CAViaR: CONDITIONAL AUTOREGRESSIVE VALUE AT RISK BY REGRESSION QUANTILES

Robert F. Engle and Simone Manganelli

University of California, San Diego

July 1999

1. INTRODUCTION

Recent financial disasters have emphasized the importance of effective risk management for financial institutions. The use of quantitative risk measures has become an essential management tool to be placed in parallel with the models of returns. These measures are used for investment decisions, supervisory decisions, risk capital allocation and external regulation. In the fast paced financial world, effective risk measures must be as responsive to news as are other forecasts and must be easy to grasp even in complex situations. Many financial institutions have switched from management based on accrual accounting (a practice according to which transactions are booked at historical costs plus or minus accruals) to management based on daily marking-to-market. This switch has caused an increase in the volatility of the apparent value of overall positions held by financial institutions, which now reflects the volatility of the underlying markets and the effectiveness of hedging strategies.

Value at Risk (VaR) has become the standard measure of risk employed by financial institutions and their regulators. VaR is an estimate of how much a certain portfolio can lose within a given time period and at a given confidence level. More precisely VaR is defined so that the probability that a portfolio will lose more than its VaR over a particular time horizon is equal to θ, a prespecified number. The great popularity that this instrument has achieved among financial practitioners is essentially due to its conceptual simplicity: VaR reduces the (market) risk associated with any portfolio to just one dollar amountThe summary of many complex bad outcomes in a single number, naturally represents a compromise between the needs of different users. This compromise has received the blessing of a wide range of users and regulators.

Despite its conceptual simplicity, the measurement of VaR is a very challenging statistical problem and none of the methodologies developed so far gives satisfactory solutions. Since VaR can be computed as the quantile of future portfolio returns, conditional on current information, and since the distribution of portfolio returns typically changes over time, the challenge is to find a suitable model for time varying order statistics.

The problem is to forecast a value each period that will be exceeded with probability *(1-θ)* by the current portfolio. That is, for $\{y_t\}_{t=1}^{n}$, find VaR_t such that

(1) $$\Pr\left[y_t < -VaR_t \mid \Omega_{t-1}\right] = \theta ,$$

where Ω_{t-1} denotes the information set at time *t-1*. Any reasonable model should solve the following three issues:

1) provide a formula for calculating VaR_t as a function of variables known at time *t-1* and a set of parameters that need to be estimated;
2) provide a procedure (namely, a loss function and a suitable optimization algorithm) to estimate the set of unknown parameters;
3) provide a test to establish the quality of the estimate.

In this paper we address each of these issues. We propose a conditional autoregressive specification for VaR_t, which we call Conditional Autoregressive Value at Risk (CAViaR). The unknown parameters of the CAViaR models are estimated using Koenker and Bassett's (1978) regression quantile framework. Building on White (1994) and Weiss (1991), we extend the results of the linear regression quantile to the nonlinear dynamic case, providing the asymptotic distribution of the estimator and a procedure to estimate the variance-covariance matrix. We also show how to construct the Wald and LM statistics to test for significance of the coefficients of the CAViaR process. Since the regression quantile objective function is not differentiable and has many local optima (in the nonlinear case), we use a genetic algorithm for the numerical optimization. Finally, we propose a new test, based on an artificial regression, to evaluate the quality of the estimated CAViaR processes.

The paper is structured as follows. In section 2, we quickly review the current approaches to Value at Risk estimation. Section 3 introduces the CAViaR models. In section 4 we discuss the issue of how to evaluate a quantile estimate. In sections 5 and 6 we review the literature on regression quantiles and hypothesis testing. Section 7 contains a brief description of the genetic algorithm we use for the numerical optimization. Sections 8 and 9 present a Monte Carlo simulation and some empirical applications to real data of our methodology. Section 10 concludes the paper.

2. VALUE AT RISK MODELS

VaR estimates can be used for many purposes. The natural first field of application is risk management. Setting position limits in terms of VaR can help management estimate the cost of positions in term of risk. This allows managers to allocate risk in a more efficient way. Second, VaR can be applied to evaluate the performance of the risk takers on a risk/return basis. Rewarding risk takers only on a return basis can bias their behavior toward taking excessive risk. Hence, if the performance (in terms of returns) of the risk takers is not properly adjusted for the amount of risk effectively taken, the overall risk of the firm may exceed its optimal level. Third, the European Community and the Basel Committee on Banking Supervision at the Bank for International Settlements require financial institutions such as banks and investment firms to meet capital requirements to cover the market risks that they incur as a result of their normal operations. However, if the underlying risk is not properly estimated, these requirements may lead financial institutions to overestimate (or underestimate) their market risks and consequently to maintain excessively high (low) capital levels. The result is an inefficient allocation of financial resources that ultimately could induce firms to move their activities into jurisdictions with more liberal financial regulations.

The existing models for calculating VaR differ in the methodology they use, the assumptions they make and the way they are implemented. However, all the existing models follow a common general structure, which can be summarized in three points: 1) the portfolio is marked-to-market daily, 2) the distribution of the portfolio's returns is estimated, 3) the VaR of the portfolio is computed.

The main differences among VaR models are related to the second point, namely the way they address the problem of the portfolio distribution estimation. Existing models can be classified initially into two broad categories: a) factor models such as RiskMetrics, b) portfolio models such as historical quantiles. In the first case, the universe of assets is projected onto a limited number of factors whose volatilites and correlations have been forecast. Thus time variation in the risk of a portfolio is associated with time variation in the volatility or correlation of the factors. The VaR is assumed to be proportional to the computed standard deviation of the portfolio, often assuming normality.

The portfolio models construct historical returns that mimic the past performance of the current portfolio. From these historical returns, the current VaR is constructed based on a statistical model. Thus changes in the risk of a particular portfolio are associated with the historical experience of this portfolio. Although there may be issues in the construction of the historical returns, the interesting modeling question is how to forecast the quantiles. Several different approaches have been employed. Some first estimate the volatility of the portfolio, perhaps by GARCH or exponential smoothing, and then compute VaR from this, often assuming normality. Others use rolling historical quantiles under the assumption that any return in a particular period is equally likely. A third appeals to extreme value theory.

It is easy to criticize each of these methods. The volatility approach assumes that the negative extremes follow the same process as the rest of the returns and that the distribution of the returns divided by standard deviations will be independent and identically distributed if not normal. The rolling historical quantile method assumes that for a certain window, such as a year, any return is equally likely, but a return more than a year old has zero probability of occurring. It is easy to see that the VaR of a portfolio will drop dramatically just one year after a very bad day. Implicit in this methodology is the assumption that the distribution of returns does not vary over time at least within a year.

An interesting variation of the historical simulation method is the hybrid approach proposed by Boudoukh, Richardson and Whitelaw (1998). The hybrid approach combines volatility and historical simulation methodologies, by applying exponentially declining weights to past returns of the portfolio. This approach constitutes a significant improvement over the existing methodologies, since it drastically simplifies the assumptions needed in the traditional VaR methodology and solves part of the contradictions implicit in the historical estimation. However, both the choice of the parameters of interest and the procedure behind the computation of the VaR seem to be ad hoc and based on empirical justifications rather than on a sound statistical theory.

Applications of extreme quantile estimation methods to VaR have been recently proposed by Danielsson and de Vries (1998) and Gourieroux and Jasiak (1998). The intuition here is to exploit results from statistical extreme value theory and to concentrate the attention on the asymptotic form of the tail, rather than modeling the whole distribution. There are two problems with this approach. First it works only for very low probability quantiles. As shown by Danielsson and de Vries (1998), the approximation may be very poor at very common probability levels (such as 5%), because they are not "extreme" enough. Second, and most importantly, these models are nested in a framework of i.i.d. variables, which is not consistent with the characteristics of most financial datasets and consequently, the risk of a portfolio may not vary with the conditioning information set.

Beder (1995) applies eight common VaR methodologies to three hypothetical portfolios. The results show that the differences among these methods can be very large, with VaR estimates varying by more that 14 times for the same portfolio! Clearly, there is a need for a statistical approach to estimation and model selection.

3. CAVIAR

We propose another approach to quantile estimation. Instead of modeling the whole distribution, we model directly the quantile. The choice of the best functional form is mainly an empirical problem and will be determined by the data set under study. The first thing to keep in mind is the empirical fact that volatilities of stock market returns tend to cluster over time. This fact may be translated in statistical words by saying that the distribution of stock market returns tends to be autocorrelated. Consequently, the VaR, which is tightly linked to the standard deviation of the distribution, must exhibit a similar behavior. A natural way to formalize this characteristic is to use some type of autoregressive specification. We propose a conditional autoregressive quantile specification, which we call Conditional Autoregressive Value at Risk (CAViaR).

A very general specification for the CAViaR might be the following:

(2)
$$\begin{aligned} VaR_t &= f\left(x_t, \beta_\theta\right) \\ &= \beta_0 + \sum_{i=1}^{p} \beta_i VaR_{t-1} + l\left(\beta_{p+1},\ldots,\beta_{p+q};\Omega_{t-1}\right) \end{aligned}$$

where Ω_{t-1} is the information set available at time t and we suppressed the θ subscript for notational convenience.

In most practical cases the above formulation might reduce to a first order model:

(3)
$$VaR_t = \beta_0 + \beta_1 VaR_{t-1} + l(\beta_2, y_{t-1}, VaR_{t-1})$$

The autoregressive term $\beta_1 VaR_{t-1}$ ensures that the VaR changes "smoothly" over time. The role of $l(\beta_2, y_{t-1}, VaR_{t-1})$, instead, is that of linking the level of VaR_t to the level of y_{t-1}. That is, it measures how much the VaR should change based on the new information in *y*. This term thus has much the same role as the News Impact Curve for GARCH models introduced by Engle and Ng(1993). Indeed, we would expect VaR_t to increase as y_{t-1} becomes very negative, as one bad day makes the probability of the next somewhat greater. It might be that very good days also increase VaR as would be the case for volatility models. Hence VaR could depend symmetrically upon $|y_{t-1}|$.

Note that in order for the process in (2) not to be explosive, the roots of

(4) $$1 - \beta_1 z - \beta_2 z^2 - ... - \beta_p z^p = 0$$

must lie outside the unit circle.

Here are some examples of CAViaR processes which will be estimated. Obviously these merely scratch the surface. Throughout we use the notation $(x)^+ = max(x,0), (x)^- = min(x,0)$.

1. ADAPTIVE: $VaR_t = VaR_{t-1} + \beta\left[I\left(y_{t-1} \le -VaR_{t-1}\right) - \theta\right]$

In terms of the general specification, we set

$$\beta_0 = 0,\ \beta_1 = 1,\ l(\beta_2, y_{t-1}, VaR_{t-1}) = \beta_2\left[I(y_{t-1} \le -VaR_{t-1}) - \theta\right]$$

This model incorporates the very simple rule: whenever you exceed your VaR you should immediately increase it, but when you don't exceed it, you should decrease it very slightly. This strategy will obviously reduce the probability of sequences of hits and will also make it unlikely that there will never be hits. It however learns nothing from returns which are close to the VaR or which are extremely positive.

2. PROPORTIONAL SYMMETRIC ADAPTIVE:

$$VaR_t = VaR_{t-1} + \beta_1(|y_{t-1}| - VaR_{t-1})^+ - \beta_2(|y_{t-1}| - VaR_{t-1})^-$$

3. SYMMETRIC ABSOLUTE VALUE: $VaR_t = \beta_0 + \beta_1 VaR_{t-1} + \beta |y_{t-1}|$

4. ASYMMETRIC ABSOLUTE VALUE: $VaR_t = \beta_0 + \beta_1 VaR_{t-1} + \beta_2 |y_{t-1} - \beta_3|$

5. ASYMMETRIC SLOPE:

   $$VaR_t = \beta_0 + \beta_1 VaR_{t-1} + \beta_2 (y_{t-1})^+ - \beta_3 (y_{t-1})^-$$

6. INDIRECT GARCH(1,1): $VaR_t = \left(\beta_1 + \beta_2 VaR_{t-1}^2 + \beta_3 y_{t-1}^2\right)^{1/2}$

The INDIRECT GARCH model would be correctly specified if the underlying data were truly a GARCH(1,1) with an i.i.d. error distribution. It is therefore a useful model for simulations. However if this model is correctly specified, then it would be more efficient to estimate the GARCH model directly by Maximum Likelihood and then infer the VaR from the distribution of the standardized residuals.

4. TESTING VALUE AT RISK MODELS

If a model is correctly specified, then $\Pr(y_t < -VaR_t) = \theta \quad \forall t$, at the true parameter. This is equivalent to requiring that the sequence of indicator functions $\{I(y_t < -VaR_t)\}_{t=1}^{T}$ be independent and identically distributed. Hence, a property that any VaR estimate should satisfy is that of providing a filter to transform a (possibly) serially correlated and heteroskedastic time series into a serially independent sequence of indicator functions. A natural way to test the validity of the forecast model is to check whether the sequence $\{I(y_t < -VaR_t)\}_{t=1}^{T} \equiv \{I_t\}_{t=1}^{T}$ is i.i.d.

Several statistical procedures are available to check the i.i.d. assumption. At least three possibilities have been discussed in the literature for general dynamic Bernoulli random variables: Cowles and Jones (1937), the runs test by Mood (1940) and straightforward application of Ljung and Box (1978).

All these tests can detect the presence of serial correlation in the sequence of indicator functions $\{I_t\}_{t=1}^{T}$. However, this is not enough to assess the performance of a VaR estimate. Indeed, it is not difficult to generate a sequence of independent $\{I_t\}_{t=1}^{T}$ from a given sequence of $\{y_t\}_{t=1}^{T}$. It suffices to define a sequence of independent random variables $\{z_t\}_{t=1}^{T}$, such that

(5) $$z_t = \begin{cases} 1 & \textit{with probability}\ \theta \\ -1 & \textit{with probability}\ (1-\theta) \end{cases}$$

Then, setting $VaR_t = Kz_t$, for K large, will do the job. Notice however, that once z is observed, the probability of exceeding the Value at Risk is known to be almost zero or one. Thus the unconditional probabilities are correct and serially uncorrelated, but the conditional probabilities given VaR are not. This example is an extreme case of measurement error in *VaR.* Any noise introduced into the Value at Risk will change the conditional probability of a hit *given VaR*.

Therefore, none of these tests has power against conditional bias and none can be simply extended to examine other explanatory variables. We propose a new test which can be easily extended to incorporate a variety of alternatives.

Define:

(6) $$Hit(y_t, x_t, \theta) \equiv Hit_{\theta t} \equiv I(y_t < -VaR_t) - \theta .$$

The *Hit* function assumes value $(1-\theta)$ every time y_t is less than VaR_t (i.e., every time a "hit" is realized) and $(-\theta)$ otherwise. Clearly the expected value of *Hit* is zero. Furthermore, from the definition of the quantile function, the conditional expectation of *Hit* given any information known at *t-1* must also be zero. A simple application of the law of iterated expectations shows that *Hit* must be uncorrelated with anything that belongs to the information set Ω:

(7) $$E(Hit_t\, \omega_{t-1}) = \omega_{t-1}\, E(Hit_t \,|\, \omega_{t-1}) = 0 \qquad \forall\, \omega_{t-1} \in \Omega_{t-1}\ .$$

In particular, Hit_t must be uncorrelated with any lagged Hit_{t-k}, with the forecasted VaR_t and with a constant. If Hit_t satisfies these moment conditions, then it is sure that there will be no autocorrelation in the hits, there will be no measurement error as in (5), and there will be the correct fraction of exceedences. If it is desired to check whether there are the right proportion of hits in each calendar year, then this can be measured by checking the correlation of *Hit* with annual dummy variables. If other functions of the past information set are suspected of being informative such as rolling standard deviations or a GARCH volatility estimate, these can be incorporated. A very convenient way to construct a test is to regress *Hit* on these independent variables[1]:

(8) $$\begin{aligned} Hit_t = \delta_0 + \delta_1 Hit_{t-1} + ... + \delta_p Hit_{t-p} + \delta_{p+1} VaR_t + \\ \delta_{p+2} I_{year1,t} + ... + \delta_{p+2+n} I_{yearn,t} + u_t \end{aligned}$$

Rewriting this artificial regression in matrix form, we get:

(9) $$Hit_t = X\delta + u_t \qquad u_t = \begin{cases} -\theta & prob\,(1-\theta) \\ (1-\theta) & prob\,\theta \end{cases}$$

A good model should produce a sequence of unbiased and uncorrelated hits, so that the explanatory power of this artificial regression should be zero. Hence, what we want to test is the null hypothesis H_0: $\delta=0$. Noticing that the terms in X are measurable-Ω_{t-1}, the asymptotic distribution of the OLS estimator under the null can be easily established, invoking an appropriate central limit theorem:

(10) $$\hat{\delta}_{OLS} = (X'X)^{-1} X'Hit \overset{a}{\sim} N\left(0, \theta(1-\theta)(X'X)^{-1}\right)$$

It is now straightforward to derive the Dynamic Quantile test statistic:

(11) $$\frac{\hat{\delta}_{OLS}' X' X \hat{\delta}_{OLS}}{\theta(1-\theta)} \overset{a}{\sim} \chi^2(p+n+2)$$

While this measure of performance is quite useful, its distribution in-sample is affected by the fact that the Hits are functions of estimated parameters. We will discuss this problem in section 6.

5. Regression Quantiles

Regression quantiles models were introduced by Koenker and Bassett (1978). They show how a simple minimization problem yielding the ordinary sample quantiles in the location model can be generalized to the linear regression model. Consider a sample of independent observations on random variables $y_1,\ldots,y_t$ distributed according to

(12) $$\Pr(y_t < \tau | x_t) = F_y(\tau | x_t) \qquad t = 1,\ldots,T$$

where x_t is a $(k,1)$ vector of regressors. Let $x_t'\beta_\theta$ be the θ-quantile. Then the model can be rewritten as

(13) $$\theta = \int_{-\infty}^{x_t'\beta_\theta} f_y(s|x_t)ds$$

or, following the convention established by the literature, as

(14) $$y_t = x_t'\beta_\theta + u_{\theta t} \qquad Quant_\theta(y_t|x_t) = x_t'\beta_\theta$$

[1] An analogous procedure to evaluate interval forecasts was proposed by Christoffersen (1998).

where $Quant_\theta(y_t|x_t) = x_t'\beta_\theta$ is the θ-quantile of y_t conditional on x_t.

When $x_t=1$, $t= 1,\ldots,T$, we get as a special case the location model, in which β_θ simply represents the sample θ-quantile. It is straightforward to show that the sample θ-quantile of a random sample $\{y_t, t = 1,\ldots,T\}$ on a random variable y is defined as any solution to

(15) $$\min_b \frac{1}{T}\left\{\sum_{t:y_t \geq b} \theta \,|\, y_t - b \,| + \sum_{t:y_t < b} (1-\theta)\,|\, y_t - b \,|\right\}.$$

Koenker and Bassett (1978) show that a direct generalization of this objective function extends the notion of sample quantile to the linear model. The θ^{th} regression quantile is defined as any $\hat{\beta}_\theta$ that solves:

(16) $$\min_\beta \frac{1}{T}\left\{\sum_{t:y_t \geq x_t'\beta} \theta \,|\, y_t - x_t'\beta \,| + \sum_{t:y_t < x_t'\beta} (1-\theta)\,|\, y_t - x_t'\beta \,|\right\}.$$

Rewriting this expression in terms of the indicator function yields the equivalent objective function:

(17) $$\min_\beta \frac{1}{T}\left\{-\sum_t \left[I\left(y_t < x_t'\beta\right) - \theta\right]\left[y_t - x_t'\beta\right]\right\}$$

Regression quantiles include as a special case the least absolute deviation (LAD) model. The properties of LAD have been discussed for many years and it is very well known that they are more robust than OLS estimators whenever the errors have a long tailed distribution. Koenker and Bassett (1978) ran a simple Monte Carlo experiment and show how the empirical variance of the median, compared to the variance of the mean, is slightly higher under the normal distribution, but it is much lower under all the other distributions taken into consideration.[2] The result is particularly striking in the Cauchy case.

A very important generalization of the basic linear model is the one proposed by Powell (1986), who introduced the censored regression quantiles model. Newey and Powell (1990) show that a slight modification of the quantile regression objective function is able to deliver efficient estimates. They address the issue of attainable asymptotic efficiency for a linear regression model with error term restricted to have a zero quantile, conditional on the regressors. They prove that weighting the terms of the typical objective function with the conditional density at zero of the errors produces estimators whose variance covariance matrix attains (asymptotically) the theoretical lower bound.

In the nonlinear case, in the context of time series, the most important contributions are those by White (1991, 1994 p. 75) who proves the consistency of the nonlinear regression quantile, both in the i.i.d. and stationary mixing or ergodic cases. Weiss (1991) shows consistency, asymptotic normality and asymptotic equivalence of LM and Wald tests for LAD estimators for nonlinear dynamic models. Using the consistency results provided by White, Weiss's proofs of asymptotic normality and asymptotic equivalence of the LM and Wald tests can be easily modified to accommodate the more general case of nonlinear regression quantiles. In the rest of the paper, we rely heavily on Weiss's assumptions and results. For convenience, we report here White's consistency result and the quantile generalization of Weiss's asymptotic normality result.

Consider the model

[2] They consider Gaussian mixture, Laplace and Cauchy distributions.

(18) $\quad y_t = f(x_t, \beta_\theta) + \varepsilon_{t\theta} \qquad\qquad Quant_\theta(\varepsilon_t \mid \Omega_{t-1}) = 0$

The following assumptions are needed to guarantee the consistency of the regression quantile estimator:

Consistency Assumptions

C0. (*Ω*, *F*, *P*) is a complete probability space and $\{\varepsilon_t, x_t\}$, *t = 1, 2,…*, are random vectors on this space.

C1. The function $f(x_t, \beta_\theta): \Re^{k_t} \times B \to \Re$ is such that for each β_θ in *B*, a compact subset of $\Re^p$, $f(x_t, \beta_\theta)$ is measurable with respect to the Borel set B^{k_t} and $f(x_t, \cdot)$ is continuous in *B*, a.s.-*P*, *t=1,2,…* for a given choice of explanatory variables $X=\{x_t\}$.

C2. (a) $E([I(y_t < f(x_t, \beta_\theta)) - \theta][y_t - f(x_t, \beta_\theta)])$ exists and is finite for each β_θ in *B*.

(b) $E([I(y_t < f(x_t, \beta_\theta)) - \theta][y_t - f(x_t, \beta_\theta)])$ is continuous in β_θ.

(c) $\{[I(y_t < f(x_t, \beta_\theta)) - \theta][y_t - f(x_t, \beta_\theta)]\}$ obeys the strong (weak) law of large numbers. For example, we could assume that $\{\varepsilon_t, x_t\}$ are α-mixing. That is, $\alpha(m)$ satisfies $\alpha(m) \to 0$ as $m \to \infty$. See, for example, Andrews (1988) or White and Domowitz (1984) for further details.

C3. $\{n^{-1}E\{[I(y_t < f(x_t, \beta_\theta)) - \theta][y_t - f(x_t, \beta_\theta)]\}\}$ has identifiably unique maximizers.

Theorem 1 *(Consistency, White (1994) page 75) - In model (18), under C0, C1, C2 and C3,* $\hat{\beta}_\theta \to \beta_\theta$ *as* $n \to \infty$ *a.s.-*P_0*, where* $\hat{\beta}_\theta$ *is the solution to:*

$$\max_{\beta_\theta} n^{-1} \sum_{t=1}^{n} \{[I(y_t < f(x_t, \beta_\theta)) - \theta][y_t - f(x_t, \beta_\theta)]\}.$$

To prove the asymptotic normality of $\hat{\beta}_\theta$, we need to introduce some extra notation. Following Weiss, let v_t be a $(r\times 1)$ vector of variables that determine the shape of the conditional distribution of ε_t. Associated with v_t is a set of parameters ϕ.[3] Denote the density of ε_t, conditional on all the past information, as $h_t(\varepsilon;\phi,v_t)$, $\varepsilon\in\Re$. Whenever the dependence on v_t and ϕ is not relevant, we'll denote the conditional density of ε_t simply by $h_t(\varepsilon)$. Let $u_t(\phi, \beta_\theta, s)$ be an unconditional density of $s_t = (\varepsilon_t, x_t, v_t)$. Finally, define the operators $\nabla\equiv\partial/\partial\beta$, $\nabla_i\equiv\partial/\partial\beta_i$, where β_i is the i[th] element of β, $\nabla_i f_t(\beta)\equiv\nabla_i f(x_t,\beta)$ and $\nabla f_t(\beta)\equiv\nabla f(x_t,\beta)$.

Asymptotic Normality Assumptions

AN1. $\nabla_i f_t(\beta_\theta)$ is A-smooth with variables A_{it} and functions ρ_i, $i=1,\ldots,p$. In addition, $\max_i \rho_i(d)\leq d$ for $d>0$ small enough.[4]

AN2. (i) $h_t(\varepsilon)$ is Lipschitz continuous in ε uniformly in t.

(ii) For each t and (ε, v), $h_t(\varepsilon;\phi, v)$ is continuous in ϕ.

AN3. For each t and s, $u_t(\phi, \beta_\theta, s)$ is continuous in (ϕ, β_θ).

AN4. $\{\varepsilon_t, x_t\}$ are α-mixing, with parameter $\alpha(m)$, and there exist $\Delta<\infty$ and $r > 2$ such that $\alpha(m)\leq \Delta m^\lambda$ for some $\lambda<-2r/(r-2)$.

---

[3] For example, v_t might include the conditional variance and ϕ might be the vector of parameters that define a GARCH model.

[4] $f(x_t,\beta)$ is A-smooth with variables A_{0t} and function ρ if, for each $\beta\in B$, there is a constant $\tau>0$ such that $||\beta^*-\beta||\leq\tau$ implies that $|f(x_t,\beta^*)- f(x_t,\beta)|\leq A_{0t}(x_t)\rho(||\beta^*-\beta||)$ for all t, a.s.-P, where A_{0t} and ρ are nonrandom functions such that $A_{0t}(x_t)$ is a random variable, $\lim\sup_{n\to\infty} n^{-1}\sum E[A_{0t}(x_t)]<\infty$, $\rho(v)>0$ for $v>0$, $\rho(v)\to 0$ as $v\to 0$ and τ, A_{0t}, ρ and the null set may depend on β.

AN5. For some $r>2$, $\nabla_i f_t(\beta)$ is uniformly r-dominated by functions a_{1t}.

AN6. For all t and i, $E|sup_\beta A_{it}|^r \leq \Delta_1 < \infty$. There exist measurable functions a_{2t} such that $|u_t| \leq a_{2t}$ and for all t, $\int a_{2t} dv < \infty$ and $\int (a_{1t})^3 a_{2t} dv < \infty$.

AN7. There exists a matrix A such that $n^{-1} \sum_{t=a+1}^{a+n} E[\nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)] \to A$, as $n \to \infty$, uniformly in a.

Theorem 2 *(Asymptotic Normality) - In the model (18), if AN1-AN7 hold and if the estimator is consistent, then:*

$$\sqrt{\frac{n}{\theta(1-\theta)}} A_n^{-1/2} D_n \left(\hat{\beta}_\theta - \beta_\theta\right) \xrightarrow{d} N(0, I)$$

where $A_n = n^{-1} \sum E[\nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)]$, $D_n = n^{-1} \sum E[h_t(0) \nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)]$ *and* $\hat{\beta}_\theta$ *is computed as in Theorem 1.*

Proof - Substituting $\psi(x)=sign(x)=2[I(x<0)-1/2]$ with $Hit(x) \equiv [I(x<0)-\theta]$ in theorem 3 of Weiss (1991) doesn't affect the validity of the argument.

Note that the Adaptive model does not satisfy White's assumptions for consistency, since the quantile function, $VaR_t(\beta)$, is not continuous in β. Note also that the only two models for which the gradient of the quantile function is defined for all β are the Symmetric Absolute Value and the Indirect GARCH models. Hence, strictly speaking, the asymptotic normality results apply only to these models. However, each of the six models we take into consideration (including the Adaptive) can be approximated arbitrarily well by continuous and differentiable functions. Since taking these approximations don't affect the nature of the models (in the sense that the autoregressive mechanism still applies), we can treat them as if they satisfy all the necessary assumptions to give consistency and asymptotic normality results.

6. HYPOTHESIS TESTING

Weiss (1991) proves the asymptotic equivalence of the Wald and LM test, in the case of nonlinear dynamic LAD models. As for the asymptotic normality theorem, his proofs can be easily modified to get the analogous tests in the more general case of regression quantile models.

The problem with these tests is related to the estimation of the variance-covariance matrix of the estimator $\hat{\beta}_\theta$, and more precisely to the fact that we need an estimate of the density function of the distribution of the errors, $h_t(\varepsilon)$, at zero. Under the homogeneity assumption, i. e. assuming that $h_t(0) = h(0)$ for all t, Powell (1984) and Weiss (1991) propose to estimate the density function with kernel estimation techniques:

$$\hat{h}(0) = (\hat{c}_n n)^{-1} \sum k\left(\hat{\varepsilon}_{t,n} / \hat{c}_n\right), \tag{19}$$

where $k(\cdot)$ is a kernel, $\hat{\varepsilon}_{t,n}$ is the t$^{\text{th}}$ residual and $\hat{c}_n$ is the bandwidth. The most common and simplest kernel is $k(u) = I(|u|\leq 1)/2$. The idea is that if $c_n \to 0$ as $n \to \infty$, then $\hat{h}(0) \xrightarrow{p} h(0)$.

The homogeneity assumption is clearly too much restrictive for our setting, since as the quantile is changing over time it seems highly implausible that the density function of the returns at the quantile stays fixed. We propose, instead, a less restrictive and, in our opinion, more realistic assumption to estimate $h_t(0)$.

The strategy we adopt is the following. We estimate the CAViaR model to obtain an estimate of the quantile $f\left(x_t, \hat{\beta}_\theta\right)$, and then construct the series of standardized quantile residuals:

(20) $$\hat{\varepsilon}_{t\theta} / f\left(x_t, \hat{\beta}_\theta\right) = y_t / f\left(x_t, \hat{\beta}_\theta\right) - 1 .$$

Let $g_t(\cdot)$ be the density function of the standardized quantile residuals. To get an estimate of $h_t(\varepsilon)\big|_{\varepsilon=0}$ we impose the following assumptions.

Variance-Covariance Estimation Assumptions

VC1 - $g_t\left({}^{\varepsilon_{t\theta}} \Big/ {}_{f(x_t, \beta_\theta)} \right)\Big|_0 = g\left({}^{\varepsilon_{t\theta}} \Big/ {}_{f(x_t, \beta_\theta)} \right)\Big|_0$, for all t, provided that $f(x_t, \beta_\theta) \neq 0$ for all t.

VC2 - $\hat{c}_n / c_n \xrightarrow{p} 1$, where the nonstochastic sequence c_n satisfies $c_n = o(1)$ and $c_n^{-1} = o\left(n^{1/2}\right)$.

VC3 - The elements of $\nabla f_t(\beta_\theta)$ are uniformly 4-dominated.

The crucial assumption is VC1. This assumption states that the density function of the standardized quantile residuals at zero is not time varying.

To understand the implications of this assumption, consider the following. Assume that the underlying model is $y_t = \varepsilon_t = \sigma_t s_t$ where $s_t \sim (0,1)$. By setting the mean of $s_t=0$ we ensure that y_t is a Martingale difference sequence. Even if assuming that $E(s_t)=0$ is not necessary to get the desired result, imposing this assumption clarifies the plausibility of our setting.

Clearly, $f(x_t, \beta_\theta) = \sigma_t \kappa_t$, where κ_t is the θ-quantile of s_t. If s_t is i.i.d., then κ_t is constant and our assumption trivially holds. However, when s_t is not iid, κ_t may change over time and our assumption has force. The assumption can be reformulated in the following way:

VC1' - Let q_t be the density function of s_t / κ_t. Then $q_t(1)=q(1)$ for all t.

Using the jacobian of the transformation, we can get the estimate of the density function of the errors at zero:

(21) $$h_t(0) = \frac{g(0)}{|\ f(x_t, \beta_\theta)\ |}.$$

There is a third way to evaluate this assumption. Assume that model (18) is *locally* correctly specified, where by this we mean that there is a neighborhood of θ in which model (18) holds. Consider θ_1, θ_2 in this neighborhood and let $f(x_t, \beta_{\theta_1})$ and $f(x_t, \beta_{\theta_2})$ be the corresponding quantile estimates. An estimator of the density function at the θ-quantile ($\theta = (\theta_1 + \theta_2)/2$) is:

(22) $$h_t(0) \approx \frac{|\ \theta_1 - \theta_2\ |}{|\ f(x_t, \beta_{\theta_1}) - f(x_t, \beta_{\theta_2})\ |}.$$

As the difference between θ_1 and θ_2 goes to zero, this should give a consistent estimate of the density function. The problem of this approach is that we need to estimate $f(x_t, \beta_{\theta_1})$ and $f(x_t, \beta_{\theta_2})$ together with $f(x_t, \beta_\theta)$. However, with an extra assumption we can avoid the problem of re-estimation. Rewrite $f(x_t, \beta_\theta)$, $f(x_t, \beta_{\theta_1})$ and $f(x_t, \beta_{\theta_2})$ as $\sigma_t \kappa_t$, $\sigma_t \kappa_t^1$ and $\sigma_t \kappa_t^2$. Then:

$$f(x_t, \beta_{\theta_i}) = f(x_t, \beta_\theta) \frac{\kappa_t^i}{\kappa_t} \qquad \text{i = 1, 2} \tag{23}$$

If we assume that $\dfrac{|\theta_1 - \theta_2|}{\left|\left(\dfrac{\kappa_t^1 - \kappa_t^2}{\kappa_t}\right)\right|} \equiv \tau$, for θ_1 and θ_2 in a neighborhood of θ, then:

$$h_t(0) \approx \frac{|\theta_1 - \theta_2|}{\left| f(x_t, \beta_{\theta_1}) \left(\dfrac{\kappa_t^1 - \kappa_t^2}{\kappa_t}\right)\right|} \equiv \frac{\tau}{|f(x_t, \beta_{\theta_1})|}, \tag{24}$$

which is exactly the same result we got previously.

With assumption VC1 and using the Jacobian of the transformation as in (21), it is possible to rewrite the *D* matrix that enters the variance-covariance matrix in Theorem 2 as:

$$\begin{aligned} D_n &= n^{-1} \sum E\left[h_t(0) \nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)\right] \\ &= n^{-1} \sum E\left[\frac{g(0)}{|f(x_t, \beta_\theta)|} \nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)\right] \\ &= n^{-1} g(0) \sum E\left[\frac{\nabla f_t(\beta_\theta) \nabla' f_t(\beta_\theta)}{|f(x_t, \beta_\theta)|}\right] \end{aligned}$$

A natural estimate of this matrix is:

$$\begin{aligned} \hat{D}_n &= (\hat{c}_n n)^{-1} \sum k\left(\frac{\hat{\varepsilon}_{t,n}}{f(x_t, \hat{\beta}_\theta)} / \hat{c}_n\right) \sum \frac{\nabla f_t(\hat{\beta}_\theta) \nabla' f_t(\hat{\beta}_\theta)}{f(x_t, \hat{\beta}_\theta)} \\ &\equiv \hat{g}(0) \hat{C}_n \end{aligned} \tag{25}$$

where *k(u)=I(|u|≤1)/2* is the kernel for *g(·)*, $\hat{g}(0) \equiv \hat{c}_n^{-1} \sum k\left(\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} / \hat{c}_n \right)$ and

$$\hat{C}_n \equiv n^{-1} \sum \frac{\nabla f_t\left(\hat{\beta}_\theta\right) \nabla' f_t\left(\hat{\beta}_\theta\right)}{f\left(x_t, \hat{\beta}_\theta\right)}.$$

We can now state the theorem.

Theorem 3 *(Estimation of the Asymptotic Variance-Covariance Matrix) - Under VC1-VC3 and the same conditions of Theorem 2,*

$$\hat{D}_n - D_n \xrightarrow{p} 0,$$

where $\hat{D}_n$ *is defined as in (25).*

Proof - The proof that $\hat{C}_n \xrightarrow{p} C_n$ is standard and will be omitted. To prove that $\hat{g}(0) \xrightarrow{p} g(0)$, we show that

$$\Delta_n \equiv (c_n n)^{-1} \sum \left[I\left(0 \leq \frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} \leq \hat{c}_n \right) - I\left(0 \leq \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)} \leq c_n \right) \right] \xrightarrow{p} 0$$

Then, the Lebesgue Dominated Convergence Theorem implies that

$$(c_n n)^{-1} \sum I\left(0 \leq \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)} \leq c_n \right) \xrightarrow{p} g(0).$$

By a simple application of the absolute value inequality, we have:

$$\Delta_n \le (c_n n)^{-1}\left|\sum\left[I\left(0 \le \frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} \le \hat{c}_n\right) - I\left(0 \le \frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} \le c_n\right)\right]\right|$$

(26)

$$+ (c_n n)^{-1}\left|\sum\left[I\left(0 \le \frac{\hat{\varepsilon}_{t,n}}{f(x_t,\hat{\beta}_\theta)} \le \hat{c}_n\right) - I\left(0 \le \frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} \le \hat{c}_n\right)\right]\right|$$

Arguing as in Powell (1984), equation A.27, it is possible to show that the two terms on the right hand side of (26) are $o_p(1)$. The first term can be rewritten as:

$$(c_n n)^{-1}\sum I\left(\left|\frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} - \hat{c}_n\right| \le |\hat{c}_n - c_n|\right)$$

For any $\eta>0$,

$$\Pr\left\{(c_n n)^{-1}\sum I\left(\left|\frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} - \hat{c}_n\right| \le |\hat{c}_n - c_n|\right) > \eta\right\}$$
$$\le \Pr\left\{(c_n n)^{-1}\sum I\left(\left|\frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} - \hat{c}_n\right| \le z\right) > \eta\right\} + \Pr\left\{c_n^{-1}|\hat{c}_n - c_n| > z\right\}$$
$$\le \eta^{-1}(c_n n)^{-1}\sum \Pr\left(\left|\frac{\varepsilon_{t,n}}{f(x_t,\beta_\theta)} - \hat{c}_n\right| \le z\right) + \Pr\left\{c_n^{-1}|\hat{c}_n - c_n| > z\right\}$$
$$\le \eta^{-1}c_n^{-1}L_0 z + \Pr\left\{c_n^{-1}|\hat{c}_n - c_n| > z\right\}$$

by Markov's inequality and the Lipschitz continuity of $h(\cdot)$. Since $c_n^{-1}|\hat{c}_n - c_n| = o_p(1)$ and $c_n^{-1} = o_p(n^{1/2})$, it is possible to choose z sufficiently small to make the last term of the above expression arbitrarily small for large n.

For the second term in (26), note that

$$\left| I\left(0 \le \frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} \le \hat{c}_n\right) - I\left(0 \le \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)} \le \hat{c}_n\right)\right|$$

$$\le I\left(\left|\frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right| < \left|\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} - \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right|\right) + I\left(\left|\frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)} - \hat{c}_n\right| < \left|\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} - \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right|\right)$$

$$\le I\left(\left|\frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right| < \left|\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} - \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right|\right)$$

$$+ I\left(\left|\frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)} - c_n\right| < \left|\hat{c}_n - c_n\right| + \left|\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} - \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right|\right)$$

Since $c_n^{-1}\left|\frac{\hat{\varepsilon}_{t,n}}{f\left(x_t, \hat{\beta}_\theta\right)} - \frac{\varepsilon_{t,n}}{f\left(x_t, \beta_\theta\right)}\right| = o_p(1)$, we can apply the same reasoning as before to show that also this term is $o_p(1)$.

Q.E.D

It is now possible to perform hypothesis testing on CAViaR models. We will concentrate our attention only on Wald and LM test, since the Likelihood Ratio test has different asymptotic distribution when the homogeneity assumption is not satisfied (see Weiss). As usual, let R denote the matrix of restrictions with rank q. We want to test the null hypothesis H_0: $R\beta_\theta=r$.

Theorem 4 (Wald and LM Test) - If the same conditions as Theorem 3 hold, under H_0

$$W_n \equiv \frac{n}{\theta(1-\theta)}\left[\hat{g}(0)\right]^2\left(R\hat{\beta}_\theta - r\right)'\left[R\hat{C}_n^{-1}\hat{A}_n\hat{C}_n^{-1}R'\right]^{-1}\left(R\hat{\beta}_\theta - r\right) \overset{a}{\sim} \chi_q^2$$

$$LM_n \equiv d_n(\tilde{\beta}_\theta)'\tilde{C}_n^{-1}R'\left[R\tilde{C}_n^{-1}\tilde{A}_n\tilde{C}_n^{-1}R'\right]^{-1}R\tilde{C}_n^{-1}d_n(\tilde{\beta}_\theta) \overset{a}{\sim} \chi_q^2$$

where $\hat{A}_n = n^{-1}\sum \nabla f_t(\hat{\beta}_\theta)\nabla' f_t(\hat{\beta}_\theta)$,

$$\hat{C}_n = n^{-1}\sum \left[f(x_t, \hat{\beta}_\theta)\right]^{-1}\nabla f_t(\hat{\beta}_\theta)\nabla' f_t(\hat{\beta}_\theta),$$

$$d_n(\tilde{\beta}_\theta) = n^{-½}\sum \nabla f_t(\tilde{\beta}_\theta)Hit_t(\tilde{\beta}_\theta)$$

$\hat{g}(0)$ *is computed as in Theorem 3 and ~ denotes that the variables are evaluated at the restricted estimates.*

Proof - The proof is again a straightforward extension of Weiss's proofs of theorems 6 and 8. It suffices to replace $\psi(x)=sign(x)=2[I(x<0)-1/2]$ with $Hit(x)\equiv [I(x<0)-\theta]$.

Note that to compute the LM test under assumption VC, we don't need to estimate the density of the standardized quantile residuals, since the *g(0)* term drops out of the expression.

The LM and Wald tests are appropriate if we want to test the null hypothesis $R\beta=0$, where β are the parameters of the CAViaR model. For example, these tests can be used to evaluate whether more terms should be included in the CAViaR specification, such as extra lagged *VaR* or *y*.

If we adopt the point of view that any model is necessarily misspecified, given the complexity of the real world, then the Dynamic Quantile test should be regarded as complementary to the LM test. What we want to check is whether the chosen model satisfies some basic requirements a good quantile estimate must have, such as unbiasedeness, independent hits and independence of the quantile estimate. With the DQ test, we are testing the null $\delta=0$, where δ are the coefficient of the artificial regression

(8). If we cannot reject the null that the estimated hits are distributed as a *bernoulli(θ)*, then we have "some" evidence that the model under study provides a satisfactory description of the real world.

The main problem with the DQ test is that we don't know its correct distribution when β is estimated with the same data being used for the test. However, if all we care about is whether the hits are uncorrelated and unbiased, this can be tested by constructing the chi square statistic proposed in expression (11). That is, we can interpret the DQ test as testing the hit sequence conditional on the estimated betas. Hence, the DQ test can be used as a model diagnostic or preliminary screening device to distinguish between good and bad models. For example, the DQ test could be used to evaluate the performance of the different VaR methodologies. If, for a given time series of (in sample) VaR estimates, the DQ statistic falls into the rejection region, then we must conclude that the data provide evidence against the model that produced those estimates. If the DQ statistic falls into the rejection region for an out of sample test, then this is further evidence against the model and its stability over time.

7. DIFFERENTIAL EVOLUTIONARY GENETIC ALGORITHM

The main problem of nonlinear regression quantiles estimation is that the objective function is not differentiable. Consequently, traditional algorithms based on differentiation will not work. We use, instead, a genetic algorithm that in theory is able to locate the global optimum, even for very complicated problems.

Genetic algorithms have been the subject of increasing interest in the past few years, since they provide a robust search procedure to solve very difficult problems. For large problems, random search or stochastic algorithms may be the only feasible alternative. Randomization gives a search algorithm the ability to break the curse of dimensionality that makes nonrandom and exhaustive search methods increasingly inefficient for functions with many parameters.

The idea behind this optimization routine is based on the process of natural selection and on some principles of genetics. The genetic algorithm starts with a population of initial trials for the parameter vector to be optimized, and interprets the value of the objective function at each of these trials as a measure of these points' "fitness" as an optimum. To develop a new population from this initial trial values there are three steps to follow:

1. *Reproduction based on fitness* - The members of the population are chosen for reproduction on the basis of their fitness, defined according to some specific criterion. At this stage some sort of "survival of the fittest" principle applies, so that the fittest members of the old population are given a higher probability of survival and/or reproduction.

2. *Crossover* - Crossover (or recombination) resembles the actual process of mating and establishes the rules of the reproduction. In this step, new parameter combinations are built from the components of existing vectors. Many different recombination methods exist and each combines parameter values from two or more parents in its own peculiar way.

3. *Mutation* - Some genes are given the chance of randomly changing, so that there is a possibility of improving the characteristics of the population (in the case the

mutation increases the fitness of the members of the population). Mutation is crucial for maintaining diversity in a population, even if excessive mutation may be harmful.

These three features make genetic algorithms radically different from the traditional search procedures. They allow the algorithm to develop generations that explore the region of interest and avoid getting stuck at a particular local optimum. This characteristic is useful for difficult optimization problems and in particular for those with multiple local minima and maxima. While traditional minimization routines tend to find only a local optimum, genetic algorithms are generally able to locate the global optimum.[5]

The type of genetic algorithm we use, Differential Evolutionary Genetic Algorithm (DEGA) is based on Price and Storn (1997).[6] This kind of algorithm has been proved to be much faster than traditional genetic algorithm, when applied to numerical optimization problems, and more robust at finding global optima. There are three factors that determine the evolutionary process of DEGA: the population size (NP), the crossover parameter (CR) and the mutation parameter (F). Suppose we want to maximize a real-valued function, with D parameters. DEGA starts by randomly generating NP, D-dimensional, real-valued vectors within the user-given intervals and evaluating the objective function at each of these initial trials. These values are stored in an (NP, D+1) array, called the *target* vector population. There are three steps to follow to develop a new generation.

1) An (NP, D+1) array of *trial* vectors is created.

[5] For a broad overview of the argument, see Goldberg (1989). Among the results stated in this book, there is the *fundamental theorem of genetic algorithms*, which provides a scientific justification of the ability of this class of algorithms to find global optima.
[6] The Matlab code was provided by Rick Baker, from the MathWorks, Inc.

2) The fitness of each trial vector is compared with the fitness of the corresponding target vector.
3) The fittest vectors survive and are stored in a new (NP, D+1) array that becomes the target vector population for the next generation.

The peculiarity of DEGA is related to the construction of the trial vector population. Each trial vector has two parents. The first parent is the target vector with which it has to compete in point 2) above. The second parent is constructed from three other randomly chosen target vectors. If we denote the second parent as P_2 and the three randomly chosen target vectors as T_1, T_2, T_3, then DEGA imposes $P_2=T_1+F(T_2-T_3)$, where F is the given mutation parameter. Note that this feature allows the trial vector to assume values outside the initial parameter range. Finally, the crossover parameter (CR) determines which genes of the trial vector are taken from which parent, by a series of D binomial experiments. D uniform random numbers are generated from the interval [0,1). If the d^{th} random number (d=1, …, D) is greater than CR, the trial vector gets the d^{th} parameter from the target vector parent, otherwise the parameter is inherited from P_2. For more details and sample codes, see Price and Storn (1997).

8. Monte Carlo Simulation

To check the ability of the nonlinear regression quantile function and genetic algorithm to produce consistent parameter estimates, we ran a few Monte Carlo simulations. First we generated 1000 samples of 3,000 observations using a GARCH(1,1) process with parameters (0.3, 0.05, 0.90). Then we estimated the GARCH parameters indirectly, by minimizing the nonlinear regression quantile objective function using the

Indirect GARCH CAViaR process as quantile specification. The probability levels of the quantile were set at 0.1%, 1%, 5% and 25%.

To implement DEGA, we generated 5000 random vectors, uniformly distributed in the interval [0, 2] for the 1%, 5% and 25% quantiles and in the interval [0,5] for the 0.1% quantile. We computed the value of the regression quantile criterion for each of these vectors. The best 50 vectors, that is the 50 vectors that yielded the lowest criterion value, were used as the starting population in DEGA. This selection process of the starting population reduces the number of necessary generations to achieve convergence and should make the final results more reliable. We set the population size (NP) equal to 50, the crossover parameter (CR) at 0.5, the mutation parameter (F) at 0.8 and the number of generations equal to 200. It is possible to increase the accuracy of the minimizing parameters by choosing a higher number of generations. We believe that the number we chose represents an acceptable trade off between precision of the estimate and computing time. Note that DEGA is able to locate a global optimum even if it lies outside the initial parameter range.[7] This is due to the way the trial vector population is generated, as explained in the previous section.

The results are shown in table 1. For each quantile, we report the value of the parameters of the true DGP, and the mean, the median and the variance-covariance matrix of the 1000 vectors of estimated parameters. For the mean, we computed also the t-statistic, using the empirical variance-covariance matrix. Note how in all the cases the median is a much better measure of location than the mean.

As we could expect, the worst results were those for the 0.1% quantile. In a sample of 3,000 observations, the 0.1% quantile is expected to be exceeded only three times and it may be hard, if not impossible, to get precise estimates. The mean of the estimated parameters is significantly different from their true values in most of the cases. The high variances confirm that estimation at such low confidence levels is very noisy. Perhaps, introducing extreme value theory in the CAViaR framework might be a better strategy to accomplish such a task.

The estimates at the other confidence levels are more reliable, as shown by the big drop in the variance of the first parameter. However, for some samples, the resulting estimated processes showed very little persistence (the coefficient of the autoregressive term of the GARCH process was close to zero), with the estimated quantile tending to the unconditional quantile. This result would arise naturally if the extremes are not clustered in the sample. With low probability events, there is the possibility that the timing would not reflect the predictability of extremes even though the DGP incorporated this feature. Under these circumstances, the lack of precision of the estimate might have no practical consequences, since the relevant properties of the quantile (unbiasedness and independence of lagged hits) are preserved. Moreover, this problem is very likely to be related to the sample size and should disappear as the number of observations in the sample becomes larger.

In table 2 we compute the mean, the median and the variance-covariance matrix, after excluding the estimates with GAMMA2 (the coefficient of the autoregressive lag) less than 0.5. The sample size after the trimming was 906 for the 0.1% quantile estimate, 986 for the 1%, 996 for the 5% and 953 for the 25%. The accuracy of the estimates improves dramatically, as shown by the reduction in their variances.

[7] The initial parameter range was [0, 2], but some of the optimal parameters were greater than 3.

9. EMPIRICAL RESULTS

To implement our methodology on real data, the researcher needs to construct the historical series of portfolio returns and to choose a specification of the functional form of the quantile. We took a sample of 3392 daily prices from Datastream for General Motors, IBM and S&P 500, and we computed the daily returns as the difference of the log of the prices. The samples range from April 7 1986 to April 7 1999. Note that our samples include the crash of the 1987. We used the first 2892 observations to estimate the model and the last 500 for out of sample testing. Figure 1 reports the plot of the returns of the three assets for the full sample.

We estimated 0.1%, 1%, 5% and 25% one day VaR, using the six CAViaR specifications described above. The estimated 5% VaR for the three assets are plotted in Figures 2, 3 and 4.

The results of the estimates are reported in tables 3 to 8. In each table, we report the value of the estimated parameters, the corresponding standard errors and (one-sided) p-values, the value of the regression quantile objective function at the optimum, the percentage of the times the VaR is exceeded, and the p-value of the Dynamic Quantile test, both in and out-of-sample. The standard errors were computed using the kernel described in theorem 3, with a bandwidth of 0.1 for all the assets and for all the confidence levels. A data dependent choice of the bandwidth would be preferable, since it would probably increase the precision of the estimate. We didn't report any standard error and any DQ test for the 0.1% VaR because the sample we have is not large enough to provide reliable estimates.

To identify the causes of a rejection in the DQ test, we used four different sets of regressors in the Dynamic Quantile artificial regression (8): 1) the constant and the first five lagged hits, 2) the VaR estimate, 3) the constant, the first lagged hit and the VaR estimate, 4) the constant, the first five lagged hits and the VaR estimate. Note that the last test encompasses all the previous ones.

Finally, the genetic algorithm was implemented by generating 5,000 random vectors within a given interval[8] and then selecting the fittest ones as starting population in DEGA. The size of the population was set at 20 times the number of parameters to be estimated and the VaR was initialized at the (in-sample) empirical VaR.

In figure 5 we report a plot of the CAViaR news impact curve for the 5% VaR estimate of S&P 500. Notice how the Adaptive and the Asymmetric Slope news impact curves differ from the others. In particular, the sharp difference between the impact of positive and negative returns in the Asymmetric Slope model suggests that there are relevant asymmetries in the behavior of the 5% quantile of this asset. As we discuss below, other tests confirm this finding also at the 1% quantile for GM and S&P 500.

Our results show that all the models but the Adaptive and the Proportional Symmetric Adaptive perform well according to the number of hits and to the DQ test for all the three assets and at all confidence levels, both in and out-of-sample. The Proportional Symmetric Adaptive model is consistently rejected at the 1% and 5% confidence levels. The graphs reported in figures 2 and 3 give a visual confirmation of the clearly different pattern generated by this model. This seems to us enough evidence to discard the model.

The performance of the Adaptive model is more controversial. This model performs very well if we just look at the number of hits it produces, both in and out-of-sample.

However, the DQ test reveals that these hits tend to be autocorrelated. In other words, the unconditional performance of the Adaptive model is good, but the conditional one might be seriously biased. By eyeballing the graphs of figures 2, 3 and 4, we can infer that the main drawback of the Adaptive model is that it is not flexible enough to adapt to sudden changes in volatility, like the one that occurred in the fall 1987. This defect must be attributed to the simplicity of the model, which depends on only one parameter.

The other four models under study do extremely well for all assets and at all confidence levels. The only exception is the 5% VaR for S&P 500, whose out-of-sample performance is rejected in all the cases. The DQ test reveals that there is some unexplained autocorrelation among the hits. It may be the case that we need to look for other CAViaR specifications that can provide a better fit for the 5% quantile of this asset, or it may simply be a feature of this set of data.

To fully appreciate the performance of the CAViaR models, recall that the samples over which the models are estimated include the crash of October 1987 and that the out-of-sample period includes the days of high volatility of the summer 1998. Moreover, the length of the out of sample period is 500 trading days. This roughly corresponds to two calendar years! It is likely that financial institutions will re-estimate their models on a more frequent basis (monthly, weekly, or even daily) and that this procedure of re-estimation will improve the performance of CAViaR models.

The issue of model selection is a critical one. Ideally, a good model should have stable parameters over time, so that it doesn't need to be re-estimated very often. A model with this feature would very likely have a good performance out-of-sample, which is what practitioners are interested in.

---

[8] The intervals were [0,2] for the 1%, 5% and 25% VaR confidence levels, and [0,4] for the 0.1%.

One possible strategy to choose among the models is to discard all the models rejected by the DQ test, either in-sample or out-of-sample. Among the surviving models, we choose the one with the lowest out-of-sample RQ criterion. We can think of the model with the minimum RQ criterion as the specification closest to the true quantile process. Clearly, using in-sample results would bias the choice towards the largest model. Looking at the out-of-sample values avoids this problem.

An alternative strategy could be to compute an Akaike Information Criterion for CAViaR models and choose the model with the lowest AIC. Clearly, the topic of model selection deserves a more rigorous and systematic treatment, which we leave for future research.

According to the RQ criterion, the Asymmetric Slope model is the best CAViaR specification for GM and S&P 500. Note that the coefficient of the positive slope term is never significantly different from zero for S&P 500, suggesting that there might be no impact news from positive returns.

The best model for IBM, instead, was GARCH at 1% and 25% confidence levels, and the Asymmetric Absolute Value at 5%. This finding can be taken as a further indication that different confidence levels might require different models.

Finally, based on the results of our Monte Carlo simulation, we believe that the estimates for the 0.1% quantile must be taken with extreme caution. Even if the performance in terms of number of hits is acceptable (both in and out-of-sample), it is very challenging to get a reliable estimate of events that should happen only once every 4 years. Such an estimate will be, to say the least, very noisy. We believe that a better strategy to estimate these extreme quantiles might be to incorporate the extreme value theory into the CAViaR modeling approach.

To have a preliminary comparison of the performance of CAViaR models relative to the existing methodologies, we computed the four quantiles of the three assets by estimating a plain GARCH (1,1) using the in-sample daily returns. The quantile was then computed by finding the empirical quantile of the standardized residuals and multiplying it by the square root of the estimated variance. The results are reported in table 9. The overall performance of this approach seems good. The difficulty of getting a good out-of-sample performance of the 5% VaR for S&P 500 is confirmed. The out-of-sample estimates of the 1% VaR for S&P 500 are also rejected at a confidence level of 5%. Note how the overall performance of this procedure is similar to the performance of the Indirect GARCH CAViaR. It is important however to stress that the assumptions of the CAViaR are much weaker, since there is no need to assume that the standardized residuals are i.i.d. like in the GARCH framework.

10. CONCLUSION

We propose a new approach to Value at Risk estimation. All the existing models try to estimate the distribution of the returns and then recover its quantile in an indirect way. On the contrary we try to model directly the quantile. To do this we introduce a new class of models, the Conditional Autoregressive Value at Risk or CAViaR models, which specify the evolution of the quantile over time using a special type of autoregressive process. The parameters of the CAViaR model are estimated by minimizing the regression quantiles objective function. Since this function is not differentiable, we use a genetic algorithm for the numerical optimization. A Monte Carlo experiment shows how both regression quantiles and genetic algorithm are able to produce unbiased estimates. We also introduce a new test based on an artificial regression to evaluate the performance of the CAViaR models. Applications to real data provide empirical support to our methodology and illustrate the ability of CAViaR models to adapt to new risk environments.

REFERENCES

Andrews, D.W.K. (1988), Laws of large numbers for dependent non-identically distributed random variables, *Econometric Theory*, 4: 458-467.

Beder, T. S. (1995), VaR: seductive but dangerous, *Financial Analyst Journal*, Sep-Oct, 12-24.

Boudoukh, J., M. Richardson and R.F. Whitelaw (1998), The best of both worlds, *Risk*, 11: 64-67.

Cowles, A. and H. Jones (1937), 'Some a posteriori probabilities in stock market action', *Econometrica*, 5: 280-294.

Christoffersen, P. F. (1998), Evaluating interval forecasts, *International Economic Review*, 39: 841-862.

Danielsson, J. and C.G. de Vries (1998), Beyond the sample: extreme quantile and probability estimation, London School of Economics, Discussion Paper 298.

Engle, R. F. and V. Ng (1993), Measuring and Testing the Impact of News On Volatility, *Journal of Finance*, 48: 1749-1778.

Foresi, S. and F. Peracchi (1995), The conditional distribution of excess returns: and empirical Analysis, *Journal of the American Statistical Association*, 90: 451-466.

Goldberg, D.E. (1989), *Genetic algorithms in search, optimization, and machine learning*, Reading: Addison-Wesley Publishing Corporation, Inc.

Gourieroux, C. and J. Jasiak (1998), Truncated maximum likelihood, goodness of fit tests and tail analysis, Unpublished manuscript.

Granger, C.W.J., H. White and M. Kamstra (1989), Interval forecasting. An analysis based upon ARCH-quantile estimators, *Journal of Econometrics*, 40: 87-96.

Koenker, R. and G. Bassett (1978), Regression quantiles, *Econometrica*, 46: 33-50.

Ljung, G. and G. Box (1979), 'On a measure of lack of fit in time series models, *Biometrica*, 66: 265-270.

Mood, A. (1940), 'The distribution theory of runs', *Annals of Mathematical Statistics*, 11: 367-392.

Newey, W. and J. Powell (1990), Efficient estimation of linear and type I censored regression models under conditional quantile restrictions, *Econometric Theory*, 6: 295-317.

Powell, J. (1984), Least absolute deviations estimation for the censored regression model, *Journal of Econometrics*, 25: 303-325.

Powell, J. (1986), Censored regression quantiles, *Journal of Econometrics*, 32: 143-155.

Price, K. and R. Storn (1997), Differential Evolution, *Dr. Dobb's Journal*, April, 18-24.

Weiss, A. (1991), Estimating nonlinear dynamic models using least absolute error estimation, *Econometric Theory*, 7: 46-68.

White, H. (1991), Nonparametric estimation of conditional quantiles using neural networks, Unpublished paper.

White, H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge Universty Press.

White, H. and I. Domowitz (1984), Nonlinear regression with dependent observations, *Econometrica*, 52: 143-161.

Figure 1 - Returns for GM, IBM and S&P 500 form April 1986 to April 1999

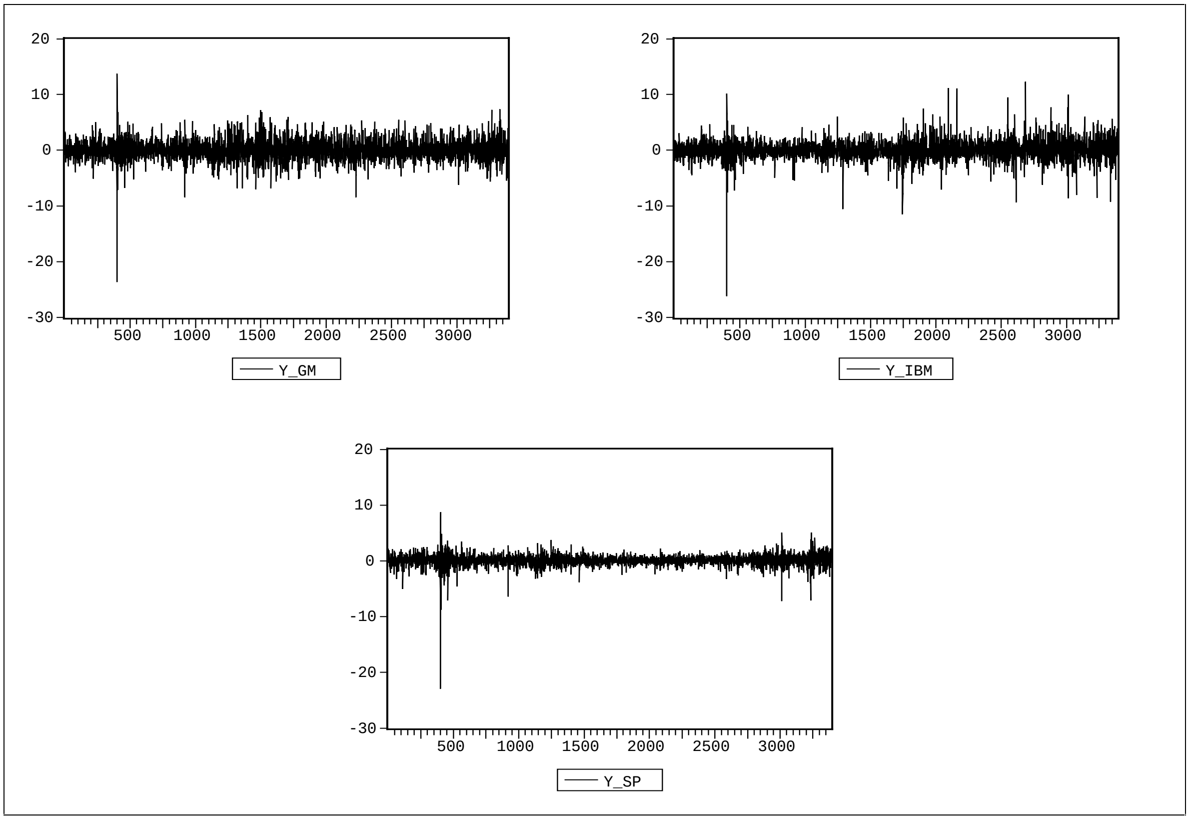

Figure 2 - 5% CAViaR plots for General Motors

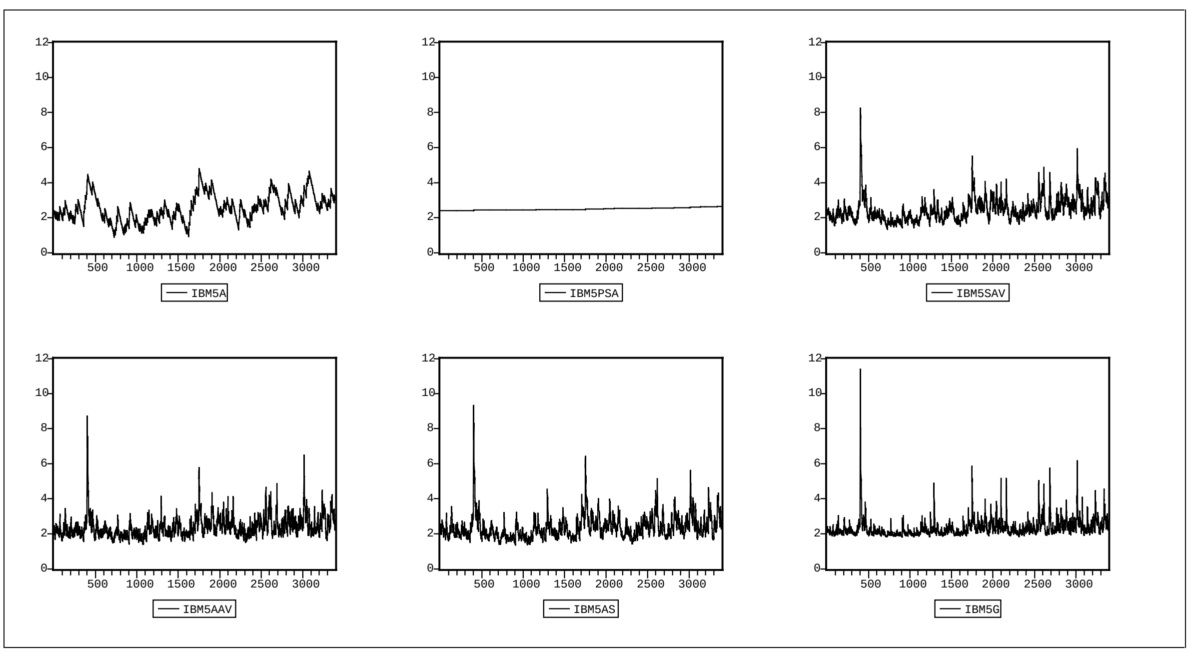

Figure 3 - 5% CAViaR plots for IBM

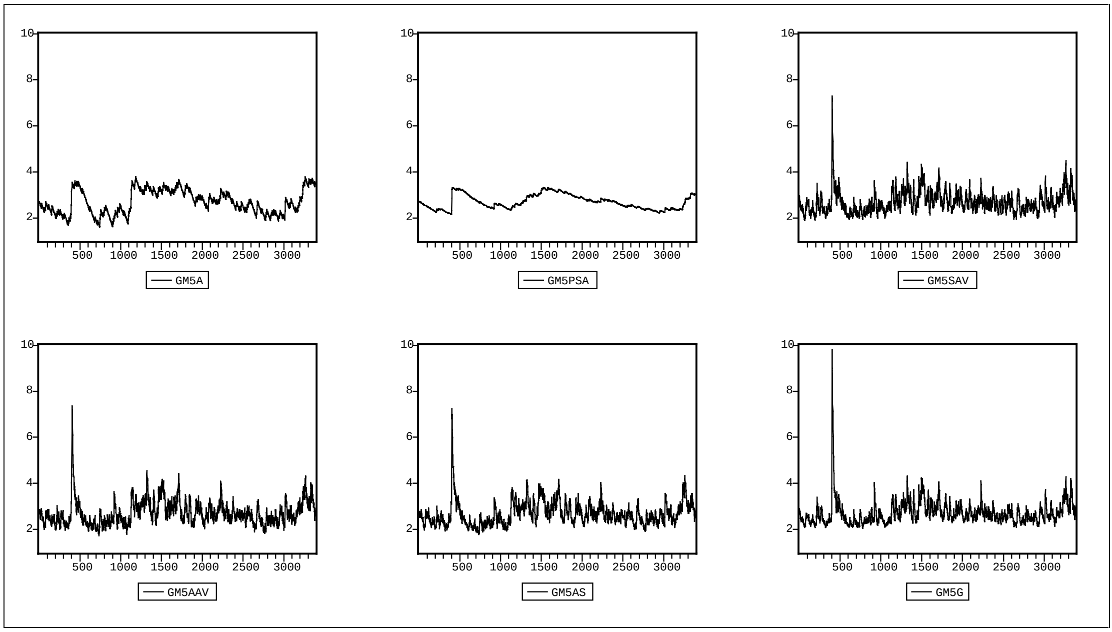

Figure 4 - 5% CAViaR plots for S&P 500

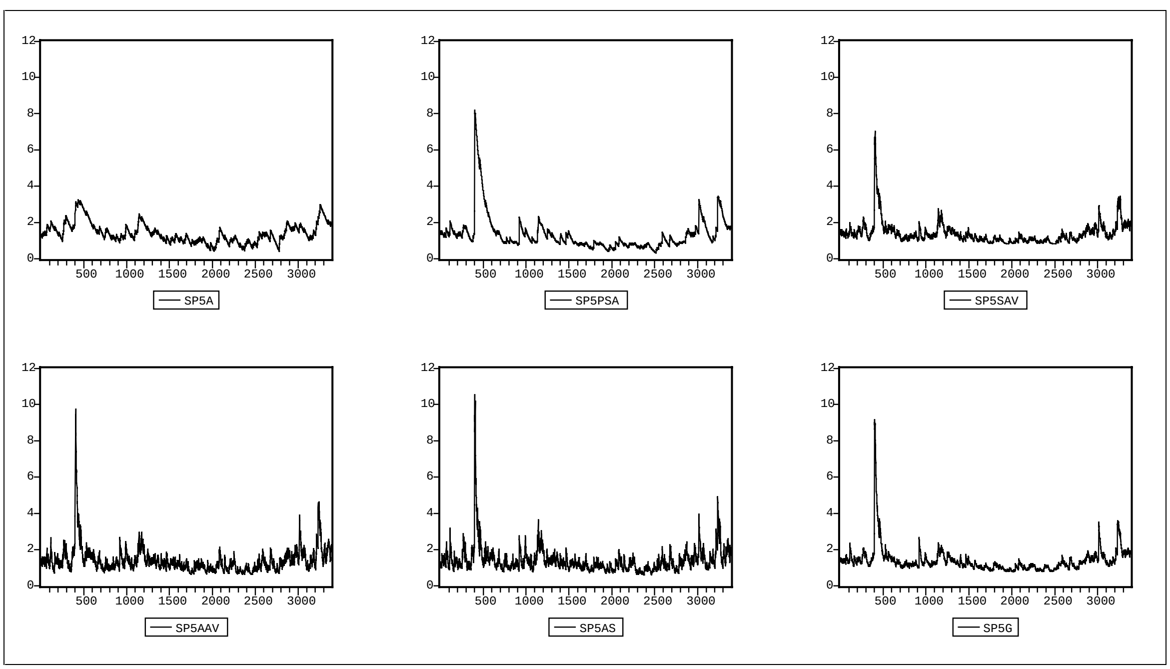

Figure 5 - 5% CAViaR news impact curves for S&P 500

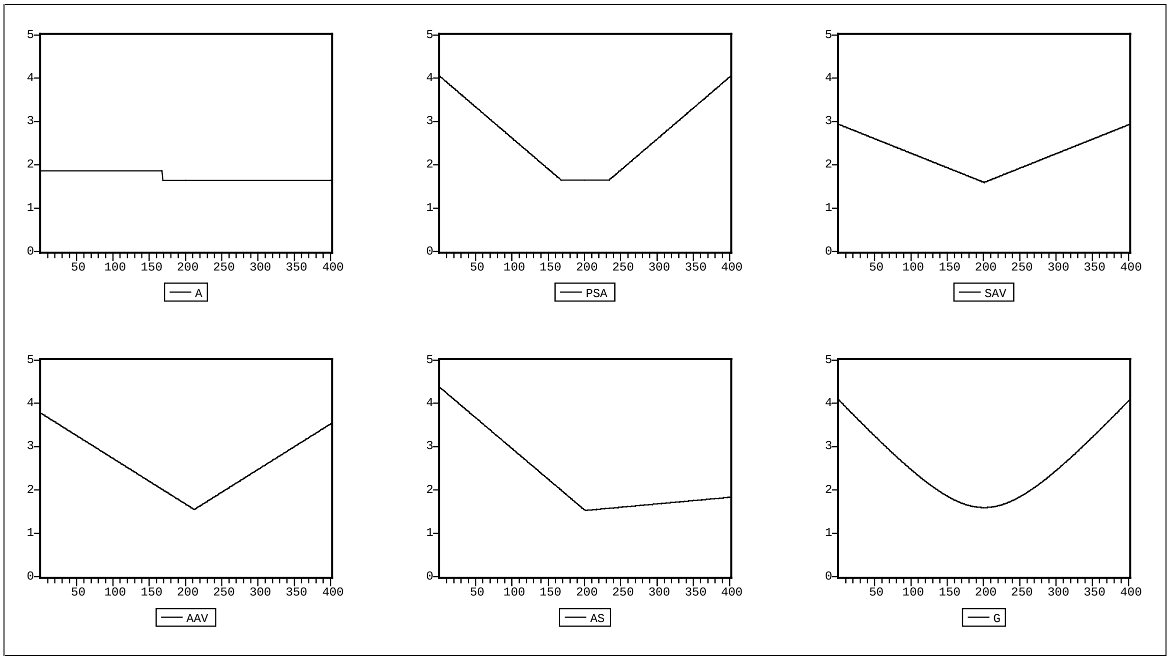

Table 1 - Summary statistics of the Monte Carlo experiment

0.1%	GAMMA1	GAMMA2	GAMMA3
True mean	4.15	0.90	0.69
Mean	7.16	0.80	0.67
t-statistic	*8.54*	*-13.60*	*-0.95*
Median	2.90	0.89	0.53
Var-Cov matrix	125.16	-2.45	2.60
	-2.45	0.05	-0.07
	2.60	-0.07	0.32

1%	GAMMA1	GAMMA2	GAMMA3
True mean	1.62	0.90	0.27
Mean	2.28	0.87	0.30
t-statistic	*7.59*	*-7.91*	*6.01*
Median	1.57	0.90	0.27
Var-Cov matrix	7.79	-0.28	0.19
	-0.28	0.01	-0.01
	0.19	-0.01	0.02

5%	GAMMA1	GAMMA2	GAMMA3
True mean	0.81	0.90	0.14
Mean	1.02	0.88	0.14
t-statistic	*7.59*	*-7.59*	*4.74*
Median	0.81	0.90	0.14
Var-Cov matrix	0.79	-0.06	0.02
	-0.06	0.00	0.00
	0.02	0.00	0.00

25%	GAMMA1	GAMMA2	GAMMA3
True mean	0.13	0.90	0.03
Mean	0.26	0.84	0.03
t-statistic	*9.80*	*-10.12*	*1.58*
Median	0.13	0.90	0.02
Var-Cov matrix	0.17	-0.07	0.00
	-0.07	0.03	0.00
	0.00	0.00	0.00

Table 2 - Monte Carlo summary statistics after excluding the samples with GAMMA2<0.5

0.1%	GAMMA1	GAMMA2	GAMMA3
True mean	4.15	0.90	0.69
Trimmed Mean	4.04	0.87	0.60
Trimmed Median	2.49	0.90	0.50
Trimmed Var-Cov matrix	18.36	-0.40	0.81
	-0.40	0.01	-0.03
	0.81	-0.03	0.23

1%	GAMMA1	GAMMA2	GAMMA3
True mean	1.62	0.90	0.27
Trimmed Mean	2.02	0.88	0.29
Trimmed Median	1.55	0.90	0.27
Trimmed Var-Cov matrix	2.72	-0.11	0.12
	-0.11	0.00	-0.01
	0.12	-0.01	0.02

5%	GAMMA1	GAMMA2	GAMMA3
True mean	0.81	0.90	0.135
Trimmed Mean	0.99	0.89	0.14
Trimmed Median	0.81	0.90	0.14
Trimmed Var-Cov matrix	0.49	-0.04	0.02
	-0.04	0.00	0.00
	0.02	0.00	0.00

25%	GAMMA1	GAMMA2	GAMMA3
True mean	0.13	0.90	0.027
Trimmed Mean	0.18	0.88	0.03
Trimmed Median	0.13	0.90	0.02
Trimmed Var-Cov matrix	0.03	-0.01	0.00
	-0.01	0.01	0.00
	0.00	0.00	0.00

Table 3 - Parameter estimates and relevant statistics for the Adaptive model

ADAPTIVE *** 0.1%	GM	IBM	S&P 500
Gamma 1	**0.0000**	**0.0000**	**2.7409**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	39.40	46.84	33.99
RQ out of sample	4.21	5.33	7.01
Hits in sample (%)	0.0692	0.1037	0.0692
Hits out of sample (%)	0.0000	0.0000	0.4000
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-

ADAPTIVE *** 1%	GM	IBM	S&P 500
Gamma 1	**0.26**	**0.00**	**2.11**
Standard Errors	*0.11*	*0.08*	*0.22*
P-values	*0.01*	*0.50*	*0.00*
RQ in sample	179.66	191.79	114.9
RQ out of sample	29.56	42.11	29.10
Hits in sample (%)	0.90	1.00	1.00
Hits out of sample (%)	1.80	2.00	1.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.00	0.00	0.82
2) [VaR]	0.53	0.98	0.07
3) [c, hit(-1), VaR]	0.68	0.02	0.00
4) [c, hit(-1 to -5), VaR]	0.00	0.00	0.01
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.00	0.00	0.01
2) [VaR]	0.06	0.02	0.79
3) [c, hit(-1), VaR]	0.22	0.30	0.37
4) [c, hit(-1 to -5), VaR]	0.00	0.00	0.01

ADAPTIVE *** 5%	GM	IBM	S&P 500
Gamma 1	**0.22**	**0.44**	**0.23**
Standard Errors	*0.03*	*0.05*	*0.02*
P-values	*0.00*	*0.00*	*0.00*
RQ in sample	553.26	527.45	312.65
RQ out of sample	100.84	120.20	72.41
Hits in sample (%)	4.91	5.01	5.08
Hits out of sample (%)	6.40	5.20	5.00
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.31	0.47	0.46
2) [VaR]	0.52	0.34	0.48
3) [c, hit(-1), VaR]	0.12	0.01	0.10
4) [c, hit(-1 to -5), VaR]	0.06	0.01	0.07
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.40	0.98	0.01
2) [VaR]	0.26	0.90	0.80
3) [c, hit(-1), VaR]	0.40	0.21	0.55
4) [c, hit(-1 to -5), VaR]	0.45	0.56	0.01

ADAPTIVE *** 25%	GM	IBM	S&P 500
Gamma 1	**0.021**	**0.012**	**0.017**
Standard Errors	*0.004*	*0.003*	*0.003*
P-values	*0.000*	*0.000*	*0.000*
RQ in sample	1507	1368	752
RQ out of sample	291.62	312.44	184
Hits in sample (%)	24.86	25.31	25.07
Hits out of sample (%)	27.00	24.80	27.40
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.94	0.25	0.59
2) [VaR]	0.73	0.80	0.68
3) [c, hit(-1), VaR]	0.58	0.64	0.30
4) [c, hit(-1 to -5), VaR]	0.81	0.23	0.32
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.57	0.67	0.45
2) [VaR]	0.43	0.97	0.29
3) [c, hit(-1), VaR]	0.23	0.28	0.39
4) [c, hit(-1 to -5), VaR]	0.64	0.30	0.41

Table 4 - Parameter estimates and relevant statistics for the Proportional Symmetric Adaptive model

PROP SYM ADAPT *** 0.1%	GM	IBM	S&P 500
Gamma 1	**13.3772**	**0.0000**	**15.6603**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 2	**0.0049**	**0.0001**	**0.0017**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	24.77	46.52	21.43
RQ out of sample	6.63	4.78	4.08
Hits in sample (%)	0.07	0.10	0.10
Hits out of sample (%)	0.40	0.20	0.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-

PROP SYM ADAPT *** 1%	GM	IBM	S&P 500
Gamma 1	**0.0594**	**0.0023**	**0.0116**
Standard Errors	*0.0365*	*0.0275*	*0.0074*
P-values	*0.0519*	*0.4674*	*0.0588*
Gamma 2	**0.0004**	**0.0000**	**0.0002**
Standard Errors	*0.0002*	*0.0003*	*0.0001*
P-values	*0.0326*	*0.5000*	*0.0035*
RQ in sample	180.21	191.67	123.27
RQ out of sample	30.08	40.86	34.43
Hits in sample (%)	1.00	0.93	1.28
Hits out of sample (%)	2.20	1.80	5.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.02	0.00	0.00
2) [VaR]	0.98	0.73	0.12
3) [c, hit(-1), VaR]	0.62	0.01	0.03
4) [c, hit(-1 to -5), VaR]	0.03	0.00	0.00
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.00	0.01	0.00
2) [VaR]	0.00	0.07	0.00
3) [c, hit(-1), VaR]	0.01	0.24	0.00
4) [c, hit(-1 to -5), VaR]	0.00	0.01	0.00

PROP SYM ADAPT *** 5%	GM	IBM	S&P 500
Gamma 1	**0.0261**	**0.0007**	**0.2868**
Standard Errors	*0.0086*	*0.0047*	*0.0504*
P-values	*0.0013*	*0.4413*	*0.0000*
Gamma 2	**0.0012**	**0.0000**	**0.0108**
Standard Errors	*0.0004*	*0.0002*	*0.0016*
P-values	*0.0007*	*0.5000*	*0.0000*
RQ in sample	559.88	542.11	322.33
RQ out of sample	102.62	118.52	77.25
Hits in sample (%)	4.94	4.60	6.98
Hits out of sample (%)	6.80	6.60	5.60
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.01	0.00	0.00
2) [VaR]	0.85	0.35	0.21
3) [c, hit(-1), VaR]	0.41	0.21	0.00
4) [c, hit(-1 to -5), VaR]	0.02	0.00	0.00
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.03	0.49	0.01
2) [VaR]	0.10	0.09	0.66
3) [c, hit(-1), VaR]	0.04	0.26	0.01
4) [c, hit(-1 to -5), VaR]	0.05	0.59	0.00

PROP SYM ADAPT *** 25%	GM	IBM	S&P 500
Gamma 1	**0.0148**	**0.0001**	**0.0011**
Standard Errors	*0.0043*	*0.0006*	*0.0004*
P-values	*0.0003*	*0.4501*	*0.0025*
Gamma 2	**0.0110**	**0.0000**	**0.0009**
Standard Errors	*0.0030*	*0.0003*	*0.0002*
P-values	*0.0001*	*0.5000*	*0.0001*
RQ in sample	1507.1	1369.8	753.5
RQ out of sample	292.63	312.01	189.2
Hits in sample (%)	25.52	24.24	25.17
Hits out of sample (%)	25.20	25.40	34.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.54	0.43	0.64
2) [VaR]	0.95	0.37	0.93
3) [c, hit(-1), VaR]	0.03	0.35	0.85
4) [c, hit(-1 to -5), VaR]	0.06	0.33	0.68
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.91	0.63	0.00
2) [VaR]	0.62	0.74	0.00
3) [c, hit(-1), VaR]	0.30	0.87	0.00
4) [c, hit(-1 to -5), VaR]	0.62	0.69	0.00

Table 5 - Parameter estimates and relevant statistics for the Symmetric Absolute Value model

SYM ABS VALUE *** 0.1%	GM	IBM	S&P 500
Gamma 1	**6.7241**	**1.2797**	**1.3240**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 2	**0.0000**	**0.7139**	**0.6360**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 3	**2.4084**	**3.0075**	**2.5713**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	28.63	35.71	18.17
RQ out of sample	4.64	7.95	3.78
Hits in sample (%)	0.10	0.07	0.10
Hits out of sample (%)	0.00	0.00	0.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
LM test for VaR(t-2)	-	-	-

SYM ABS VALUE *** 1%	GM	IBM	S&P 500
Gamma 1	**0.4784**	**0.1260**	**0.2040**
Standard Errors	*0.2199*	*0.0930*	*0.1490*
P-values	*0.0148*	*0.0876*	*0.0856*
Gamma 2	**0.8168**	**0.9477**	**0.8733**
Standard Errors	*0.0690*	*0.0332*	*0.0869*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.3503**	**0.1130**	**0.3817**
Standard Errors	*0.1216*	*0.0665*	*0.2711*
P-values	*0.0020*	*0.0445*	*0.0796*
RQ in sample	172.12	182.46	109.66
RQ out of sample	28.99	40.81	26.23
Hits in sample (%)	1.00	0.97	0.97
Hits out of sample (%)	1.20	1.60	1.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.60	0.25	0.81
2) [VaR]	0.97	0.88	0.83
3) [c, hit(-1), VaR]	0.96	0.58	0.96
4) [c, hit(-1 to -5), VaR]	0.71	0.33	0.88
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	1.00	0.05	0.05
2) [VaR]	0.88	0.22	0.18
3) [c, hit(-1), VaR]	0.48	0.42	0.15
4) [c, hit(-1 to -5), VaR]	0.92	0.05	0.03
LM test for VaR(t-2)	0.92	0.94	0.97

SYM ABS VALUE *** 5%	GM	IBM	S&P 500
Gamma 1	**0.1865**	**0.1192**	**0.0512**
Standard Errors	*0.0897*	*0.0462*	*0.0217*
P-values	*0.0188*	*0.0050*	*0.0092*
Gamma 2	**0.8936**	**0.9053**	**0.9369**
Standard Errors	*0.0441*	*0.0294*	*0.0250*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.1135**	**0.1481**	**0.1339**
Standard Errors	*0.0428*	*0.0437*	*0.0549*
P-values	*0.0040*	*0.0004*	*0.0074*
RQ in sample	551.02	522.58	306.51
RQ out of sample	99.23	120.47	73.85
Hits in sample (%)	4.98	4.98	4.98
Hits out of sample (%)	4.60	6.00	5.60
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.45	0.10	0.45
2) [VaR]	1.00	1.00	0.95
3) [c, hit(-1), VaR]	0.85	0.65	0.99
4) [c, hit(-1 to -5), VaR]	0.56	0.15	0.55
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.89	0.23	0.00
2) [VaR]	0.55	0.55	0.88
3) [c, hit(-1), VaR]	0.95	0.05	0.29
4) [c, hit(-1 to -5), VaR]	0.94	0.09	0.00
LM test for VaR(t-2)	0.90	0.93	0.95

SYM ABS VALUE *** 25%	GM	IBM	S&P 500
Gamma 1	**0.0440**	**0.0235**	**0.0111**
Standard Errors	*0.0250*	*0.0138*	*0.0056*
P-values	*0.0393*	*0.0445*	*0.0234*
Gamma 2	**0.9272**	**0.9587**	**0.9469**
Standard Errors	*0.0350*	*0.0211*	*0.0247*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.0318**	**0.0197**	**0.0338**
Standard Errors	*0.0139*	*0.0093*	*0.0163*
P-values	*0.0112*	*0.0166*	*0.0193*
RQ in sample	1500.61	1363.02	746.90
RQ out of sample	289.58	311.80	183.89
Hits in sample (%)	25.03	25.03	24.97
Hits out of sample (%)	26.00	23.20	27.60
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.65	0.74	0.53
2) [VaR]	0.92	0.95	0.99
3) [c, hit(-1), VaR]	0.98	0.88	0.99
4) [c, hit(-1 to -5), VaR]	0.75	0.83	0.64
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.85	0.59	0.43
2) [VaR]	0.72	0.41	0.33
3) [c, hit(-1), VaR]	0.66	0.82	0.30
4) [c, hit(-1 to -5), VaR]	0.91	0.70	0.39
LM test for VaR(t-2)	0.80	0.89	0.48

Table 6 - Parameter estimates and relevant statistics for the Asymmetric Absolute Value model

ASYM ABS VALUE *** 0.1%	GM	IBM	S&P 500
Gamma 1	**0.9233**	**1.0330**	**0.5634**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 2	**0.7078**	**0.6884**	**0.7888**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 3	**1.4313**	**2.4417**	**1.9190**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 4	**0.8916**	**0.8468**	**0.6005**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	25.26	31.40	17.91
RQ out of sample	5.19	6.33	3.86
Hits in sample (%)	0.10	0.10	0.14
Hits out of sample (%)	0.20	0.00	0.00
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-

ASYM ABS VALUE *** 1%	GM	IBM	S&P 500
Gamma 1	**0.4304**	**0.2746**	**0.1776**
Standard Errors	*0.1340*	*0.1421*	*0.1033*
P-values	*0.0007*	*0.0266*	*0.0428*
Gamma 2	**0.8271**	**0.8942**	**0.8631**
Standard Errors	*0.0428*	*0.0475*	*0.0537*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.3597**	**0.2191**	**0.3766**
Standard Errors	*0.0826*	*0.0931*	*0.1402*
P-values	*0.0000*	*0.0093*	*0.0036*
Gamma 4	**0.1877**	**0.0855**	**0.6402**
Standard Errors	*0.1738*	*0.3125*	*0.1530*
P-values	*0.1401*	*0.3922*	*0.0000*
RQ in sample	170.01	181.63	105.63
RQ out of sample	28.83	40.20	23.75
Hits in sample (%)	1.04	0.86	1.07
Hits out of sample (%)	1.20	1.60	1.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.62	0.23	0.64
2) [VaR]	0.86	0.49	0.65
3) [c, hit(-1), VaR]	0.95	0.38	0.89
4) [c, hit(-1 to -5), VaR]	0.73	0.30	0.74
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	1.00	0.05	0.05
2) [VaR]	0.83	0.20	0.11
3) [c, hit(-1), VaR]	0.66	0.52	0.29
4) [c, hit(-1 to -5), VaR]	0.97	0.07	0.07

ASYM ABS VALUE *** 5%	GM	IBM	S&P 500
Gamma 1	**0.0908**	**0.1800**	**0.0582**
Standard Errors	*0.0542*	*0.0628*	*0.0291*
P-values	*0.0470*	*0.0021*	*0.0228*
Gamma 2	**0.9275**	**0.8591**	**0.9059**
Standard Errors	*0.0284*	*0.0358*	*0.0271*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.1031**	**0.1763**	**0.2105**
Standard Errors	*0.0366*	*0.0413*	*0.0581*
P-values	*0.0024*	*0.0000*	*0.0001*
Gamma 4	**0.7512**	**0.6940**	**0.5681**
Standard Errors	*0.3121*	*0.1754*	*0.1002*
P-values	*0.0080*	*0.0000*	*0.0000*
RQ in sample	547.52	518.24	300.95
RQ out of sample	99.26	118.80	72.84
Hits in sample (%)	4.94	5.05	4.94
Hits out of sample (%)	5.00	7.00	6.20
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.97	0.47	0.68
2) [VaR]	0.93	0.85	0.89
3) [c, hit(-1), VaR]	0.98	0.93	0.96
4) [c, hit(-1 to -5), VaR]	0.99	0.58	0.76
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.92	0.16	0.00
2) [VaR]	0.82	0.08	0.65
3) [c, hit(-1), VaR]	0.96	0.06	0.02
4) [c, hit(-1 to -5), VaR]	0.95	0.16	0.00

ASYM ABS VALUE *** 25%	GM	IBM	S&P 500
Gamma 1	**0.0500**	**0.0193**	**0.0121**
Standard Errors	*0.0291*	*0.0127*	0.0064
P-values	*0.0427*	*0.0648*	0.0303
Gamma 2	**0.9196**	**0.9566**	**0.9433**
Standard Errors	*0.0399*	*0.0189*	0.0279
P-values	*0.0000*	*0.0000*	0.0000
Gamma 3	**0.0345**	**0.0237**	**0.0366**
Standard Errors	*0.0156*	*0.0092*	0.0188
P-values	*0.0135*	*0.0052*	0.0260
Gamma 4	**0.0274**	**0.6532**	**0.0490**
Standard Errors	*0.4417*	*0.3670*	0.1592
P-values	*0.4753*	*0.0375*	0.3790
RQ in sample	1500.30	1361.60	746.19
RQ out of sample	289.50	310.92	183.51
Hits in sample (%)	24.86	25.21	24.93
Hits out of sample (%)	26.00	23.60	26.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.66	0.89	0.55
2) [VaR]	0.93	0.79	0.98
3) [c, hit(-1), VaR]	0.98	0.92	0.95
4) [c, hit(-1 to -5), VaR]	0.75	0.94	0.66
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.85	0.56	0.30
2) [VaR]	0.72	0.57	0.54
3) [c, hit(-1), VaR]	0.66	0.94	0.49
4) [c, hit(-1 to -5), VaR]	0.91	0.68	0.31

Table 7 - Parameter estimates and relevant statistics for the Asymmetric Slope model

ASYM SLOPE *** 0.1%	GM	IBM	S&P 500
Gamma 1	**2.7753**	**1.0863**	**0.4325**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 2	**0.4342**	**0.6587**	**0.6871**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 3	**0.6130**	**1.1402**	**1.8655**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 4	**2.0416**	**2.8743**	**2.2849**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	25.01	29.27	18.15
RQ out of sample	4.15	5.93	3.65
Hits in sample (%)	0.10	0.10	0.14
Hits out of sample (%)	0.00	0.00	0.00
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
LM test for VaR(t-2)	-	-	-

ASYM SLOPE *** 1%	GM	IBM	S&P 500
Gamma 1	**0.3928**	**0.0572**	**0.1473**
Standard Errors	*0.2216*	*0.0580*	*0.0833*
P-values	*0.0381*	*0.1623*	*0.0385*
Gamma 2	**0.7983**	**0.9427**	**0.8699**
Standard Errors	*0.0676*	*0.0227*	*0.0484*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.2725**	**0.0512**	**0.0001**
Standard Errors	*0.1148*	*0.0616*	*0.1168*
P-values	*0.0088*	*0.2029*	*0.4997*
Gamma 4	**0.4437**	**0.2474**	**0.5045**
Standard Errors	*0.1589*	*0.1006*	*0.2403*
P-values	*0.0026*	*0.0070*	*0.0179*
RQ in sample	169.30	179.54	105.84
RQ out of sample	28.48	40.54	22.69
Hits in sample (%)	1.00	0.97	0.97
Hits out of sample (%)	1.40	1.60	1.60
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.60	0.81	0.56
2) [VaR]	0.98	0.89	0.96
3) [c, hit(-1), VaR]	0.96	0.96	0.94
4) [c, hit(-1 to -5), VaR]	0.71	0.88	0.68
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.96	0.05	0.05
2) [VaR]	0.46	0.21	0.13
3) [c, hit(-1), VaR]	0.67	0.53	0.45
4) [c, hit(-1 to -5), VaR]	0.97	0.07	0.07
LM test for VaR(t-2)	0.92	0.92	0.96

ASYM SLOPE *** 5%	GM	IBM	S&P 500
Gamma 1	**0.0704**	**0.0951**	**0.0410**
Standard Errors	*0.0425*	*0.0444*	*0.0221*
P-values	*0.0488*	*0.0161*	*0.0316*
Gamma 2	**0.9353**	**0.8916**	**0.9026**
Standard Errors	*0.0222*	*0.0272*	*0.0239*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.0411**	**0.0597**	**0.0307**
Standard Errors	*0.0285*	*0.0335*	*0.0469*
P-values	*0.0745*	*0.0372*	*0.2565*
Gamma 4	**0.1182**	**0.2110**	**0.2841**
Standard Errors	*0.0399*	*0.0558*	*0.0895*
P-values	*0.0015*	*0.0001*	*0.0008*
RQ in sample	548.63	515.72	300.76
RQ out of sample	99.20	121.05	72.05
Hits in sample (%)	4.98	4.91	4.98
Hits out of sample (%)	5.20	7.40	6.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.83	0.74	0.69
2) [VaR]	0.98	0.87	0.94
3) [c, hit(-1), VaR]	0.97	0.97	0.64
4) [c, hit(-1 to -5), VaR]	0.89	0.82	0.74
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.92	0.03	0.00
2) [VaR]	0.97	0.06	0.20
3) [c, hit(-1), VaR]	0.96	0.00	0.13
4) [c, hit(-1 to -5), VaR]	0.95	0.01	0.00
LM test for VaR(t-2)	0.96	0.77	0.94

ASYM SLOPE *** 25%	GM	IBM	S&P 500
Gamma 1	**0.0404**	**0.0125**	**0.0014**
Standard Errors	*0.0298*	*0.0104*	*0.0047*
P-values	*0.0877*	*0.1151*	*0.3820*
Gamma 2	**0.9132**	**0.9605**	**0.9481**
Standard Errors	*0.0393*	*0.0169*	*0.0212*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.0415**	**0.0108**	**0.0288**
Standard Errors	*0.0193*	*0.0098*	*0.0192*
P-values	*0.0157*	*0.1349*	*0.0664*
Gamma 4	**0.0290**	**0.0297**	**0.0288**
Standard Errors	*0.0170*	*0.0127*	*0.0175*
P-values	*0.0441*	*0.0097*	*0.0502*
RQ in sample	1500.88	1360.53	746.90
RQ out of sample	289.41	311.51	183.48
Hits in sample (%)	25.00	25.14	24.93
Hits out of sample (%)	25.60	23.40	25.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.69	0.83	0.49
2) [VaR]	0.97	0.85	0.93
3) [c, hit(-1), VaR]	0.97	0.90	0.99
4) [c, hit(-1 to -5), VaR]	0.79	0.90	0.60
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.88	0.64	0.29
2) [VaR]	0.88	0.45	0.77
3) [c, hit(-1), VaR]	0.67	0.79	0.67
4) [c, hit(-1 to -5), VaR]	0.94	0.70	0.32
LM test for VaR(t-2)	0.99	0.89	0.60

Table 8 - Parameter estimates and relevant statistics for the Indirect GARCH model

GARCH *** 0.1%	GM	IBM	S&P 500
Gamma 1	**36.8548**	**2.7007**	**3.6423**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 2	**0.0000**	**0.4997**	**0.5719**
Standard Errors	-	-	-
P-values	-	-	-
Gamma 3	**10.2205**	**30.3776**	**14.6629**
Standard Errors	-	-	-
P-values	-	-	-
RQ in sample	26.20	33.80	18.22
RQ out of sample	4.14	7.31	3.59
Hits in sample (%)	0.10	0.07	0.10
Hits out of sample (%)	0.00	0.00	0.00
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	-	-	-
2) [VaR]	-	-	-
3) [c, hit(-1), VaR]	-	-	-
4) [c, hit(-1 to -5), VaR]	-	-	-

GARCH *** 1%	GM	IBM	S&P 500
Gamma 1	**1.4965**	**1.3288**	**0.2329**
Standard Errors	*0.7152*	*0.8869*	*0.1552*
P-values	*0.0182*	*0.0670*	*0.0668*
Gamma 2	**0.7803**	**0.8740**	**0.8350**
Standard Errors	*0.0630*	*0.0728*	*0.0651*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.9363**	**0.3374**	**1.0575**
Standard Errors	*0.2906*	*0.2238*	*0.4964*
P-values	*0.0006*	*0.0659*	*0.0166*
RQ in sample	171.04	183.49	108.33
RQ out of sample	29.36	40.17	24.99
Hits in sample (%)	0.97	1.04	1.04
Hits out of sample (%)	1.20	1.60	1.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.56	0.28	0.82
2) [VaR]	0.87	0.77	0.76
3) [c, hit(-1), VaR]	0.96	0.64	0.92
4) [c, hit(-1 to -5), VaR]	0.67	0.36	0.87
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.99	0.05	0.05
2) [VaR]	0.87	0.21	0.16
3) [c, hit(-1), VaR]	0.61	0.51	0.24
4) [c, hit(-1 to -5), VaR]	0.96	0.06	0.05

GARCH *** 5%	GM	IBM	S&P 500
Gamma 1	**0.3336**	**0.6529**	**0.0262**
Standard Errors	*0.1788*	*0.2184*	*0.0153*
P-values	*0.0310*	*0.0014*	*0.0428*
Gamma 2	**0.9042**	**0.7930**	**0.9287**
Standard Errors	*0.0390*	*0.0582*	*0.0250*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.1220**	**0.1784**	**0.1407**
Standard Errors	*0.0530*	*0.0594*	*0.0591*
P-values	*0.0107*	*0.0013*	*0.0086*
RQ in sample	552.31	524.86	305.83
RQ out of sample	99.79	119.66	74.08
Hits in sample (%)	4.98	5.01	5.05
Hits out of sample (%)	4.60	7.60	5.80
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.31	0.34	0.39
2) [VaR]	0.93	0.98	0.88
3) [c, hit(-1), VaR]	0.98	0.84	1.00
4) [c, hit(-1 to -5), VaR]	0.32	0.39	0.50
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.89	0.03	0.00
2) [VaR]	0.55	0.02	0.77
3) [c, hit(-1), VaR]	0.94	0.01	0.11
4) [c, hit(-1 to -5), VaR]	0.93	0.04	0.00

GARCH *** 25%	GM	IBM	S&P 500
Gamma 1	**0.0505**	**0.0203**	**0.0019**
Standard Errors	*0.0316*	*0.0123*	*0.0014*
P-values	*0.0549*	*0.0501*	*0.0890*
Gamma 2	**0.8921**	**0.9570**	**0.9429**
Standard Errors	*0.0511*	*0.0225*	*0.0296*
P-values	*0.0000*	*0.0000*	*0.0000*
Gamma 3	**0.0183**	**0.0056**	**0.0077**
Standard Errors	*0.0093*	*0.0031*	*0.0049*
P-values	*0.0241*	*0.0363*	*0.0576*
RQ in sample	1500.42	1365.02	747.31
RQ out of sample	289.59	310.51	183.65
Hits in sample (%)	25.03	25.03	25.03
Hits out of sample (%)	26.20	24.20	28.00
DQ in sample (p-values)			
1) [c, hit(-1 to -5)]	0.56	0.58	0.31
2) [VaR]	0.94	0.96	0.93
3) [c, hit(-1), VaR]	0.94	0.92	0.97
4) [c, hit(-1 to -5), VaR]	0.68	0.70	0.40
DQ out of sample (p-values)			
1) [c, hit(-1 to -5)]	0.81	0.75	0.15
2) [VaR]	0.64	0.86	0.18
3) [c, hit(-1), VaR]	0.70	0.78	0.37
4) [c, hit(-1 to -5), VaR]	0.88	0.77	0.20

Table 9 - Plain GARCH(1,1) estimates

GM			
0.10%	1%	5%	25%
Hit in Sample			
0.07	0.97	4.98	24.97
Hit out of Sample			
0.00	1.20	4.40	26.20
DQ in Sample			
1.00	0.03	0.88	0.54
0.86	0.92	0.76	0.78
0.55	0.94	0.64	0.66
0.91	0.05	0.78	0.53
DQ out of Sample			
1.00	0.99	0.96	0.73
0.49	0.82	0.40	0.72
1.00	0.60	0.84	0.40
1.00	0.96	0.98	0.81

IBM			
0.10%	1%	5%	25%
Hit in Sample			
0.07	0.97	4.98	24.97
Hit out of Sample			
0.00	1.60	5.80	23.00
DQ in Sample			
1.00	0.81	0.46	0.98
0.76	0.60	0.62	0.31
0.90	0.71	0.24	0.03
1.00	0.77	0.32	0.20
DQ out of Sample			
1.00	0.05	0.16	0.52
0.50	0.23	0.78	0.31
1.00	0.52	0.07	0.78
1.00	0.06	0.06	0.61

S&P 500			
0.10%	1%	5%	25%
Hit in Sample			
0.07	0.97	4.98	24.97
Hit out of Sample			
0.20	1.40	6.00	28.00
DQ in Sample			
1.00	0.81	0.32	0.45
0.88	0.83	0.60	0.67
0.59	0.96	0.77	0.84
0.94	0.88	0.27	0.47
DQ out of Sample			
1.00	0.03	0.00	0.13
0.57	0.51	0.68	0.23
0.89	0.71	0.09	0.31
1.00	0.04	0.00	0.16