hat{Z}_j^* = (\hat{z}_1^*, \dots, \hat{z}_n^*)^T$  and the one-day-ahead volatility forecasts  $\hat{\sigma}_{t+1|t}$  for the EWMA and the GARCH models.
2. Use the Ljung-Box test statistic, the empirical autocorrelation function of residuals/filtered losses and of their squared values to check if  $\hat{Z}_j$  and  $\hat{Z}_j^*$  are *iid*. Evaluate the presence of fat tails by means of Q-Q plots, histograms and formal (Jarque-Bera) test statistics. Test whether the GARCH parameters estimates  $\hat{\theta}$  are significantly different from zero.
3. Apply the POT procedure of Sect. 3 both to  $\hat{Z}_j$  and  $\hat{Z}_j^*$ . In each  $\mathbb{T}_j$  set the threshold at the empirical 87-th quantile  $(\hat{u}, \hat{u}^*)$ , obtaining the estimated excesses  $\hat{Y}_j = (\hat{y}_1, \dots, \hat{y}_{N_u})^T$  and  $\hat{Y}_j^* = (\hat{y}_1^*, \dots, \hat{y}_{N_u}^*)^T$ . The estimation of the GPD parameters  $(\xi, \beta_u)$  uses the MLE method. The quantile estimate from the tail of the filtered loss distribution at coverage level  $q$  is:

$$\hat{z}_q = \hat{u} + \frac{\hat{\beta}_u}{\hat{\xi}} \left\{ \left[ \frac{1-q}{\hat{F}_z(u)} \right]^{-\hat{\xi}} - 1 \right\}, \quad \text{with } \hat{F}_z(u) = \sum_{i=1}^n \mathbf{1}_{\{\hat{z}_i > \hat{u}\}}.$$

The standardized Student- $t$  and the GED quantile estimators require the specification of  $\nu$  and  $\nu$ ,

respectively the degrees of freedom for the Student- $t$  fitted residuals/filtered losses and the GED shape parameter. For the GARCH residuals, both  $\nu$  and  $\upsilon$  are estimated within the GARCH estimation procedure, while in the EWMA case an ad hoc estimation procedure is required.

4. For each day in  $\mathbb{T}_{bp}$ , compute the estimated filtered loss quantile  $\hat{z}_q$  for each of the aforementioned distributional hypotheses, procedures and models.
5. For each model in Table 1, compute all the VaR estimates for the backtesting period  $\mathbb{T}_{bp}$ , adapting (2) to each specific formulation. Therefore, for each  $\mathbb{T}_j$ , an out-of-sample VaR estimate is computed, in order to get a VaR estimate for each trading day in the backtesting period  $\mathbb{T}_{bp}$ :  $\widehat{VaR}(q)_1, \dots, \widehat{VaR}(q)_m$ .

#### 4.4 The backtesting procedure

A backtesting analysis requires the evaluation of the VaR models over a sufficiently long backtesting period. Accuracy is generally assessed by means of statistical testing procedures that rely on the comparison of the expected number of VaR violations implied by the model in use and the realized violation series. In this analysis we employ an unconditional coverage test, the so-called *Proportion of Failures (PoF)* test proposed by Kupiec (1995) and Christoffersen (1998).

Let  $I_t$  be an indicator variable equal to one if the VaR value estimated on day  $t$  for day  $t + 1$  is smaller than the corresponding  $t + 1$  loss and zero otherwise:

$$I_{t+i} = \begin{cases} 1; & \text{if } x_{t+i} > VaR(q)_{t+i}, \\ 0; & \text{if } x_{t+i} \leq VaR(q)_{t+i}, \end{cases} \quad \forall i = 1, \dots, m,$$

where  $t$  denotes the  $t$ -th trading day in  $\mathbb{T}_{bp}$ .  $(I_t)_{t=1, \dots, m}$  is a sequence of independently and identically distributed Bernoulli random variables:

$$I_{t+i} \stackrel{iid}{\sim} \text{Be}(\alpha), \quad \alpha = \mathbb{P}(x_{t+i} > \widehat{VaR}(q)_{t+i}), \quad \forall i = 1, \dots, m.$$

It follows that the sum of the  $I_t$ s is a binomial random variable with  $m$  trials and probability of success  $\alpha$ , whose expected number of successes is equal to  $m \cdot \alpha$ :

$$Y_m = \sum_{i=1}^m I_{t+i} \sim \text{Bin}(m, \alpha).$$

The PoF unconditional coverage test is a likelihood ratio test. The null hypothesis states that the failure rate  $\hat{\pi} = Y_m/m$  is equal to  $\alpha$ , while the alternative hypothesis is  $\hat{\pi} \neq \alpha$ . Formally, the PoF test statistic is defined as follows:

$$PoF = -2 \log \left[ \frac{l(\alpha; \widehat{I}_{t+1}, \dots, \widehat{I}_{t+m})}{l(\hat{\pi}; \widehat{I}_{t+1}, \dots, \widehat{I}_{t+m})} \right],$$

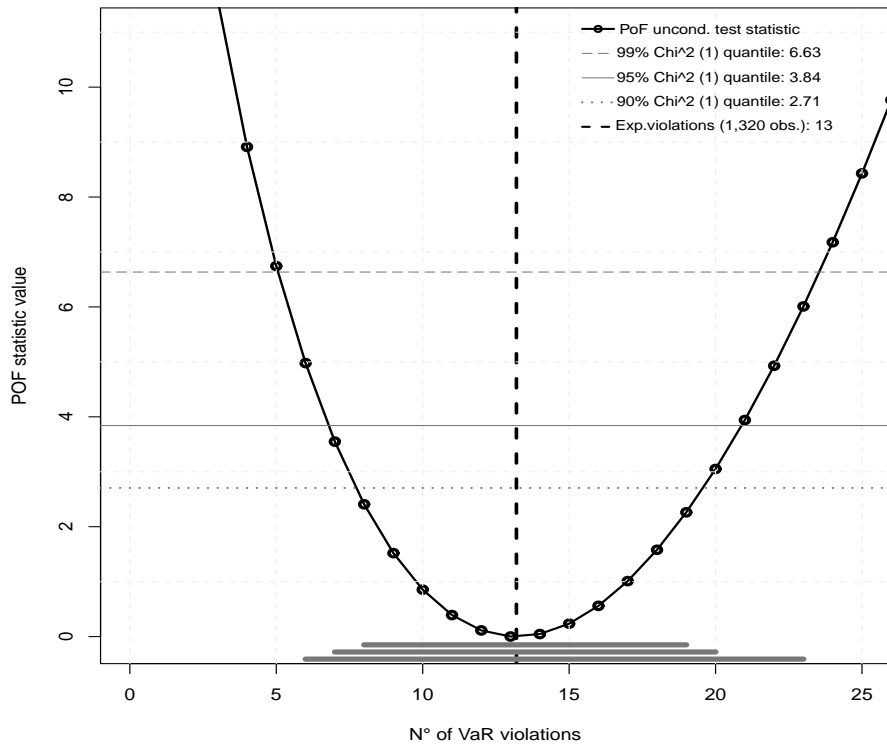
where  $l(\alpha; \widehat{I}_{t+1}, \dots, \widehat{I}_{t+m}) = (1 - \alpha)^{N_0} \alpha^{N_1}$ ,  $l(\hat{\pi}; \widehat{I}_{t+1}, \dots, \widehat{I}_{t+m}) = (1 - \hat{\pi})^{N_0} \hat{\pi}^{N_1}$ ,  $N_1$  is the number of

violations in the backtesting period  $\mathbb{T}_{bp}$  and  $N_0 = m - N_1$ . The test statistic tends to be small when  $N_1$  is “close” to the expected number of violations ( $PoF = 0$  when  $\hat{\pi} = \alpha$ ) and large when  $\hat{\pi}$  is either larger or smaller than  $\alpha$ . Thus, the  $PoF$  testing procedure allows to reject both a too conservative VaR model (too large VaR estimates) and a poor VaR model (too small VaR estimates).

The  $PoF$  test statistic, for  $m \mapsto \infty$ , is distributed as a  $\chi^2$  random variable with 1 degree of freedom. If the test statistic is smaller than the  $\gamma$ -quantile of the  $\chi^2_1$  distribution, where  $1 - \gamma$  is the size of the test, the VaR model under examination is accepted and therefore can be considered reliable and effective.

The test was computed at sizes  $1 - \gamma = 10\%$  and  $1 - \gamma = 5\%$ . This choice is motivated by the trade-off between a small size and its costs in terms of Type II error: the smaller the size, the larger the probability of correctly choosing the null hypothesis but also the chance of a Type II error. In other words, setting a high probability (for example, 99%) of correctly choosing an effective VaR model, we also increase the probability of erroneously selecting an inappropriate VaR model, i.e. a VaR model with too many or too few violations. This can be seen from Fig. 1, which represents the  $PoF$  test statistic with  $m = 1,320$  (approximately five years of daily observations) and a 99% coverage level. The thick black curve represents the values of the test statistic as a function of the number of violations ( $x$ -axis), while the vertical dashed line is the expected number of violations. The horizontal lines represent three different critical values for the test, i.e. the  $\gamma$ -quantiles of the  $\chi^2_1$  distribution with  $\gamma = 90, 95$  and 99%. With a 1% size, the acceptance region of the test, represented by the lowermost thick grey line and including the numbers of violations such that the curve lies under the dashed horizontal line, is significantly larger than in the cases corresponding to the 5 and 10% levels, causing a Type II error that we deem too large.

**Figure 1:** The  $PoF$  test statistic as a function of the number of VaR violations (at 99% coverage level) and the acceptance region as a function of the critical values corresponding to the 1, 5 and 10% sizes.



## 5 Results

Since the end of the second quarter 2007, equity markets became increasingly volatile due to the emerging financial crisis culminated, in September 2008, in the default of Lehman Brothers and continued until the end of 2008. In this period, the markets were characterized by unusually high volatility levels and huge losses due to deleveraging and flight to quality. The equity indexes of most industrialized countries wiped off almost all the capitalization accumulated since mid 2004 and, in some cases, even more. Similarly, the first quarter 2009 was dominated by high volatility and large losses worldwide.

The models are implemented with three VaR probability levels ( $\alpha$ ) equal to 5, 1 and 0.5%, thus assuming three coverage levels (95, 99 and 99.5%). The comparison of the overall performance of the models during the backtesting period takes into consideration the number of rejections of the null hypothesis for each VaR model: we count how many times each VaR model applied to the 15 loss time series (see Sect. 4.1) is rejected by the PoF test. This approach is justified because a model never rejected by the PoF test is certainly preferable to a rejected (one or more times) model: thus, a VaR model with fewer rejections is considered more accurate. More specifically, we adopt two different rankings. The first one considers the number of joint rejections with respect to the 99 and 99.5% VaR, so that it focuses on the behavior of VaR estimates in the tail of the loss distribution. The alternative ranking counts the number of rejections at 95% and is used along with the first one to complete the empirical evidence concerning VaR accuracy.

### 5.1 Preliminary results

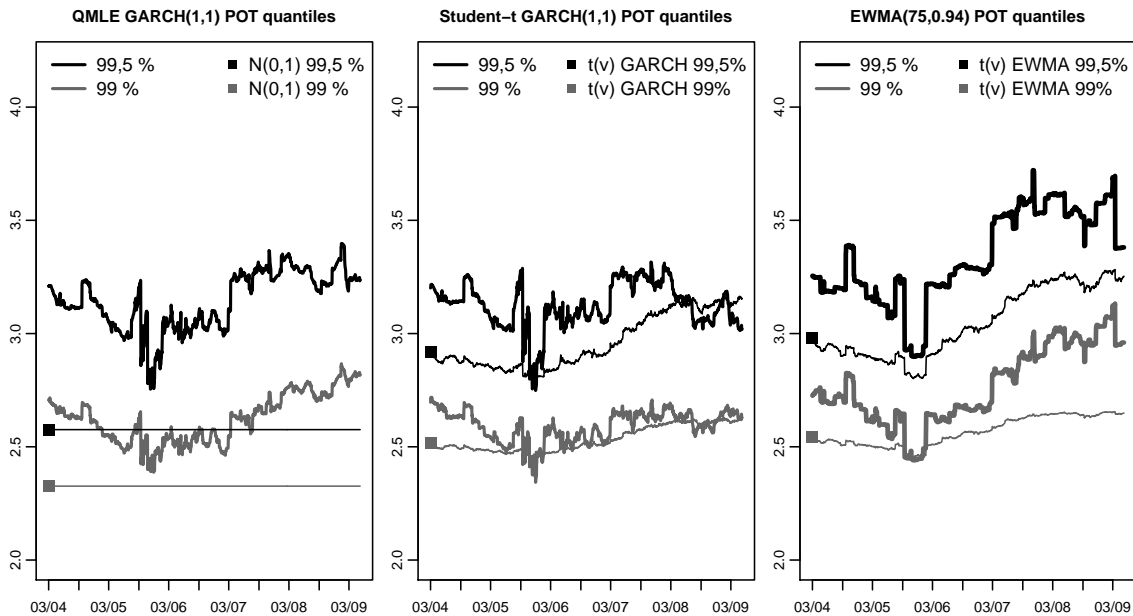
Due to space limitations, we cannot show here all the results. We limit ourselves to detail the most important outcomes of the analysis of filtered losses. We use both formal testing procedures and graphical representations to evaluate the statistical properties of volatility-filtered losses. As for the graphs, three equally-sized filtered losses sub-samples are used for each asset, starting respectively on July 6, 2002, June 15, 2004 and December 27, 2005. The main findings can be summarized as follows.

- \* The Ljung-Box test with 20 lags applied to each filtered loss sample ( $\hat{Z}_j$  and  $\hat{Z}_j^*$ ) and to each asset, does not produce, in almost all cases, any evidence of serial correlation. Two exceptions are the FORD and Luxottica GARCH/EWMA residuals for which, at the extremes of the backtesting period, the test accepts the presence of autocorrelation.
- \* The Q-Q plot of the filtered losses (with both EWMA and GARCH filters) with respect to the standard normal distribution shows the presence of heavy tails, in particular on the right tail (losses). This is more evident for the first and the third EWMA/GARCH filtered loss sub-sample.
- \* The mean excess estimator plot used for choosing the threshold of the filtered losses sub-samples gives ambiguous results. In general, for most assets and sample periods, the expected linear pattern starts approximately at the threshold. In some cases, however, it is hard to identify a linear trend due to a large dispersion. In other cases, the linear trend seems to start above the threshold suggesting that the threshold should be larger.

## 5.2 Estimation results

The first issue of interest is the difference between the various estimates  $\hat{z}_q$  of the quantiles of the residuals obtained with the models under consideration. For Citigroup, Fig. 2(a) shows the estimates  $\hat{z}_q$  for the QMLE-GARCH POT model and the quantiles of the standard normal, all with a 99% coverage level. Similarly, Fig. 2(b) and (c) display  $\hat{z}_q$  for the remaining two POT models ( $t$ -GARCH and EWMA) and the estimated quantiles for dynamic models with the standardized Student- $t$  distribution. The results for the other assets are similar and therefore not reported here.

**Figure 2:** Citigroup. Estimates of the quantiles of the filtered losses  $z_q$  with  $q=99$  and 99.5%. Dynamic POT procedures.



As can be seen from Fig. 2(a), the normal quantiles are significantly lower than the POT (GARCH-QMLE,  $t$ -GARCH, EWMA) for the smallest VaR probabilities (1% and 0.5%). Thus, the assumption of a normal-driven dynamic VaR model leads to a larger risk of VaR underestimation, in particular at high coverage levels. This result is in line with the majority of VaR backtesting analyses in the literature (Angelidis *et al.*, 2004; McNeil, Frey, 2000; Hung, *et al.* 2008). In figures 2(b) and (c) the POT  $t$ -GARCH and POT EWMA filtered losses quantiles are plotted against the standardized Student- $t$  quantile obtained from the estimation of dynamic VaR models. The POT  $t$ -GARCH quantiles (Fig. 2(b)) are larger than the Student- $t$  quantiles. In the last year of analysis, the POT  $t$ -GARCH quantiles declines slightly, whereas the  $t$  quantiles rise sharply; however, this feature is specific to Citigroup. The most interesting results can be seen by looking at Fig. 2(c): the POT quantiles estimated from EWMA filtered losses are significantly larger than the standardized Student- $t$  quantiles.

A second issue of interest is the relationship between the method used for volatility estimation and the estimated quantiles of the filtered losses. In this case the main messages of the empirical analysis are as follows.

- For most assets and models, towards the end of  $\mathbb{T}_{bp}$  (from August 2007, say), the numerical values

of the estimated quantiles increase significantly.

- The EWMA estimated quantiles are generally larger than all other filtered losses estimated quantiles. This is probably due to both the weight structure of the EWMA filter with  $\lambda = 0.94$  and the accuracy of the POT quantile estimator. With respect to a typical GARCH, the weight structure of the EWMA filter gives less weight to the very last observation but is more persistent: unusually high squared losses, frequently observed during the entire backtesting period, are not immediately overweighted, but the weight structure spreads its effects over a relatively long time horizon, contributing to constantly large and less volatile estimated quantiles with respect to the estimated quantiles of the GARCH residuals.
- As a consequence of fewer excesses, the 99.5% estimated quantiles are more volatile than the 99% quantiles.

### 5.3 Backtesting results

Tables 4-10 in appendix A.1 show the full set of backtesting results for each asset involved in the analysis. The tables include two patterns: each horizontal box contains three rows with the results of the PoF test with three different VaR probabilities (5%, 1% and 0.5%). The first column shows the number of VaR violations, the second column the PoF test statistic, and the third the  $p$ -value associated with the numerical value of the PoF test statistic.

For the 10 and 5% sizes, we evaluate the models with VaR probability levels  $\alpha$  set to 5, 1 and 0.5% for all assets (see Sect. 4.1). It follows that, for either size of the test, the number of rejections for each model with a certain VaR probability  $\alpha$  ranges from 0 to 15, namely the number of assets under consideration.

#### 5.3.1 Results 1: test size 5%

Table 2 summarizes the backtesting results when the size of the test is equal to 5%. The first column of Table 2 describes the VaR model tested, the central columns report the number of violations for each VaR probability and the last column shows the model ranking, as determined by the total number of rejections: the models are listed in ascending order, according to the total number of rejections at 99 and 99.5% coverage levels.

Let us start with the worst performing VaR models, namely dynamic normal-driven VaR models. At 99 and 99.5%, for N(0,1)-GARCH, the number of rejections is 24 out of a maximum possible value of 30, while for the RiskMetrics<sup>TM</sup> VaR we get 28 rejections. Furthermore, as the VaR coverage level increases, the performance of both models gets worse. This is not surprising, given the well-known inadequacy of the normal distribution at high coverage levels in presence of fat tails in the filtered loss distributions. At 95%, the situation is quite different: the two dynamic normal-driven VaR models yield respectively 1 and 3 rejections and are at the top of the ranking.

The two static VaR models (unconditional POT and unconditional HS) perform badly regardless of the VaR probability  $\alpha$ : for all the three VaR probabilities, the number of PoF rejections ranges from 12 to

13, putting the two models at the bottom of the ranking (24 and 25 out of 30). This result is consistent with other backtesting studies (Bhattacharyya and Ritolia, 2008; McNeil and Frey, 2000) and is largely expected because of the inappropriateness of a static VaR model in a highly volatile period. A notable exception are the backtesting results of currencies, for which the two models perform well in 2 cases out of 3 (ZAR Rand and BRA Real).

Fat-tailed dynamic VaR models (GARCH and EWMA models with Student- $t$  and GED stochastic component) do not yield unequivocal performances. Looking at the total number of rejections with 1 and 0.5% VaR probabilities, the worst models are the EWMA fat-tailed specifications, both with 11 rejections out of 30. Furthermore, considering separately each VaR ranking, as the coverage level increases, the EWMA GED model performs progressively worse (3, 4 and 7 rejections out of 15), while the EWMA Student- $t$  improves its performance (9, 7 and 4 rejections). However, at 95% VaR coverage, the EWMA GED model is in the top 5 of the ranking.

The GARCH fat-tailed dynamic models achieve globally better results than the EWMA models, with 4 (GARCH Student- $t$ ) and 6 (GARCH GED) rejections out of 30; moreover, the GARCH  $t$  model is among the top 5 of the total rejection ranking. As in the previous case, the GARCH GED model worsens as the coverage level increases (1, 2 and 4 rejections), while the Student- $t$  model improves (4, 2 and 2 rejections). Summarizing, GARCH fat-tailed models perform slightly better than EWMA models, but the use of a GED stochastic component seems to produce worse results as the coverage level gets larger.

In general, DHS models work better at large VaR coverage levels than at low levels. The performance of the GED-GARCH and the  $N(0,1)$ -GARCH DHS models is identical at 5 and 1% VaR probability (4 and 3 rejections), but the GED model is preferable at 0.5%, with only one rejection. However, looking at the total number of rejections, these two DHS models perform similarly to fat-tailed dynamic models (4 and 6 rejections out of 30), which is acceptable but put the models out of the top 5 of the ranking. Consistently with the total rejection ranking, the best DHS models are the EWMA DHS and the  $t$ -GARCH DHS model, respectively with 0 and 2 rejections out of 30. In particular, despite its simplicity and questionable non-parametric methodology, the EWMA DHS model is at the top of the ranking. The  $t$ -GARCH DHS is sometimes rejected (3 times out of 15) at 5% coverage, but is the third best model at 1 and 0.5% (1 rejection).

One of the most interesting results is related to the POT models. At the top of Table 2 are listed the best VaR models according to the PoF test. On the basis of the overall ranking, dynamic POT models are the best-performing VaR approaches: the EMWA POT and the  $t$ -GARCH POT model are never rejected. The QMLE-GARCH POT models also perform well, with just 1 rejection out of 30: this result is in line with the outcomes obtained by McNeil and Frey (2000), Fernandez (2003), and Battacharya and Ritolia (2008). At 95% VaR coverage, both the QMLE and the  $t$ -GARCH POT models do not perform as well as at larger coverage levels, yielding respectively 3 and 4 rejections out of 15. Finally, the two GARCH POT VaR models turn out to be similar to the other dynamic fat-tailed VaR models.

Figures 3 (DAX30 Index) – 6 (ZAR Rand in €) in appendix A.2 show the hit sequences with a 99% coverage level: each symbol on the graph represents a VaR violation, while the grey line shows the loss time series. As can be seen by looking at the top of each figure, violations originated by the two static



**Table 2: Proportion of failures (POF) unconditional test. Confidence level: 95%.**

This table summarizes, for all assets analyzed (equities, indexes and currencies) and for each VaR model applied, the total number of rejection of the null hypothesis for correct VaR coverage using the POF unconditional coverage test (a.s. Chi squared) test with a confidence probability level of 95%

<b>VaR model</b>	<b>N° rejections VaR 5% (max. 15)</b>	<b>N° rejections VaR 1% (max. 15)</b>	<b>N° rejections VaR 0,5% (max. 15)</b>	<b>Total reject. (1%, 0.5%) (max. 30)</b>
EWMA(75,0.94) POT	0	0	0	<b>0</b>
EWMA(75,0.94) DHS	0	0	0	<b>0</b>
$t$ -GARCH(1,1) POT	4	0	0	<b>0</b>
QMLE-GARCH(1,1) POT	3	1	0	<b>1</b>
$t$ -GARCH(1,1) DHS	3	1	1	<b>2</b>
$t$ -GARCH(1,1) Student $t$	4	2	2	<b>4</b>
GED-GARCH(1,1) DHS	4	3	1	<b>4</b>
GED-GARCH(1,1) GED	1	2	4	<b>6</b>
N(0,1)-GARCH(1,1) DHS	4	3	3	<b>6</b>
EWMA(75,0.94) GED	3	4	7	<b>11</b>
EWMA(75,0.94) Student $t$	9	7	4	<b>11</b>
N(0,1)-GARCH(1,1) N(0,1)	1	11	13	<b>24</b>
Unconditional POT	12	12	12	<b>24</b>
Unconditional HS	12	13	12	<b>25</b>
EWMA(75,0.94) N(0,1)	3	13	15	<b>28</b>

models are concentrated in the last 18 months, when equity markets have become very volatile. Even the two normal-driven dynamic models perform poorly during the financial crisis. Conversely, the EWMA DHS and EWMA POT models are associated to violations spread over the entire backtesting period. Differently from most of the models tested, these two VaR models are reliable both in quiet and extreme market conditions. This is an important feature: in case of highly volatile market conditions, they react more timely and provide better VaR estimates, minimizing the slippage of VaR-based capital charges over time.

### 5.3.2 Results 2: test size 10%

Table 3 shows the results of the backtesting analysis when the size of the test is equal to 10%. Again, the worst VaR models according to the global ranking are the two normal-driven dynamic models (28 and 29 rejections) and the static VaR models. The first two result worse than in the previous ranking, falling in the last two positions of the global ranking. At 5% VaR probability, the performance of the two normal dynamic models deteriorates as well: in particular, the RiskMetrics™ model is rejected 8 times out of 15 with respect to 3 rejections at the 5% size. On the other hand, the GARCH-QMLE N(0,1) is significantly better at 5% VaR probability, with only 3 rejections out of 15.

Dynamic fat-tailed VaR models perform similarly to the preceding section: results are slightly less satisfactory for EWMA-filtered models than for GARCH models and, as the coverage levels increase, GED VaR models (both EWMA and GARCH) are rejected more often than Student- $t$  based models. The only exception to the global failure of fat-tailed dynamic models is the Student- $t$  GARCH model, whose performance is similar to the 5% size case (fourth out of 15 models).

According to the total rejection-based ranking, dynamic historical models perform not very differently from the previous section. However, the Student- $t$  DHS model now drops in the middle of the global ranking (6 rejections out of 30).

The POT-based dynamic model is again at the top of the ranking. The first place is taken by the EWMA POT model with zero rejections out of 30, the second place by the Student- $t$  POT model with 1 rejection and the third by the GARCH QMLE POT (3 rejections). At 5% VaR probability, the EWMA POT model is rejected 2 times out of 15, but this was the second smallest number of rejections of all models. Finally, the Student- $t$  GARCH is not as efficient as the other POT models at 5% VaR probability, yielding 5 rejections out of 15.

POT models are not the only well-performing approaches: the EWMA DHS model also reaches the top of both rankings (overall and at 5%), contrary to other DHS models, which turn out to be sufficiently accurate, but not at the top of the ranking. The EWMA DHS model is rejected twice and results the second best model at 5% probability level, while on the basis of the total rejection ranking it is rejected just once, similarly to the Student- $t$  POT model. These outcomes put it at the second place of the ranking.

Finally, the POT dynamic-based models are the most effective VaR estimation techniques, according to the total rejection ranking, also at the largest size employed. Only at 5% VaR probability the POT EWMA is sometimes rejected, but much more rarely than most of the other models. The other two POT-based models tested in our analysis achieve quite satisfactory results, even though the QMLE POT model is the worst POT model at both sizes.

In conclusion, the main outcomes of the backtesting analysis are the following.

- \* According to the PoF test at both 5 and 10% size, the POT EWMA and EWMA DHS models are the most reliable and effective ones when the VaR coverage level is 99 or 99.5%. At 95% coverage they are rejected in single cases, but are still preferable to all other models.
- \* Dynamic fat-tailed models do not perform particularly well at the largest coverage level, lying mostly in the middle of the two rejection-based rankings. Only the  $t$ -GARCH approach has a performance comparable to the top-ranked models.
- \* EWMA fat-tailed models are slightly less efficient than GARCH fat-tailed approaches. The number of rejections of the GED-based dynamic models (both EWMA and GARCH) gets larger as the VaR coverage level increases, contrary to the Student- $t$  models.

From the risk management perspective, the message is quite clear: the POT and DHS approaches with the EWMA volatility specification should be preferred, especially when tail risk is of paramount importance.

**Table 3: Proportion of failures (PoF) test. Confidence level: 90%.**

This table summarizes, for all assets analyzed (equities, indexes and currencies) and for each VaR model applied, the total number of rejection of the null hypothesis for correct VaR coverage using the PoF unconditional coverage test (a.s. Chi squared) test with a confidence level of 90%

<b>VaR model</b>	<b>N° rejections VaR 5% (max. 15)</b>	<b>N° rejections VaR 1% (max. 15)</b>	<b>N° rejections VaR 0,5% (max. 15)</b>	<b>Total reject. (1%, 0.5%) (max. 30)</b>
EWMA(75,0.94) POT	2	0	0	<b>0</b>
EWMA(75,0.94) DHS	2	1	0	<b>1</b>
$t$ -GARCH(1,1) POT	5	0	1	<b>1</b>
QMLE-GARCH(1,1) POT	6	1	2	<b>3</b>
$t$ -GARCH(1,1) Student $t$	5	2	2	<b>4</b>
$t$ -GARCH(1,1) DHS	5	3	3	<b>6</b>
GED-GARCH(1,1) DHS	5	4	3	<b>7</b>
GED-GARCH(1,1) GED	1	4	4	<b>8</b>
EWMA(75,0.94) Student $t$	11	7	5	<b>12</b>
N(0,1)-GARCH(1,1) DHS	6	7	6	<b>13</b>
EWMA(75,0.94) GED	7	6	8	<b>14</b>
Unconditional HS	13	13	12	<b>25</b>
Unconditional POT	12	13	13	<b>26</b>
EWMA(75,0.94) N(0,1)	8	13	15	<b>28</b>
N(0,1)-GARCH(1,1) N(0,1)	3	15	14	<b>29</b>

## 6 Conclusions

This paper compares several dynamic VaR models commonly employed in market risk measurement. The time horizon includes the extreme market conditions observed during the financial crisis caused by the sub-prime mortgage crisis. The focus is primarily on the performance of a dynamic Peaks over Threshold procedure, which not only allows to take into account heavy-tailed risk factor distributions, but also has a strong theoretical justification.

The results show that the POT model and the Dynamic Historical Simulation combined with the EWMA volatility specification are preferable, in particular at high VaR coverage levels. Whereas the good performance of the POT approach was not unexpected giving the features of the time series at hand and the sound theoretical foundations of the methodology, the success of DHS and EWMA is somewhat surprising in two respects. First, it is well-known that historical simulation does not perform particularly well in presence of extreme market movements. Second, the non-parametric EWMA estimator has weak theoretical justifications. A possible explanation for the first issue is that DHS, unlike standard historical simulation, uses a volatility filter, and this filter is likely to be the crucial tool for improving the performance of the method. As for the second one, the EWMA empirical performance is known to be rather good, and the weight structure captures reasonably well (better than typical GARCH models)

persistence effects and autocorrelation of squared losses. It should also be noted that EWMA works better than GARCH when associated to POT and DHS, but worse than GARCH when used in conjunction with ad hoc fat-tailed models.

Several issues remain open for further investigations. First, we have only used “portfolios” consisting entirely of single assets, taking an index fund as a single asset. A generalization of the analysis to more realistic portfolios (for example, portfolios containing derivatives) does not appear straightforward but would be quite important. Second, it is well known that it is difficult to beat GARCH(1,1). In particular, moving either  $p$  or  $q$  to 2 does little to improve fit. On the other hand, FIGARCH models have recently been proposed as a tool for investigating long memory in squared returns. It would be of interest to check whether models of this family can provide better results than standard GARCH.

**Acknowledgements.** We would like to thank Christopher L. Gilbert for a useful discussion and pointed remarks.

## References

- [1] Angelidis T, Benos A and Degiannakis S (2004) *The Use of GARCH Models in VaR Estimation*. Statistical Methodology 1:105-128
- [2] Baillie, RT and Bollerslev, T and Mikkelsen, HO (1996) *Fractionally integrated generalized autoregressive conditional heteroskedasticity*. J. of Econometrics 74:3-30
- [3] Baixauli JS, Alvarez S (2006) *Evaluating effects of excess kurtosis on VaR estimates: Evidence for international stock indices*. R. of Quantitative Finance and Accounting 27:27-46
- [4] Bank for International Settlements (2006) *Amendment to the Capital Accord to incorporate market risks*. Basel Committee on Banking Supervision
- [5] Bhattacharyya M, Ritolia G (2008) *Conditional VaR using EVT - Towards a planned margin scheme*. International Review of Financial Analysis 17:382-395
- [6] Bollerslev T (1986) *Generalized autoregressive conditional heteroskedasticity*. J. of Econometrics 31:307-327
- [7] Boudoukh J, Richardson M and Whitelaw R.F. (1998) *The best of both worlds*. Risk, 11: 64-67
- [8] Christodoulou-Volos, C. and Siokis, F.M. (2006) *Long range dependence in stock market returns*. Applied Financial Economics, 16, 1331-1338.
- [9] Christoffersen PF (1998) *Evaluating Interval Forecasts*. International Economic Review 39:841-862
- [10] Cont R (2001) *Empirical properties of asset returns: stylized facts and statistical issues*. Quantitative Finance 1:223-236
- [11] Ding Z and Granger CWJ (1996) *Modeling volatility persistence of speculative returns: A new approach*. J. of Econometrics 73:185-1215
- [12] Embrechts P, Klüppelberg C, Mikosch T (1997) *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin
- [13] Engle RF (1982) *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica 50:987-1007
- [14] Engle RF and Manganelli S (1999) *CAViaR: Conditional Value at Risk by Quantile Regression*. NBER, Working Paper 7341.
- [15] Fama EF (1965) *The Behavior of Stock Market Prices*. J. of Business 38:34-105
- [16] Fernandez V (2003) *Extreme Value Theory and Value at Risk*. Revista de Analisis Economico – Economic Analysis Review 18:57-85

- [17] Gil-Alana, L (2006) *Fractional integration in daily stock market indexes*. Review of Financial Economics 15, 28-48
- [18] Kupiec P (1995) *Techniques for Verifying the Accuracy of Risk Management Models*. J. of Derivatives 3, 73-84
- [19] Hung, JC, MC Lee and HC Liu (2008) *Estimation of Value-at-Risk for energy commodities via fat-tailed GARCH models*. Energy Economics, 30:1173-1191.
- [20] Mandelbrot B (1963) *The Variations of Certain Speculative Prices*. J. of Business, 36:394-419.
- [21] McNeil AJ and Frey R (2000) *Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach*. J. of Empirical Finance 7:271-300
- [22] McNeil AJ, Frey R & Embrechts P, (2005) *Quantitative Risk Management: Concepts, Techniques, and Tools* Princeton University Press, Princeton and Oxford
- [23] McNeil AJ and Saladin T (1997) *The Peaks over Threshold Method for Estimating High Quantiles of Loss Distributions*. Proceedings of XXVIIth International ASTIN Colloquium 23-43, Cairns, Australia.
- [24] RiskMetrics, (1996) *RiskMetrics<sup>TM</sup>– Technical Document*. Fourth Edition, J.P. Morgan/Reuters, New York
- [25] R Development Core Team, (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria
- [26] Würtz D *et al.* (2009) *fExtremes: Rmetrics - Extreme Financial Market Data*. R package version 290.76
- [27] Würtz D, Chalabi Y and Miklovic M (2009) *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. R package version 2100.78

# A Appendix

## A.1 Tables

**Table 4: VaR backtesting – DAX30 (DAX) index and MIBTEL (MIB) index. VaR probabilities: 5%, 1% and 0.5%.**

VaR model	DAX			MIB		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE-GARCH(1,1)	72	0,521	( 47,053 % )	87	6,524	( 1,065 % )
POT	18	1,554	( 21,225 % )	19	2,293	( 13 % )
	10	1,498	( 22,091 % )	7	0,026	( 87,251 % )
<i>t</i> -GARCH(1,1)	69	0,123	( 72,577 % )	87	6,524	( 1,065 % )
POT	18	1,554	( 21,255 % )	19	2,293	( 13 % )
	9	0,773	( 37,941 % )	8	0,286	( 59,287 % )
EWMA(75,0.94)	72	0,521	( 47,053 % )	76	1,572	( 20,986 % )
POT	15	0,227	( 63,406 % )	14	0,052	( 82 % )
	9	0,773	( 37,941 % )	9	0,798	( 37,163 % )
N(0,1)-GARCH(1,1)	74	0,933	( 33,398 % )	93	10,504	( 0,119 % )
N(0,1)	24	7,120	( 0,762 % )	39	33,533	( 0 % )
	20	17,602	( 0,003 % )	24	27,479	( 0 % )
<i>t</i> -GARCH(1,1)	78	2,099	( 14,735 % )	94	11,25	( 0,08 % )
Student <i>t</i>	21	3,9	( 4,829 % )	29	14,316	( 0,02 % )
	9	0,773	( 37,941 % )	14	6,331	( 1,186 % )
GED-GARCH(1,1)	75	1,183	( 27,677 % )	92	9,782	( 0,176 % )
GED	21	3,9	( 4,829 % )	27	11,254	( 0,08 % )
	13	4,817	( 2,818 % )	14	6,331	( 1,186 % )
EWMA(75,0.94)	82	3,702	( 5,436 % )	88	7,127	( 0,759 % )
N(0,1)	28	12,589	( 0,039 % )	38	31,349	( 0 % )
	17	11,388	( 0,074 % )	27	35,686	( 0 % )
EWMA(75,0.94)	84	4,661	( 3,087 % )	94	11,25	( 0,08 % )
Student <i>t</i>	22	4,882	( 2,714 % )	29	14,316	( 0,015 % )
	13	4,817	( 2,818 % )	16	9,647	( 0,19 % )
EWMA(75,0.94)	82	3,702	( 5,436 % )	88	7,127	( 0,759 % )
GED	19	2,231	( 13,528 % )	29	14,316	( 0,015 % )
	13	4,817	( 2,818 % )	16	9,647	( 0,19 % )
N(0,1)-GARCH(1,1)	75	1,183	( 27,677 % )	84	4,862	( 2,745 % )
DHS	20	3,015	( 8,252 % )	18	1,605	( 20,51 % )
	11	2,426	( 11,932 % )	12	3,595	( 5,795 % )
<i>t</i> -GARCH(1,1)	76	1,461	( 22,684 % )	83	4,359	( 3,682 % )
DHS	19	2,231	( 13,528 % )	18	1,605	( 20,51 % )
	10	1,498	( 22,091 % )	9	0,798	( 37,163 % )
GED-GARCH(1,1)	75	1,183	( 27,677 % )	84	4,862	( 2,745 % )
DHS	19	2,231	( 13,528 % )	18	1,605	( 20,51 % )
	11	2,426	( 11,932 % )	10	1,535	( 21,541 % )
EWMA(75,0.94)	73	0,713	( 39,861 % )	76	1,572	( 20,986 % )
DHS	18	1,554	( 21,255 % )	17	1,031	( 31 % )
	9	0,773	( 37,941 % )	9	0,798	( 37,163 % )
Unconditional	78	2,099	( 14,735 % )	103	18,968	( 0,001 % )
POT	28	12,589	( 0,039 % )	30	15,954	( 0,01 % )
	14	6,252	( 1,241 % )	17	11,499	( 0,07 % )
Unconditional	81	3,261	( 7,093 % )	105	20,92	( 0 % )
HS	29	14,145	( 0,017 % )	30	15,954	( 0,01 % )
	11	2,426	( 11,932 % )	17	11,499	( 0,07 % )

Backtesting length: 1320. POT excesses: 104. Sample length: 800. Expected VaR violations: 66 (5%), 13 (1%), 7 (0,5%).  
 Backtesting length: 1317; POT excesses: 108; Samples size: 825. Expected VaR violations: 66 (5%), 13 (1%), 7 (0,5%).

**Table 5: VaR backtesting – NIKKEI 225 index (N225) and Credit Suisse Group (CS) (NYSE).**  
 VaR probabilities: 5%,1% and 0.5%.

VaR model	N225			CS		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE–GARCH(1,1)	70	0,647	(42,125 %)	74	1,144	(28,481 %)
POT	14	0,124	(72,477 %)	15	0,272	(60,211 %)
	8	0,391	(53,222 %)	10	1,582	(20,845 %)
<i>t</i> –GARCH(1,1)	69	0,461	(49,702 %)	71	0,492	(48,319 %)
POT	16	0,785	(37,575 %)	16	0,615	(43,301 %)
	8	0,391	(53,222 %)	10	1,582	(20,845 %)
EWMA(75,0.94)	67	0,183	(66,916 %)	59	0,681	(40,935 %)
POT	16	0,785	(37,575 %)	15	0,272	(60,211 %)
	8	0,391	(53,222 %)	8	0,306	(58,033 %)
N(0,1)–GARCH(1,1)	75	2,02	(15,522 %)	56	1,492	(22,193 %)
N(0,1)	21	4,538	(3,315 %)	23	6,199	(1,278 %)
	15	8,506	(0,354 %)	16	9,778	(0,177 %)
<i>t</i> –GARCH(1,1)	75	2,02	(15,522 %)	63	0,094	(75,941 %)
Student– <i>t</i>	17	1,309	(25,257 %)	14	0,064	(80,044 %)
	10	1,776	(18,266 %)	10	1,582	(20,845 %)
GED–GARCH(1,1)	73	1,383	(23,955 %)	58	0,915	(33,885 %)
GED	17	1,309	(25,257 %)	14	0,064	(80,044 %)
	13	5,332	(2,093 %)	11	2,534	(11,139 %)
EWMA(75,0.94)	77	2,771	(9,602 %)	55	1,836	(17,542 %)
N(0,1)	25	9,326	(0,226 %)	21	4,093	(4,306 %)
	19	16,414	(0,005 %)	17	11,644	(0,064 %)
EWMA(75,0.94)	80	4,102	(4,282 %)	60	0,482	(48,747 %)
Student– <i>t</i>	21	4,538	(3,315 %)	17	1,084	(29,775 %)
	14	6,847	(0,888 %)	11	2,534	(11,139 %)
EWMA(75,0.94)	76	2,381	(12,279 %)	55	1,836	(17,542 %)
GED	20	3,573	(5,873 %)	17	1,084	(29,775 %)
	15	8,506	(0,354 %)	13	4,974	(2,572 %)
N(0,1)–GARCH(1,1)	73	1,383	(23,955 %)	78	2,413	(12,031 %)
DHS	15	0,387	(53,401 %)	19	2,375	(12,332 %)
	11	2,783	(9,528 %)	8	0,306	(58,033 %)
<i>t</i> –GARCH(1,1)	72	1,108	(29,247 %)	78	2,413	(12,031 %)
DHS	17	1,309	(25,257 %)	20	3,183	(7,441 %)
	11	2,783	(9,528 %)	8	0,306	(58,033 %)
GED–GARCH(1,1)	72	1,108	(29,247 %)	75	1,419	(23,353 %)
DHS	17	1,309	(25,257 %)	20	3,183	(7,441 %)
	11	2,783	(9,528 %)	8	0,306	(58,033 %)
EWMA(75,0.94)	70	0,647	(42,125 %)	68	0,107	(74,305 %)
DHS	17	1,309	(25,257 %)	17	1,084	(29,775 %)
	8	0,391	(53,222 %)	8	0,306	(58,033 %)
Unconditional	84	6,251	(1,242 %)	110	26,811	(0 %)
POT	27	12,223	(0,047 %)	34	23,459	(0 %)
	16	10,3	(0,133 %)	24	27,721	(0 %)
Unconditional	86	7,479	(0,624 %)	113	30,239	(0 %)
HS	27	12,223	(0,047 %)	36	27,463	(0 %)
	17	12,221	(0,047 %)	22	22,641	(0 %)

Backtesting length: 1273; POT excesses: 101; Samples size: 776. Expected VaR violations 64 (5%), 13 (1%), 6 (0,5%).  
 Backtesting length: 1308; POT excesses: 103; sample size: 791. Expected VaR violations: 65 (5%), 13 (1%), 6 (0,5%).



**Table 6: VaR backtesting – Italcementi S.p.a. (IT) and Luxottica S.p.a. (LUX).**

VaR probabilities: 5%,1% and 0.5%.

VaR model	IT			LUX		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE-GARCH(1,1)	82	3,752	( 5,273 % )	78	2,137	( 14,375 % )
POT	17	1,001	( 31,699 % )	15	0,232	( 63,004 % )
	10	1,509	( 21,933 % )	7	0,023	( 88,027 % )
t-GARCH(1,1)	81	3,309	( 6,891 % )	72	0,539	( 46,269 % )
POT	18	1,569	( 21,042 % )	16	0,553	( 45,695 % )
	10	1,509	( 21,933 % )	7	0,023	( 88,027 % )
EWMA(75,0.94)	73	0,734	( 39,145 % )	69	0,132	( 71,625 % )
POT	15	0,232	( 63,004 % )	9	1,532	( 21,574 % )
	10	1,509	( 21,933 % )	5	0,431	( 51,174 % )
N(0,1)-GARCH(1,1)	80	2,892	( 8,904 % )	63	0,155	( 69,347 % )
N(0,1)	22	4,909	( 2,672 % )	20	3,035	( 8,147 % )
	16	9,575	( 0,197 % )	10	1,509	( 21,993 % )
t-GARCH(1,1)	83	4,222	( 3,99 % )	67	0,013	( 90,977 % )
Student-t	18	1,569	( 21,042 % )	9	1,532	( 21,574 % )
	9	0,78	( 37,718 % )	5	0,431	( 51,174 % )
GED-GARCH(1,1)	79	2,501	( 11,377 % )	60	0,611	( 43,453 % )
GED	16	0,553	( 45,695 % )	10	0,865	( 35,234 % )
	9	0,78	( 37,718 % )	6	0,058	( 80,905 % )
EWMA(75,0.94)	82	3,752	( 5,273 % )	66	0,001	( 98,993 % )
N(0,1)	18	1,569	( 21,042 % )	21	3,924	( 4,761 % )
	16	9,575	( 0,197 % )	13	4,837	( 2,786 % )
EWMA(75,0.94)	91	8,881	( 0,288 % )	77	1,801	( 17,961 % )
Student-t	16	0,553	( 45,695 % )	13	0,004	( 95,138 % )
	8	0,275	( 59,987 % )	5	0,431	( 51,174 % )
EWMA(75,0.94)	81	3,309	( 6,891 % )	65	0,019	( 88,931 % )
GED(v)	15	0,232	( 63,004 % )	12	0,117	( 73,192 % )
	8	0,275	( 59,987 % )	5	0,431	( 51,174 % )
N(0,1)-GARCH(1,1)	77	1,801	( 17,96 % )	81	3,309	( 6,891 % )
DHS	20	3,035	( 8,147 % )	18	1,569	( 21,042 % )
	10	1,509	( 21,933 % )	10	1,509	( 21,933 % )
t-GARCH(1,1)	79	2,501	( 11,377 % )	77	1,801	( 17,961 % )
DHS	19	2,248	( 13,375 % )	19	2,248	( 13,375 % )
	11	2,44	( 11,831 % )	9	0,781	( 37,718 % )
GED-GARCH(1,1)	79	2,501	( 11,377 % )	77	1,801	( 17,961 % )
DHS	18	1,569	( 21,042 % )	19	2,248	( 13,375 % )
	11	2,44	( 11,831 % )	10	1,509	( 21,993 % )
EWMA(75,0.94)	68	0,057	( 81,134 % )	68	0,057	( 81,134 % )
DHS	15	0,232	( 63,004 % )	13	0,004	( 95,138 % )
	10	1,509	( 21,933 % )	7	0,023	( 88,027 % )
Unconditional	105	20,604	( 0,001 % )	98	14,202	( 0,016 % )
POT	33	21,116	( 0 % )	23	5,986	( 1,442 % )
	20	17,643	( 0,003 % )	13	4,837	( 2,786 % )
Unconditional	108	23,662	( 0 % )	102	17,732	( 0,003 % )
HS	33	21,116	( 0 % )	28	12,634	( 0,038 % )
	21	19,927	( 0,001 % )	16	9,575	( 0,197 % )

Backtesting length: 1322; POT excesses: 108, Samples size: 825. Expected VaR violations: 66 (5%), 13 (1%), 7 (0,5%).  
 Backtesting length: 1322; POT excesses: 108, Samples size: 825. Expected violations: 66 (5%), 13 (1%), 7 (0,5%).

**Table 7: VaR backtesting – Citigroup Inc. (C) and Vodafone Plc (VOD)**

*VaR probabilities: 5%, 1% and 0.5%.*

VaR model	C			VOD		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE–GARCH(1,1)	80	3,213	( 7,306 % )	85	5,673	( 1,723 % )
POT	21	4,093	( 4,306 % )	15	0,272	( 60,211 % )
	12	3,672	( 5,539 % )	7	0,032	( 85,851 % )
<i>t</i> –GARCH(1,1)	86	6,241	( 1,248 % )	83	4,611	( 3,176 % )
POT	18	1,673	( 19,585 % )	16	0,615	( 43,301 % )
	10	1,582	( 20,845 % )	7	0,032	( 85,851 % )
EWMA(75,0.94)	81	3,653	( 5,597 % )	77	2,054	( 15,178 % )
POT	15	0,272	( 60,211 % )	12	0,093	( 76,085 % )
	8	0,306	( 58,033 % )	7	0,032	( 85,851 % )
N(0,1)–GARCH(1,1)	74	1,144	( 28,481 % )	71	0,492	( 48,319 % )
N(0,1)	32	19,695	( 0,001 % )	22	5,101	( 2,393 % )
	21	20,328	( 0,001 % )	15	8,039	( 0,458 % )
<i>t</i> –GARCH(1,1)	80	3,213	( 7,306 % )	71	0,492	( 48,319 % )
Student– <i>t</i>	18	1,673	( 19,585 % )	16	0,615	( 43,301 % )
	13	4,974	( 2,572 % )	8	0,306	( 58,033 % )
GED–GARCH(1,1)	72	0,681	( 40,964 % )	69	0,205	( 65,066 % )
GED	20	3,183	( 7,441 % )	15	0,272	( 60,211 % )
	14	6,434	( 1,119 % )	9	0,832	( 36,174 % )
EWMA(75,0.94)	78	2,413	( 12,031 % )	68	0,107	( 74,305 % )
N(0,1)	26	10,014	( 0,155 % )	19	2,375	( 12,332 % )
	19	15,727	( 0,007 % )	13	4,974	( 2,572 % )
EWMA(75,0.94)	93	10,904	( 0,096 % )	79	2,8	( 9,429 % )
Student– <i>t</i>	21	4,093	( 4,306 % )	14	0,064	( 80,041 % )
	11	2,534	( 11,139 % )	7	0,032	( 85,851 % )
EWMA(75,0.94)	78	2,413	( 12,031 % )	67	0,041	( 83,975 % )
GED	21	4,093	( 4,306 % )	13	0	( 98,225 % )
	12	3,671	( 5,539 % )	7	0,032	( 85,851 % )
N(0,1)–GARCH(1,1)	80	3,213	( 7,306 % )	85	5,673	( 1,723 % )
DHS	21	4,093	( 4,306 % )	20	3,183	( 7,441 % )
	13	4,974	( 2,572 % )	11	2,534	( 11,139 % )
<i>t</i> –GARCH(1,1)	85	5,673	( 1,723 % )	81	3,653	( 5,597 % )
DHS	19	2,375	( 12,332 % )	20	3,183	( 7,441 % )
	13	4,974	( 2,572 % )	11	2,534	( 11,139 % )
GED–GARCH(1,1)	83	4,611	( 3,176 % )	86	6,241	( 1,248 % )
DHS	19	2,375	( 12,332 % )	21	4,093	( 4,306 % )
	14	6,434	( 1,119 % )	11	2,534	( 11,139 % )
EWMA(75,0.94)	81	3,653	( 5,597 % )	74	1,144	( 28,481 % )
DHS	20	3,183	( 7,441 % )	15	0,272	( 60,211 % )
	10	1,582	( 20,845 % )	7	0,032	( 85,851 % )
Unconditional	142	71,809	( 0 % )	96	13,254	( 0,027 % )
POT	60	90,672	( 0 % )	20	3,183	( 7,441 % )
	34	57,756	( 0 % )	13	4,974	( 2,572 % )
Unconditional	146	78,646	( 0 % )	94	11,665	( 0,064 % )
HS	56	78,477	( 0 % )	26	10,014	( 0,155 % )
	35	61,126	( 0 % )	13	4,974	( 2,572 % )

*Backtesting length: 1308; POT excesses: 103; Sample size: 792; Expected VaR violations: 65 (5%), 13 (1%), 6 (0,5%).*

*Backtesting length: 1308; POT excesses: 103; Samples size: 792. Expected VaR violations: 65 (5%), 13 (1%), 6 (0,5%).*

**Table 8: VaR backtesting – Ford Motor Company (F) and General Motors (GM).**

VaR probabilities: 5%,1% and 0.5%.

VaR model	F			GM		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE–GARCH(1,1)	67	0,041	( 83,975 % )	84	5,129	( 2,353 % )
POT	18	1,673	( 19,585 % )	16	0,615	( 43,301 % )
	12	3,671	( 5,539 % )	8	0,306	( 58,033 % )
<i>t</i> –GARCH(1,1)	72	0,681	( 40,964 % )	84	5,129	( 2,353 % )
POT	18	1,673	( 19,585 % )	13	0,01	( 98,225 % )
	12	3,671	( 5,539 % )	8	0,306	( 58,033 % )
EWMA(75,0.94)	67	0,041	( 83,975 % )	81	3,653	( 5,597 % )
POT	13	0,001	( 98,225 % )	13	0,01	( 98,225 % )
	10	1,582	( 20,845 % )	6	0,046	( 83,001 % )
N(0,1)–GARCH(1,1)	57	1,185	( 27,635 % )	75	1,419	( 23,353 % )
N(0,1)	20	3,183	( 7,441 % )	28	12,995	( 0,032 % )
	17	11,644	( 0,064 % )	18	13,629	( 0,022 % )
<i>t</i> –GARCH(1,1)	69	0,205	( 65,066 % )	83	4,611	( 3,176 % )
Student– <i>t</i>	17	1,084	( 29,775 % )	12	0,093	( 76,085 % )
	9	0,832	( 36,174 % )	7	0,032	( 85,851 % )
GED–GARCH(1,1)	60	0,482	( 48,747 % )	76	1,723	( 18,934 % )
GED	18	1,673	( 19,585 % )	15	0,272	( 60,211 % )
	11	2,534	( 11,139 % )	9	0,832	( 36,174 % )
EWMA(75,0.94)	72	0,681	( 40,964 % )	81	3,653	( 5,597 % )
N(0,1)	29	14,537	( 0,014 % )	29	14,537	( 0,014 % )
	16	9,778	( 0,177 % )	24	27,721	( 0 % )
EWMA(75,0.94)	78	2,413	( 12,031 % )	91	9,425	( 0,211 % )
Student– <i>t</i>	15	0,272	( 60,211 % )	22	5,1	( 2,393 % )
	11	2,534	( 11,139 % )	8	0,306	( 58,033 % )
EWMA(75,0.94)	72	0,681	( 40,964 % )	80	3,213	( 7,306 % )
GED	16	0,615	( 43,301 % )	20	3,183	( 7,441 % )
	13	4,974	( 2,572 % )	8	0,306	( 58,033 % )
N(0,1)–GARCH(1,1)	67	0,041	( 83,975 % )	84	5,129	( 2,353 % )
DHS	20	3,183	( 7,441 % )	26	10,014	( 0,155 % )
	12	3,671	( 5,539 % )	14	6,434	( 1,119 % )
<i>t</i> –GARCH(1,1)	73	0,897	( 34,348 % )	83	4,611	( 3,176 % )
DHS	19	2,375	( 12,332 % )	21	4,093	( 4,306 % )
	12	3,671	( 5,539 % )	9	0,832	( 36,174 % )
GED–GARCH(1,1)	72	0,681	( 40,964 % )	84	5,129	( 2,353 % )
DHS	19	2,375	( 12,332 % )	23	6,199	( 1,278 % )
	12	3,671	( 5,539 % )	10	1,582	( 20,845 % )
EWMA(75,0.94)	68	0,107	( 74,305 % )	80	3,213	( 7,306 % )
DHS	19	2,375	( 12,332 % )	15	0,272	( 60,211 % )
	10	1,582	( 20,845 % )	9	0,832	( 36,174 % )
Unconditional	100	16,702	( 0,004 % )	130	52,841	( 0 % )
POT	26	10,014	( 0,155 % )	38	31,697	( 0 % )
	14	6,434	( 1,119 % )	19	15,727	( 0,007 % )
Unconditional	102	18,557	( 0,002 % )	132	55,837	( 0 % )
HS	33	21,546	( 0 % )	44	45,658	( 0 % )
	17	11,644	( 0,064 % )	23	25,137	( 0 % )

Backtesting length: 1308; POT excesses: 103; Samples size: 792 Expected violations: 65 (5%), 13 (1%), 6 (0,5%).  
 Backtesting length: 1308; POT excesses: 103; Samples size: 792. Expected violations: 65 (5%), 13 (1%), 6 (0,5%).

**Table 9: VaR backtesting – Microsoft Corp. (MSFT) and Unicredit Group (UCG).**  
 VaR probabilities: 5%,1% and 0.5%.

VaR model	MSFT			UCG		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE–GARCH(1,1)	74	1,144	( 28,481 % )	81	2,984	( 8,407 % )
POT	19	2,375	( 12,332 % )	12	0,145	( 70,356 % )
	8	0,306	( 58,033 % )	6	0,072	( 78,838 % )
<i>t</i> –GARCH(1,1)	77	2,054	( 15,178 % )	79	2,221	( 13,612 % )
POT	15	0,272	( 60,211 % )	11	0,448	( 50,325 % )
	7	0,032	( 85,851 % )	5	0,465	( 49,515 % )
EWMA(75,0.94)	72	0,681	( 40,964 % )	80	2,591	( 10,757 % )
POT	15	0,272	( 60,211 % )	13	0,010	( 92,079 % )
	6	0,046	( 83,001 % )	5	0,465	( 49,515 % )
N(0,1)–GARCH(1,1)	67	0,041	( 83,975 % )	73	0,589	( 44,291 % )
N(0,1)	22	5,1	( 2,393 % )	20	2,892	( 8,902 % )
	17	11,644	( 0,064 % )	12	3,44	( 6,363 % )
<i>t</i> –GARCH(1,1)	61	0,318	( 57,254 % )	76	1,279	( 25,821 % )
Student– <i>t</i>	14	0,064	( 80,044 % )	12	0,145	( 70,356 % )
	6	0,046	( 83,001 % )	5	0,465	( 49,515 % )
GED–GARCH(1,1)	58	0,915	( 33,885 % )	74	0,791	( 37,395 % )
GED	12	0,093	( 76,085 % )	10	0,935	( 33,359 % )
	8	0,306	( 58,033 % )	5	0,465	( 49,515 % )
EWMA(75,0.94)	51	3,599	( 5,781 % )	80	2,591	( 10,757 % )
N(0,1)	21	4,093	( 4,306 % )	26	9,464	( 0,209 % )
	14	6,434	( 1,119 % )	16	9,377	( 0,221 % )
EWMA(75,0.94)	69	0,205	( 65,066 % )	83	3,852	( 4,967 % )
Student– <i>t</i>	14	0,064	( 80,044 % )	16	0,496	( 48,143 % )
	8	0,306	( 58,033 % )	7	0,015	( 90,198 % )
EWMA(75,0.94)	52	3,099	( 7,834 % )	80	2,591	( 10,757 % )
GED	14	0,064	( 80,044 % )	15	0,196	( 65,831 % )
	11	2,534	( 11,139 % )	7	0,015	( 90,198 % )
N(0,1)–GARCH(1,1)	98	14,934	( 0,011 % )	79	2,221	( 13,612 % )
DHS	26	10,014	( 0,155 % )	15	0,196	( 65,831 % )
	13	4,974	( 2,572 % )	10	1,438	( 23,052 % )
<i>t</i> –GARCH(1,1)	75	1,419	( 23,353 % )	81	2,984	( 8,407 % )
DHS	19	2,375	( 12,332 % )	13	0,010	( 92,079 % )
	9	0,832	( 36,174 % )	9	0,731	( 39,289 % )
GED–GARCH(1,1)	77	2,054	( 15,178 % )	81	2,984	( 8,407 % )
DHS	21	4,093	( 4,306 % )	13	0,010	( 92,079 % )
	10	1,582	( 20,845 % )	10	1,438	( 23,052 % )
EWMA(75,0.94)	71	0,492	( 48,319 % )	78	1,881	( 17,033 % )
DHS	18	1,673	( 19,585 % )	16	0,496	( 48,143 % )
	8	0,306	( 58,033 % )	8	0,246	( 61,956 % )
Unconditional	103	19,516	( 0,001 % )	129	48,492	( 0 % )
POT	28	12,955	( 0,032 % )	41	37,249	( 0 % )
	12	3,671	( 5,539 % )	23	24,434	( 0 % )
Unconditional	103	19,516	( 0,001 % )	135	57,297	( 0 % )
HS	33	21,546	( 0 % )	43	41,918	( 0 % )
	17	11,644	( 0,064 % )	24	26,976	( 0 % )

Backtesting length: 1308; POT excesses: 103; Samples size: 792. Expected VaR violations: 65 (5%), 13 (1%), 6 (0,5%).  
 Backtesting length: 1336; POT excesses: 108; Samples size: 825. Expected VaR violations: 67 (5%), 13 (1%), 7 (0,5%).

**Table 10: VaR backtesting – GBP British Pound (in €) and ZAR South African Rand (in €).**

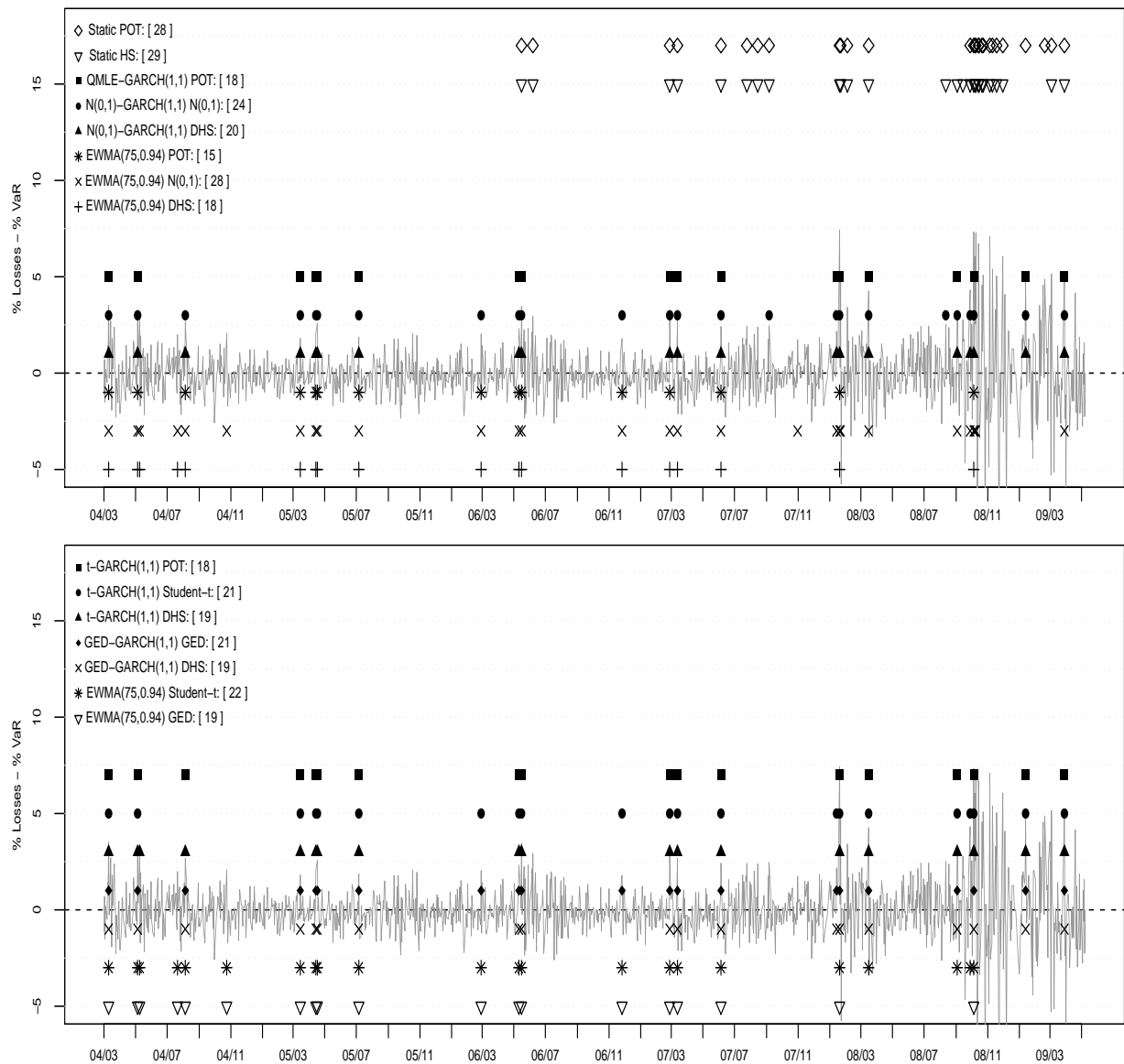
VaR probabilities: 5%,1% and 0.5%.

VaR model	GBP			ZAR		
	Violations	LRT <sub>unc</sub>	p-value	Violations	LRT <sub>unc</sub>	p-value
QMLE–GARCH(1,1)	70	0,196	( 65,768 % )	63	0,192	( 66,149 % )
POT	13	0,006	( 93,605 % )	12	0,131	( 71,767 % )
	7	0,019	( 89,114 % )	7	0,019	( 89,114 % )
<i>t</i> –GARCH(1,1)	74	0,872	( 35,032 % )	68	0,038	( 84,588 % )
POT	12	0,131	( 71,767 % )	12	0,131	( 71,767 % )
	7	0,019	( 89,114 % )	8	0,261	( 60,97 % )
EWMA(75,0.94)	78	2,006	( 15,664 % )	63	0,192	( 66,149 % )
POT	11	0,423	( 51,524 % )	15	0,213	( 64,414 % )
	6	0,065	( 79,868 % )	6	0,065	( 79,868 % )
N(0,1)–GARCH(1,1)	74	0,872	( 35,032 % )	80	2,738	( 9,797 % )
N(0,1)	22	4,815	( 2,821 % )	24	7,037	( 0,798 % )
	13	4,769	( 2,898 % )	19	15,328	( 0,009 % )
<i>t</i> –GARCH(1,1)	79	2,359	( 12,457 % )	86	5,563	( 1,835 % )
Student– <i>t</i>	15	0,213	( 64,414 % )	18	1,518	( 21,795 % )
	11	2,393	( 12,188 % )	11	2,393	( 12,188 % )
GED–GARCH(1,1)	74	0,872	( 35,032 % )	79	2,359	( 12,457 % )
GED	16	0,524	( 46,912 % )	20	2,963	( 8,518 % )
	11	2,393	( 12,188 % )	11	2,393	( 12,188 % )
EWMA(75,0.94)	84	4,519	( 3,353 % )	87	6,122	( 1,335 % )
N(0,1)	27	11	( 0,091 % )	30	15,644	( 0,008 % )
	15	7,768	( 0,532 % )	22	22,145	( 0 % )
EWMA(75,0.94)	90	7,947	( 0,482 % )	96	12,234	( 0,047 % )
Student– <i>t</i>	19	2,187	( 13,918 % )	24	7,037	( 0,798 % )
	9	0,755	( 38,5 % )	15	7,768	( 0,532 % )
EWMA(75,0.94)	84	4,519	( 3,353 % )	87	6,122	( 1,335 % )
GED	19	2,187	( 13,918 % )	24	7,037	( 0,798 % )
	11	2,393	( 12,188 % )	15	7,768	( 0,532 % )
N(0,1)–GARCH(1,1)	67	0,005	( 94,488 % )	67	0,005	( 94,488 % )
DHS	15	0,213	( 64,414 % )	13	0,006	( 93,605 % )
	10	1,473	( 22,488 % )	10	1,473	( 22,488 % )
<i>t</i> –GARCH(1,1)	69	0,102	( 74,971 % )	68	0,038	( 84,588 % )
DHS	16	0,524	( 46,912 % )	12	0,131	( 71,767 % )
	10	1,473	( 22,488 % )	9	0,755	( 38,5 % )
GED–GARCH(1,1)	67	0,005	( 94,488 % )	68	0,038	( 84,588 % )
DHS	15	0,213	( 64,414 % )	12	0,131	( 71,767 % )
	10	1,473	( 22,488 % )	10	1,473	( 22,488 % )
EWMA(75,0.94)	73	0,659	( 41,675 % )	67	0,005	( 94,488 % )
DHS	14	0,038	( 84,614 % )	15	0,213	( 64,414 % )
	9	0,755	( 38,5 % )	9	0,755	( 38,5 % )
Unconditional	103	18,254	( 0,002 % )	70	0,196	( 32,751 % )
POT	34	22,783	( 0 % )	14	0,038	( 84,614 % )
	17	11,309	( 0,077 % )	9	0,755	( 38,5 % )
Unconditional	104	19,2	( 0,001 % )	68	0,038	( 84,588 % )
HS	35	24,724	( 0 % )	14	0,038	( 84,614 % )
	22	22,145	( 0 % )	8	0,261	( 60,97 % )

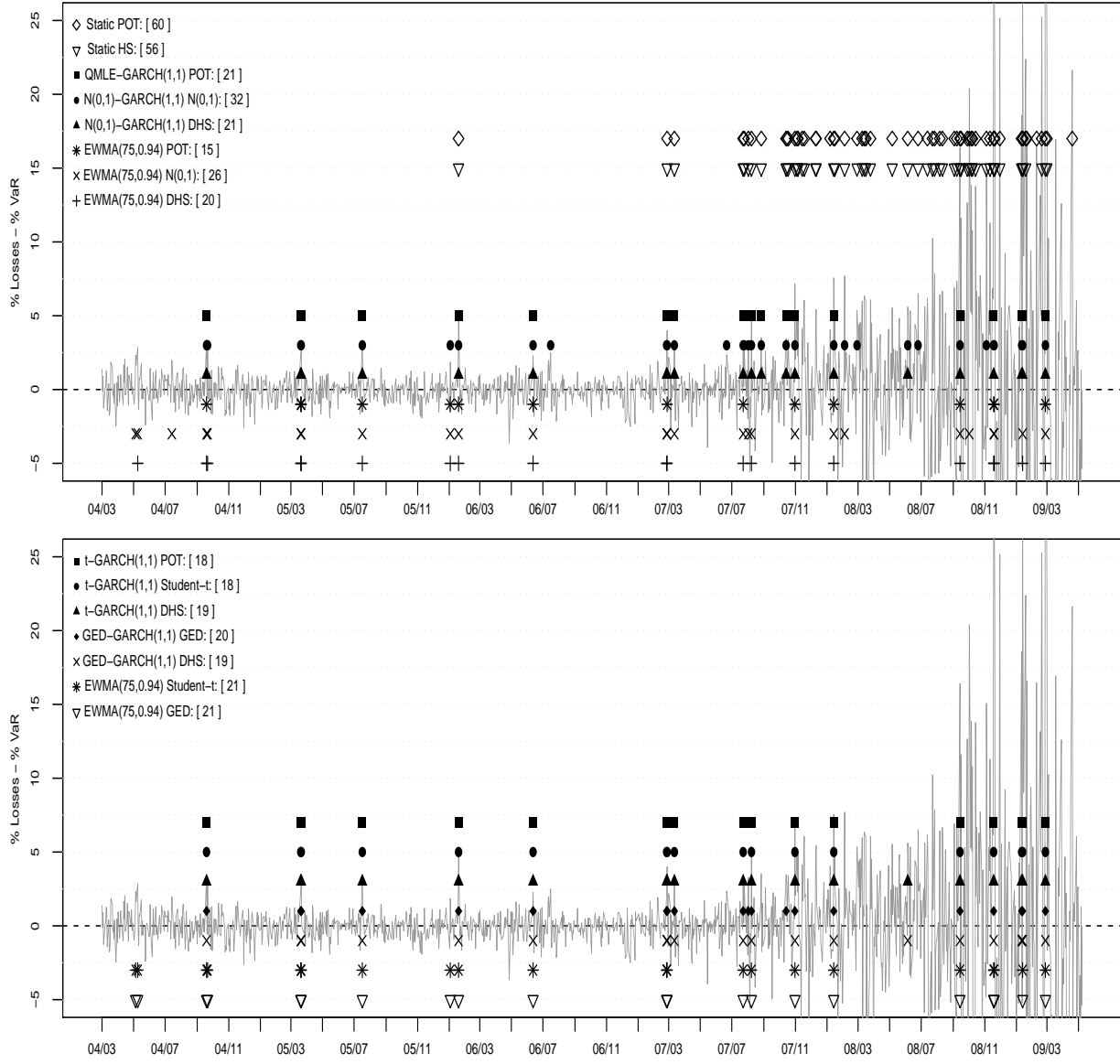
Backtesting length: 1329; POT excesses: 105; Samples size: 806. Expected VaR violations: 66 (5%), 13 (1%), 7 (0,5%).  
Backtesting length: 1329; POT excesses: 105; Samples size: 806. Expected violations: 66 (5%), 13 (1%), 7 (0,5%).

## A.2 Figures

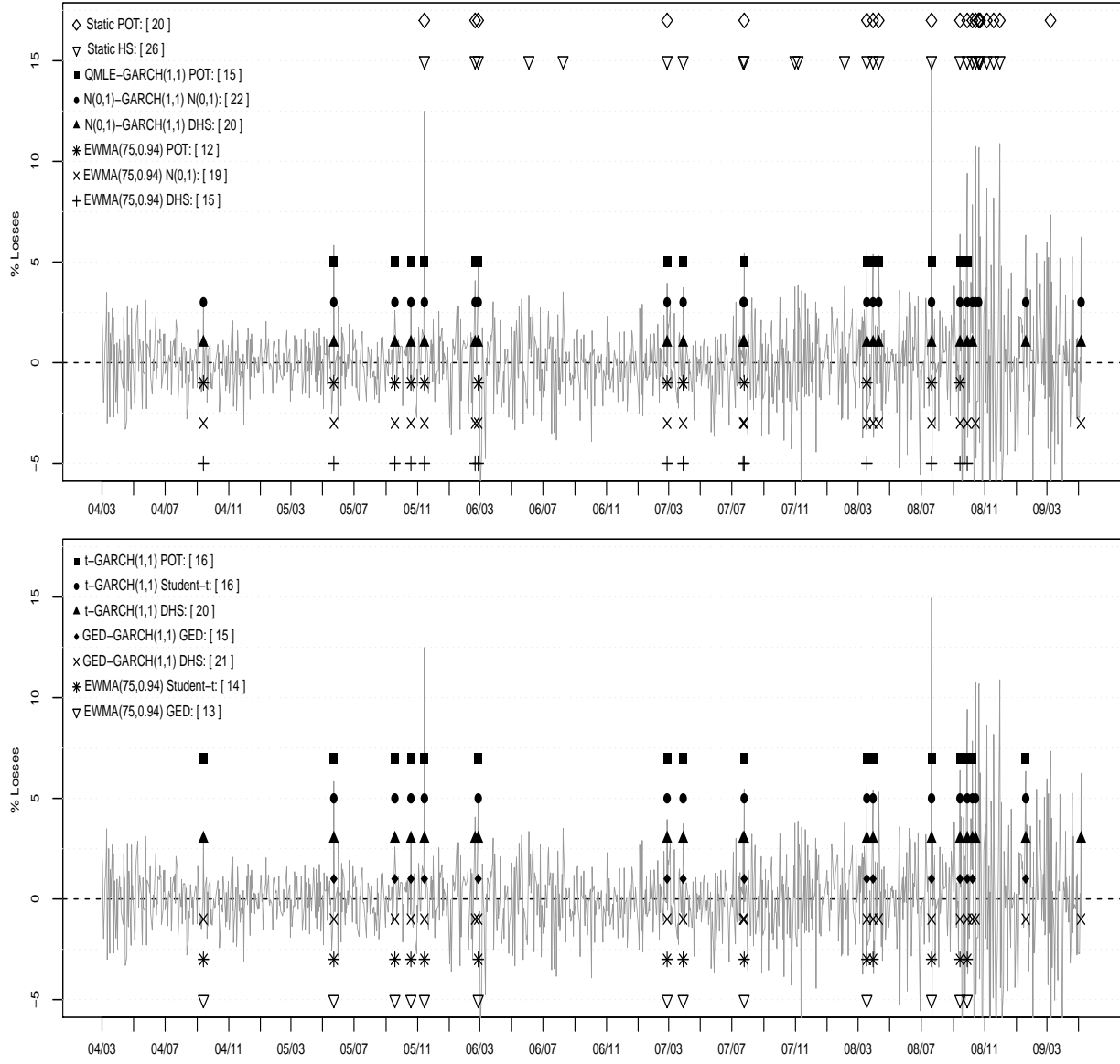
**Figure 3: DAX30 Index. VaR 99% backtesting from March, 2 2004 to May, 9 2009. Each symbol represents a VaR violation for the VaR models listed in the legend.**



**Figure 4: Citigroup Corp. VaR 99% backtesting from March, 2 2004 to May, 9 2009.** Each symbol represents a VaR violation for the VaR models listed in the legend.

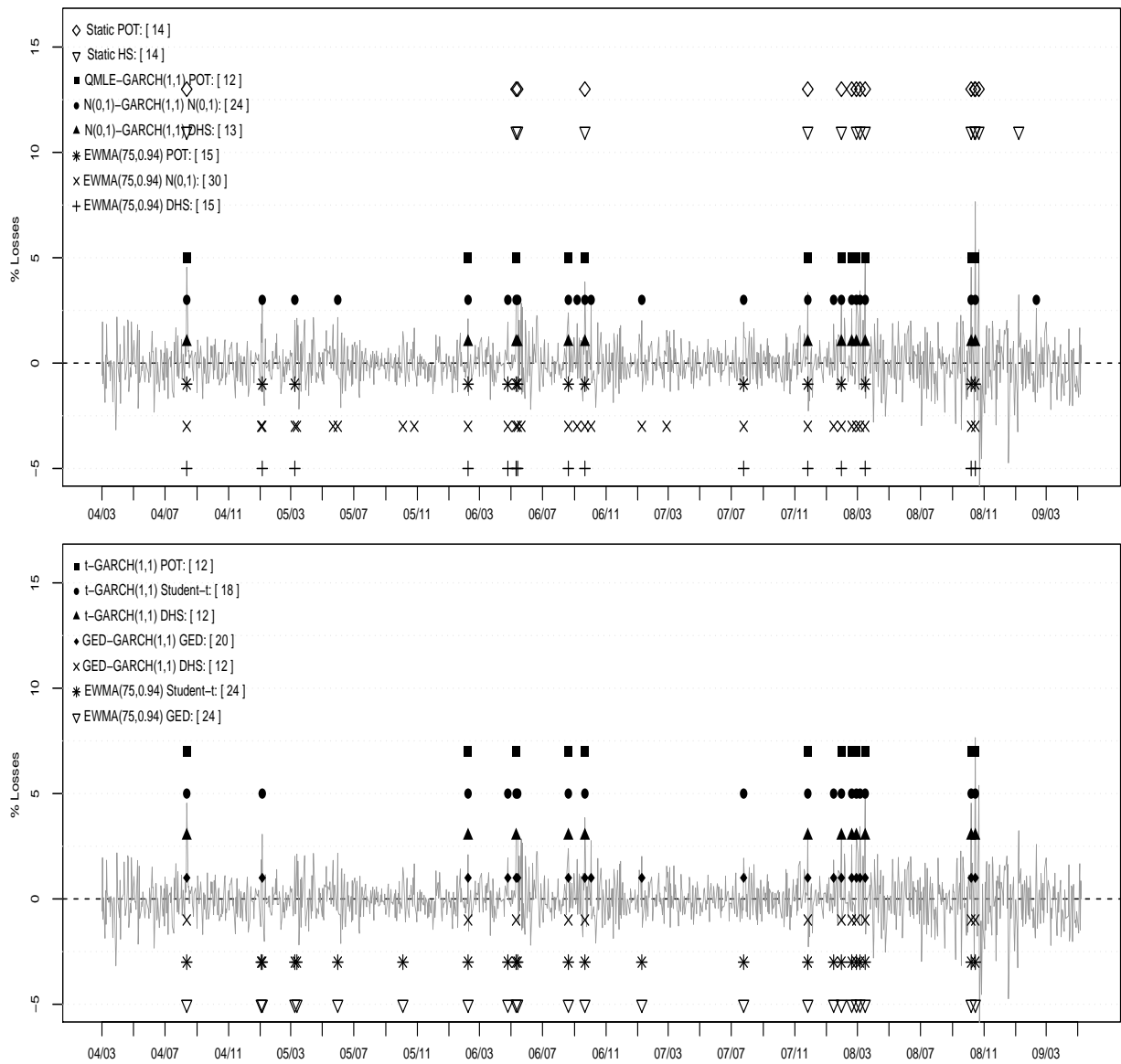


**Figure 5: Vodafone Plc. VaR 99% backtesting from March, 2 2004 to May, 9 2009. Each symbol represents a VaR violation for the VaR models listed in the legend.**





**Figure 6: ZAR Rand (in €). VaR 99% backtesting from March, 2 2004 to May, 9 2009. Each symbol represents a VaR violation for the VaR models listed in the legend.**





Elenco dei papers del Dipartimento di Economia  
pubblicati negli ultimi due anni

2008.1 *A Monte Carlo EM Algorithm for the Estimation of a Logistic Autologistic Model with Missing Data*, by Marco Bee and Giuseppe Espa.

2008.2 *Adaptive microfoundations for emergent macroeconomics*, Edoardo Gaffeo, Domenico Delli Gatti, Saul Desiderio, Mauro Gallegati.

2008.3 *A look at the relationship between industrial dynamics and aggregate fluctuations*, Domenico Delli Gatti, Edoardo Gaffeo, Mauro Gallegati.

2008.4 *Demand Distribution Dynamics in Creative Industries: the Market for Books in Italy*, Edoardo Gaffeo, Antonello E. Scorcu, Laura Vici.

2008.5 *On the mean/variance relationship of the firm size distribution: evidence and some theory*, Edoardo Gaffeo, Corrado di Guilmi, Mauro Gallegati, Alberto Russo.

2008.6 *Uncomputability and Undecidability in Economic Theory*, K. Vela Velupillai.

2008.7 *The Mathematization of Macroeconomics: A Recursive Revolution*, K. Vela Velupillai.

2008.8 *Natural disturbances and natural hazards in mountain forests: a framework for the economic valuation*, Sandra Notaro, Alessandro Paletto

2008.9 *Does forest damage have an economic impact? A case study from the Italian Alps*, Sandra Notaro, Alessandro Paletto, Roberta Raffaelli.

2008.10 *Compliance by believing: an experimental exploration on social norms and impartial agreements*, Marco Faillo, Stefania Ottone, Lorenzo Sacconi.

2008.11 *You Won the Battle. What about the War? A Model of Competition between Proprietary and Open Source Software*, Riccardo Leoncini, Francesco Rentocchini, Giuseppe Vittucci Marzetti.

2008.12 *Minsky's Upward Instability: the Not-Too-Keynesian Optimism of a Financial Cassandra*, Elisabetta De Antoni.

2008.13 *A theoretical analysis of the relationship between social capital and corporate social responsibility: concepts and definitions*, Lorenzo Sacconi, Giacomo Degli Antoni.

2008.14 *Conformity, Reciprocity and the Sense of Justice. How Social Contract-based Preferences and Beliefs Explain Norm Compliance: the Experimental Evidence*, Lorenzo Sacconi, Marco Faillo.

- 2008.15 *The macroeconomics of imperfect capital markets. Whither saving-investment imbalances?* Roberto Tamborini
- 2008.16 *Financial Constraints and Firm Export Behavior*, Flora Bellone, Patrick Musso, Lionel Nesta and Stefano Schiavo
- 2008.17 *Why do firms invest abroad? An analysis of the motives underlying Foreign Direct Investments* *Financial Constraints and Firm Export Behavior*, Chiara Franco, Francesco Rentocchini and Giuseppe Vittucci Marzetti
- 2008.18 *CSR as Contractarian Model of Multi-Stakeholder Corporate Governance and the Game-Theory of its Implementation*, Lorenzo Sacconi
- 2008.19 *Managing Agricultural Price Risk in Developing Countries*, Julie Dana and Christopher L. Gilbert
- 2008.20 *Commodity Speculation and Commodity Investment*, Christopher L. Gilbert
- 2008.21 *Inspection games with long-run inspectors*, Luciano Andreozzi
- 2008.22 *Property Rights and Investments: An Evolutionary Approach*, Luciano Andreozzi
- 2008.23 *How to Understand High Food Price*, Christopher L. Gilbert
- 2008.24 *Trade-imbalances networks and exchange rate adjustments: The paradox of a new Plaza*, Andrea Fracasso and Stefano Schiavo
- 2008.25 *The process of Convergence Towards the Euro for the VISEGRAD – 4 Countries*, Giuliana Passamani
- 2009.1 *Income Shocks, Coping Strategies, and Consumption Smoothing. An Application to Indonesian Data*, Gabriella Berloffia and Francesca Modena
- 2009.2 *Clusters of firms in space and time*, Giuseppe Arbia, Giuseppe Espa, Diego Giuliani e Andrea Mazzitelli
- 2009.3 *A note on maximum likelihood estimation of a Pareto mixture*, Marco Bee, Roberto Benedetti e Giuseppe Espa
- 2009.4 *Job performance and job satisfaction: an integrated survey*, Maurizio Pugno e Sara Depedri
- 2009.5 *The evolution of the Sino-American co-dependency: modeling a regime switch in a growth setting*, Luigi Bonatti e Andrea Fracasso
- 2009.6 *The Two Triangles: What did Wicksell and Keynes know about macroeconomics that modern economists do not (consider)?* Ronny Mazzocchi, Roberto Tamborini e Hans-Michael Trautwein

2009.7 *Mobility Systems and Economic Growth: a Theoretical Analysis of the Long-Term Effects of Alternative Transportation Policies*, Luigi Bonatti e Emanuele Campiglio

2009.8 *Money and finance: The heterodox views of R. Clower, A. Leijonhufvud and H. Minsky*, Elisabetta de Antoni

2009.9 *Development and economic growth: the effectiveness of traditional policies*, Gabriella Berloff, Giuseppe Folloni e Ilaria Schnyder

2009.10 *Management of hail risk: insurance or anti-hail nets?*, Luciano Pilati, Vasco Boatto

2010.1 *Monetary policy through the "credit-cost channel". Italy and Germany pre- and post-EMU*, Giuliana Passamani, Roberto Tamborini

2010.2 *Has food price volatility risen?* Christopher L. Gilbert and C. Wyn Morgan

2010.3 *Simulating copula-based distributions and estimating tail probabilities by means of Adaptive Importance Sampling*, Marco Bee

2010.4 *The China-US co-dependency and the elusive costs of growth rebalancing*, Luigi Bonatti and Andrea Fracasso

2010.5 *The Structure and Growth of International Trade*, Massimo Riccaboni and Stefano Schiavo

2010.6 *The Future of the Sino-American co-Dependency*, Luigi Bonatti and Andrea Fracasso

2010.7 *Anomalies in Economics and Finance*, Christopher L. Gilbert

2010.8 *Trust is bound to emerge (In the repeated Trust Game)*, Luciano Andreozzi

2010.9 *Dynamic VaR models and the Peaks over Threshold method for market risk measurement: an empirical investigation during a financial crisis*, Marco Bee e Fabrizio Miorelli