CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*

Robert F. Engle
University of California, San Diego and New York University
rengle@stern.nyu.edu

and

Simone Manganelli
European Central Bank
simone.manganelli@ecb.int

First Draft: July 1999
This Draft: July 2002

Abstract – Value at Risk (VaR) has become the standard measure of market risk employed by financial institutions for both internal and regulatory purposes. VaR is defined as the value that a portfolio will lose with a given probability, over a certain time horizon (usually one or ten days). Interpreting the VaR as the quantile of future portfolio values conditional on current information, we propose a new approach to quantile estimation that does not require any of the extreme assumptions invoked by existing methodologies (such as normality or i.i.d. returns). The Conditional Autoregressive Value at Risk or CAViaR model moves the focus of attention from the distribution of returns directly to the behavior of the quantile. We specify the evolution of the quantile over time using an autoregressive process and use the regression quantile framework introduced by Koenker and Bassett to determine the unknown parameters. Utilizing the criterion that each period the probability of exceeding the VaR must be independent of all the past information, we introduce a new test of model adequacy, the Dynamic Quantile (DQ) test. Applications to real data provide empirical support to this methodology.

Keywords: Non-linear Regression Quantile, Value at Risk, Risk Management

* We would like to thank Hal White, Moshe Buchinsky, Wouter Den Haan, Joel Horowitz, Jim Powell, the participants of the UCSD, Yale, MIT-Harvard, Wharton, Iowa, and NYU econometric seminars, for their valuable comments.

CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles

Abstract – Value at Risk (VaR) has become the standard measure of market risk employed by financial institutions for both internal and regulatory purposes. VaR is defined as the value that a portfolio will lose with a given probability, over a certain time horizon (usually one or ten days). Interpreting the VaR as the quantile of future portfolio values conditional on current information, we propose a new approach to quantile estimation that does not require any of the extreme assumptions invoked by existing methodologies (such as normality or i.i.d. returns). The Conditional Autoregressive Value at Risk or CAViaR model moves the focus of attention from the distribution of returns directly to the behavior of the quantile. We specify the evolution of the quantile over time using an autoregressive process and use the regression quantile framework introduced by Koenker and Bassett to determine the unknown parameters. Utilizing the criterion that each period the probability of exceeding the VaR must be independent of all the past information, we introduce a new test of model adequacy, the Dynamic Quantile (DQ) test. Applications to real data provide empirical support to this methodology.

Keywords: Non-linear Regression Quantile, Value at Risk, Risk Management

1. Introduction

The importance of effective risk management has never been greater. Recent financial disasters have emphasized the needs for accurate risk measures for financial institutions. As the nature of the risks has changed over time, the methods to measure them must adapt to recent experience. The use of quantitative risk measures has become an essential management tool to be placed in parallel with 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.

Value at Risk (VaR) has become the standard measure of market 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, for a given confidence level. 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 monetary amount. The 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 is simply a particular quantile of future portfolio values, 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 conditional quantiles. The problem is to forecast a value each period that will be exceeded with probability (1-*q*) by the current portfolio, where $q \in (0,1)$ represents the confidence level associated to the VaR. Let $\{y_t\}_{t=1}^{T}$ denote the time series of portfolio returns and *T* the sample size. We want to find VaR_t such that $\Pr[y_t < -VaR_t \mid \Omega_t] = q$, where Ω_t denotes the information set at the end of time *t*-1. Any reasonable methodology 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 are estimated using Koenker and Bassett's (1978) regression quantile framework. Consistency and asymptotic results build on existing contribution of the regression quantile literature. We propose a new test, the Dynamic Quantile (DQ) test, which can be interpreted as an overall goodness of fit test for the estimated CAViaR processes. This test, which has been independently derived by Chernozhukov (1999), is new in the literature on regression quantiles.

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 sections 4, we review the literature on regression quantiles. Section 5 introduces the DQ test. Section 6 presents an empirical application to real data. Section 7 concludes the paper.

2. Value at Risk Models

VaR was developed at the beginning of the 90's in the financial industry to provide senior management a single number that could quickly and easily incorporate information about the risk of a portfolio. Today it is part of every risk manager's toolbox. Indeed, VaR can help management estimate the cost of positions in terms of risk, allowing them to allocate risk in a more efficient way. Also the Basel Committee on Banking Supervision (1996) at the Bank for International Settlements uses VaR to 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 less restrictive financial regulations.

The existing models for calculating VaR differ in many aspects. Nevertheless they all follow a common structure, which can be summarized in three points: 1) the portfolio is marked-to-market daily, 2) the distribution of the portfolio returns is estimated, 3) the VaR of the portfolio is computed. The main differences among VaR models are related to the second point. VaR methodologies can be classified initially into two broad categories: a) factor models such as RiskMetrics (1996), 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. 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 (see, for example, Danielsson and de Vries (2000)). 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 (2000), 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. Recently, McNeil and Frey (2000) suggested fitting a GARCH model to the time series of returns and then applying the extreme value theory to the standardized residuals, which are assumed to be i.i.d. Although it is an improvement over existing applications, this approach still suffers from the same criticism applied to the volatility models.

3. CAViaR

We propose another approach to quantile estimation. Instead of modeling the whole distribution, we model directly the quantile. The empirical fact that volatilities of stock market returns cluster over time may be translated in statistical words by saying that their distribution is 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).

Suppose we observe a vector of portfolio returns $\{y_t\}_{t=1}^{T}$. Let q be the probability associated with VaR, x_t a vector of time t observable variables and b_q a p-vector of unknown parameters. Finally, let $f_t(b) \equiv f(x_{t-1}, b_q)$ denote the time t q-quantile of the distribution of portfolio returns formed at time t-1, where we suppressed the q subscript from b_q for notational convenience. A very general CAViaR specification might be the following:

$$(1) \qquad f_t(b) = g_0 + \sum_{i=1}^{q} g_i f_{t-i}(b) + \sum_{i=1}^{p} a_i l(x_{t-i}, j)$$

where $b' = (a', g', j')$ and l is a function of a finite number of lagged values of observables. The autoregressive terms $g_i f_{t-i}(b)$, i=1,…,q ensure that the quantile changes "smoothly" over time. The role of $l(x_{t-i}, j)$, is to link $f_t(b)$ to observable variables that belong to the information set. This term thus has much the same role as the News Impact Curve for GARCH models introduced by Engle and Ng (1993). A natural choice for x_{t-1} is lagged returns. Indeed, we would expect the VaR 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 $|y_{t-1}|$.

Here are some examples of CAViaR processes that will be estimated. Throughout we use the notation $(x)^+ = \max(x,0)$, $(x)^- = -\min(x,0)$.

ADAPTIVE: $$f_t(b_1) = f_{t-1}(b_1) + b_1\left\{\left[1+\exp\left(G\left[y_{t-1} - f_{t-1}(b_1)\right]\right)\right]^{-1} - q\right\}$$

where G is some positive finite number. Note that as $G \to \infty$, the last term converges almost surely to $b_1[I(y_{t-1} \le f_{t-1}(b_1)) - q]$, where $I(\cdot)$ represents the indicator function; for finite G this model is a smoothed version of a step function. The adaptive model incorporates the following 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 little from returns which are close to the VaR or which are extremely positive. It increases the VaR by the same amount regardless of whether the returns exceeded the VaR by a small or a large margin. This model has a unit coefficient on the lagged VaR.

Other alternatives are:

SYMMETRIC ABSOLUTE VALUE: $f_t(b) = b_1 + b_2 f_{t-1}(b) + b_3 |y_{t-1}|$

ASYMMETRIC SLOPE: $f_t(b) = b_1 + b_2 f_{t-1}(b) + b_3 (y_{t-1})^{+} + b_4 (y_{t-1})^{-}$

INDIRECT GARCH(1,1): $f_t(b) = \left(b_1 + b_2 f_{t-1}^2(b) + b_3 y_{t-1}^2\right)^{1/2}$

The first and third respond symmetrically to past returns while the second allows the response to positive and negative returns to be different. All three are mean reverting in the sense that the coefficient on the lagged VaR is not constrained to be one.

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. The Symmetric Absolute Value and Asymmetric Slope quantile specifications, would be correctly specified by a GARCH process in which the standard deviation, rather than the variance, is modeled either symmetrically or asymmetrically with i.i.d. errors. This model has been introduced and estimated by Taylor (1986) and Schwert (1988) and is analyzed by Engle (2002). The CAViaR specifications are however more general than these GARCH models. Various forms of non-i.i.d. error distributions can be modeled in this way. In fact these models can be used for situations with constant volatilities, but changing error distributions, or situations where both error densities and volatilities are changing.

4. Regression Quantiles

The parameters of CAViaR models are estimated by regression quantiles, as introduced by Koenker and Bassett (1978). Koenker and Bassett show how to extend the notion of a sample quantile to a linear regression model. Consider a sample of observations $y_1,\ldots,y_T$ generated by the following model:

$$(2) \qquad y_t = x_t'\beta^0 + \varepsilon_{\theta t} \qquad Quant_\theta(\varepsilon_{\theta t}|x_t) = 0$$

where x_t is a p-vector of regressors and $Quant_\theta(\varepsilon_{\theta t}|x_t)$ is the θ-quantile of $\varepsilon_{\theta t}$ conditional on x_t. Let $f_t(\beta) \equiv x_t\beta$. Then the θ^{th} regression quantile is defined as any $\hat{\beta}$ that solves:

$$(3) \qquad \min_\beta \frac{1}{T}\left[\theta - I\left(y_t < f_t(\beta)\right)\right]\left[y_t - f_t(\beta)\right]$$

Regression quantiles include as a special case the least absolute deviation (LAD) model. It is well known that LAD is more robust than OLS estimators whenever the errors have a fat tailed distribution. Koenker and Bassett (1978), for example, ran a simple Monte Carlo experiment and showed 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.[1]

Analysis of linear regression quantile models has been extended to cases with heteroskedastic (Koenker and Bassett (1982)) and non-stationary dependent errors (Portnoy (1991)), time series models (Bloomfield and Steiger (1983)), simultaneous equations models (Amemiya (1982) and Powell (1983)) and censored regression models (Powell (1986) and Buchinsky and Hahn (1998)). Extensions to the autoregressive

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

quantiles have been proposed by Koenker and Zhao (1996), and Koul and Saleh (1995). These approaches differ from the one proposed in this paper in that all the variables are observable and the models are linear in the parameters. In the nonlinear case, asymptotic theory for models with serially independent (but not identically distributed) errors have been proposed, for example, by Oberhofer (1982), Dupacova (1987), Powell (1991) and Jureckova and Prochazka (1993). There is relatively little literature that considers non-linear quantile regressions in the context of time series. The most important contributions are those by White (1994, Corollary 5.12), who proves the consistency of the nonlinear regression quantile, both in the i.i.d. and stationary dependent cases, and by Weiss (1991), who shows consistency, asymptotic normality and asymptotic equivalence of LM and Wald tests for LAD estimators for nonlinear dynamic models. Finally, Mukherjee (1999) extends the concept of regression and autoregression quantiles to nonlinear time series models with i.i.d. error terms.

Consider the model

(4) $$y_t = f\left(y_{t-1}, x_{t-1}, \ldots, y_1, x_1; b^0\right) + e_{tq} \qquad Quant_q\left(e_{tq} \mid \Omega_t\right) = 0$$

$$\equiv f_t(b^0) + e_{tq} \qquad t = 1, \ldots, T$$

where $f_1(b^0)$ is some given initial condition, x_t is a vector of exogenous or predetermined variables, $b^0 \in \Re^p$ is the vector of true unknown parameters that need to be estimated and $\Omega_t = [y_{t-1}, x_{t-1}, \ldots, y_1, x_1, f_1(b^0)]$ is the information set available at time t. Let $\hat{b}$ be the parameter vector that minimizes (3).

Theorems 1 and 2 below show that the non-linear regression quantile estimator $\hat{b}$ is consistent and asymptotically normal. Theorem 3 provides a consistent estimator of the variance-covariance matrix. In Appendix A, we give sufficient conditions on f in (4), together with technical assumptions, for these results to hold. The proofs are simple extensions of Weiss (1991) and Powell (1984, 1986, 1991) and will be omitted. We denote the conditional density of e_{tq} evaluated at 0 by $h_t(0|\Omega_t)$, the (1,*p*) gradient of $f_t(b)$ by $\nabla f_t(b)$ and define $\nabla f(b)$ to be a (*T*,*p*) matrix with typical row $\nabla f_t(b)$.

Theorem 1 (Consistency) - *In model (4), under C0-C7 in Appendix A,* $\hat{b} \xrightarrow{p} b^0$, *where* $\hat{b}$ *is the solution to (3).*

Theorem 2 (Asymptotic Normality) - *In model (4), under AN1-AN4 in Appendix A and the conditions of Theorem 1,* $\sqrt{\frac{T}{q(1-q)}} A_T^{-1/2} D_T \left(\hat{b} - b^0\right) \xrightarrow{d} N(0,I)$, *where* $A_T = E\left[T^{-1}\nabla' f(b^0)\nabla f(b^0)\right]$, $D_T = E\left[T^{-1}\nabla' f(b^0) H \nabla f(b^0)\right]$ *and* H *is a diagonal matrix with typical entry* $h_t(0|\Omega_t)$.

Theorem 3 (Variance-Covariance Matrix Estimation) - *Under VC1-VC2 in Appendix A and the conditions of Theorems 1 and 2,* $\hat{A}_T \xrightarrow{p} A_T$ *and* $\hat{D}_T \xrightarrow{p} D_T$, *where* $\hat{A}_T = T^{-1}\nabla' f(\hat{b})\nabla f(\hat{b})$ *and* $\hat{D}_T = (2T\hat{c}_T)^{-1}\sum_{t=1}^{T} I\left(\left|y_t - f_t(\hat{b})\right| < \hat{c}_T\right)\nabla' f_t(\hat{b})\nabla f_t(\hat{b})$.

Derivation of the asymptotic distribution builds on the approximation of the discontinuous gradient of the objective function with a smooth differentiable function, so that the usual Taylor expansion can be performed. The device to obtain such an approximation is provided by an extension of Huber's (1967) Theorem 3.[2] This technique is standard in the regression quantile and LAD literature (see Powell 1984, 1991 and Weiss 1991). The interested reader is referred to these papers for further details.

Regarding the variance-covariance matrix, note that $\hat{A}_T$ is simply the outer product of the gradient. Estimation of the D_T matrix is less straightforward, as it involves the $h_t(0\,|\,\Omega_t)$

[2] Alternative strategies to derive the asymptotic distribution are the one suggested by Amemiya (1982), based on the approximation of the regression quantile objective function by a continuously differentiable function, and the one based on the empirical processes approach as suggested, for example in van de Geer (2000).

term. Following Powell (1984, 1986, 1991), we propose an estimator that combines kernel density estimation with the heteroskedasticity-consistent covariance matrix estimator of White (1980). Our theorem 3 is a generalisation of Powell's (1991) theorem 3 that accommodates the (stationary) nonlinear dependent case.

5. Testing quantile models

If model (4) is the true DGP, then $\Pr\left(y_t < f_t(b^0)\right) = q, \ \forall t$. This is equivalent to requiring that the sequence of indicator functions $\left\{I\left(y_t < f_t(b^0)\right)\right\}_{t=1}^{T}$ 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 $\left\{I\left(y_t < f_t(b^0)\right)\right\}_{t=1}^{T} \equiv \{I_t\}_{t=1}^{T}$ is i.i.d., as done for example by Granger, White and Kamstra (1989) and Christoffersen (1998). Although these tests can detect the presence of serial correlation in the sequence of indicator functions $\{I_t\}_{t=1}^{T}$, this is only a necessary, but not sufficient condition to assess the performance of a quantile model. 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) $$z_t = \begin{cases} 1 & \text{with probability } q \\ -1 & \text{with probability } (1-q) \end{cases}$$

Then, setting $f_t(b^0) = Kz_t$, for K large, will do the job. Notice however, that once z_t is observed, the probability of exceeding the quantile is known to be almost zero or one. Thus the unconditional probabilities are correct and serially uncorrelated, but the conditional probabilities given the quantile are not. This example is an extreme case of quantile measurement error. Any noise introduced into the quantile estimate will change the conditional probability of a hit *given* the estimate itself.

Therefore, none of these tests has power against this form of misspecification and none can be simply extended to examine other explanatory variables. We propose a new test that can be easily extended to incorporate a variety of alternatives. Define:

(6) $$Hit_t(b^0) \equiv I\left(y_t < f_t(b^0)\right) - q \ .$$

The $Hit_t(b^0)$ function assumes value (1-q) every time y_t is less than the quantile and $-q$ otherwise. Clearly the expected value of $Hit_t(b^0)$ is zero. Furthermore, from the definition of the quantile function, the conditional expectation of $Hit_t(b^0)$ given any information known at t-1 must also be zero. In particular, $Hit_t(b^0)$ must be uncorrelated with its own lagged values and with $f_t(b^0)$, and must have expected value equal to zero. If $Hit_t(b^0)$ 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 exceptions. If it is desired to check whether there is the right proportion of hits in each calendar year, then this can be measured by checking the correlation of $Hit_t(b^0)$ 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.

Let's denote with T the number of in sample observations and with N the number of out of sample observations. Let's make explicit the dependence of the relevant variables on the number of observations, using appropriate subscripts. A natural way to set up a test is to check whether the test statistic $X_N'(\hat{b}_T)Hit_N(\hat{b}_T)$ is significantly different from zero, where $X_n(\hat{b})$, $n = T+1,\ldots, T+N$, the typical row of $X_N(\hat{b})$ (possibly depending on $\hat{b}$), is a q-vector measurable-Ω_n and $Hit_N(\hat{b}) = \left[Hit_{T+1}(\hat{b}),\ldots, Hit_{T+N}(\hat{b})\right]'$. A simple out of sample version of the Dynamic Quantile (DQ) test statistic would be:

(7) $$DQ_0 \equiv \frac{Hit_N'(\hat{b}_T)X_N(\hat{b}_T)\left[X_N'(\hat{b}_T)X_N(\hat{b}_T)\right]^{-1}X_N'(\hat{b}_T)Hit_N(\hat{b}_T)}{q(1-q)} \underset{\substack{T\to\infty \\ N\to\infty}}{\overset{d}{\to}} c_q^2$$

provided that $X_N'(\hat{b}_T)X_N(\hat{b}_T)$ is nonsingular. The limit for $T\to\infty$ is required to ensure that $\hat{b}_T \overset{p}{\to} b^0$. Than a simple application of a suitable central limit theorem yields the result. While this measure of performance can be quite useful, when applied in sample, its distribution is affected by the fact that $Hit_T(\hat{b}_T)$ is a function of estimated parameters.

Let $M_T = \left(X_T'(b^0) - E\left[T^{-1}X_T'(b^0)H_T\nabla f_T(b^0)\right]D_T^{-1}\nabla' f_T(b^0)\right)$, where H has been defined in theorem 2. Theorem 4 provides the correct distribution of the in-sample DQ test.

Theorem 4 (In-Sample Dynamic Quantile Test) - *Under DQ2-DQ4 in Appendix A and the same conditions of theorems 1-2,* $T^{-1/2}X_T'(\hat{b})Hit_T(\hat{b}) \overset{d}{\sim} N\left(0, q(1-q)E\left(T^{-1}M_T M_T'\right)\right)$. *If also Assumption DQ1 holds,*

$$DQ_1 \equiv \frac{T^{-1}Hit_T'(\hat{b})X_T(\hat{b})E\left(T^{-1}M_T M_T'\right)^{-1}X_T'(\hat{b})Hit_T(\hat{b})}{q(1-q)} \overset{d}{\sim} c_q^2.$$ *Moreover, under the conditions of theorem 3,* $T^{-1}\hat{M}_T\hat{M}_T' \overset{p}{\to} E\left(T^{-1}M_T M_T'\right)$,

where $\hat{M}_T = X_T'(\hat{b}) - \left\{(2T\hat{c}_T)^{-1}\sum_{t=1}^{T} I\left(\left|y_t - f_t(\hat{b})\right| < \hat{c}_T\right)X_t'(\hat{b})\nabla f_t(\hat{b})\right\}\hat{D}_T^{-1}\nabla' f_T(\hat{b})$.[3]

[3] If $X_T(b)$ contains $m<q$ lagged $Hit_{t-i}(b)$ ($i = 1,\ldots, m$), then $X_T(b)$, $Hit_T(b)$, H_T and $\nabla f_T(b)$ are not conformable, as $X_T(b)$ contains only (T-m) elements. Here we implicitly assume, without loss of generality,

Proof – See Appendix B.

Note that if we choose $X_T(\hat{b}) = \nabla f_T(\hat{b})$, then M_T = **0**, where **0** is a (p, p) matrix of zeroes. This is consistent with the fact that $T^{-1/2}\nabla' f_T(\hat{b}) Hit_T(\hat{b}) = o_p(1)$, by the first order conditions of the regression quantile framework.

The in-sample DQ test is a specification test for the particular CAViaR process under study and it can be very useful for model selection purposes. The simpler version of the out-of-sample DQ test (7), instead, can be used by regulators to check whether the VaR estimates submitted by a financial institution satisfy some basic requirements every good quantile estimates must have, such as unbiasedeness, independent hits and independence of the quantile estimates. The nicest features of the out-of-sample DQ test are its simplicity and the fact that it does not depend on the estimation procedure: to implement it, the evaluator (either the regulator or the risk manager) just needs a sequence of VaR's and the corresponding values of the portfolio.

6. 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 computed the daily returns as 100 times the difference of the log of the prices. The samples range from April 7 1986 to April 7 1999. We used the first 2892 observations to estimate the model and the last 500 for out-of-sample testing. We estimated 1% and 5% one day VaR, using the four CAViaR specifications described in

---

that the matrices are made conformable by deleting the first m rows of $\nabla f_T(b)$ and $Hit_T(b)$, and the first m rows and columns of H_T.

section 3.[4] Note that all the models considered in section 3 satisfy assumptions C1, C7 and AN1 of Appendix A. In particular, all the models are both continuous and continuously differentiable in b. The others are technical assumptions that are impossible to verify in finite samples. The 5% VaR estimates for GM are plotted in Figure 1, and all the results are reported in Table 1. 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 (equation 3), the percentage of times the VaR is exceeded, and the p-value of the Dynamic Quantile test, both in and out-of-sample. In order to compute the VaR series with the CAViaR models, we initialize $f_1(b)$ to the empirical q-quantile of the first 300 observations. The instruments used in the out-of-sample DQ test were a constant, the VaR forecast and the first four lagged hits. For the in-sample DQ test, we didn't include the constant and the VaR forecast because for some models there was collinearity with the matrix of derivatives.[5] The standard errors and the variance covariance matrix of the in-sample DQ test were computed as described in theorems 3 and 4. The formulae to compute $\hat{D}_T$ and $\hat{M}_T$ were implemented using k-nearest neighbor estimators, with k = 40 for 1% VaR and k = 60 for 5% VaR.

As optimization routines, we used the Nelder-Mead Simplex Algorithm and a quasi-Newton method.[6] The models were optimized using the following procedure. We generated n vectors using a uniform random number generator between 0 and 1. We computed the regression quantile (RQ) function described in equation (3) for each of these vectors and selected the m vectors that produced the lowest RQ criterion as initial values

[4] For the Adaptive model, we set G=10, where G entered the definition of the Adaptive model in section 3.

[5] Lagged hit variables contain the indicator function. Since the indicator function is Lipschitz continuous, it satisfies condition DQ3 of theorem 4.

[6] All the computations were done in Matlab 6.1. We used the functions *fminsearch* and *fminunc* as optimization algorithms. The loops to compute the recursive quantile functions were coded in *C*.

for the optimization routine.[7] For each of these initial values, we ran first the simplex algorithm. We then fed the optimal parameters to the quasi-Newton algorithm and the new optimal parameters were chosen as the new initial conditions for the simplex. We repeated this procedure until the convergence criterion was satisfied.[8] Finally, we selected the vector that produced the lowest RQ criterion. An alternative optimization routine is the interior point algorithm for nonlinear regression quantiles suggested by Koenker and Park (1996).

In Figure 2 we report a plot of the CAViaR news impact curve for the 1% VaR estimates of S&P 500. Notice how the Adaptive and the Asymmetric Slope news impact curves differ from the others. For both Indirect GARCH and Symmetric Absolute Value models, past returns (either positive or negative) have a symmetric impact on VaR. In the case of the Adaptive model, instead, the most important news is whether past returns exceeded the previous VaR estimate or not. Finally, the sharp difference between the impact of positive and negative returns in the Asymmetric Slope model suggests that there might be relevant asymmetries in the behavior of the 1% quantile of this portfolio.

Turning now our attention to Table 1, the first striking result is that the coefficient of the autoregressive term (b_2) is always very significant. This confirms that the phenomenon of clustering of volatilities is relevant also in the tails. A second interesting point is the precision of all the models, as measured by the percentage of in-sample hits. This is not surprising as the objective function of regression quantile models is designed exactly to achieve this kind of result. The results for the 1% VaR show that the Symmetric Absolute Value, the Asymmetric Slope and the Indirect GARCH models do a good job at describing the evolution of the left tail for the three assets under study. The results are particularly good for GM, producing a rather accurate percentage of hits out-of-sample

[7] We set $n = [10^4, 10^5, 10^4, 10^4]$ and $m = [10, 15, 10, 5]$ respectively for the symmetric absolute value, asymmetric slope, Indirect GARCH and the adaptive models.

(1.4% for the Symmetric Absolute Value and the Asymmetric Slope and 1.2% for the Indirect GARCH). The performance of the Adaptive model is inferior both in and out-of-sample, even though the percentage of hits seems reasonably close to 1. This shows that looking only at the number of exceptions (as suggested by the Basle Committee on Banking Supervision (1996)) may be a very unsatisfactory way of evaluating the performance of a VaR model. 5% results present a different picture. All the models perform well with GM. Note the remarkable precision of the percentage of out-of-sample hits generated by the Asymmetric Slope model (5.0%). Notice further that this time also the Adaptive model is not rejected by the DQ tests. For IBM, instead, the Asymmetric Slope, of which the Symmetric Absolute Value is a special case, tends to overfit in-sample, providing a very poor performance out-of-sample. Finally, for S&P 500 5% VaR, only the Adaptive model survives the DQ test at 1% confidence level, producing a rather accurate number of out-of-sample hits (4.6%). The poor out-of-sample performance of the other models can probably be explained by the fact that the last part of the sample of S&P 500 is characterized by a sudden spur of volatility and roughly coincides with our out-of-sample period. Finally it is interesting to note in the Asymmetric Slope model how the coefficients of the negative part of lagged returns are always strongly significant, while those associated to positive returns are sometimes not significantly different from zero. This indicates the presence of strong asymmetric impacts on VaR of lagged returns.

The fact that the DQ tests select different models for different confidence levels suggests that the process governing the tail behavior might change as we move further out in the tail. In particular, this contradicts the assumption behind GARCH and RiskMetrics, since these approaches implicitly assume that the tails follow the same process as the rest of the returns. While GARCH might be a useful model to describe the evolution of

---

[8] Tolerance levels for the function and the parameters values were set to 10^{-10}.

volatility, the results in this paper show that it might provide an unsatisfactory approximation when applied to tail estimation.

7. Conclusion

We proposed a new approach to Value at Risk estimation. Most existing methods estimate the distribution of the returns and then recover its quantile in an indirect way. On the contrary we modeled directly the quantile. To do this we introduced 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 unknown parameters were estimated by minimizing the regression quantiles loss function. We also introduced the Dynamic Quantile test, a new test to evaluate the performance of quantile models. Applications to real data illustrated the ability of CAViaR models to adapt to new risk environments. Moreover, our findings suggested that the process governing the behavior of the tails might be different from that of the rest of the distribution.

Appendix A: Assumptions

Consistency Assumptions

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

C1. The function $f_t(b): \Re^{k_t} \times B \to \Re$ is such that for each $b \in B$, a compact subset of $\Re^p$, $f_t(b)$ is measurable with respect to the information set Ω_t and $f_t(\cdot)$ is continuous in B, t=1,2,… for a given choice of explanatory variables $y_{t-1}, x_{t-1}, \ldots, y_1, x_1$.

C2. Conditional on all the past information Ω_t, the error terms e_{tq} form a stationary process, with continuous conditional density $h_t(e \mid \Omega_t)$, and continuous joint density $h_t^{e,\Omega}(e, y, x)$.

C3. There exists h>0 such that for all t $h_t(0 \mid \Omega_t) \geq h$.

C4. $|f_t(b)| < K(\Omega_t)$ for each $b \in B$ and for all t, where $K(\Omega_t)$ is some (possibly) stochastic function of variables that belong to the information set, such that $E\left(|K(\Omega_t)|^r\right) \leq K_0 < \infty$, r>1 for some constant K_0.

C5. $E(|e_{tq}|) < \infty$ for all t.

C6. $\{[q - I(y_t < f_t(b))][y_t - f_t(b)]\}$ obeys the law of large numbers.

C7. For every d>0, there exists a t>0 such that if $\left\|b - b^0\right\| \geq d$, $\liminf_{T\to\infty} T^{-1} \sum P\left[\left|f_t(b) - f_t(b^0)\right| > t\right] > 0$.

Asymptotic Normality Assumptions

AN1. $f_t(b)$ is differentiable in B and for all b and g in a neighborhood u_0 of b^0, such that $\|b - g\| \leq d$ for d sufficiently small and for all t:

(i) $\|\nabla f_t(b)\| \leq F(\Omega_t)$, where $F(\Omega_t)$ is some (possibly) stochastic function of variables that belong to the information set and $E\left(|F(\Omega_t)|^3\right) \leq F_0 < \infty$, for some constant F_0

(ii) $\nabla f_t(b)$ satisfies the Lipschitz condition $\|\nabla f_t(b) - \nabla f_t(g)\| \leq M \|b - g\|$, where M<¥

AN2. (i) $h_t(e \mid \Omega_t) \leq N < \infty$, $\forall\ t$

(ii) $h_t(e \mid \Omega_t)$ satisfies the Lipschitz condition $|h_t(l_1 \mid \Omega_t) - h_t(l_2 \mid \Omega_t)| \leq L |l_1 - l_2|$, where L<∞, ∀ t.

AN3. There exists $d>0$, such that for any $b \in B$ and T sufficiently large $\det\left(T^{-1} \sum_{t=1}^{T} E[\nabla' f_t(b) \nabla f_t(b)]\right) > d$.

AN4. $\{[q - I(y_t < f_t(b))]\nabla' f_t(b)\}$ satisfies a central limit theorem.

Variance Covariance Matrix Estimation Assumptions.

VC1. $\hat{c}_T / c_T \xrightarrow{p} 1$, where the nonstochastic sequence c_T satisfies c_T=o(1) and $c_T^{-1} = o(T^{1/2})$.

VC2. $E\left(|F(\Omega_t)|^4\right) \leq F_1 < \infty$, for all t and for some constant F_1, where $F(\Omega_t)$ has been defined in AN1 (i).

Dynamic Quantile Test Assumption

DQ1 - $\lim_{T\to\infty} E\left(T^{-1} M_T M_T'\right)$ converges to a non-singular matrix, where $M_T = \left(X'(b^0) - E\left(T^{-1} X'(b^0) H \nabla f(b^0)\right) D_T^{-1} \nabla' f(b^0)\right)$.

DQ2 – $X_t(b)$ is different element wise from $\nabla f_t(b)$, is measurable-Ω_t, $\|X_t(b)\| \leq W(\Omega_t)$, where $W(\Omega_t)$ is some (possibly) stochastic function of variables that belong to the information set, $E\left[\left|W(\Omega_t)[F(\Omega_t)]^2\right|\right] < \infty$ and $F(\Omega_t)$ has been defined in AN1(i).

DQ3 – $X_t(b)$ satisfies the Lipschitz condition $\|X_t(b) - X_t(g)\| \leq S \|b - g\|$, where S<∞.

DQ4 – $\left\{\left[X_t'(b) - \left[T^{-1} X_t'(b) h_t(0|\Omega_t) \nabla f_t(b)\right] D_T^{-1} \nabla' f_t(b)\right] Hit_t(b)\right\}$ satisfies a central limit theorem.

Appendix B: Proofs

Proof of Theorem 4 – To simplify the notation, in the proof we drop the dependence of $\hat{b}_T$, $X_T(\hat{b})$ and $Hit_T(\hat{b})$ on the (in sample) size *T*. Approximate the discontinuous function $Hit_t(\hat{b})$ with $Hit_t^{\oplus}(\hat{b}) = \left[1+\exp\left\{\hat{e}_t / \hat{c}_T\right\}\right]^{-1} - q \equiv I^*(\hat{e}_t) - q$, where $\hat{e}_t \equiv y_t - f_t(\hat{b})$ and $\hat{c}_T$ is defined in VC1. Then, $\nabla_b Hit_t^{\oplus}(\hat{b}) = \frac{1}{\hat{c}_T}\exp\left\{\hat{e}_t / \hat{c}_T\right\}\left[1+\exp\left\{\hat{e}_t / \hat{c}_T\right\}\right]^{-2} \nabla f_t(\hat{b}) \equiv k_c(\hat{b})\nabla f_t(\hat{b})$. Note that $k_c(\hat{b})$ is the pdf of a logistic with zero mean and parameter $\hat{c}_T$ and hence it is a kernel of order 2. In matrix form we write: $\nabla_b Hit^{\oplus}(\hat{b}) = K(\hat{b})\nabla f(\hat{b})$, where $K(\hat{b})$ is a diagonal matrix with typical entry $k_c(\hat{e}_t)$. Rewriting the DQ test in terms of this approximation and applying the mean value theorem:

(A1) $$T^{-1/2} X'(\hat{b}) Hit^{\oplus}(\hat{b}) = T^{-1/2} X'(\hat{b})\left[Hit^{\oplus}(b^0) + K(\hat{b})\nabla f(\ddot{b}) D_T^{-1} D_T(\hat{b} - b^0)\right]$$

where D_T *has been defined in Theorem 2, and* $\ddot{b}$ lies between $\hat{b}$ and b^0. In proving the asymptotic normality result, one can show that

$$T^{1/2} D_T(\hat{b} - b^0) = -T^{-1/2}\nabla' f(b^0) Hit(b^0) + o_p(1)$$

Substituting in (A1):

$$T^{-1/2} X'(\hat{b}) Hit^{\oplus}(\hat{b}) = T^{-1/2} X'(\hat{b})\left[Hit^{\oplus}(b^0) - T^{-1} K(\hat{b})\nabla f(\ddot{b}) D_T^{-1}\nabla' f(b^0) Hit(b^0)\right] + o_p(1)$$

$$= T^{1/2}\Big\{ T^{-1}\left[X'(b^0) - T^{-1}X'(b^0) H\nabla f(b^0) D_T^{-1}\nabla' f(b^0)\right] Hit(b^0) +$$
$$+ T^{-1}\left(X'(\hat{b}) Hit^{\oplus}(b^0) - X'(b^0) Hit^{\oplus}(b^0)\right) +$$
$$+ T^{-1}\left[\left(T^{-1}X'(\hat{b})K(\hat{b})\nabla f(\ddot{b})\right) - \left(T^{-1}X'(b^0)K(\hat{b})\nabla f(b^0)\right)\right] D_T^{-1}\nabla' f(b^0) Hit(b^0) +$$
$$+ T^{-1}\left(X'(b^0) Hit^{\oplus}(b^0) - X'(b^0) Hit(b^0)\right) +$$
$$T^{-1}\left[\left(T^{-1}X'(b^0)K(\hat{b})\nabla f(b^0)\right) - \left(T^{-1}X'(b^0)H\nabla f(b^0)\right)\right] D_T^{-1}\nabla' f(b^0) Hit(b^0)\Big\} + o_p(1)$$

Since the terms in the last four rows are all $o_p(1)$, application of a suitable central limit theorem gives the desired result. Here we show only that the last two terms of the above equation are $o_p(1)$, as the others are easily checked, using assumptions AN1, DQ2 and DQ3.

To prove that $Hit^{\oplus}(b^0) \xrightarrow{p} Hit(b^0)$, note that since $I^*(|e_{tq}|) = 1 - I^*(-|e_{tq}|)$, we have for each *t*:

$$\left|Hit_t^{\oplus}(b^0) - Hit_t(b^0)\right| = \left|I^*(|e_{tq}|)\left[I(e_{tq} \geq 0) - I(e_{tq} < 0)\right]\right|$$
$$= I^*(|e_{tq}|)\left[I(|e_{tq}| \geq T^{-d}) + I(|e_{tq}| < T^{-d})\right] \equiv C + D$$

where d is a positive number such that $(\hat{c}_T T^d)^{-1} \to \infty$. For D simply note that $\Pr\left(I(|e_{tq}| < T^{-d}) = 1\right) = O_p(T^{-d}) = o_p(1)$. For *C* note that $I^*(|e_{tq}|)$ is decreasing in $|e_{tq}|$. Hence:

$$C \equiv I^*(|e_{tq}|) I\left(|e_{tq}| \geq T^{-d}\right) \leq I^*(T^{-d}) = \left[1+\exp\left\{(\hat{c}_T T^d)^{-1}\right\}\right]^{-1} \to 0$$

Let's look at the other term. We need to show that $T^{-1}X'(b^0)K(\hat{b})\nabla f(b^0) \xrightarrow{p} T^{-1}X'(b^0)H\nabla f(b^0)$.

We show first that $K(\hat{b}) \xrightarrow{p} K(b^0)$. Applying the mean value theorem to each term of the diagonal matrix $K(\hat{b})$, we get $k_c(\hat{e}_t) = k_c(e_{tq}) + \frac{\partial}{\partial b} k_c\left(e_{tq} - d_t(\ddot{b})\right)(\hat{b} - b^0)$, where $d(\ddot{b}) \equiv f_t(\ddot{b}) - f_t(b^0)$ and $\ddot{b}$ lies between $\hat{b}$ and b^0. Since $\left|\frac{\partial}{\partial b} k_c\left(e_{tq} - d_t(\ddot{b})\right)\right| < \infty$ and $(\hat{b} - b^0) = o_p(1)$, we have $k_c(\hat{e}_t) = k_c(e_{tq}) + o_p(1)$.

It remains to show that $\left(T^{-1}X'(b^0)K(b^0)\nabla f(b^0)\right) - \left(T^{-1}X'(b^0)H\nabla f(b^0)\right) = o_p(1)$. Rewrite this term as:

(A2) $$T^{-1}\sum_{t=1}^{T}\left[k_c(e_{tq})-E\left(k_c(e_{tq})|\Omega_t\right)\right]X'_t(b^0)\nabla f_t(b^0)+$$
$$+T^{-1}\sum_{t=1}^{T}\left[E\left(k_c(e_{tq})|\Omega_t\right)-h_t\left(0|\Omega_t\right)\right]X'_t(b^0)\nabla f_t(b^0)$$

First we show that the expected value of $k_c(e_{tq})$ given Ω_t is equal to $h_t\left(0|\Omega_t\right)$. Let $k(u)\equiv e^u\left[1+e^u\right]^{-2}$. Then:

$$E\left[k_c(e_{tq})|\Omega_t\right]=\int_{-\infty}^{\infty}k(u)h_t\left(u\hat{c}_T|\Omega_t\right)du$$
$$=h_t\left(0|\Omega_t\right)+\tfrac{1}{2}h''_t\left(0|\Omega_t\right)\hat{c}_T^2\int_{-\infty}^{\infty}k(u)u^2du+o(\hat{c}_T^2)$$
$$=h_t\left(0|\Omega_t\right)+o(\hat{c}_T^2)$$

where in the first equality we performed a change of variables, in the second we applied the Taylor expansion to $h_t\left(u\hat{c}_T|\Omega_t\right)$ around 0 and used the fact that $k(u)$ is a density function with first moment equal to 0 and finite second moment. Hence, the last terms of the r.h.s. of (A2) is zero. The first term has obviously zero expectation. If also its variance converges to zero, the result follows from Chebyshev inequality:

$$E\left[T^{-1}\sum_{t=1}^{T}\left[k_c(e_{tq})-E\left(k_c(e_{tq})|\Omega_t\right)\right]X'_t(b^0)\nabla f_t(b^0)\right]^2$$
$$=E\left[T^{-2}\sum_{t=1}^{T}\left[k_c(e_{tq})-E\left(k_c(e_{tq})|\Omega_t\right)\right]^2\left[X'_t(b^0)\nabla f_t(b^0)\right]^2\right]$$
$$\leq (T\hat{c}_T)^{-2}\sum_{t=1}^{T}E\left[J\left[X'_t(b^0)\nabla f_t(b^0)\right]^2\right]\leq T^{-1}\hat{c}_T^{-2}O_p(1)=o_p(1)$$

where the first equality follows because all the cross product are zero by the law of iterated expectations, the term J in the first inequality is some finite real number, and the rest follows from assumptions DQ2 and VC1.

Finally, the proof of the last part of the theorem is analogous to the proof of theorem 3 (see Powell's (1991) theorem 3.

–

References

Amemiya, T. (1982), Two Stage Least Absolute Deviations Estimators, *Econometrica*, 50: 689-711.

Basle Committee on Banking Supervision (1996), Amendment to the Capital Accord to Incorporate Market Risks.

Bloomfield, P. and W.L. Steiger (1983), *Least Absolute Deviations: Theory, Applications and Algorithms*, Boston: Birkhauser.

Boudoukh, J., M. Richardson and R.F. Whitelaw (1998), The Best of Both Worlds, *Risk*, 11: 64-67.

Buchinsky, M. and J. Hahn (1998), An Alternative Estimator for the Censored Quantile Regression Model, *Econometrica*, 66: 653-71.

Chernozhukov, V. (1999), Specification and Other Test Processes for Quantile Regression, mimeo, Stanford University.

Christoffersen, P. F. (1998), Evaluating Interval Forecasts, *International Economic Review*, 39: 841-62.

Danielsson, J. and C.G. de Vries (2000), Value-at-Risk and Extreme Returns, *Annales d'Economie et de Statistique*, 60: 239-70.

Dupacova, J. (1987), Asymptotic Properties of Restricted L_1-Estimates of Regression, in: Y. Dodge, eds., *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, North-Holland, Amsterdam.

Engle, R.R. (2002), "New Frontiers for ARCH Models", *forthcoming Journal of Applied Econometrics*

Engle, R. F. and V. Ng (1993), Measuring and Testing the Impact of News On Volatility, *Journal of Finance*, 48: 1749-78.

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.

Huber, P.J. (1967), The Behaviour of Maximum Likelihood Estimates under Nonstandard Conditions, *Proceedings of the Fifth Berkeley Symposium*, 4: 221-33.

Jureckova, J. and B. Prochazka (1993), Regression Quantiles and Trimmed Least Squares Estimators in Nonlinear Regression Models, *Journal of Nonparametric Statistics*, 3: 202-22.

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

Koenker, R. and G. Bassett (1982), Robust Tests for Heteroscedasticity Based on Regression Quantiles, *Econometrica*, 50: 43-61.

Koenker, R. and B.J. Park (1996), An Interior Point Algorithm for Nonlinear Quantile Regression, *Journal of Econometrics*, 71: 265-83.

Koenker, R. and Q. Zhao (1996), Conditional Quantile Estimation and Inference for ARCH models, *Econometric Theory*, 12: 793-813.

Kould, H.L. and A.K.E. Saleh (1995), Autoregression Quantiles and Related Rank-Scores Processes, *The Annals of Statistics*, 23: 670-89.

McNeil, A.J. and R. Frey (2000), Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach, *Journal of Empirical Finance*, 7: 271-300.

Mukherjee, K. (1999), Asymptotics of Quantiles and Rank Scores in Nonlinear Time Series, *Journal of Time Series Analysis*, 20: 173-92.

Oberhofer, W. (1982), The Consistency of Nonlinear Regression Minimizing the L_1 Norm, *The Annals of Statistics*, 10: 316-19.

Portnoy, S. (1991), Asymptotic Behavior of Regression Quantiles in Non-stationary, Dependent Cases, *Jorunal of Multivariate Analysis*, 38: 100-13.

Powell, J. (1983), The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators, *Econometrica*, 51: 1596-75.

Powell, J. (1984), Least Absolute Deviations Estimation for the Censored Regression Model, *Journal of Econometrics*, 25: 303-25.

Powell, J. (1986), Censored Regression Quantiles, *Journal of Econometrics*, 32: 143-55.

Powell, J. (1991), Estimation of Monotonic Regression Models Under Quantile Restriction, in: *Nonparametric and Semiparametric Methods in Economics and Statistics*, Barnett et al., eds., Cambridge University Press.

RiskMetrics (1996), *Technical Document*, Morgan Guarantee Trust Company of New York.

Schwert, G.W., "Why does stock market volatility change over time?," *Journal of Finance* 44 (1988), 1115–53.

Taylor, Stephen J., *Modelling Financial Time Series*, John Wiley and Sons Ltd. (1986).

van de Geer, S. (2000), *Empirical Processes in M-Estimation*, Cambridge University Press.

Weiss, A. (1991), Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation, *Econometric Theory*, 7: 46-68.

White, H. (1980), A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*, 48: 817-38.

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

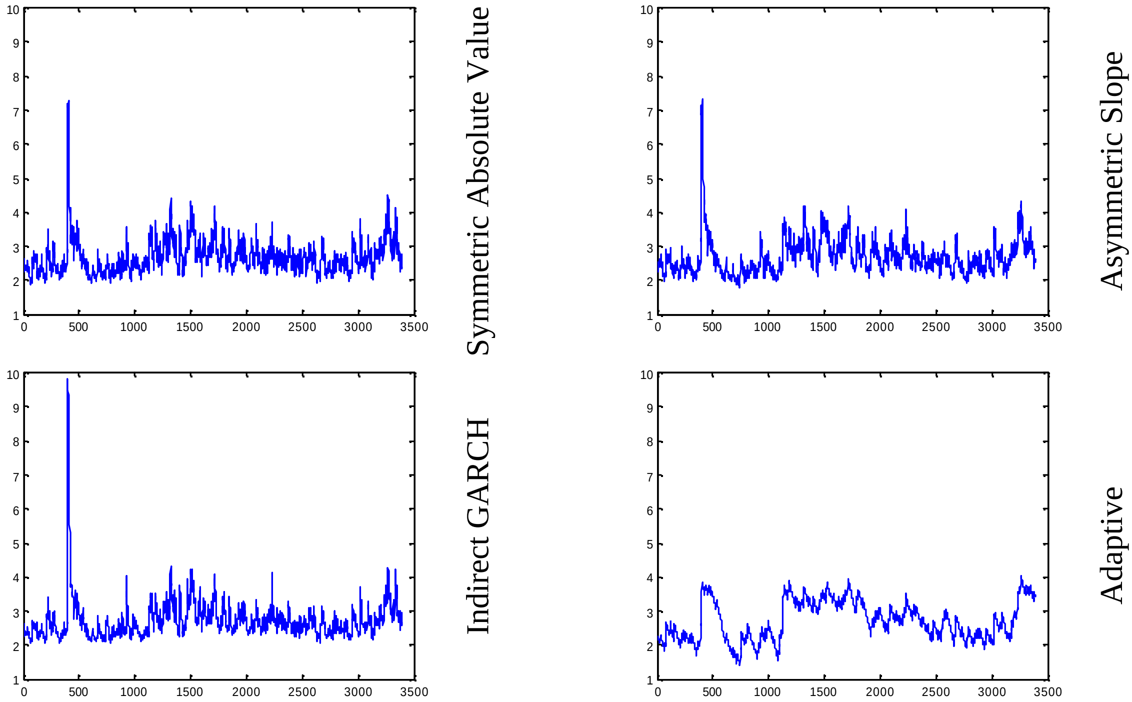

Figure 1 – 5% estimated CAViaR plots for GM. Since VaR is usually reported as a positive number, we set $\hat{VaR}_{t-1} = -f_{t-1}(\hat{b})$. The sample ranges from April 7 1986 to April 7 1999. The spike at the beginning of the sample is the 1987 crash. The increase in the quantile estimates towards the end of the sample reflects the increase in overall volatility following the Russian and Asian crises.

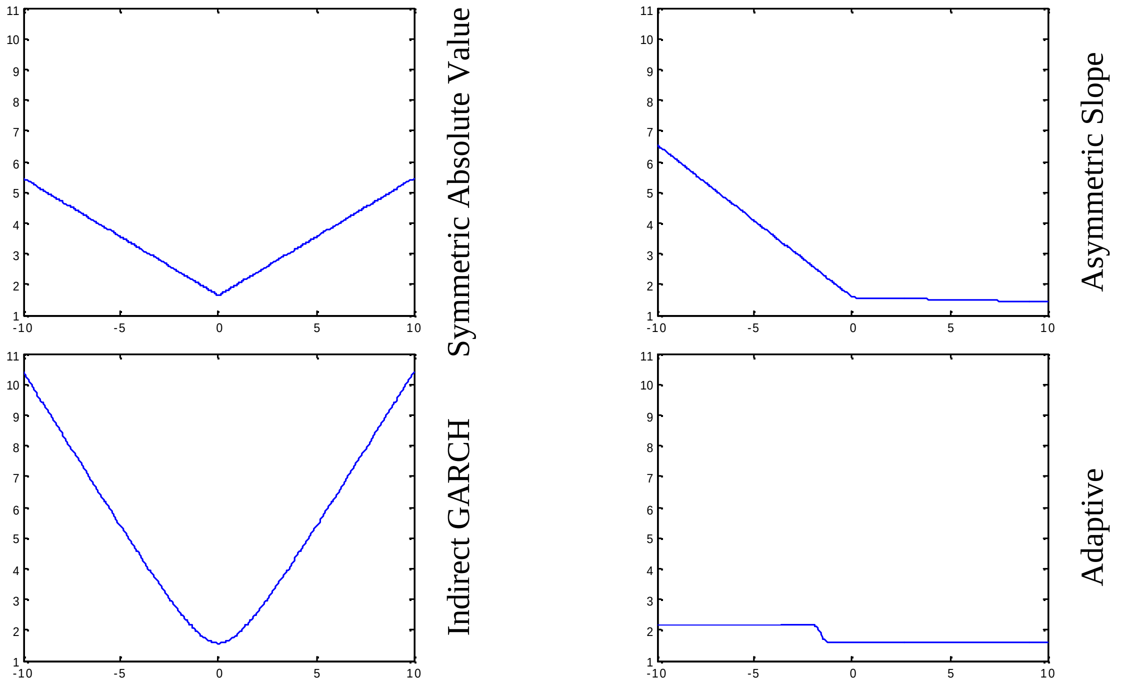

Figure 2 – 1% CAViaR News Impact Curve for S&P 500. For given estimated parameter vector $\hat{b}$ and setting (arbitrarily) $\hat{VaR}_{t-1} = -1.645$, the CAViaR News Impact Curve shows how $\hat{VaR}_t$ changes as lagged portfolio returns y_{t-1} vary. The strong asymmetry of the Asymmetric Slope News Impact Curve suggests that negative returns might have a much stronger effect on the VaR estimate than positive returns.

Table 1 – Estimates and relevant statistics for the four CAViaR specification (significant coefficients at 5% formatted in bold; shaded boxes denote rejection from the DQ test at 1% significance level).

1% Value at Risk	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	**0.4511**	0.1261	**0.2039**	0.3734	0.0558	**0.1476**	1.4959	1.3289	**0.2328**	**0.2968**	**0.1626**	**0.5562**
Standard Errors	*0.2028*	*0.0929*	*0.0604*	*0.2418*	*0.0540*	*0.0456*	*0.9252*	*1.9488*	*0.1191*	*0.1109*	*0.0736*	*0.1150*
P-values	*0.0131*	*0.0872*	*0.0004*	*0.0613*	*0.1509*	*0.0006*	*0.0530*	*0.2477*	*0.0253*	*0.0037*	*0.0136*	*0.0000*
Beta2	**0.8263**	**0.9476**	**0.8732**	**0.7995**	**0.9423**	**0.8729**	**0.7804**	**0.8740**	**0.8350**			
Standard Errors	*0.0826*	*0.0501*	*0.0507*	*0.0869*	*0.0247*	*0.0302*	*0.0590*	*0.1133*	*0.0225*			
P-values	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*			
Beta3	**0.3305**	0.1134	0.3819	**0.2779**	0.0499	-0.0139	0.9356	**0.3374**	1.0582			
Standard Errors	*0.1685*	*0.1185*	*0.2772*	*0.1398*	*0.0563*	*0.1148*	*1.2619*	*0.0953*	*1.0983*			
P-values	*0.0249*	*0.1692*	*0.0842*	*0.0235*	*0.1876*	*0.4519*	*0.2292*	*0.0002*	*0.1676*			
Beta4				**0.4569**	**0.2512**	**0.4969**						
Standard Errors				*0.1787*	*0.0848*	*0.1342*						
P-values				*0.0053*	*0.0015*	*0.0001*						
RQ	172.04	182.32	109.68	169.22	179.40	105.82	170.99	183.43	108.34	179.61	192.20	117.42
Hits in-sample (%)	1.0028	0.9682	1.0028	0.9682	1.0373	0.9682	1.0028	1.0028	1.0028	0.9682	1.2448	0.9336
Hits out-of-sample (%)	1.4000	1.6000	1.8000	1.4000	1.6000	1.6000	1.2000	1.6000	1.8000	1.8000	1.6000	1.2000
DQ in-sample (P-values)	0.6349	0.5375	0.3208	0.5958	0.7707	0.5450	0.5937	0.5798	0.7486	0.0117	0.0000	0.1697
DQ out-of-sample (P-values)	0.8965	0.0326	0.0191	0.9432	0.0431	0.0476	0.9305	0.0350	0.0309	0.0017	0.0009	0.0035

5% Value at Risk	Symmetric Absolute Value			Asymmetric Slope			Indirect GARCH			Adaptive		
	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500	GM	IBM	S&P 500
Beta1	**0.1812**	0.1191	**0.0511**	**0.0760**	**0.0953**	**0.0378**	**0.3336**	**0.5387**	**0.0262**	**0.2871**	**0.3969**	**0.3700**
Standard Errors	*0.0833*	*0.0839*	*0.0083*	*0.0249*	*0.0532*	*0.0135*	*0.1039*	*0.1569*	*0.0100*	*0.0506*	*0.0812*	*0.0767*
P-values	*0.0148*	*0.0778*	*0.0000*	*0.0011*	*0.0366*	*0.0026*	*0.0007*	*0.0003*	*0.0043*	*0.0000*	*0.0000*	*0.0000*
Beta2	**0.8953**	**0.9053**	**0.9369**	**0.9326**	**0.8892**	**0.9025**	**0.9042**	**0.8259**	**0.9287**			
Standard Errors	*0.0361*	*0.0500*	*0.0224*	*0.0194*	*0.0385*	*0.0144*	*0.0134*	*0.0294*	*0.0061*			
P-values	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*	*0.0000*			
Beta3	**0.1133**	**0.1481**	**0.1341**	0.0398	**0.0617**	**0.0377**	0.1220	0.1591	0.1407			
Standard Errors	*0.0122*	*0.0348*	*0.0517*	*0.0322*	*0.0272*	*0.0224*	*0.1149*	*0.1152*	*0.6198*			
P-values	*0.0000*	*0.0000*	*0.0047*	*0.1088*	*0.0117*	*0.0457*	*0.1441*	*0.0836*	*0.4102*			
Beta4				**0.1218**	**0.2187**	**0.2871**						
Standard Errors				*0.0405*	*0.0465*	*0.0258*						
P-values				*0.0013*	*0.0000*	*0.0000*						
RQ	550.83	522.43	306.68	548.31	515.58	300.82	552.12	524.79	305.93	553.79	527.72	312.06
Hits in-sample (%)	4.9793	5.0138	5.0484	4.9101	4.9793	5.0138	4.9793	5.0484	5.0138	4.9101	4.8409	4.7372
Hits out-of-sample (%)	4.8000	6.0000	5.6000	5.0000	7.4000	6.4000	4.6000	7.4000	5.8000	6.0000	5.0000	4.6000
DQ in-sample (P-values)	0.3609	0.0824	0.3685	0.9132	0.6149	0.9540	0.1037	0.1727	0.2661	0.0543	0.0032	0.0380
DQ out-of-sample (P-values)	0.9855	0.0884	0.0005	0.9235	0.0071	0.0007	0.8770	0.1208	0.0001	0.3681	0.5021	0.0240