Comparing Value-at-Risk Methodologies*

Luiz Renato Lima**
Breno Pinheiro Néri**

Abstract

In this paper, we compare four different Value-at-Risk (VaR) methodologies through Monte Carlo experiments. Our results indicate that the method based on quantile regression with ARCH effect dominates other methods that require distributional assumption. In particular, we show that the non-robust methodologies have higher probability of predicting VaRs with too many violations. We illustrate our findings with an empirical exercise in which we estimate VaR for returns of São Paulo stock exchange index, IBOVESPA, during periods of market turmoil. Our results indicate that the robust method based on quantile regression presents the least number of violations.

Keywords: Time Series, Value-at-Risk, Quantile Regression.

JEL Codes: C52, C53, G15.

*Submitted in November 2005. Revised in November 2006.
**Graduate School of Economics – Getulio Vargas Foundation, 190 Praia de Botafogo, 11th Floor, Rio de Janeiro, Brazil – 22250-900. E-mail: luizr@fgv.br and bneri@fgvmail.br

Brazilian Review of Econometrics
v. 27, nº 1, pp. 1–25 May 2007

1. Introduction

Every day, financial institutions (e.g.: banks) estimate measures of market risk exposure, which are analyzed by the institutions' decision makers. These estimates are also analyzed by internal and external auditors and regulatory agencies, which enforce that those institutions set aside enough capital to cover their risk exposures. This concern about market risk exposure has been increasing since the stock market crash in 1987, when 1 trillion dollars (23% drop in value) was lost in a single day, known as the Black Monday. The recent turbulence in emerging markets, starting in Mexico in 1995, continuing in Asia in 1997, and spreading to Russia and Latin America in 1999, has further extended the interest in risk management.

Imprecise measures of risk cause inefficiencies: on the one hand, if the measure is too conservative, then too much capital, which could be used in a more profitable way, is set aside; on the other hand, if it is too risky, then it yields a large number of violations, which may lead the institution to bankruptcy. Hence, researching into more and more reliable and accurate risk measurement methodologies is an active and growing literature.

Value-at-Risk (VaR) is probably the most used measure of risk since the 1996 amendment to the Basel Capital Accord which proposed that commercial banks with significant[1] trade activity could use their own VaR measure to define how much capital they should set aside to cover their market risk exposure, and U.S. bank regulatory agencies could audit the VaR methodology employed by the banks. This amendment was adopted in 1998 (Lopez, 1999b). In Brazil, article 59 of resolution No. 2.829, March 2001, of the Brazilian Central Bank mandates the use of VaR for some markets.

Value-at-Risk is the loss in market value over a given time horizon that is exceeded with probability τ. That is, for a time series of returns r_t, find VaR_t such that

$$P\left[r_t < -VaR_t | I_{t-1}\right] = \tau \tag{1}$$

where I_{t-1} denotes the information set at time $t-1$. From this definition, it is clear that finding a VaR is essentially the same as finding a $100\tau\%$ conditional quantile. Note that, by convention, the sign is changed to avoid negative number in the $VaR_t(\tau)$ time series. For regulatory purpose, τ is generally set to 1%. It does not mean that the banks may not estimate VaRs under different significance levels for their risk managers.

Although VaR is a relatively simple concept, its robust estimation is often ignored in practice. Indeed, one popular approach to estimate VaR assumes a

[1] Any bank or bank holding company whose trading activity equals greater than 10 percent of its total assets or whose trading activity equals greater than $1 billion must hold regulatory capital against their market risk exposure.

conditionally normal return distribution. The estimation of VaR is, in this case, equivalent to estimating conditional volatility of returns. Another popular method is to compute the empirical quantile nonparametrically, for example, rolling historical quantiles or Monte Carlo simulations based on an estimated model.[2]

However, these models are based on restricted assumptions about the distribution of returns. There has been accumulated evidence that portfolio returns (or log returns) are usually not normally distributed. In particular, it is frequently found that market returns display structural shifts, negative skewness and excess kurtosis in the distribution of the time series. This is particularly true in periods of market stress such as the financial crises faced by the Brazilian economy from 1997 to 2000. These market characteristics suggest that more robust methods are needed to estimate VaR.

In this paper, we estimate VaR by using a robust method based on quantile regression model that allows for ARCH effect, and we compare it to three other non-robust VaR methodologies that are based on GARCH type volatility models. It is important to mention that Engle and Manganelli (1999) consider a different quantile-regression-based method. In particular, they consider an autoregression of the estimated VaRs. Our approach, however, has the advantage of pursuing a well-developed distributional theory which facilitates statistical inference and computational optimization.

We are not the first ones to compute VaR using a quantile regression model that allows for ARCH effect. In fact, Wu and Xiao (2002) used this model to estimate VaR and left-tail measures that were next employed to construct a risk-managed index fund. The performance of the ARCH Quantile method was then evaluated according to the capacity of the risk-managed index fund to track the S&P500 index.

There are, however, other ways to assess the quality of a VaR methodology. In this paper, we follow Engle and Manganelli (2001), who compare VaR methodologies using descriptive statistics of the distributions of violations obtained via Monte Carlo simulations. Specifically, we simulate many trajectories of the return time series assuming different innovation distributions, and compute the number of violations[3] using four different VaR methodologies. For each simulated trajectory of the return series, we save the number of violations. At the end of the experiment, we will have a distribution of the number of violations for each VaR methodology. Hence, we can compute descriptive statistics of the various distributions of violations and evaluate the quality of a VaR methodology according to these statistics.

Our Monte Carlo simulations indicate that the robust model based on quantile regression dominates other models that require distributional assumptions. In

[2]This approach includes the weighted moving average method by J.P. Morgan's Riskmetrics and the hybrid method by Boudoukh et al. (1998).

[3]The number of violations is defined as the number of losses greater than $VaR(\tau)$.

particular, the distribution of violations generated from non-robust models is right-skewed and presents excess kurtosis, meaning that these non-robust models have high probability of presenting trajectories with too many violations. We illustrate our findings with an empirical application. We consider returns of the São Paulo stock exchange index, IBOVESPA, and show that the VaR estimated by the quantile regression approach tends to predict VaRs more accurately during periods of market stress.

The outline of this paper is as follows: In Section 2, we describe the general framework and present the competing models. We describe our Monte Carlo experiment in Section 3. An empirical illustration is provided in Section 4, and Section 5 concludes.

2. The Competing Models

Most of the VaR methodologies are GARCH type models. Hence, they can be described using a GARCH framework (Giot and Laurent, 2004). GARCH models are designed to model the conditional heteroskedasticity in the time series of returns y_t, that is,

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= \sigma_t z_t \\ \mu_t &= c\left(\eta | I_{t-1}\right) \\ \sigma_t &= h\left(\eta | I_{t-1}\right) \end{aligned} \tag{2}$$

Where $c\left(\eta | I_{t-1}\right)$ and $h\left(\eta | I_{t-1}\right)$ are functions of the vector of parameters η and of the information set I_{t-1}; z_t is an independent and identically distributed process, independent of I_{t-1}, with $E\left[z_t\right] = 0$ and $Var\left[z_t\right] = 1$; μ_t is the conditional mean for y_t and σ_t^2 is its conditional variance. The volatility model (2) encompasses a family of methodologies used to predict VaRs. Next, we describe some members of such family.

2.1 RiskMetrics[4]

RiskMetrics (Morgan, 1996) is the most simple analyzed methodology. However, it is still one of the most used model to compute VaR, and it is available for free from J.P. Morgan. In fact, RiskMetrics is a Gaussian Integrated GARCH(1,1) model where the autoregressive parameter is set at a prespecified value of 0.94 (for daily VaR, in the United States) and the decay parameter (it can be viewed as an exponential filter in volatility) is set at 0.06, that is,

$$\sigma_t^2 = 0.06\varepsilon_{t-1}^2 + 0.94\sigma_{t-1}^2 \tag{3}$$

[4]RiskMetrics is a trademark of J.P. Morgan.

The conditional mean μ_t is estimated by OLS, running y_t against its own lags,[5] and $z_t \sim N(0,1)$.

2.2 Gaussian GARCH(1,1)

Rather than using RiskMetrics, we could use the same GARCH(1,1) model but, instead of setting prespecified parameter values, we estimate them. In other words, we estimate the model

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{4}$$

and $z_t \sim N(0,1)$.

This is the second model to be analyzed in our Monte Carlo experiment. The Gaussian (or Normal) GARCH(1,1) is expected to generate better forecasts than RiskMetrics, because the parameters are estimated rather than prespecified.

Observe that these two first models capture neither the asymmetric dynamics[6] nor all the leptokurtosis that is generally present in macroeconomics and financial time series, due to the fact that they assume normality for z_t. Indeed, in VaR applications, the choice of an appropriate distribution for the innovation process z_t is an important issue as it directly affects the quality of the estimation of the required quantiles. One way to weaken the assumption on the distribution of z_t is to consider the Skewed Student-t APARCH model, which we describe next.

2.3 Skewed Student-t APARCH(1,1)

The APARCH (Ding et al., 1993) is an extension of the GARCH model that nests at least seven GARCH specifications. It can be described as

$$\sigma_t^\delta = \omega + \alpha_1 \left(|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1} \right)^\delta + \beta_1 \sigma_{t-1}^\delta \tag{5}$$

where ω, α_1, γ_1, β_1 and δ ($\delta > 0$) are parameters to be estimated. δ plays the role of a Box-Cox transformation of σ_t, while γ_1 ($-1 < \gamma_1 < 1$) reflects the so-called leverage effect: the stylized fact that negative shocks impact volatility more than positive shocks.

Giot and Laurent (2003) and Giot (2003) use the above model considering a standardized version of the Skewed Student-t distribution – introduced by

[5] The number of lags in the OLS regression can be chosen using Information Criteria. One can also add other conditioning variables.

[6] Beaudry and Koop (1993) showed that positive shocks to U.S. GDP are more persistent than negative shocks, indicating asymmetric business cycle dynamics. More recently, Nam et al. (2005) identified asymmetric dynamics for daily return on the S&P 500 and used that to develop optimal technical trading strategies. In 1992, Brock, Lakonishok and LeBaron showed that two of the simplest and most popular trading rules – moving average and trading range break – consistently generate buy signals with higher returns than sell signals, and further, the returns following buy signals are less volatile than returns following sell signals.

Fernández and Steel (1998) – for the z_t process. They show that such standardized version provides more accurate VaR forecasts than the GARCH model. This result is somehow expected, because the Skewed Student-t APARCH(1,1) nests the Gaussian GARCH(1,1).[7]

According to Lambert and Laurent (2001) and provided that the degrees of freedom $\nu > 2$, the innovation process z_t is said to follow a standardized Skewed Student-t distribution, i.e. $z_t \sim SKST(0, 1, \xi, \nu)$ if:

$$f(z_t|\xi,\nu) = \begin{cases} \frac{2}{\xi+\frac{1}{\xi}} sg\left[\xi\left(sz_t+m\right)|\nu\right], \text{ if } z_t < -\frac{m}{s} \\ \frac{2}{\xi+\frac{1}{\xi}} sg\left[\frac{sz_t+m}{\xi}|\nu\right], \text{ if } z_t \geq -\frac{m}{s} \end{cases} \tag{6}$$

where $g[\cdot|\nu]$ is a symmetric (unit variance) Student-t density and $\xi > 0$ is the asymmetry coefficient. The parameters m and s^2 are, respectively, the mean and the variance of the nonstandardized Skewed Student-t:

$$m = \frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(\xi - \frac{1}{\xi}\right) \tag{7}$$

and

$$s^2 = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2 \tag{8}$$

In short, ξ models the asymmetry, while ν accounts for the tail thickness. See Lambert and Laurent (2001) for a discussion on the link between these two parameters and the skewness and the kurtosis.

2.4 ARCH(q) quantile

Koenker and Zhao (1996) introduced the quantile regression model that allows for ARCH effect. The ARCH(q) Quantile methodology uses the OLS estimator to estimate the conditional mean μ_t, but this is the only similarity with the first three methodologies. The ARCH(q) Quantile does not assume any particular distribution to the process z_t. The model can be described as follows

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \varepsilon_t &= \left(\gamma_0 + \gamma_1\left|\varepsilon_{t-1}\right| + \ldots + \gamma_q\left|\varepsilon_{t-q}\right|\right) z_t \end{aligned} \tag{9}$$

Thus, the ARCH(q) Quantile specification assumes that the errors follow an

[7] Indeed, the GARCH(1,1) is an APARCH(1,1) with $\delta = 2$ and $\gamma_1 = 0$, and the Skewed Student-t distribution with the asymmetry coefficient $\xi = 1$ (no asymmetry) converges to the Gaussian distribution when the degrees of freedom ν tend to the infinity.

ARCH(q) type model,[8] in which the fundamental innovation z_t is drawn from an unknown distribution F_z.

In all the models presented in this paper, the $VaR(\tau)$ is defined as the τth conditional quantile of the return, that is

$$-VaR_t(\tau) = \mu_t + Q_\varepsilon(\tau|I_{t-1}) \tag{10}$$

where $Q_\varepsilon(\tau|I_{t-1})$ is the conditional quantile function of ε_t.

Given a known distribution for the process z_t, the computation of (10) is straightforward. When the distribution of z_t is unknown, we are led to the problem of quantile regression. The quantile regression method is an extension of the empirical quantile methods. While classical linear regression methods, based on minimization of the sum of squared residuals, enable one to estimate models for conditional mean functions, quantile regression methods offer a mechanism for estimating models for the conditional quantile functions, as the one appearing in (10). Thus, quantile regression is capable of providing a complete statistical analysis of the stochastic relationships among random variables.

Moreover, the quantile regression method has the important property that it is robust to distributional assumptions. This property is inherited from the robustness property of the ordinary sample quantiles. Quantile estimation is only influenced by the local behavior of the conditional distribution of the response variable near the specified quantile. As a result, the estimated conditional quantile function is not sensitive to outlier observations. Such a property is especially attractive in financial applications, since many financial data such as IBOVESPA returns are usually heavy-tailed and thus are not (conditional) normally distributed

2.4.1 Quantile regression

As we stated above, the idea of quantile regression provides a natural way of estimating Value-at-Risk. Quantile regression was introduced by Koenker and Basset (1978) and has received a lot of attention in econometrics research in the past two decades. To introduce quantile regression, let Y be a random variable with distribution function $F(y)$. The τ-th quantile of Y is defined by

$$Q_Y(\tau) = \inf\{y|F(y) \geq \tau\} \tag{11}$$

Similarly, if we have a random sample $\{y_1, y_2, ..., y_n\}$ from the distribution F, the τ-th sample quantile is:

$$\hat{Q}_y(\tau) = \inf\left\{y|\hat{F}(y) \geq \tau\right\} \tag{12}$$

[8] Observe both the similarities to and the differences from the classical ARCH specification introduced by Engle (1982).

where $\hat{F}$ is the empirical distribution function of the random sample. This sample quantile may be found by solving the minimization problem:

$$\min_{b\in\mathbb{R}}\left[\sum_{t\in\{t:y_t\geq b\}}\tau\left|y_t-b\right|+\sum_{t\in\{t:y_t<b\}}(1-\tau)\left|y_t-b\right|\right] \tag{13}$$

Generalizing, if we consider the model:

$$y_t=x_t'b+\eta_t \tag{14}$$

where x_t is a kx1 vector of regressors including an intercept term. Then, conditional on the regressor x_t, the τ-th quantile of y:

$$Q_Y\left(\tau|x_t\right)=\inf\left\{y|F\left(y|x_t\right)\geq\tau\right\} \tag{15}$$

is a linear function of x_t:

$$b_1+x_{2t}'b_2+...+x_{kt}'b_k+F_{\varepsilon}^{-1}\left(\tau\right) \tag{16}$$

where $F_{\varepsilon}\left(\cdot\right)$ is the cumulative distributional function of the residual. The τ-th conditional quantile of y can be estimated by an analogue of equation (13):

$$\hat{Q}_Y\left(\tau|x_t\right)=x_t'\hat{b}\left(\tau\right) \tag{17}$$

where

$$\hat{b}\left(\tau\right)=\arg\min_{b\in\mathbb{R}^k}\left[\sum_{t\in\{t:y_t\geq x_t'b\}}\tau\left|y_t-x_t'b\right|+\sum_{t\in\{t:y_t<x_t'b\}}(1-\tau)\left|y_t-x_t'b\right|\right] \tag{18}$$

is called the quantile regression. As a special case, the least absolute deviation (LAD) estimator (or l_1 regression) is the median regression, *i.e.*, the quantile regression for $\tau=0.5$.[9]

2.4.2 Estimating the ARCH quantile VaR

Given equation (9), we denote $(1,\left|\varepsilon_{t-1}\right|,...,\left|\varepsilon_{t-q}\right|)'$ as X_t and the corresponding coefficient vector as γ. Then,

$$Q_{\varepsilon}\left(\tau|I_{t-1}\right)=X_t'\alpha\left(\tau\right) \tag{19}$$

where

[9] For more details on quantile regression, see Koenker (2005).

$$\alpha(\tau) = (\gamma_0 Q_z(\tau), \gamma_1 Q_z(\tau), ..., \gamma_q Q_z(\tau))' \tag{20}$$

and $Q_z(\tau) = F_z^{-1}(\tau)$ is the quantile function of z. By definition, $VaR_t(\tau)$, the conditional Value-at-Risk at the τ-th quantile is just

$$-VaR_t(\tau) = \mu_t + X_t'\alpha(\tau) \tag{21}$$

So we need to estimate $\widehat{\alpha}(\tau)$.[10] It can be achieved by solving the problem:

$$\widehat{\alpha}(\tau) = \arg\min_{\gamma \in \mathbb{R}^{q+1}} \left[\sum_{t \in \{t: u_t \geq Z_t'\gamma\}} \tau |\varepsilon_t - X_t'\gamma| + \sum_{t \in \{t: u_t < Z_t'\gamma\}} (1-\tau) |\varepsilon_t - X_t'\gamma| \right] \tag{22}$$

In practice, we can replace ε_t by their OLS estimators $\widehat{\varepsilon}_t = y_t - \widehat{\mu}_t$. For example, if $\mu_t = \beta_o + \beta_1 y_{t-1}$, then $\widehat{\varepsilon}_t = y_t - \widehat{\beta}_o - \widehat{\beta}_1 y_{t-1}$, where $\widehat{\beta}_o$ and $\widehat{\beta}_1$ are estimated by OLS. Under mild regularity conditions, Koenker and Zhao (1996) show that $\widehat{\alpha}(\tau)$ estimated based on $\widehat{\varepsilon}_t$ is still a consistent estimator of $\alpha(\tau)$.

3. Monte Carlo Simulations

The objective of this section is to compare the four aforementioned VaR methodologies. We perform Monte Carlo Simulations in which we generate 1000 time series with 1250 observations. We use a rolling window of 250 observations to estimate the parameters of the four methodologies and forecast the $VaR(1\%)$ associated to the 251^{st} observation. The result is a 1-day-ahead VaR time series, one for each methodology. At the end, we will have 1001 forecast observations for each methodology. We decide for a 250-observation window because it is the number of observations required to compute the multiplication factor F_t in the Basel capital charge formula and because it is approximately 1 year, a reasonable time to be used by the banks, containing enough information for estimation of the parameters, without losing too many observations. To find the violations, we need to compare the last 1000 observations of the generated series with the first 1000 observations of the VaR forecasts. The choice for the $VaR(1\%)$ is due to regulatory purpose.

The DGPs used in this experiment are

$$y_t = 0.5 y_{t-1} + \varepsilon_t \tag{23}$$

$$\varepsilon_t = \sigma_t z_t \tag{24}$$

[10] Remember that $\widehat{\mu}_t$ is estimated by OLS.

$$\sigma_t^2 = 1 + 0.5\varepsilon_{t-1}^2 + \psi\sigma_{t-1}^2 \tag{25}$$

where z_t are independent and identically distributed fundamental innovations, $z_t \sim i.i.d.$ There are five different innovation distributions, and we repeat the experiment with GARCH effect ($\psi = 0.5$). Therefore, there are 10 DGPs that are described in the following table.

Table 1
Data generating processes

DGP	Distribution of z_t	ψ
1	$N(0,1)$	0
2	$t_{(3)}$	0
3	$\chi^2_{(1)} - \delta_{(1)}$	0
4	$\delta_{(2)} - Gamma\,(2,1)$	0
5	$\chi^2_{(1)} I_{\{\nu_t \leq 0.2\}} + \left(\chi^2_{(1)} + \delta_{(-4)}\right) I_{\{0.2<\nu_t\leq 0.8\}} + \delta_{(-4)} I_{\{\nu_t > 0.8\}}$	0
6	$N(0,1)$	0.5
7	$t_{(3)}$	0.5
8	$\chi^2_{(1)} - \delta_{(1)}$	0.5
9	$\delta_{(2)} - Gamma\,(2,1)$	0.5
10	$\chi^2_{(1)} I_{\{\nu_t \leq 0.2\}} + \left(\chi^2_{(1)} + \delta_{(-4)}\right) I_{\{0.2<\nu_t\leq 0.8\}} + \delta_{(-4)} I_{\{\nu_t > 0.8\}}$	0.5

where $\delta_{(x_0)}$ is the Dirac's Delta density, whose distribution $F_{\delta_{(x_0)}}(x)$ is given by

$$F_{\delta_{(x_0)}}(x) = \begin{cases} 1, & \text{if } x \geq x_0 \\ 0, & \text{if } x < x_0 \end{cases} \tag{26}$$

$I_{\{\cdot\}}$ is an indicator function that values 1 if the condition inside the braces is true and 0 otherwise, and ν_t is an independent and identically distributed standard uniform distribution, $\nu_t \sim U[0,1]$.

Therefore, DGP_1 corresponds to the standard Gaussian distribution. In the DGP_2, z_t are drawn from student-t distribution with 3 degrees of freedom (thus it has finite expectation and variance), so z_t presents leptokurtosis. The distribution of z_t in DGP_3 is no longer symmetric. The distribution of z_t in DGP_4 and DGP_5 are nonstandard (with mass points) and they are considered to verify the robustness of VaR methodologies against distributional misspecification. GARCH effect is introduced in DGP_6 to DGP_{10}.

3.1 Results

For each replication (with 1000 daily forecasts), the ideal number of violations of a $VaR\,(1\%)$ is 10, but there are replications with more violations and there are replications with fewer violations. Hence, we shall analyze the distribution of the number of violations. Since there are 1000 replications, such a distribution of violations will have 1000 points (each point represents the number of violations that occurred in each trajectory). Recall that there are four VaR methodologies labelled as $VaR\ i$, $i = 1, 2, 3, 4$. Hence, VaR 1, VaR 2, VaR 3, and VaR 4 correspond to the RiskMetrics, GARCH(1,1), APARCH(1,1), and ARCH(1) Quantile VaR methodologies, respectively. We assess the performance of each VaR methodology under the 10 aforementioned DGPs. Tables 2, 3 and 4 present location and scale parameter estimates of the various distributions of violations.

Table 2
Distributions of violations: estimated mean and bias

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
		Mean		
DGP_1	19.4	12.7	12.8	14.7
DGP_2	19.8	13.2	13.5	14.7
DGP_3	19.8	12.9	13.3	14.6
DGP_4	19.7	13.2	13.5	14.8
DGP_5	20.0	13.4	13.5	14.6
DGP_6	20.0	13.2	13.5	14.5
DGP_7	19.8	13.1	13.6	14.7
DGP_8	19.5	12.8	13.2	14.7
DGP_9	20.5	14.0	14.4	14.7
DGP_{10}	19.7	13.1	13.3	14.6
		Bias		
DGP_1	9.4	2.7	2.8	4.7
DGP_2	9.8	3.2	3.5	4.7
DGP_3	9.8	2.9	3.3	4.6
DGP_4	9.7	3.2	3.5	4.8
DGP_5	10.0	3.4	3.5	4.6
DGP_6	10.0	3.2	3.5	4.5
DGP_7	9.8	3.1	3.6	4.7
DGP_8	9.5	2.8	3.2	4.7
DGP_9	10.5	4.0	4.4	4.7
DGP_{10}	9.7	3.1	3.3	4.6

On the one hand, we notice in Table 2 that all four methodologies present positive estimated biases. The RiskMetrics methodology is the most biased and the Gaussian GARCH(1,1) is the least biased. The robust ARCH(1) Quantile method exhibits a very stable estimated bias across different innovation distributions. On the other hand, Table 3 shows that the variance is much higher (one order of magnitude higher) in the first three methodologies than in the ARCH(1) Quantile.

In Table 3, we compute the range of the distribution of the number of violations, *i.e.*, the difference between the maximum and the minimum number of violations. We notice that the fourth methodology has the lowest range. Indeed, as shown in Table 4, its maximum value never exceeds 27 violations, which can be considered a good performance in a $VaR\,(1\%)$. The non-robust methodologies have maximum number of violations at least three times as large as the ARCH(1) Quantile method. This excess dispersion invalidates the first three VaR methodologies, since they jeopardize the bank or institution that uses them to compute VaR measures. It is not acceptable for a measure of risk to be too *risky*, in the sense that its probability of having trajectories with too many violations is too high. A bank may go belly-up if this trajectory is the true (realized) one.

Table 3
Distributions of violations: estimated variance and range

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Variance				
DGP_1	62.1	88.1	80.5	7.4
DGP_2	72.0	105.4	102.2	7.5
DGP_3	72.5	103.3	100.1	7.5
DGP_4	66.6	102.7	93.6	7.3
DGP_5	69.0	102.1	98.7	8.1
DGP_6	68.5	103.9	97.1	6.8
DGP_7	74.7	104.6	104.6	8.1
DGP_8	59.9	89.2	87.4	7.5
DGP_9	89.2	133.2	127.2	7.7
DGP_{10}	61.5	100.5	90.4	6.7
Range (Max – Min)				
DGP_1	70	81	72	19
DGP_2	68	65	68	18
DGP_3	72	76	76	19
DGP_4	62	65	63	20
DGP_5	69	74	64	20
DGP_6	61	62	65	16
DGP_7	61	67	73	18
DGP_8	70	71	67	18
DGP_9	63	70	68	17
DGP_{10}	62	74	70	17

On the other hand, the Gaussian GARCH(1,1) and the Skewed Student-t APARCH(1,1) present some trajectories with very few violations. In fact, the former has trajectories with no violations at all.

Table 4
Distributions of violations: estimated minimum and maximum

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Minimum Value (Min)				
DGP_1	6	2	3	7
DGP_2	7	3	3	7
DGP_3	8	3	2	6
DGP_4	7	1	2	7
DGP_5	6	3	2	3
DGP_6	7	1	2	7
DGP_7	7	3	3	7
DGP_8	6	0	2	5
DGP_9	8	2	3	7
DGP_{10}	6	3	1	7
Maximum Value (Max)				
DGP_1	76	83	75	26
DGP_2	75	68	71	25
DGP_3	80	79	78	25
DGP_4	69	66	65	27
DGP_5	75	77	66	23
DGP_6	68	63	67	23
DGP_7	68	70	76	25
DGP_8	76	71	69	23
DGP_9	71	72	71	24
DGP_{10}	68	77	71	24

The ARCH(1) Quantile VaR exhibits the second greatest bias, but displays the lowest variance and range. To assess the trade-off between bias and variance, we adopt the Mean Squared Error (MSE), abiding by the formula (see Engle and Manganelli (2001))

$$MSE\left(\hat{X}\right) := \frac{1}{1000}\sum_{i=1}^{1000}\left(X_i - 10\right)^2 \tag{27}$$

where X_i is the number of violation in the i-th replication, 1000 is the total number of replications and 10 is the ideal number of violations, for a $VaR(1\%)$, at each replication. It can be shown that the $MSE\left(\hat{X}\right) = Var\left(\hat{X}\right) + Bias\left(\hat{X}\right)^2$, where $Bias\left(\hat{X}\right) = \bar{X} - 10$ and $\bar{X} = \frac{1}{1000}\sum_{i=1}^{1000} X_i$. The bias and the MSE are shown in Table 5.

Notice that, even considering the bias, the ARCH(1) Quantile VaR methodology has, by far, the lowest MSE. We can see that the estimated MSE is pretty stable across the different DGPs for the robust ARCH(1) method.

Table 5
Distributions of violations: Estimated mean squared error

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Mean Squared Error				
DGP_1	150.7	95.3	88.1	29.6
DGP_2	168.8	115.4	114.6	29.5
DGP_3	168.0	110.9	110.9	29.0
DGP_4	160.9	112.7	106.1	30.3
DGP_5	169.3	113.5	111.0	29.2
DGP_6	167.9	113.7	109.0	27.0
DGP_7	170.3	114.1	117.3	29.8
DGP_8	151.0	97.1	97.6	29.2
DGP_9	199.8	149.0	146.7	29.8
DGP_{10}	155.0	109.8	101.3	28.0

In Table 6, we show estimates of skewness and excess kurtosis of the distributions of number of violations. We observe that the non-robust VaR methodologies yield distributions of violations that are skewed to the right and possess excess kurtosis. Again, the robust ARCH(1) Quantile method gives rise to a well-behaved distribution of violations, with almost no skewness or no excess kurtosis.

Table 6
Distributions of violations: Estimated skewness and excess kurtosis

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Skewness				
DGP_1	2.8	3.2	3.3	0.3
DGP_2	2.6	2.8	3.0	0.3
DGP_3	2.7	3.0	3.1	0.3
DGP_4	2.6	2.8	2.8	0.3
DGP_5	2.6	2.8	2.7	0.1
DGP_6	2.6	2.8	2.8	0.2
DGP_7	2.7	2.9	2.9	0.2
DGP_8	2.6	2.9	2.9	0.2
DGP_9	2.4	2.6	2.6	0.2
DGP_{10}	2.6	2.9	2.9	0.1
Excess Kurtosis				
DGP_1	10.1	11.6	11.9	0.2
DGP_2	8.2	7.8	9.2	0.1
DGP_3	8.5	8.9	10.4	0.2
DGP_4	7.8	7.5	7.7	0.6
DGP_5	8.3	7.8	7.2	0.1
DGP_6	8.0	7.5	7.8	0.0
DGP_7	8.5	8.2	8.8	0.1
DGP_8	8.8	8.7	8.9	0.0
DGP_9	6.1	6.2	6.4	0.0
DGP_{10}	8.5	8.9	9.1	0.2

For completeness, we present the histograms of the number of violations for the four methodologies under the five DGPs with GARCH effect in Figures 1 to 5. It is clear that the non-robust methods present some trajectories with too many violations. Moreover, the estimated distribution of the number of violations is more concentrated in the fourth methodology, under all DGPs.

In sum, our Monte Carlo experiment suggests that the robust method dominates the other methods, since the former yields a distribution of number of violations that present very low MSE, almost no skewness and no excess kurtosis. More importantly, these nice properties are preserved over a wide range of innovation distributions, with or without GARCH effect. This result is expected because the robust method does not depend on distributional assumption.

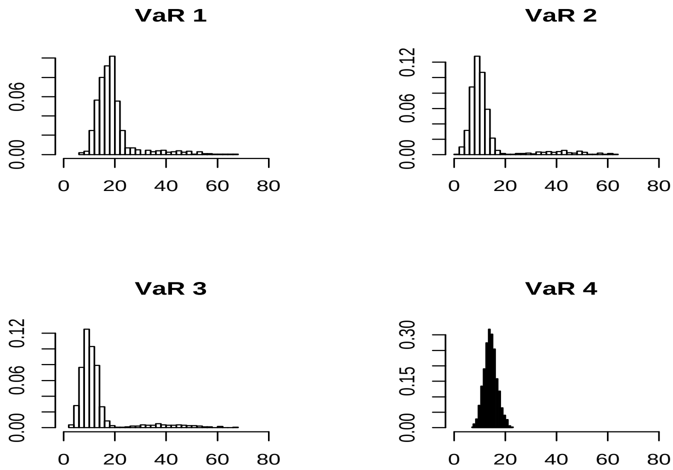

Figure 1
Histograms of the number of violations under DGP 6

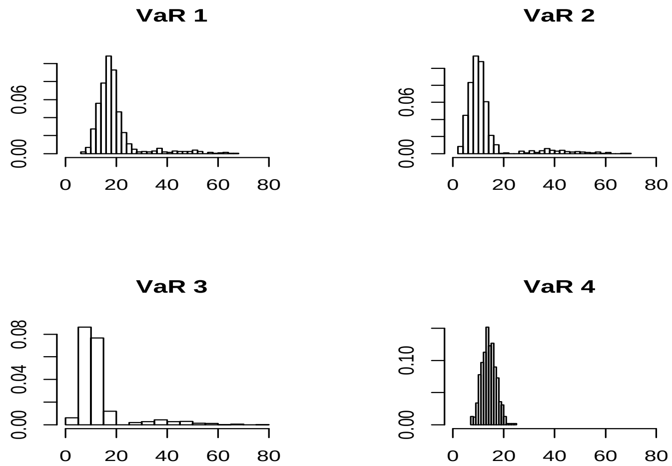

Figure 2
Histograms of the number of violations under DGP 7

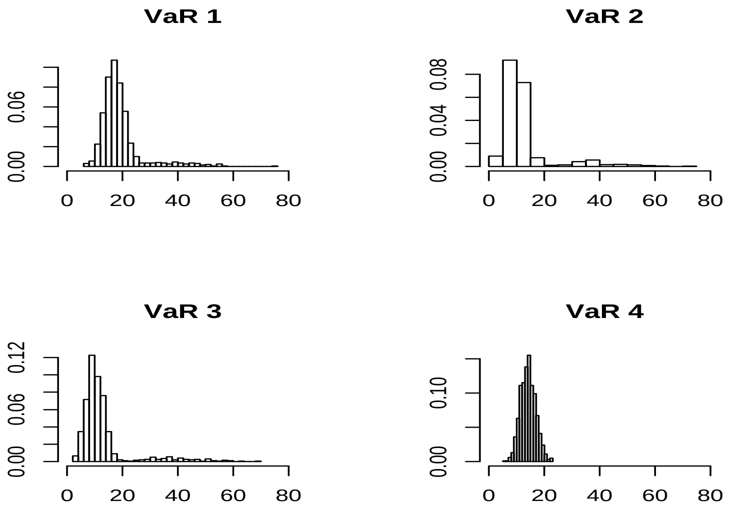

Figure 3
Histograms of the number of violations under DGP 8

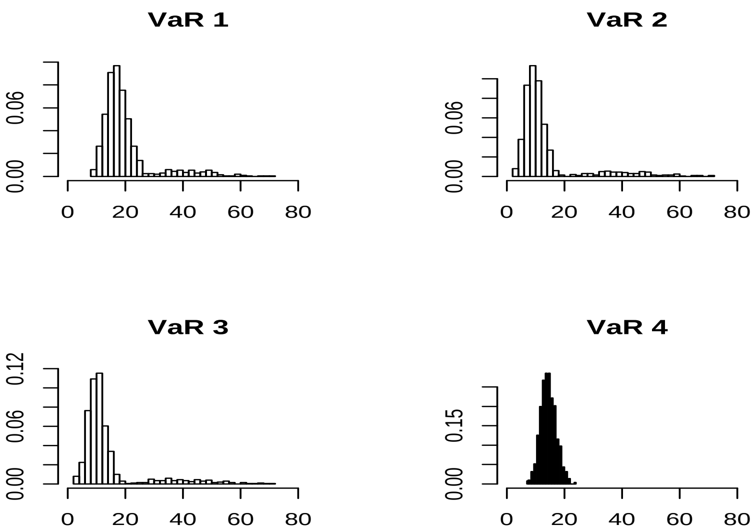

Figure 4
Histograms of the number of violations under DGP 9

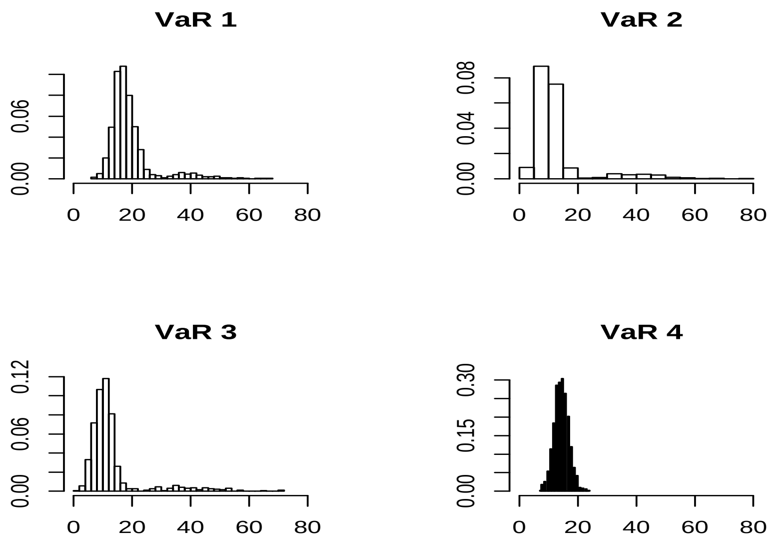

Figure 5
Histograms of the number of violations under DGP 10

3.2 Backtest

The Unconditional Coverage backtest was proposed by Kupiec (1995). Under the null hypothesis that $P\left[r_t < -VaR_t\left(\tau\right)|I_{t-1}\right] = \tau$, $\forall t$, i.e., that the probability of occurrence of a violation is indeed τ, the number of violations, X, in a given time span, T, follows a binomial distribution: $X \sim Binomial\left(T, \tau\right)$. Define $\hat{\tau} := \frac{X}{T}$. Then, the likelihood ratio test statistic[11] is

$$LR_{uc} = 2\ln\left(\frac{\hat{\tau}^X\left(1-\hat{\tau}\right)^{T-X}}{\tau^X\left(1-\tau\right)^{T-X}}\right) \tag{28}$$

Under the null hypothesis that $\tau = \hat{\tau}$, $LR_{uc} \sim \chi^2_{(1)}$.

Table 7 shows the probability of rejection, at 1% level of significance, of the null hypothesis of correct unconditional coverage, *i.e.*, the hypothesis that the probability of the occurrence of a violation is indeed 1%. All the four methodologies present an oversized test. However, the RiskMetric model is rejected in one third of the trajectories, approximately, while the ARCH Quantile VaR is rejected in 4% or 5% of the replications.

[11] This is the uniformly most powerful (UMP) test for a given T.

Table 7
Empirical size for the unconditional coverage test

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Size at 1% significance level				
DGP_1	0.342	0.082	0.080	0.050
DGP_2	0.361	0.100	0.097	0.045
DGP_3	0.346	0.094	0.093	0.042
DGP_4	0.339	0.100	0.098	0.042
DGP_5	0.369	0.106	0.108	0.054
DGP_6	0.370	0.100	0.098	0.037
DGP_7	0.349	0.101	0.101	0.052
DGP_8	0.345	0.100	0.098	0.043
DGP_9	0.359	0.126	0.126	0.047
DGP_{10}	0.327	0.099	0.097	0.034

4. An Empirical Illustration

4.1 The data

We perform an empirical exercise using daily returns, in U.S. dollars, of the Brazilian São Paulo Stock Exchange Index (IBOVESPA) from 07/08/1996 to 03/24/2000, summing up 920 observations. We choose this sample because we want to check the performance of each VaR methodology during periods of market turmoil. Indeed, the above sample period covers the Korean Crisis in 1997, the Russian crisis in 1999, and the blast of the technology-stock market bubble in 2000. Figure 6 displays the behavior of the IBOVESPA return over the above sample period.

It is well known that GARCH volatility models tend to predict implausible high VaR during periods of market turmoil. This happens because GARCH models treat both large positive and large negative return shocks as indicators of high volatility, whose only large negative return shocks indicate higher Value-at-Risk. In other words, volatility and VaR are not the same thing, and this is not taken into account by the non-robust GARCH volatility models.[12] In contrast, the robust ARCH Quantile, while predicting higher volatility in the ARCH component, assigns a much larger weight to a big negative return shock than to a big positive return shock and, thus, we expect the resulting estimated VaRs to be closer to reality during periods of market turmoil (see similar argument in Wu and Xiao (2002)).

[12] It is true, however, that the APARCH model assigns different weights to negative and positive shocks, which helps avoiding estimation of high VaRs during periods of market stress.

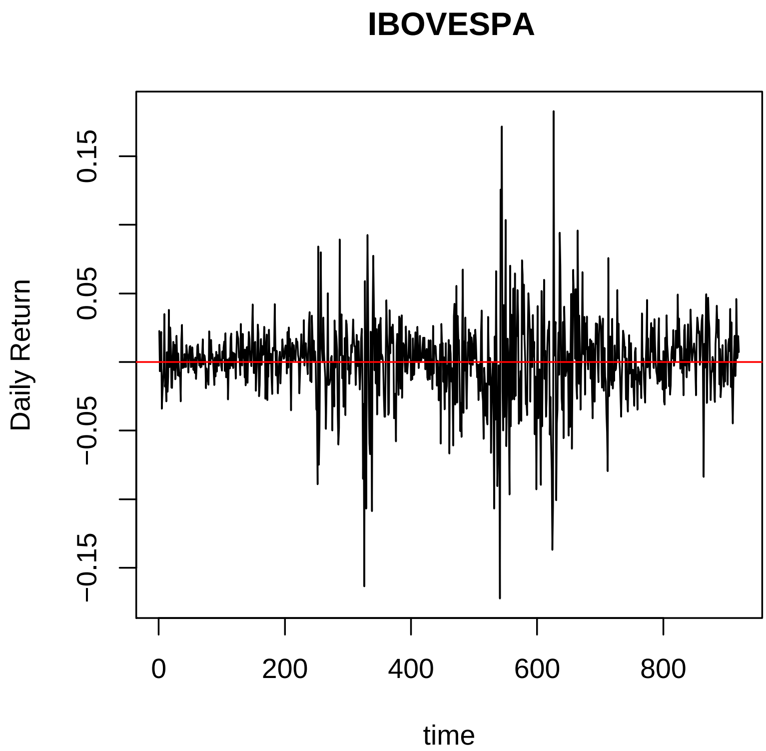

Figure 6
IBOVESPA return time series

Next, we examine the distribution of the Ibovespa return. It was argued in this paper that, unlike the non-robust methods, the ARCH Quantile method has no need to specify the distribution of the innovation process, z_t. The importance of this robustness aspect is revealed by the Quantile-Quantile plots (Q-Q plot). Recall that the Q-Q plot graphs the quantiles of the observed variable (IBOVESPA return) against the quantiles of a specified distribution. Hence, if the returns are distributed according to that specified distribution, then the points in the Q-Q plots should lie alongside a straight line. The next two Figures show the Q-Q plots against all the five innovation distributions used in our Monte Carlo experiment.

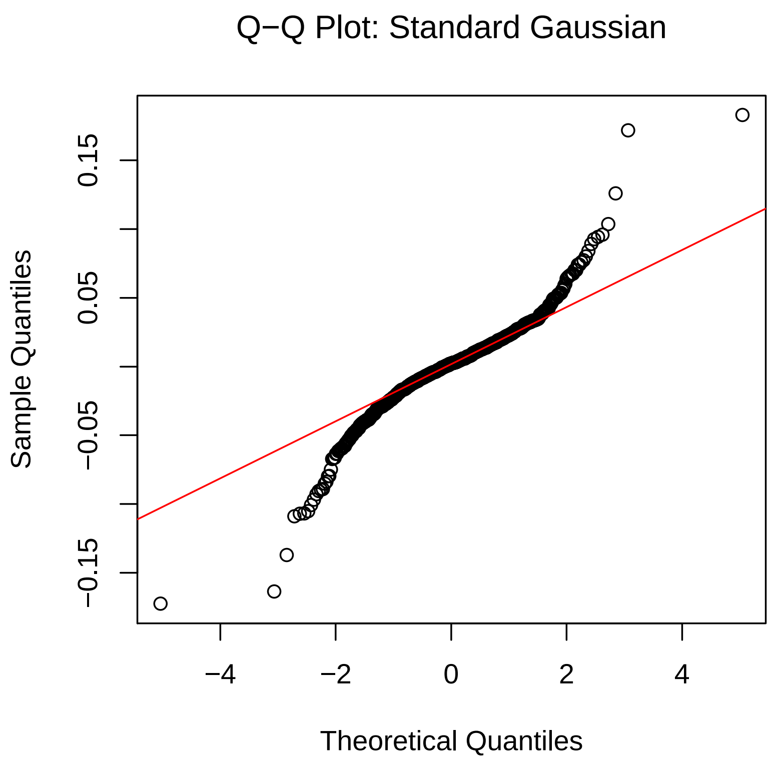

Figure 7
Q-Q Plot of the IBOVESPA return versus the normal distribution

Figure 7 indicates that it is primarily large negative and positive shocks that are driving the departure from normality. In other words, the tail behavior of the distribution of the IBOVESPA return is far different from the tail behavior of a Gaussian distribution. Figure 8 exhibits the Q-Q plots against the four remaining distributions. It seems that the student-t distribution with 3 degrees of freedom approximates the data distribution reasonably well, but there could still be extreme positive and negative observations that lie off the straight line suggesting that the student-t distribution with 3 degrees of freedom does not fit the tail of the data distribution pretty well, which is particularly bad for risk measures. Figure 8 also shows that the data distribution departs from the other three distributions considered in our Monte Carlo experiment, but that they fit the data distribution even worse than the previous two distributions.

Thus, given this uncertainty about the specification of the innovation distribution, how could we go about computing Value-at-Risk correctly? A natural answer to it is to use a robust method against distribution misspecification, such as the method based on the ARCH Quantile model.

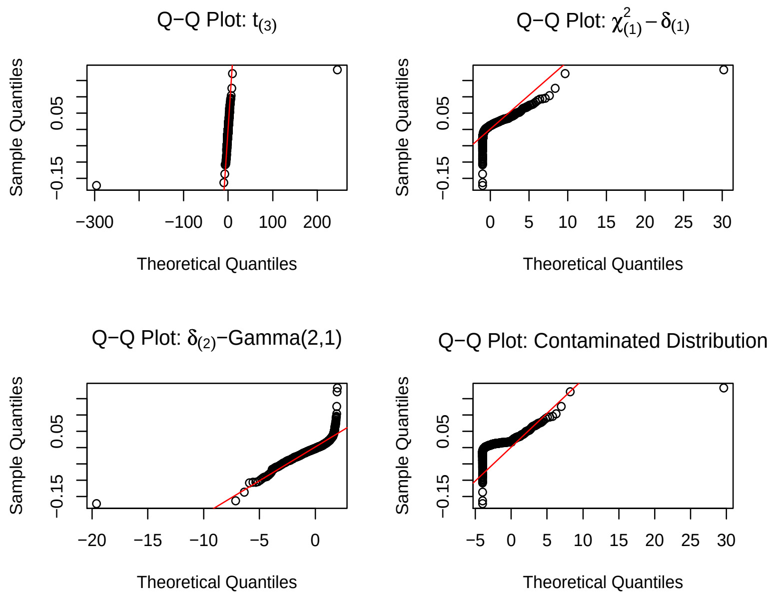

Figure 8
Q-Q Plots of the IBOVESPA return versus the other distributions

4.2 The estimated VaRs

Notice that there is no VaR forecast for the first 250 observations, due to the first temporal window. Hence, there are 670 1-day-ahead forecast observations for each $VaR(1\%)$ methodology, ranging from 07/11/1997 to 03/24/2000. Since it is a 1% Value-at-Risk, we expect about seven violations. Before we move to the results of our empirical illustration, it is important to mention that the comparison of different VaR methodologies depends on the specification of μ_t (the conditional mean) and the specifications of the conditional volatility. In the empirical example that follows, we consider $\mu_t = \beta_o + \beta_1 y_{t-1}$.[13] As for the number of lags appearing in the definition of ε_t in equation (9), we follow Wu and Xiao (2002) and use a Wald

[13]First-order serial correlation in returns is not necessarily at odds with the efficient market hypothesis. See Campbell et al. (1997) for a detailed discussion.

test to determine the optimal lag choice.[14] As for the order of the GARCH and APARCH models, we follow Enders (2003), pp 136) and use adjusted information criteria.

Table 8 shows the results of the Unconditional Coverage Test.

Table 8
Unconditional coverage test

Methodology	VaR 1	VaR 2	VaR 3	VaR 4
Number of Violations	14	12	13	11
Test Statistic LR_{uc}	6.115232	3.429641	4.693915	2.335267
P-Value	0.013402	0.064036	0.030270	0.126473

Note that the number of violations in the RiskMetrics methodology is the greatest, while the ARCH(1) Quantile presents the greatest p-value in the Unconditional Coverage test, that is to say, we do not reject, even at a 10% significance level, the null hypothesis that the conditional probability of occurrence of a violation in this 1% Value-at-Risk estimated time series is indeed 1%. The GARCH(1,1) does not present a bad result, since we do not reject, at least at 5% significance level, the null hypothesis of correct unconditional coverage.

5. Conclusion

We perform a Monte Carlo experimet to compare four different Value-at-Risk methodologies, RiskMetrics, Gaussian GARCH(1,1), Generalized Student-t APARCH(1,1), and ARCH(1) Quantile, under 10 different data generating processes. The ARCH(1) Quantile methodology does not assume any distribution for the returns, and this robustness is shown to avoid trajectories with too many violations. The number of violations tends to be higher in the non-robust methodologies.

We also perform an empirical exercise by applying the four Value-at-Risk methodologies to the daily return of the IBOVESPA (measured in dollar values) in a period of market turmoil (1996-2000), when the Korean crisis, the Russian crisis, and the blast of the technology-stock market bubble occur. We show again that the ARCH(1) Quantile methodology dominates the non-robust methodologies, in the sense that it presents the least number of violations.

[14] As all the other R and Ox codes used in this paper, this test is available at www.fgv.br/aluno/bneri.

References

BACEN (2001). Resolution 2.829. Brazilian Central Bank.

Beaudry, P. & Koop, G. (1993). Do recessions permanently change output? *Journal of Monetary Economics*, 31:149–163.

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

Box, G. E. P. & Muller, M. E. (1958). A note on the generation of normal random deviates. *Annals of Mathematical Statistics*, 29:610–611.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stocastic properties of stock returns. *The Journal of Finance*, XLVII(5).

Campbell, J. Y., Lo, A. W., & MacKinlay, C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106.

Doornik, J. A. (2002). *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press and Oxford, London, 3rd. edition.

Enders, W. (2003). *Applied Econometric Time Series*. Wiley Series in Probability and Statistics, 2nd. edition.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50:987–1008.

Engle, R. F. & Manganelli, S. (1999). CAViaR: Conditional autoregressive value at risk by regression quantiles. University of California, San Diego, Working Paper.

Engle, R. F. & Manganelli, S. (2001). Value-at-risk models in finance. Working Paper Series, European Central Bank.

Fernández, C. & Steel, M. (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association*, 93:359–371.

Giot, P. (2003). The information content of implied volatility in agricultural commodity markets. *Journal of Future Markets*, 23:441–454.

Giot, P. & Laurent, S. (2003). Value-at-risk for long and short positions. *Journal of Applied Econometrics*, 18:641–664.

Giot, P. & Laurent, S. (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11:379–398.

Koenker, R. (2005). Quantile regression. Econometric Society Monographs.

Koenker, R. & Basset, G. (1978). Regression quantiles. *Econometrica*, 46:33–50.

Koenker, R. & Zhao, Q. (1996). Conditional quantile estimation and inference of ARCH models. *Econometric Theory*, 12:793–813.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3:73–84.

Lambert, P. & Laurent, S. (2001). Modelling financial time series using GARCH-Type models and a skewed student density. Mimeo, Université de Liège.

Laurent, S. & Peters, J. P. (2005). *G@RCH 4.0, Estimating and Forecasting ARCH Models*. Timberlake Consultants Press and Oxford, London.

L'Ecuyer, P. (1999). Tables of maximally-equidistributed combined LFSR generators. *Mathematics of Computation*, 68:261–269.

Lopez, J. A. (1999a). Methods for evaluating value-at-risk estimates. *Economic Review*, 2:3–17. Federal Reserve Bank of San Francisco.

Lopez, J. A. (1999b). Regulatory evaluation of value-at-risk models. *Journal of Risk*, 1:37–64.

Marsaglia, G. (1997). A random number generator for C. Discussion Paper, Posting on usenet newsgroup sci.stat.math.

Marsaglia, G. & Zaman, A. (1994). Some portable very-long-period random number generators. *Computers in Physics*, 8:117–121.

Matsumoto, M. & Nishimura, T. (1998). Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modelling and Computer Simulation*, 8:3–30.

Morgan, J. (1996). Riskmetrics. J. P. technical document.

Nam, K., Washer, K. M., & Chu, Q. C. (2005). Asymmetric return dynamics and technical trading strategies. *Journal of Banking & Finance*, 29:391–418.

Park, S. & Muller, K. (1988). Random number generators: Good ones are hard to find. *Communications of the ACM*, 31:1192–1201.

Wu, G. & Xiao, Z. (2002). An analysis of risk measures. *Journal of Risk*, 4:53–75.

Appendix

Computational Details

We use R and Ox to conduct this experiment. The former is an open-source computer-programming language. Hence, it (and its source codes) can be freely downloaded from the Internet.[15] This availability keeps R always updated with the most recent techniques in Statistics, Econometrics and Computer Science. Ox can also be downloaded from the Internet for research purposes.[16]

The time series are generated in R because their default[17] Random Number Generator (RNG) is the Mersenne-Twister (see Matsumoto and Nishimura (1998)), an impressive RNG with period $2^{19937} - 1$ and equidistribution in 623 consecutive dimensions (over the whole period).

This Monte Carlo experiment is extremely computational intensive. For each observation in the VaR forecast, there are three likelihood maximizations – RiskMetrics, Gaussian GARCH(1,1) and Skewed Student-t APARCH(1,1) – with 250 observations (the window length) each. The third maximization occurs in a 7-dimensional hyperplane within an 8-dimensional space (5 parameters for the APARCH(1,1) specification and 2 parameters for the Skewed Student-t distribution). The R is supposed to take several months to conclude all the Monte Carlo, even in our server with 4 Intel Pentium IV Xeon at 2.8 GHz, a 4 GB RAM and a 100 GB SCSI Hard Disk running Linux Debian as Operating System. R is not so fast since it is an interpreted language: the interpreter executes the code line by line, so the user can enter a single line and see the results, which makes it more interactive and user-friendly. Ox is one order of magnitude faster than R since it is a compiled language: the compiler analyzes the code as a whole, really optimizing it before executing it, which makes it much faster in large computations.

However, the ARCH(1) Quantile VaR must be estimated in R because the *quantreg* package for R, version 3.82, May 15, 2005, developed mostly by Roger Koenker himself, is very complete and operational.[18] Thus, we proceed as follows: R generates the time series, then it calls Ox to estimate the first three VaR methodologies.[19] Next, Ox returns these VaR forecasts to R, which estimates the ARCH(1) Quantile VaR, computes the descriptive statistics, and saves the results in the Hard Disk. R then generates another time series and the next replication begins. Using this hybrid solution (Ox and R), all the Monte Carlo experiment takes a couple of months. Every written code, for both R and Ox, used in this paper are available at www.fgv.br/aluno/bneri.

[15] http://www.r-project.org

[16] http://www.doornik.com

[17] Alternatively, the user can select one of the eight RNGs available, or supply another one.

[18] Ox has also a code called *rq* to, at least, estimate quantile regression, but it is quite incomplete. It was written by a Roger Koenker's student, Daniel Morillo, but it has been abandoned in its version 1.0, August 1999.

[19] Our Ox code uses some function from the package G@RCH 4.0, by Laurent and Peters (see Laurent and Peters (2005)), for Ox.