A DARE for VaR*

Benjamin Hamidi†, Christophe Hurlin‡, Patrick Kouontchou§ and Bertrand Maillet¶

March 2015

Abstract
This paper introduces a new class of models for the Value-at-Risk (VaR) and Expected Shortfall (ES), called the Dynamic AutoRegressive Expectiles (DARE) models. Our approach is based on a weighted average of expectile-based VaR and ES models, i.e. the Conditional Autoregressive Expectile (CARE) models introduced by Taylor (2008a) and Kuan et al. (2009). First, we briefly present the main non-parametric, parametric and semi-parametric estimation methods for VaR and ES. Secondly, we detail the DARE approach and show how the expectiles can be used to estimate quantile risk measures. Thirdly, we use various backtesting tests to compare the DARE approach to other traditional methods for computing VaR forecasts on the French stock market. Finally, we evaluate the impact of several conditional weighting functions and determine the optimal weights in order to dynamically select the more relevant global quantile model.

Résumé
Cet article introduit une nouvelle classe de modèles pour la Value-at-Risk (VaR) et l’Expected Shortfall (ES), appelés modèles Dynamic AutoRegressive Expectiles (DARE). Notre approche est fondée sur une moyenne pondérée de modèles de VaR et d’ES, eux-mêmes basés sur la notion d’expectiles, i.e. les modèles Conditional Autoregressive Expectile (CARE) introduits par (2008a) et Kuan et al. (2009). Premièrement, nous recensons brièvement les principales approches non paramétriques, paramétriques et semi paramétriques d’estimation de la VaR et de l’ES. Deuxièmement, nous détaillons l’approche DARE et montrons comment les expectiles peuvent être utilisés pour estimer ces mesures de risque. Troisièmement, nous utilisons différents tests de validation (backtesting) afin de comparer l’approche DARE à différentes méthodes alternatives de prévision de la VaR. Finalement, nous évaluons l’impact du choix des pondérations sur la qualité des prévisions et déterminons les poids optimaux dans le but de sélectionner de façon dynamique le modèle de prévision le plus adapté.

Keywords : Expected Shortfall, Value-at-Risk, Expectile, Risk Measures, Backtests.

JEL Classification : C14, C15, C50, C61, G11.

*We thank Georges Bresson, Christophe Boucher, Thierry Chauveau, Gilbert Colletaz, Gregory Jannin, Jean-Philippe Médecin and Paul Merlin for their help, suggestions and encouragements when preparing this work. We also thank the coeditor and the editor of the review (Franck Moraux), as well as the anonymous referee, for their advices. We thank the support of the Risk Foundation Chair Dauphine-ENSAE-Groupama “Behavioral and Household Finance, Individual and Collective Risk Attitudes” (Louis Bachelier Institute) and the Global Risk Institute in Financial Services (www.globalriskinstitute.com). Usual disclaimers apply.

†Neuflize OBC Investissements. Email: benjamin.hamidi@fr.abnamro.com.
‡University of Orleans (LEO, UMR CNRS 7322). Email : christophe.hurlin@univ-orleans.fr.
§Variances and University of Lorraine (CEREFIGE). Email: patrick.kouontchou@univ-lorraine.fr.
¶A.A. Advisors-QCG (ABN AMRO), Variances, Univ. Paris Dauphine (LEDa-SDFi), Department of Economics (France). Correspondance to: Pr. Bertrand B. Maillet, University Paris Dauphine, Place du Maréchal de Lattre de Tassigny F-75016 Paris (France). Email: bertrand.maillet@dauphine.fr.
1 Introduction

Value-at-Risk (VaR) and Expected Shortfall (ES) are the two standard measures of financial market risk. The VaR measures the potential loss of a given portfolio over a specified holding period at a specified coverage rate, which is commonly fixed at 1% or 5%. Formally, the VaR is defined as the negative of the quantile of the underlying return (conditional or unconditional) distribution. Although VaR became the standard measure of market risk in the Basel regulation framework, it has been criticized for disregarding outcomes beyond the quantile. In addition, VaR is not a coherent risk measure (Artzner et al. (1999)) and, in particular, this measure is not subadditive.\footnote{This property concerns the principle of risk diversification: the aggregated risk of a portfolio should not be greater than the individual risk of its constituent parts.}

On the contrary, the ES is a risk measure that precisely overcomes this weakness (Acerbi and Tasche (2002)). Furthermore, the ES, defined as the conditional expectation of the return given that it falls below the VaR, gives more information than VaR about the tail of the return distribution.

Many estimation methods have been proposed for both VaR and ES. These methods can be grouped into three main approaches: parametric (ARMA-GARCH models, RiskMetrics, etc.), non-parametric (historical simulation, kernel estimator, etc.) and semi-parametric estimation methods (Cornish-Fisher for instance). The latter include, among others, the dynamic conditional quantile models, \textit{i.e.} the so-called Conditional Autoregressive Value-at-Risk (CAViaR) models proposed by Engle and Manganelli (2004). The intuition of these models consists of specifying an autoregressive dynamic of the conditional quantile, similar to that used in GARCH models for the volatility. The parameters of the CAViaR model are then estimated using a quantile regression (Koenker and Bassett (1978)). This approach has strong appeal in that it does not rely on particular distributional assumptions.

Another stream of literature departs from the formal definition of the VaR, and rather than considering a model for the quantiles, considers a model for the expectiles. An expectile is a downside risk measure that is more tail sensitive than the VaR (Newey and Powell (1987)). Formally, the estimate of the $\tau$-th expectile, with $\tau \in [0, 1]$, is the solution to the minimization of asymmetrically weighted mean squared errors, with the weights $\tau$ and $1 - \tau$ assigned to positive and negative deviations, respectively. The $\tau$ parameter can be considered as an index of prudentiality since the $\tau$-expectile corresponds to a quantile with distinct coverage rate under different distributions. Thus, the expectile may be interpreted as a flexible quantile, in the sense that its coverage rate is not specified \textit{a priori} but is determined by the underlying return distribution. The expectile is a risk measure, and Kuan et al. (2009) propose to use it as an alternative to the VaR. However, the risk to observe a loss larger than the expectile cannot be controlled \textit{ex-ante}, since the implied
coverage rate is asset specific. As a consequence, the usual backtesting tests cannot be directly used to evaluate the expectile models.

An alternative approach consists of using the expectiles in order to estimate the traditional quantile-based VaR (Taylor (2008a) and (2008b)). Since for each expectile there is a quantile (Efron (1991), Jones (1994), Abdous and Remillard (1995), Yao and Tong (1996)), modelling the dynamics of the expectile allows us to produce VaR forecasts. Hence, Kuan et al. (2009) and Taylor (2008a) propose a class of Conditional Autoregressive Expectile (CARE) models. These models are similar to the CAViaR models, except that the quantile is replaced by the expectile and the parameters are estimated using an Asymmetric Least Squares (ALS) regression.

In this context, our paper proposes a new class of models, called Dynamic AutoRegressive Expectiles (DARE), based on a weighted average of expectile-based VaR or ES (CARE models). The intuition of the DARE is related to the averaging model literature (see Hansen (2007) for a general discussion on model averaging). Model averaging is an alternative to model selection. Rather than choosing the best specification of the CARE model, the DARE allows us to consider a set of alternative specifications and to determine the optimal weights for a given criterion. Many criteria can be used here. We propose a Mallows criterion based on the distance between the coverage rate and the empirical frequency of in-sample VaR violations. So, the weight of each constituent of the DARE model is dynamically determined at each date, given the past information set, in order to deliver the best VaR or ES forecast. The logic of the DARE model is then relatively close to that of the Dynamic Additive Quantile (DAQ) model (Gouriéroux and Jasiak (2008)) which is based on the linear combination of \( K \) path-independent quantile functions. The main difference is that we consider path-dependent expectile-based VaRs.

The forecasting performance of two DARE models (one with equal weights and one with optimal weights) is compared to the usual VaR forecasting methods (historical simulation, parametric unconditional VaRs, GARCH models, RiskMetrics, CAViaR) for the French CAC40 index from 9th July, 1987 until 18th March, 2009. The VaR forecasts are evaluated with nine backtesting tests designed to test the unconditional coverage, the independence or the conditional coverage of the VaR violations (i.e. the situations where the ex-post losses are larger than the VaR). The main conclusion is that the VaR forecasts issued from the DARE models and the CAViaR models are generally not rejected by the usual backtesting tests. One advantage of the DARE is that it does not impose the choice of a particular specification for the dynamics of the quantiles or expectiles. In that sense, our approach allows us to diversify the model risk associated to each specific model included in the DARE.
The paper is organized as follows: in Section 2, we briefly review the main literature about VaR and ES measures, and estimation methods. We also introduce the concept of expectile and the CARE models. In Section 3, we define the DARE models and optimization specification. Section 4 presents the data set, the estimation method, the backtesting tests and the empirical results. The last section concludes.

2 Expectile-based VaR and ES

Risk management has become in these past few years a central object of interest for researchers, market practitioners and regulators. It involves the computation and follow-up of some risk measures that can be viewed as single statistics of asset/portfolio returns. We shall first herein focus on (i) the definitions of the VaR and ES, the main two important market risk measures currently used in the regulation framework and (ii) the corresponding estimation methods. Then, we will introduce the notion of expectile and expectile-based VaR and ES.

2.1 VaR and ES definitions

The VaR is defined as the potential loss a portfolio may suffer from, at a given confidence level over a fixed holding period, such as at time $t$:

$$\Pr \left[ r_t < -\text{VaR}_t (\alpha) \right] = \alpha,$$

where $r_t$ is the asset return over the holding period, $\Pr$ stands for the conditional probability given the past information set available at time $t-1$, $\mathcal{F}_{t-1}$, and $\alpha \in [0,1]$ is the confidence level. The VaR can thus be defined as the negative of the $\alpha$-quantile, such as:

$$\text{VaR}_t (\alpha) = -q_t (\alpha),$$

where $q_t(\cdot)$ is the quantile function associated to the conditional distribution of the returns. Although the VaR is the risk measure imposed by regulators, some criticisms have been formulated against its generalized use (Beder (1995), Cheridito and Stadje (2009)). Another issue, pointed out by Artzner et al. (1999), concerns the “non-coherence” property of this risk measure; it fails indeed to respect the subadditivity property, i.e. the VaR of a combined portfolio can be larger than the sum of the VaRs of its components (diversification principle).

2In general, banks and financial institutions define the VaR as $\text{VaR}_t (1-\alpha) = -q_t (\alpha)$. For instance the 95%-VaR corresponds to the opposite of the 5%-quantile associated to the conditional distribution of the asset returns. In this paper, for simplicity, we will define $\alpha$-VaR as the negative of the $\alpha$-quantile.

3Similarly, the unconditional (or marginal) VaR is defined as the negative of the $\alpha$-quantile of the marginal distribution of the returns, i.e. $q(\alpha) = -\text{VaR}(\alpha)$. Under the stationarity assumption, this VaR is time invariant.
An alternative risk measure is the Expected Shortfall (ES), also called C-VaR or tail-VaR. ES is defined as the conditional expectation of the return given that it exceeds the VaR. Formally, the ES is defined as such:

$$ES_t(\alpha) = -\mathbb{E}_{t-1}[r_t| r_t \leq -VaR_t(\alpha)] = -\frac{1}{\alpha} \int_0^\alpha q_t(p) \, dp,$$

where $\mathbb{E}_{t-1}[]$ is the conditional expectation given $\mathcal{F}_{t-1}$. Contrary to the VaR, this measure satisfies the subadditivity property mentioned above, and provides information about the magnitude of the loss when the VaR is exceeded.

In general the conditional distribution of the returns is assumed to be location-scale. Under this assumption, VaR and ES can be generically written as such:

$$VaR_t(\alpha) = -m_t - q(\alpha)\sigma_t$$

$$ES_t(\alpha) = -m_t + \alpha^{-1} f(q(\alpha)) \sigma_t,$$

where $m_t$ and $\sigma_t$ are respectively the conditional mean and standard deviation of $r_t$, $f(.)$ denotes the pdf of the marginal distribution of the standardized returns $r^*_t = (r_t - m_t)/\sigma_t$ and $q(.)$ the corresponding quantile function. Note that if the standardized returns are assumed to be i.i.d., both functions $f(.)$ and $q(.)$ are time-invariant.

### 2.2 Estimation Methods

Several approaches are used to estimate VaR and ES. These approaches are usually classified into three categories (see Engle and Manganelli (1999), Jorion (2006), Nieto and Ruiz (2008)): (i) non-parametric, (ii) parametric and (iii) semi-parametric methods.

The most widely used non-parametric method is the so-called historical simulation (HS) approach. HS requires no distributional assumptions and consists of estimating the VaR with the empirical quantile of past returns and a moving window of observations. The main difficulty with HS concerns the choice of a window’s width: choosing too few observations will lead to an important estimation error, whereas too many observations will slow down the reaction of estimates to changes in the true distribution of financial returns (Beder (1995), Pritsker (2001)). Many extensions of the HS have been proposed such as the Weighted-HS (Boudoukh et al. (1998)).

On the contrary, the parametric approach assumes that returns follow a specific conditional probability distribution, as for example a Normal or a $t$-Student. The conditional mean and volatility are estimated using a parametric model (typically an ARMA-GARCH model as in Engle and Rangel

---

2For non parametric estimation methods of ES, see Scaillet (2004) and (2005), and for a sensitivity analysis Gouriéroux et al. (2000) or Fermanian and Scaillet (2005).
(2008) or Brownlees and Gallo (2010)) by maximum likelihood (ML), and the quantile \( q(\alpha) \) of the innovations is simply deduced from the conditional distribution.\(^5\) Using the \( t \)-distribution, the Normal Inverse Gaussian (NIG, Barndorff-Nielsen (1998), Venter and de Jongh (2002)), the Generalized Extreme Value (GEV), the Generalized Pareto Distribution (GPD) are alternative possibilities to increase the kurtosis of the marginal distribution of the returns.\(^6\) Note that these distributions may also be used to estimate the marginal VaR and ES.

The semi-parametric estimation methods combine the two previous approaches. One of the most often used semi-parametric approaches is the quantile regression (Engle and Manganelli (2004), Koenker and Xiao (2009)) which needs mild distributional assumptions. Engle and Manganelli propose a class of CAViaR models (symmetric absolute value, asymmetric slope, indirect GARCH(1,1) and adaptive CAViaRs) which have a structure similar to the GARCH models. Formally, these models are defined as follows:

**Symmetric Absolute Value CAViaR:**

\[
q_t(\alpha) = \beta_1 + \beta_2 q_{t-1}(\alpha) + \beta_3 |r_{t-1}|
\]

**Asymmetric Slope CAViaR:**

\[
q_t(\alpha) = \beta_1 + \beta_2 q_{t-1}(\alpha) + \beta_3 (r_{t-1})^+ + \beta_4 (r_{t-1})^-
\]

**Indirect GARCH(1,1) CAViaR:**

\[
q_t(\alpha) = [\beta_1 + \beta_2 q_{2t-1}(\alpha) + \beta_3 r_{t-1}^2]^{1/2}
\]

**Adaptive CAViaR:**

\[
q_t(\alpha) = q_{t-1}(\alpha) + \beta_1 \left\{ \frac{1}{1 + \exp\{\beta_2 \times |q_{t-1}(\alpha)|\}} \right\}^{-1} - \alpha,
\]

where \((x)^+ = \max(x, 0), (x)^- = -\min(x, 0)\) and \(\beta_1, \beta_2, \beta_3\) are parameters. The asymmetric slope model is specifically designed to model the asymmetric leverage effect, i.e. the tendency for volatility to be greater following a negative return than a positive return of equal size. The indirect GARCH(1,1) CAViaR model is correctly specified if the underlying data are generated by a GARCH(1,1) model with an i.i.d. innovation process. The adaptive specification adapts itself to past errors in order to reduce the probability that the VaR is consecutively underestimated.

CAViaR parameters are estimated by using the quantile regression minimization (QR sum thereafter) introduced by Koenker and Bassett (1978):

\[
\hat{\beta} = \arg \min_\beta \sum_t \left[ \alpha - \mathbb{I}(r_t < q_t(\alpha; \beta)) \right] \times \left[ r_t - q_t(\alpha; \beta) \right],
\]

\(^5\)For instance, if we assume that the conditional distribution of the returns is normal, then:

\[
\hat{\text{VaR}}_t(\alpha) = -\hat{\mu}_t - \Phi^{-1}(\alpha)\hat{\sigma}_t
\]

\[
\hat{\text{ES}}_t(\alpha) = -\hat{\mu}_t + \alpha^{-1}\phi(\Phi^{-1}(\alpha))\hat{\sigma}_t,
\]

where \(\phi(.)\) and \(\Phi(.)\) respectively denote the pdf and cdf of the standard normal distribution.

\(^6\)A natural extension consists of estimating the model by quasi-maximum likelihood (QML). In this case, no specific assumption is made on the distribution of the innovations, except those required by the QML. The quantile \(q(\alpha)\) of the innovations is then estimated using a non-parametric approach (empirical quantile, kernel, etc.) applied on the standardized residuals of the model.
where \( q_t(\alpha; \beta) \) is the negative of the conditional \( \alpha \)-VaR, \( \beta \) is the vector of parameters of the CAViaR model (Equations (8) to (11)), and \( I(.) \) is the indicator function. When the quantile model is linear, this minimization can be formulated as a linear program for which the dual problem is conveniently solved.\(^7\) The resulting quantile estimate \( \hat{q}_t(\alpha) = q_t(\alpha; \hat{\beta}) \) partitions the \( r_t \) observations so that the proportion less than the corresponding quantile estimate is equal to \( \alpha \).

### 2.3 From Quantiles to Expectiles

An undesirable property of the existing VaR measure is that it is insensitive to the magnitude of extreme losses (Kuan et al. (2009)). When the magnitude of loss matters, a quantile-based VaR (QVaR thereafter) may be considered too liberal or too conservative, depending on the tail shape of the underlying distribution. A solution to circumvent this drawback consists of considering an expectile-based VaR based on the expectile introduced by Newey and Powell (1987).

The expectile can be defined by analogy with the quantile. The population \( \alpha \)-quantile is the parameter \( q_t(\alpha) \) that minimizes the function \( \mathbb{E}_{t-1} \{ [\alpha - I(r_t < q_t(\alpha))] \times [r_t - q_t(\alpha)] \} \), where \( \mathbb{E}_{t-1} \) denotes the conditional expectation given \( F_{t-1} \). Similarly, the population \( \tau \)-expectile for \( \tau \in [0, 1] \) is the parameter \( \mu_t(\tau) \) that minimizes the function \( \mathbb{E}_{t-1} \{ |\tau - I(r_t < \mu_t(\tau))| \times [r_t - \mu_t(\tau)]^2 \} \), where \(|.|\) is the absolute value operator. So, the \( \tau \)-expectile is the solution to the minimization of asymmetrically weighted mean squared errors, with the weights \( \tau \) and \( 1 - \tau \) assigned to positive and negative deviations, respectively. As a consequence, the expectiles are sensitive to extreme values of the distribution.\(^8\) For instance, altering the shape of the upper tail of a distribution does not change the quantiles of the lower tail, but it has an impact on all the expectiles.

For each \( \tau \)-expectile, there is a corresponding \( \alpha \)-quantile, but it is important to note that \( \alpha \) is typically not equal to \( \tau \). Another way to express the relationship between both measures consists of expressing the \( \tau \)-expectile as a function of the \( \tau \)-quantile. Yao and Tong (1996) show that:

\[
\mu_t(\tau) = \frac{\tau \cdot q_t(\tau) - \int_{-\infty}^{q_t(\tau)} r \ dF_t(r)}{m_t - 2 \int_{-\infty}^{q_t(\tau)} r \ dF_t(r) + (1 - 2\tau) \cdot q_t(\tau)},
\]

where \( F_t(.) \) is the conditional cumulative density function of the returns \( r_t \). For instance, if this distribution is a uniform distribution over \([-a, a]\), then \( q_t(\tau) = 2\tau a - a \) and \( \mu_t(\tau) = \tau^2 / (2\tau^2 - 2\tau + 1) \).

\(^7\)The procedure proposed by Engle and Manganelli (2004) to estimate their CAViaR models is to generate vectors of parameters from a uniform random number generator between zero and one, or between minus one and zero (depending on the appropriate sign of the parameters). For each of the vectors, the QR Sum is then evaluated. The ten vectors that produced the lowest values for the function are used as initial values in a quasi-Newton algorithm. The QR Sum is calculated for each of the ten resulting vectors, and the one which produces the lowest value of the QR Sum is chosen as the final parameter vector.

\(^8\)Expectile can be distinguished from ES because the latter is determined by a conditional downside mean, which depends only on the tail event and hence is much larger (more conservative) than the corresponding expectile and quantile.
Given this relationship between expectiles and quantiles, two main approaches can be distinguished. The first approach consists of defining the expectile as a new market risk measure. Thus, Kuan et al. (2009) define an expectile-based VaR (EVaR) as the negative of the $\alpha$-expectile, i.e. $EVA_{\tau}(\alpha) = -\mu_{\tau}(\alpha)$. Their EVaR can be interpreted as a flexible quantile-based VaR for the underlying return distribution. The expectile with a given $\tau$ (for instance 1%) corresponds to different quantiles with distinct $\alpha$ values (for instance 0.8% or 1.5%) under different distributions. Thus, instead of finding the QVaR with a pre-determined $\alpha$ (typically 1% or 5%), Kuan et al. (2009) propose to identify the EVaR with a given $\tau$ and to allow the data to reveal their risk in terms of tail probability $\alpha$. However, this approach is quite different from current risk management practices. In particular, the EVaR cannot be evaluated with the traditional backtesting tests (see section 4.3), since (i) all these tests are based on empirical quantiles and (ii) the coverage rate $\alpha$ that corresponds to the $\tau$-EVaR is unknown and specific to each asset or portfolio.

On the contrary, the second approach, proposed by Taylor (2008a), is compatible with current risk management practices. As usual, the coverage rate $\alpha$ is set by the regulator or the management level, typically at 1% or 5%. Then, Taylor determines numerically the corresponding index of prudentiality $\tau$ (see section 3.1), which is asset specific. Considering a model for the dynamics of the $\tau$-expectile $\mu_{\tau}(\tau)$, similar to the CAViaR models, it is then possible to estimate $\mu_{\tau}(\tau)$ given $\mathcal{F}_{t-1}$. This estimate corresponds to the $\alpha$-VaR estimate, i.e. to the negative of the $\alpha$-quantile:

$$VaR_t(\alpha) = -\mu_{\tau}(\tau).$$

Thus, in Taylor’s approach, the dynamics of the conditional VaR are not modeled directly as in the CAViaR models, but indirectly, through the dynamics of the expectile. For this reason, this VaR is also called an expectile-based VaR.

Taylor (2008a) and Kuan et al. (2009) propose four Conditional AutoRegressive Expectile (CARE) models. These models are similar to the CAViaR models (Equations 8 to 11), except that the quantile is replaced by the expectile. For instance, the Symmetric Absolute Value CARE model is defined as:

$$\mu_{\tau}(\tau) = \beta_1 + \beta_2 \mu_{\tau-1}(\alpha) + \beta_3 |r_{t-1}|. \quad (14)$$

For each of the CARE specifications, the parameters $\beta_i$ are estimated using an Asymmetric Least Squares (ALS) regression, which is the least squares analogue of quantile regression:

$$\hat{\beta} = \arg \min_{\beta} \sum_t |\tau - I(r_t < \mu_{\tau}(\tau; \beta))| \times |r_t - \mu_{\tau}(\tau; \beta)|^2,$$  

(15)

where $\mu_{\tau}(\tau; \beta)$ denotes the $\tau$-expectile and $\beta$ is the vector of parameters of the CARE model. Once the CARE parameters are estimated, the corresponding estimated expectile is given by $\hat{\mu}_{\tau}(\tau) = \mu_{\tau}(\tau; \hat{\beta})$. 

8
Finally, it is also possible to use the expectiles in order to estimate the ES. Taylor (2008a) and (2008b) and Kuan et al. (2009) show that:

\[
ES_t(\alpha) = \left[1 + \frac{\tau}{(1-2\tau)\alpha}\right] \times \mu_t(\tau) - \frac{\tau}{(1-2\tau)\alpha} m_t,
\]

(16)

where \(m_t\) denotes the conditional mean of \(r_t\).

3 Dynamic AutoRegressive Expectiles

In this section, we propose a method, called Dynamic AutoRegressive Expectiles (DARE), that allows us to aggregate well specified expectiles models in order to obtain better forecasts of VaR and ES.

3.1 DARE models

The intuition of the DARE is related to the model averaging literature. In the specific context of the quantile models, Gouriéroux and Jasiak (2008) propose a Dynamic Additive Quantile (DAQ) model based on the linear combination of \(K\) path-independent quantile functions. The DAQ is formally defined as:

\[
q_t(\alpha) = \sum_{k=1}^{K} a_k \left(r_{t-1}; \theta_k\right) \times q_k(\alpha; \gamma_k) + a_0 \left(r_{t-1}; \theta_0\right),
\]

(17)

where \(q_k(\alpha; \gamma_k)\) are path-independent quantile functions, with \(k = \{1, ..., K\}\) and \(a_k(\cdot)\) are non-negative functions of past returns, denoted \(r_{t-1}\), and a set of parameter \(\theta_k\). The trick is that, in this dynamic quantile model, the conditional (time-variant) conditional quantile \(q_t(\alpha)\) is defined as a weighted sum of \(K\) path-independent (constant) quantile functions, where the weights are time-dependent.

The DARE model is slightly different since it corresponds to a weighted average of \(K\) CARE models, i.e. \(K\) expectile-based conditional quantiles. The prefix Dynamic in DARE is related to the use of time-variant weights that may depend on the past information set \(F_{t-1}\). Formally the DARE model for conditional quantiles is defined as follows:

Definition 1 (DARE model) The DARE model for the conditional \(\alpha\)-quantile is defined as a weighted average of \(K\) CARE models, such as:

\[
q_t(\alpha) = -VaR_t(\alpha) = \sum_{k=1}^{K} w_{k,t} \mu_k(\tau; \beta_k),
\]

(18)

where \(\mu_k(\tau; \beta_k) = -VaR_k(\alpha)\) for \(k = \{1, ..., K\}\) denotes the conditional \(\tau\)-expectile issued from the \(k^{th}\) specification of the CARE model (or equivalently the negative of the expectile-based \(\alpha\)-VaR), \(w_{k,t} \geq 0\) and \(\sum_{k=1}^{K} w_{k,t} = 1\).
A similar definition can be proposed for the ES. The DARE model for the conditional $\alpha$-ES is given by:

$$ES_t(\alpha) = \sum_{k=1}^{K} \nu_{k,t} ES_{k,t}(\alpha) = \left[1 + \frac{\tau}{(1-2\tau)\alpha}\right] \sum_{k=1}^{K} \nu_{k,t} \mu_t(\tau; \beta_k) - \frac{\tau}{(1-2\tau)\alpha} m_t,$$

where $ES_{k,t}(\alpha)$ is given by Equation (16), $\nu_{k,t} \geq 0$ and $\sum_{k=1}^{K} \nu_{k,t} = 1$. For both models (quantile and ES), we consider $K = 4$ specifications for the CARE models, namely the Symmetric Absolute Value CARE, the Asymmetric Slope CARE, Indirect GARCH(1,1) CARE and the Adaptive CARE. As in Taylor (2008a), for each of the four CARE models, we use as estimator of the $\alpha$ quantile, the $\tau$ expectile for which the proportion of in-sample observations lying below the expectile is $\alpha$. To find the optimal value of $\tau$, we estimate the CARE models for different values of $\tau$ over a grid with a step size of .0001. The final optimal value of $\tau$ was derived by linearly interpolating between grid values. We used just the first moving window of observations to optimize $\tau$. Once, the prudentiality index $\tau$ is fixed for a specific CARE model, we estimate the corresponding vector of parameters $\beta_k$ using the ALS regression (Equation (15)).

### 3.2 Optimizing Weights

Notice that in both models, the weights $\omega_{k,t}$ and $\nu_{k,t}$ are time-variant, meaning that the combination of the $K$ expectile-based quantiles may change at each date $t$, given the past information set $\mathcal{F}_{t-1}$. The issue is then to determine the optimal weights for a given criterion. Many criteria can be used in this context. As usual in the averaging model literature, we consider the forecast combination of the Mallows model averaging method (Mallows (1973), Hansen (2007) and (2008)). The goal is to obtain the set of weights that minimizes the in-sample Mean-Squared Error (MSE) over the set of feasible forecast combinations. We thus select forecast weights by minimizing a Mallows criterion, which is an asymptotically unbiased estimate of the MSE.

Formally, let $W_t = (w_{1,t}, w_{2,t}, \ldots, w_{K,t})$ be the vector of weights used in the definition of the DARE model (Equation (18)) and $VaR_t(\alpha) = -q_t(\alpha)$ the corresponding conditional $\alpha$-VaR. Notice that $VaR_t(\alpha)$, which is defined as the weighted average of $K$ CARE expectile-based VaRs, implicitly depends on $W_t$ and can be denoted as $VaR_t(\alpha; W_t)$. Let us define a binary variable associated to the ex-post $\alpha$-VaR violation at the time $t$ as:

$$I_t(\alpha; W_t) = \begin{cases} 1 & \text{if } r_t < -VaR_t(\alpha; W_t) \\ 0 & \text{otherwise.} \end{cases} \quad (21)$$

If the DARE model is well specified, then $E_{t-1}[I_t(\alpha; W_t)] = \alpha$. If we denote by $\hat{\alpha}_t(W_t) = (t-1)^{-1} \sum_{s=1}^{t-1} I_s(\alpha; W_t)$ the empirical frequency of the violations (in-sample) obtained with the
DARE model and $\hat{\alpha}_{k,t}$ the empirical frequency of the violations associated to the $k^{th}$ CARE model, the Mallows criterion (or penalized sum of squared residuals), can be expressed as:

$$C_t(W_t) = [\alpha - \hat{\alpha}_t(W_t)]^2 + 2K \sum_{k=1}^{K} w_k s_k^2,$$

(22)

where $s_k^2 = (t - K)^{-1}(\alpha - \hat{\alpha}_{k,t})^2$ is an estimate of error variance for a model $k$.

The optimal weight vector is then defined as the solution of the following minimization problem:

$$W^*_t = \arg \min_{W_t} C_t(W_t).$$

(23)

under the constraints $w_{k,t} \in [0, 1]$ and $\sum_{k=1}^{K} w_{k,t} = 1$. This optimization procedure will be used hereafter for the computation of the DARE model for VaR. Since no direct equivalent criterion can be defined for the ES, we will consider the same optimal weights for the DARE models of VaR and ES, i.e. $w^*_k = v^*_k$ for $k = \{1, \ldots, K\}$.

4 Empirical Illustration

In this section, we herein describe the empirical illustration in which DARE models are used to compute VaR and ES day-ahead forecasts. The forecasting performances of the DARE models are compared to a set of usual VaR models.

4.1 Data

The data set used in this study corresponds to the daily returns of the French CAC40 Index in Euro currency (source: Datastream) from 9th July, 1987 until 18th March, 2009. This period contains 5,659 daily returns. Figure 1 represents the CAC40 Index evolution for closing prices (top of the figure) and returns (bottom of the figure). We observe in this period typically two bull markets (1998–2000 and 2003–2007), each followed by a bear market (2000–2003 and 2007–2009) corresponding respectively to the dot-com crash and the global financial crisis. These two bear markets led to the “loss decade” of stocks with a negative return over the period 1999–2009. We note also some very quick (low duration) crashes such as in 1987, 1998 and 2001. So, our sample then includes a variety of market conditions. Regarding the returns volatility, we observe both persistence and clustering phenomena that suggest the use of dynamic models.

Table 1 provides some summary statistics on our sample. The asymmetry coefficient (skewness) is negative which means that the mass of probability on the left side of the distribution (negative returns) appears slightly larger than on the right side (positive returns). The kurtosis (9.24)
indicates that the unconditional distribution of returns is leptokurtic. These results are confirmed by the Jarque-Bera test, for which the joint null hypothesis of symmetry and kurtosis equal to 3 is rejected at the 1% level of significance.

< Insert Table 1 >

4.2 Benchmark Methods

Although there are many VaR estimation methods (Cf. section 2.2), we restrict our comparison to twelve commonly used methods. First, we consider the most widely used non-parametric method, namely the HS method. Secondly, we consider a set of parametric unconditional VaRs based on normal, Student, NIG, GEV, GPD distribution respectively. Thirdly, we use the conditional VaR forecasts deduced from two standard conditional volatility models, i.e. the GARCH(1,1) and the RiskMetrics models. Finally, we consider the four CAViaR specifications (Equations 8 to 11) and the corresponding VaRs directly deduced from the quantile regressions.

For all the parametric models, we use a moving window of four years (1,044 daily returns) to re-estimate dynamically the corresponding parameters. Forecasted VaR and ES are computed for the final 4,615 days (about 18 years). These forecasts are provided for two coverage rates $\alpha$, namely 1% and 5%, since the corresponding 95% and 99% VaR and ES are the most commonly used by financial firms and regulators.

4.3 Backtesting Tests

The VaR forecasts issued from the DARE models and the alternative benchmark models are evaluated with usual backtesting tests. Traditionally, the quality of the forecast of an economic variable is assessed by comparing its ex-post realizations with the ex-ante forecast values. The comparison of the various forecast models is thus generally made by using a criterion such as the Mean Squared Error criterion or standard information criteria (AIC and BIC). However, this approach is not suitable for VaR and ES forecasts, because the true quantile of the returns distribution is not observable. That is why VaR assessment is generally based on the concept of violations. Let $I_t(\alpha)$ be a binary variable associated with an $\alpha$%-VaR violation at time $t$:

$$I_t(\alpha) = \begin{cases} 1 & \text{if } r_t < -VaR_t(\alpha) \\ 0 & \text{otherwise} \end{cases}. \quad (24)$$

We do not propose some backtests for the ES forecasts. Indeed, the test proposed by Wong (2008) is based on a normal assumption and is not sufficiently general to be used here. One solution would have been to use the test by McNeil and Frey (2000). They examine the goodness of ES predictions by means of a bootstrap test. However, rather than using a bootstrap procedure for the ES, we prefer to test the validity of the VaR models for various coverage rates. The idea is that if the VaR model is valid for various coverage rates, there is a great chance that it is also valid for the ES.
As stressed by Christoffersen (1998), VaR forecasts are valid if, and only if, the violation process $I_t(\alpha)$ satisfies the following two hypotheses:

(i) The unconditional coverage (hereafter UC) hypothesis: the probability of an ex-post return exceeding the VaR forecast must be equal to the $\alpha$ coverage rate:

$$\Pr[I_t(\alpha) = 1] = E_{t-1} [I_t(\alpha)] = \alpha. \quad (25)$$

(ii) The independence hypothesis: VaR violations observed at two different dates for the same coverage rate must be distributed independently. Formally, the variable $I_t(\alpha)$ should be independent of the variables $I_{t-k}(\alpha), \forall k \neq 0$. In other words, past VaR violations should not be informative about current and future violations.

The UC hypothesis is quite intuitive. Indeed, if the frequency of violations is significantly lower (respectively higher) than the coverage rate $\alpha$, then risk is overestimated (respectively underestimated). However, the UC hypothesis sheds no light on the possible dependence of violations. Therefore, the independence property of violations is an essential one, because it is related to the ability of a VaR model to accurately model the higher-order dynamics of the distribution of the returns. In fact, a model that does not satisfy the independence property can lead to clusters of violations even if it has the correct average number of violations. Consequently, there must be no dependence in the violations variable, whatever the coverage rate considered. When the UC and independence hypotheses are simultaneously valid, VaR forecasts are said to have a correct conditional coverage (hereafter CC), and the VaR violation process is a martingale difference with:

$$E_{t-1} [I_t(\alpha) - \alpha] = 0. \quad (26)$$

This last property is at the core of most of the validation tests for VaR models (Christoffersen (1998), Engle and Manganelli (2004), Berkowitz et al. (2011)).

Many backtesting tests have been proposed in the literature (see Hurlin and Pérignon (2012) for a survey). In this study, we apply nine tests (see appendix A for more details) on a large sample (more than 4,600 observations) because it is well known that these tests have generally a low power in small samples.\textsuperscript{10}

1. The Exception Frequency test (Kupiec (1995)), is a Likelihood Ratio (hereafter LR) test of UC based on the empirical frequency of VaR violations.

2. The Independence test (Christoffersen (1998)) is a LR test of the null of independence based on a Markov chain assumption for the violation process.

\textsuperscript{10}For more details, see Hurlin and Tokpavi (2008). For other backtesting tests, see for instance Pérignon and Smith (2008) or Colletaz et al (2013).
3. The resampling independence test (Escanciano and Olmo (2010)) is an independence test robust to the estimation risk and based on subsampling approximation techniques.

4. The conditional test (Christoffersen (1998)) is a LR test of the null of CC based on a Markov chain assumption for the violation process.

5. The exception magnitude test (Berkowitz (2001)) is a UC test based on the differences observed between the ex-ante VaR forecasts and the ex-post returns.

6. The GMM duration based test (Candelon et al. (2011)) is a backtesting test based on the durations observed between two violations. If the VaR model is well specified, these durations must have a geometric distribution (memoryless property). This framework allows us to test each of the three hypotheses: UC, independence and CC.

7. The dynamic quantile (DQ) test (Engle and Manganelli (2004)) is based on a quantile regression model of the VaR on a set of explanatory variables that belongs to $\mathcal{F}_{t-1}$. This test exploits the martingale difference hypothesis. Under the null of CC, all the parameters of the regression model (except the constant) must be equal to zero.

4.4 Empirical Results

Figure 2 displays the 99% and 95% VaR forecasts obtained from the DARE model over the period from July, 1995 to March, 2009. For comparison, we also report the HS VaR and the normal VaR. Since the DARE-VaR is a conditional quantile, it is more volatile than the unconditional VaR measures, even if these measures are computed with a rolling window. The dynamics of the DARE-VaR reflects the great periods of the French stock market index history, and especially the dot-com crash and the global financial crisis. Similar results are obtained for the ES in Figure 3.

Table 2 reports the p-values of the nine backtesting tests for 95%-VaR and 99%-VaR forecasts issued from the DARE models and the twelve competing models. Notice that, for all these backtesting tests, the null hypothesis corresponds to a well specified VaR model. First of all, we observe that the null cannot be rejected for the DARE-model, for eight out of the nine tests considered. The only exception concerns the exception magnitude test (Berkowitz (2001)). However, this test leads to the rejection of the validity of the VaR forecasts for all the models. This result is due to the large differences observed between the ex-ante VaR forecasts and the ex-post returns for some periods, especially when the coverage rate is equal to 5% (cf. Figure 2).

---

11 An extension of the DQ test, based on a Dynamic Binary (DB) regression model, is proposed in Dumitrescu and Hurlin (2012).
The VaR forecasts issued from unconditional parametric or non-parametric approaches (HS, normal, Student, NIG, GEV and GPD) are generally non valid with respect to the UC and/or independence hypothesis. For instance, the GMM duration based tests (Candelon et al. (2011)) clearly indicate that the corresponding violations are clustered, meaning that these VaR are not sufficiently volatile to capture the return dynamics. On the contrary, the VaR based on conditional variances or conditional quantiles are more flexible. Consequently, as generally observed in the literature, they are less likely to generate violations clusters than other methods. In particular, the four CAViaR models lead to valid VaR forecasts, except when we consider the exception magnitude test.

Table 2 also reports the p-values of the backtesting tests obtained for a naïve equal-weight DARE (EW-DARE thereafter) model, i.e. a model in which all the constituent (CARE) models have an equal weight. The qualitative results are quite similar to those obtained with the DARE model with optimized weights. This result is relatively usual in the averaging model literature: the simple mean of $K$ models often produces relatively robust forecasts. So, the performance of the DARE model is mainly due to the averaging approach, and less to the choice of the weights. But, the average approach is relevant when it is applied to expectile-based VaR models. Indeed, when we consider a naïve mean of the twelve VaR models (including CAViaR models), the null of validity is rejected by some of the backtesting tests.

5 Conclusion

In this paper, we present a new class of dynamic models for VaR and ES, called Dynamic AutoRegressive Expectiles (DARE). A DARE model is defined as a weighted average of expectile-based VaR or ES models (CARE models). Using expectiles has the appeal of avoiding distributional assumptions (Taylor (2008a)). Besides, the use of a model averaging approach allows us to diversify the model risk associated to each constituent. The optimal weights of each elementary component can be determined according to a Mallows-type criterion (Hansen, 2008) in order to select dynamically the more appropriate quantile model and the best VaR or ES forecasts.

An empirical application for the CAC40 Index shows that the DARE models lead to VaR forecasts for which the null hypothesis of validity is generally not rejected by the usual backtesting tests. These forecasts are compared to the (conditional or unconditional) VaR forecasts issued from usual
models or specifications (Normal, Historical, t-Student, NIG, RiskMetrics, GARCH and CAViaR). Our results confirm the intuition that the aggregation of VaR approaches is generally more robust than the classical VaR and ES computations. In that sense, the DARE approach shows that the VaR or ES model averaging may have some interest in further financial applications, such as stress test assessment or asset pricing.
Table 1: Descriptive Statistics of the CAC40 Index Returns

<table>
<thead>
<tr>
<th>Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-.0964</td>
</tr>
<tr>
<td>Median</td>
<td>.0000</td>
</tr>
<tr>
<td>Max.</td>
<td>.1118</td>
</tr>
<tr>
<td>Mean</td>
<td>.0002</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>.0139</td>
</tr>
<tr>
<td>Skewness</td>
<td>-.0058</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>9.2413</td>
</tr>
<tr>
<td>P.value of JB</td>
<td>.0000</td>
</tr>
</tbody>
</table>

Note: DataStream; daily data from July 9th 1987 until March 18th 2009 of the CAC40 Index; computations by the authors. This period contains 5,659 daily returns.
## Table 2: Evaluation of the 95% and 99% VaR forecasts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td></td>
</tr>
<tr>
<td><strong>Daily VaR 95% Methods:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>.28***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>21.34</td>
<td>.00***</td>
<td>.00***</td>
<td>.19***</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td>64.69</td>
<td>100.00</td>
<td>99.98</td>
<td>90.04</td>
<td>.00***</td>
<td>70.02</td>
<td>.00***</td>
<td>.00***</td>
<td>3.31**</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>NIG</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>.10***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>GEV</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>6.83</td>
<td>.00***</td>
<td>.00***</td>
<td>1.45**</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>GPD</td>
<td>.00***</td>
<td>.03***</td>
<td>.01***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>42.11</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>17.40</td>
<td>.00***</td>
<td>.00***</td>
<td>48.11</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Normal GARCH(1,1)</td>
<td>53.47</td>
<td>100.00</td>
<td>99.99</td>
<td>82.47</td>
<td>.00***</td>
<td>52.43</td>
<td>.00***</td>
<td>.00***</td>
<td>98.11</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Symmetric Absolute Value</td>
<td>30.80</td>
<td>100.00</td>
<td>100.00</td>
<td>59.47</td>
<td>.00***</td>
<td>31.61</td>
<td>.00***</td>
<td>.00***</td>
<td>98.72</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Asymmetric Skew</td>
<td>.00***</td>
<td>100.00</td>
<td>100.00</td>
<td>.00***</td>
<td>.00***</td>
<td>5.89</td>
<td>.00***</td>
<td>.00***</td>
<td>96.12</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Indirect GARCH</td>
<td>37.51</td>
<td>100.00</td>
<td>100.00</td>
<td>67.48</td>
<td>.00***</td>
<td>37.90</td>
<td>.00***</td>
<td>.00***</td>
<td>98.29</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Adaptive</td>
<td>82.66</td>
<td>100.00</td>
<td>100.00</td>
<td>97.63</td>
<td>.00***</td>
<td>77.14</td>
<td>.00***</td>
<td>.00***</td>
<td>93.50</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Naive Mean of VaR</td>
<td>34.84</td>
<td>100.00</td>
<td>99.91</td>
<td>64.43</td>
<td>.00***</td>
<td>39.23</td>
<td>.00***</td>
<td>.00***</td>
<td>100.00</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>DARE</td>
<td>27.83</td>
<td>100.00</td>
<td>100.00</td>
<td>55.56</td>
<td>.00***</td>
<td>25.89</td>
<td>.00***</td>
<td>.00***</td>
<td>92.49</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>EW-DARE</td>
<td>34.94</td>
<td>100.00</td>
<td>100.00</td>
<td>63.48</td>
<td>.00***</td>
<td>34.61</td>
<td>.00***</td>
<td>.00***</td>
<td>94.22</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td><strong>Daily VaR 99% Methods:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td>.03***</td>
<td>.66***</td>
<td>16.54</td>
<td>.00***</td>
<td>.00***</td>
<td>.16***</td>
<td>.00***</td>
<td>.00***</td>
<td>99.49</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>.00***</td>
<td>.00***</td>
<td>9.68</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>94.33</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td>.01***</td>
<td>.94***</td>
<td>14.20</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>99.49</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>NIG</td>
<td>.02***</td>
<td>.12***</td>
<td>14.17</td>
<td>.00***</td>
<td>.00***</td>
<td>.15***</td>
<td>.00***</td>
<td>.00***</td>
<td>99.44</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>GEV</td>
<td>.89**</td>
<td>.01***</td>
<td>8.62*</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>88.07</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>GPD</td>
<td>15.85</td>
<td>3.46**</td>
<td>6.91*</td>
<td>3.97*</td>
<td>.00***</td>
<td>18.59</td>
<td>.00***</td>
<td>.00***</td>
<td>99.81</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>RiskMetrics</td>
<td>.01***</td>
<td>.84***</td>
<td>16.98</td>
<td>.00***</td>
<td>.00***</td>
<td>.07***</td>
<td>2.97*</td>
<td>61.88</td>
<td>99.95</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Normal GARCH(1,1)</td>
<td>.07***</td>
<td>2.62**</td>
<td>20.86</td>
<td>.03***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>.00***</td>
<td>99.99</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Symmetric Absolute Value</td>
<td>5.02*</td>
<td>5.99*</td>
<td>14.02</td>
<td>5.19*</td>
<td>.00***</td>
<td>.00***</td>
<td>8.61*</td>
<td>5.99*</td>
<td>7.37*</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>CAViaR Asymmetric Skew</td>
<td>.16***</td>
<td>2.30**</td>
<td>20.96</td>
<td>.05***</td>
<td>.00***</td>
<td>.64***</td>
<td>.02***</td>
<td>.11***</td>
<td>99.88</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Indirect GARCH</td>
<td>32.22</td>
<td>15.22</td>
<td>5.39*</td>
<td>21.98</td>
<td>.00***</td>
<td>38.85</td>
<td>67.00</td>
<td>54.53</td>
<td>99.99</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>CAViaR Adaptive</td>
<td>1.82**</td>
<td>.17***</td>
<td>18.69</td>
<td>.01***</td>
<td>.00***</td>
<td>2.92**</td>
<td>.01***</td>
<td>.01***</td>
<td>99.97</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>Naive Mean of VaR</td>
<td>1.82**</td>
<td>1.18**</td>
<td>18.17</td>
<td>.20***</td>
<td>.00***</td>
<td>3.62**</td>
<td>.00***</td>
<td>.00***</td>
<td>100.00</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>DARE</td>
<td>.43***</td>
<td>5.69*</td>
<td>14.85</td>
<td>.28***</td>
<td>.00***</td>
<td>1.53**</td>
<td>2.85*</td>
<td>9.48*</td>
<td>99.99</td>
<td>.00***</td>
<td></td>
</tr>
<tr>
<td>EW-DARE</td>
<td>25.81</td>
<td>2.81**</td>
<td>6.42*</td>
<td>4.73*</td>
<td>.00***</td>
<td>32.17</td>
<td>12.98</td>
<td>9.32*</td>
<td>99.99</td>
<td>.00***</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports the p-values (in percentage) of nine backtesting tests (see section 4.3) applied to the VaR forecasts issued from DARE models and twelve benchmark models. The “EW-DARE” model is the equally-weighted average DARE model. The “Naive Mean of VaR” is the average of all the VaR benchmark models. Significant probabilities of rejection at a significance level of 10%, 5% and 1% are respectively marked with *, ** and ***. Bold entries represent rejection of the null at the 1% significance level.
Figure 1: Daily CAC40 Index Evolution

Source: DataStream; daily data from July 9th, 1987 until March 18th, 2009 of the CAC40 Index; computations by the authors. This period contains 5,659 daily returns.
Notes: Panel A displays the negative returns of the CAC40 index (from July 9th, 1995 until March 18th, 2009), the corresponding DARE-VaR, the HS-VaR and the normal-VaR, for a coverage rate of 1%. Panel B is similar for a coverage rate of 5%.
Figure 3: 99% and 95% Expected Shortfall Forecasts

Panel A: Expected Shortfall Forecasts

Panel B: 95% Expected Shortfall Forecasts

Notes: Panel A displays the negative returns of the CAC40 index (from July 9th, 1995 until March 18th, 2009), the corresponding DARE-ES, the HS-ES and the normal-ES, for a coverage rate of 1%. Panel B is similar for a coverage rate of 5%.
A Appendix A: Backtesting Tests

A.1 Unconditional Coverage (UC) Tests

Let us consider a sequence \( \{I_t(\alpha)\}_{t=1}^T \) of \( T \) violations associated to VaR and denote by \( H \) the total number of violations, \( H = \sum_{t=1}^T I_t(\alpha) \). If we assume that the variables \( I_t(\alpha) \) are i.i.d., then under the null of UC, the total number of hits has a Binomial distribution:

\[
H \sim B(T, \alpha)
\]  

(27)

with \( E(H) = \alpha T \) and \( V(H) = \alpha(1 - \alpha)T \).

Kupiec (1995) and Christoffersen (1998) propose a Likelihood Ratio (hereafter LR) test based on the process of VaR violations \( I_t(\alpha) \). Under the null of UC, the LR statistic is defined as:

\[
LR_{UC} = -2 \ln \left[ (1 - \alpha)^{T-H} \alpha^H \right] + 2 \ln \left[ \left( 1 - \frac{H}{T} \right)^{T-H} \left( \frac{H}{T} \right)^H \right] \xrightarrow{T \to \infty} \chi^2(1).
\]

(28)

Under the null of CC, the LR statistic converges to a chi-square distribution with one degree of freedom. The null of UC is not rejected if the empirical frequency of VaR violations \( H/T \) is close enough to the coverage rate \( \alpha \).

A.2 Independence and CC tests

Christoffersen (1998) proposes a LR test based on the assumption that the process of violations \( I_t(\alpha) \) is modeled with the following matrix of transition probabilities:

\[
\Pi = \begin{pmatrix}
1 - \pi_{01} & \pi_{01} \\
1 - \pi_{11} & \pi_{11}
\end{pmatrix}
\]

(29)

where \( \pi_{ij} = \Pr[I_t(\alpha) = j \mid I_{t-1}(\alpha) = i] \), i.e., probability of being in state \( j \) at time \( t \) conditioning on being in state \( i \) at time \( t-1 \). Under the null of independence, we have \( \pi_{01} = \pi_{11} = \beta \) and:

\[
H_{0,IND} : \Pi_\beta = \begin{pmatrix}
1 - \beta & \beta \\
1 - \beta & \beta
\end{pmatrix}
\]

(30)

where \( \beta \) denotes a violation probability, which can be different from the coverage rate \( \alpha \). What these transition probabilities imply is that the probability of experiencing a violation in the current period depends on the occurrence or not of a violation in the previous period. The estimated VaR violation probability is the empirical frequency of violations, \( H/T \). Under the alternative, no restriction is imposed on the \( \Pi \) matrix. The corresponding LR statistic, denoted \( LR_{IND} \), is defined by:

\[
LR_{IND} = -2 \ln \left[ \left( 1 - \frac{H}{T} \right)^{T-H} \left( \frac{H}{T} \right)^H \right] + 2 \ln \left[ (1 - \widehat{\pi}_{01})^{n_{00}} \cdot \pi_{01}^{n_{01}} (1 - \widehat{\pi}_{11})^{n_{10}} \cdot \pi_{11}^{n_{11}} \right] \xrightarrow{T \to \infty} \chi^2(1)
\]

(31)

where \( n_{ij} \) denotes the number of times we have \( I_t(\alpha) = j \) and \( I_{t-1}(\alpha) = i \), and:

\[
\widehat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}}, \quad \widehat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}.
\]

(32)

Finally, it is also possible to test the CC null hypothesis. Under CC,

\[
H_{0,CC} : \Pi_\alpha = \begin{pmatrix}
1 - \alpha & \alpha \\
1 - \alpha & \alpha
\end{pmatrix}
\]

(33)
and then:
\[
LR_{CC} = -2 \ln \left[ (1 - \alpha)^{T-H} (\alpha)^H \right] \\
+ 2 \ln \left[ (1 - \hat{\pi}_{01})^{n_{01}} \hat{\pi}_{01}^{n_{10}} (1 - \hat{\pi}_{11})^{n_{11}} \hat{\pi}_{11}^{n_{11}} \right] \\
\xrightarrow{T \to \infty} \chi^2(2)
\] (34)

The corresponding \(LR_{CC}\) statistic corresponds to the sum of the \(LR_{UC}\) and \(LR_{IND}\) statistics. Under the null of CC, it satisfies:
\[
LR_{CC} = LR_{UC} + LR_{IND} \xrightarrow{T \to \infty} \chi^2(2)
\] (35)

### A.3 Resampling test

Escanciano and Olmo (2010) alternatively approximate the critical values of the UC test by using subsampling methodology. They show that the VaR backtest may be affected by an estimation risk. Thus, they propose to use robust subsampling techniques to approximate the true sampling distribution and to determine the critical value of the test. Let us denote by \(K\) the test statistic (typically the LR test statistic) and \(G_K(x)\) the corresponding cdf, such as \(G_K(x) = \Pr(K \leq x)\), for any \(x \in \mathbb{R}\) or \(\mathbb{R}^+\) given the test statistic considered. Let \(K_{b,t} = K(r_t, r_{t+1}, \ldots, r_{t+b-1})\), with \(t = 1, 2, \ldots, T - b + 1\), be the test statistic computed with the subsample \((r_t, r_{t+1}, \ldots, r_{t+b-1})\) of size \(b\). Hence, Escanciano and Olmo (2010) approximate the sampling distribution \(G_K(x)\) using the distribution of the values of \(K_{b,t}\) computed over the \(T - b + 1\) different consecutive subsamples of size \(b\). This subsampling distribution is defined by:
\[
G_{K_b}(x) = (T - b + 1)^{-1} \sum_{t=1}^{T-b+1} \mathbb{I}(K_{b,t} \leq x).
\] (36)

The \((1 - \tau)^{th}\) sample quantile of \(G_{K_b}(x)\), is given by:
\[
c_{K_b,1-\tau} = \inf_{x \in \mathbb{R}} (x : G_{K_b}(x) \geq 1 - \tau).
\] (37)

### A.4 Exception Magnitude Test

The magnitude of VaR forecasting violations, i.e., the difference between the VaR forecasts and the ex-post returns, should be of primary interest to the users of VaR models (Colletaz et al. (2013)). The test proposed by Berkowitz (2001) is based on the following transformed series:
\[
r_t^* = \begin{cases} 
VaR_t(\alpha) & \text{if } r_t > -VaR_t(\alpha) \\
\alpha & \text{otherwise,}
\end{cases}
\] (38)

Berkowitz (2001) proposes a LR test based on an extension of the Rosenblatt transformation and a censored likelihood. Loosely speaking, the shape of the forecasted tail of the density is compared to the observed tail.
\[
LR_{\text{tail}} = -2 \left[ L_{r^*}(0,1) - L_{r^*}(\hat{\mu}, \hat{\sigma}) \right],
\] (39)

where \(L_{r^*}(\hat{\mu}, \hat{\sigma})\) is the likelihood of a censored normal. This likelihood contains only observations falling in the tail, but they are treated as continuous variables. Under the null, \(H_0 : \mu = 0\) and \(\sigma = 1\), the \(LR_{\text{tail}}\) test statistic is asymptotically distributed as a \(\chi^2(2)\) (see Berkowitz (2001) for more details).
A.5 GMM Duration-based Tests

The duration-based tests consider the duration $d_v$ between two consecutive VaR violations:

$$d_v = t_v - t_{v-1}$$

(40)

where $t_v$ denotes the date of the $v^{th}$ violation. Under CC hypothesis, the duration process $d_i$ has a geometric probability density function given by:

$$f(d_v; \alpha) = \alpha (1 - \alpha)^{d_v - 1} \quad d_v \in \mathbb{N}^*.$$ (41)

This distribution characterizes the memory-free property of the violation process $I_t(\alpha)$, which means that the probability of observing a violation today does not depend on the number of days that have elapsed since the last violation. Note that $E(d_v) = 1/\alpha$ since the CC hypothesis implies an average duration between two violations equal to $1/\alpha$. The general idea of the duration-based test consists of testing the geometric distribution.

Candelon et al. (2011) propose a test based on the Generalized Method of Moment (GMM) test framework proposed by Bontemps and Meddahi (2005) to test for the distributional assumption. The test statistics are simple J-statistics based on the moments defined by the orthonormal polynomials associated with the geometric distribution. The orthonormal polynomials associated to a geometric distribution with a success probability $\beta$ are defined by the following recursive relationship,

$$M_{j+1}(d; \beta) = \frac{(1 - \beta)(2j + 1) + \beta(j - d + 1)}{(j + 1)\sqrt{1 - \beta}} M_j(d; \beta) - \left(\frac{j}{j + 1}\right) M_{j-1}(d; \beta),$$

(42)

for any order $j \in \mathbb{N}$, with $M_{-1}(d; \beta) = 0$ and $M_0(d; \beta) = 1$. If the true distribution of $D$ is a geometric distribution with a success probability $\beta$ then, it follows that $E[M_j(d; \beta)] = 0$, $\forall j \in \mathbb{N}^*, \forall d \in \mathbb{N}^*$. Hence, the null hypothesis of CC, UC and IND can be expressed as follows:

$$H_{0,CC} : E[M_j(d_i; \alpha)] = 0, \quad j = 1, \ldots, p, \quad (43)$$

$$H_{0,UC} : E[M_1(d_i; \alpha)] = 0. \quad (44)$$

$$H_{0,IND} : E[M_j(d_i; \beta)] = 0, \quad j = 1, \ldots, p, \quad (45)$$

where $p$ denotes the number of moment conditions (fixed by the user). The $H_{0,IND}$ test consists of testing the hypothesis of a geometric distribution (implying the absence of dependence) with a success probability equalling $\beta$, where $\beta$ denotes the true violation rate that is not necessarily equal to the coverage rate $\alpha$. Candelon et al. (2011) propose three test statistics:

$$J_{CC}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i; \alpha)\right)^T \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i; \alpha)\right) \overset{d}{\underset{N \to \infty}{\to}} \chi^2(p). \quad (46)$$

$$J_{UC}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M_1(d_i; \alpha)\right)^2 \overset{d}{\underset{N \to \infty}{\to}} \chi^2(1). \quad (47)$$

$$J_{IND}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i; \beta)\right)^T \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i; \beta)\right) \overset{d}{\underset{N \to \infty}{\to}} \chi^2(p). \quad (48)$$

24
A.6 Dynamic Quantile Test

The dynamic quantile (DQ) test, proposed by Engle and Manganelli (2004), is one of the most popular backtesting tests. Let us denote by $Hit_t(\alpha) = I_t(\alpha) - \alpha$ the demeaned process of violations, that takes the value $1 - \alpha$ every time $r_t$ is less than the ex-ante VaR and $-\alpha$ otherwise. The DQ test is based on a quantile regression model of the VaR on a set of explanatory variables that belong to $F_{t-1}$:

$$Hit_t(\alpha) = \delta + \sum_{k=1}^{K} \beta_k Hit_{t-k}(\alpha) + \sum_{k=1}^{K} \gamma_k g[Hit_{t-k}(\alpha), Hit_{t-k-1}(\alpha), ..., z_{t-k}, z_{t-k-1}, ...] + \varepsilon_t,$$

where $\varepsilon_t$ is a discrete i.i.d. process and where $g(.)$ is a function of past violations and of the variables $z_{t-k}$ belonging to the entire informational set available $F_{t-1}$ (past returns $r_{t-k}$, the square of past returns $r^2_{t-k}$, the past VaR, etc.). Under CC, $Hit_t(\alpha)$ is a difference martingale and the conditional expectation of $Hit_t(\alpha)$ given any information known at $t-1$ must be zero. Then, the CC hypothesis can be written as:

$$H_0: \delta = \beta_k = \gamma_k = 0$$

Under the null hypothesis, $E[Hit_t(\alpha)] = E(\varepsilon_t) = 0$, which means that by definition $Pr[Hit_t(\alpha) = 1] = E[I_t(\alpha)] = \alpha$. Therefore, if we denote by $\Psi = [\delta, \beta_1, ..., \beta_K, \gamma_1, ..., \gamma_K]'$ the vector of the $2K + 1$ parameters of the model and by $Z$ the matrix of explanatory variables of model (49), the test statistic $DQ_{CC}$ is defined as:

$$DQ_{CC} = \frac{\hat{\Psi}'Z'Z\hat{\Psi}}{\alpha(1-\alpha)} \xrightarrow{T \to \infty} \chi^2(2K + 1)$$
References


