Volatility Forecasting before the Subprime Crisis

Felipe Araújo de Oliveira* Sinézio Fernandes Maia†

Abstract

: This paper aims to test the best model volatility forecasting using daily returns sample from Brazilian and US stock markets. This information is useful to portfolio managers and Central Bankers seeking to understand possible effects of policy interventions in financial markets. The period covered is from January of 2002 to December of 2007. The motivation to test the forecasting potency of these models comes from Engle, Patton et al. (2001), where a good volatility model must be able to predict. The path followed was the same of Cavaleri (2008) and Kuester, Mittnik e Paolella (2006). The sample period is from January of 2002 to December 2007. Using Garch Family models and VaR. The results suggest most of the models behave badly in a time of transition as preceded by the subprime crisis. The best models to predict one-step-ahead conditional variance were parsimonious iGARCH and standard GARCH models to both countries. The distribution of returns considered was crucial in results of backtesting.

Keywords: GARCH, Forecasting Volatility, VaR Backtesting, Subprime Crisis, Market Risk.

1 Introduction

The world is flat was the title of Thomas Friedman's book in 2004. The world would discover that it was not that true few years ahead. The concern of this paper is to realize if risk models were able to make financial risk managers aware of the realized risk in this period. Thus, the question to answer here is: how is the suitability of these models in transition times?

The ambition here is not to investigate if these volatility estimations or forecasting was able to predict the financial crisis. In fact, the main function of these measures have been to assess risk management and decision-making about asset allocation. The period that preceded the sub-prime crisis is featured by low volatility and to a policy of low interest rates implemented by FED since 1998 (ALEXANDER, 2008). Actually, is evidenced that high volatility does not preceed financial crisis in the real economu. Danielsson, Valenzuela e Zer (2016) built a cross-country database of 211 years time span to encounter evidences on the relationship between volatility and financial crisis. The linkage between volatility and the real economy is a

*Graduando em Economia pela UFPB <felipe.8.araujo@gmail.com>

†Professor Adjunto da Universidade Federal da Paraíba - UFPB <sineziomaia@yahoo.com.br>

negative. Apparently, low volatility incentives economic agents to establish a risk-taking posture, endogenously affecting the likelihood of future shocks.

Of course, market risk is not trivial to observe. Even though mean-variance analysis assumptions have been relaxed since Markowitz (1952) towards risk measurement, it is still a difficult task. Danielsson (2011) is emphatic when he points out: “Financial risk is a forecasting, not a measurement”. Christoffersen e Diebold (2000) state good arguments suggesting the fast decaying predictability of these measures when time horizon becomes longer.

Since Mandelbrot (1963) evidenced the presence of fat tails in financial time series, measures such Value-at-Risk or other methods as conditional heteroscedasticity have been used widely in academia and by practitioners. This tendency is still in course, as we observe in the recent literature production, there are methods trying .

Mabrouk (2016) evaluate the daily conditional volatility and h-step-ahead Value-at-Risk (VaR) using long memory GARCH family methods. The exercise is done with the indexes: Nasdaq100, Dow Jones, S&P 500, DAX30, CAC40, FTSE100 and Nikkei225. Skewed Student – t FIAPARCH(1,d,1) model provides more accurate one-day-ahead VaR forecasts than using 5 or 15-day-ahead.

Haugom e Ullrich (2012) used derivatives to improve volatility forecasting of spot electricity market, splitting it in continuous and jump components. The authors introduced the term forward realized volatility calculated from one-day-ahead forward, and used the implied volatility to feed up their model when compared to naïve measures.

For instance, some innovative method overcome the errors of GARCH models. Kristjanpoller e Minutolo (2015) uses Artificial Neural Network with GARCH models to forecast prices of gold of spot and futures. They accomplished better results than conventional GARCH forecasting for forecasts of 14-day-ahead and 18-day-ahead.

Algorithmics methods have also been used as well to improve volatility forecasting. Sermpinis et al. (2015) introduces a RG-SVR model for optimal parameter selection and identifies the optimal features of the series, finally providing a combination between them. This approach proved itself, also profitable in trading applicantions.

Implications of combination of volatilities precision is also in Cavaleri (2008). The author uses unconditional, conditional variance and stochastic volatility models and their combination. The combination using an OLS method presented better results than the other combinations and univariate models.

Kuester, Mittnik e Paolella (2006) systemically presents the literature of the theme. Doing a broad empirical exercise of forecasting using a wide number of models and, distributions to forecast VaR. Combining heavy tailed GARCH specification combined with an Extreme Value Distribution (EVT) presented the best results.

Some applications uses CAViaR model in order to avoid subadditivity problem of a VaR. Drakos, Kouretas e Zarangas (2015) aims to test the performance of alternative CAViaR specifications, splitting the sample in before, during and after the financial crisis. The authors find a relatively better performance than conventional VaR models.

An application to Brazil is found in Gaio et al. (2015), whose results suggest that models based on EVT and GPD distributions proved robust in calm and crisis periods.

The objective of this paper is to contribute to the literature of risk forecasting presenting how conditional models based in different specifications and distributions behave in times of transition to Brazil (i.e reference to an emerging market) and USA (i.e reference to a developed market). The methodology used follows Kuester, Mittnik e Paolella (2006) in the backtests used and Cavaleri (2008) in the estimation procedure. The period covered relatively calm, going from 2002 to 2008. To deliver the results proposed, are used GARCH family models to predict conditional variance and VaR using different distributions of returns, such: Normal, Skewness – Normal, t-Student, Skewness t-Student, GED, Skewness GED and JSU.

Besides the former section, the work is divided in four sections. The next present data features. The third section presents the methodology used. In the fourth section, can be accessed the results. Finally, the fifth section presents concluding remarks.

2 Methodology

2.1 Risk:Definition and Measurement

Investment decisions occur in scenarios surrounded by uncertainty. Naturally, based on their preferences the investor seeks for utility maximization. Hence, they will attempt to maximize returns while minimize risk.

In contrast, market risk is not easily measured. In **??**) definition, "financial risk is a forecasting, not a measurement". Its provocative statement is based on the wide set of factors to control, in order to measure risk effectively. Financial risk, differently of return movements, is not easily seen. Therefore, statistical inference methods are applied on price movements in order to obtain fitted estimations of risk (DANÍELSSON, 2011).

When Markowitz (1952) introduced the mean-variance analysis, introduced the trade-off between risk and return. One of his main breakthrough was a shift of the definition of markets risk from a bunch of subjective information, to standard-deviation measure [3].

For shore, price oscillations, or market risk just matters in a time dimension. While time advances, the expected deviation on wealth portfolio affects investor's satisfaction. Thus, a measure of risk based on compounding standard deviations is desirable to feed managers portfolio analysis [4].

The trivial way to forecast volatility is to measure it in a moving average standard deviation. Where the last observation is excluded and the new introduced in the measure at the pace time advances. A more formal definition is given in Hull (2006) would be the average standard deviation of annualized diary returns. Called historical volatility σ:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} u_i^2 - \frac{1}{n(n-1)}(\sum_{i=1}^{n} u_i)^2} \tag{1}$$

The biggest challenge faced by using a moving average volatility model is to use the right window of observations. A large one, could imply in unnecessary information (HULL, 2006). A short one, would imply in a two unstable measure of volatility. Hence, its measure can easily be unbiased. Moreover, an unconditional measure cannot model clusters adequately, and very sensible to extreme outcomes, as well.

To minimize the effect of an extreme event biasing the average, practitioners use the risk metrics approach. It consists of an Exponential Weighted Moving Average (EWMA), where the last observations are more important when measuring the risk. The importance

[3] Other important aspects of his framework is the quadratic form of utility functions. He also introduces an important intuition to practitioners, in a well-diversified portfolio, co-movements risk is what matters. His assumptions over distribution of returns and innovations proposed to fulfill the gaps of risk measurement in the text advancement.

[4] A broader discussion about volatility as a risk measure is found in Daníelsson (2011) and Hull (2006).

of the observations is exponential decaying. Therefore, forecasting generated by an EWMA model give less importance to irrelevant information, improving the accuracy of the estimation (ALEXANDER, 2008).

Formalizing:

$$EWMA(r_{t-1}, \cdots, x_1, \lambda) = \frac{r_{t-1} + \lambda r_{t-2} + \lambda^2 r_{t-3} + \cdots + \lambda^{t-2} r_1}{1 + \lambda + \lambda^2 + \cdots + \lambda^{t-2}} \tag{2}$$

Where,λ , is the constant $0 < \lambda < 1$, that will smooth the impact of old information. Given that asymptotically, $n \to \infty$, lambda tends to zero ($\lambda^n \to 0$). Where:

$$1 + \lambda + \lambda^2 + \cdots = (1 - \lambda)^{-1} \tag{3}$$

When time (t), becomes longer:

$$\sigma^2 = (1 - \lambda) \sum_{i=1}^{\infty} \lambda^{i-1} r_{t-i}^2 \tag{4}$$

Given the fact λ is constant and the same to every asset, EWMA models have their forecasting precision deepened. Although, they are easily used for practitioners and in applications of multivariate estimation.

Summing up, unconditional volatility measures are sensitive to extreme outcomes. Additionally, they present some problems when investors take decisions based on them. Daníelsson (2011) gives an intuitive example showing that assets that behaves differently, can have the same mean-variance measures, in other words, they have the same position in the trade-of line. Thus, decision taking based on volatility can be misleading.

On the other hand, Value-at-Risk (VaR) measures present ease in implementation and backtesting. As Jorion (1997) points out, VaR as an statistical risk measure of potential losses. Financial risk managers using it as a risk measure and managing the risk exposure, in a determined horizon time and based on a particular significance level justifies its applicability [5].

The $100\alpha\%$ h-day is a more formal definition to the loss amount that could be exceeded by a probability α, in a frozen portfolio over the next h days. Therefore, a quantile estimation α of a random variable $x_{ht,\alpha}$ over h-day of profits and losses distribution:

$$P(B_{ht} P_{t+h} \breve{} P_t < x_{ht,\alpha}) = \alpha \tag{5}$$

VaR can be estimated from a time series return distribution, when this happens is expressed as a percentage of portfolios (Q) value. Defining VaR by an h-day return of a portfolio (random variable), it can be expressed as

$$Q_{ht} = \frac{B_{ht} P_{t+h} \breve{} P_t}{P_t} \tag{6}$$

$$P(Q_{ht} < x_{ht,\alpha}) = \alpha \tag{7}$$

[5] VaR has to present some axioms to be considered a coherent risk measure, such: i) Monotonicity; ii)Subadditivity; iii) Positive Homogeneity; and, iv) Translation invariance. Depending on the application, one of them can be violated. This paper does not aims to discuss the theoretical features of the model. A complete discussion is found in Artzner et al. (1999), Alexander (2008) and Daníelsson (2011)

2.2 Conditional Volatility Models

Since Mandelbrot (1963) showed the presence of clusters in cotton time series: returns tend to peak, driving a behavior of high volatility being preceded by low volatility. The simplifying hypothesis of normal distributions are widely criticized, the distance of the theoretical framework used to reality is faced as problem. These measures of risk are useful upon the normal distribution hypothesis (multivariate normal distribution when a portfolio).

The response from academia came first from Engle (1982), whose work concern was in the adequacy of financial series to time-varying behavior. Building a functional form that introduces the lagged innovation in the set of information to explain variance, he models a conditional variance behavior. His paper introduces the Autorregressive Conditional Heteroskedasticity (ARCH).

As Tsay (2005) highlights, shocks dynamics in stock markets are serially not correlated. However, the square root of residuals is linearly dependent. Thus, the quadratic form of innovation term is desirable to estimate volatility on markets. The intuition about ARCH models suggest that modelling their squared residuals is possible to explain the conditional variance:

$$(\epsilon\Phi_{t-1}) = \sigma_t u_t, u_t \sim i.i.d(0,1) \tag{8}$$

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i \tag{9}$$

$$\epsilon_{t-1}^2 = w\alpha_i(L^q) \tag{10}$$

$(\epsilon\Phi_{t-1})$ is the set of information, ϕ, until $t-1$. The term, ϵ, is independent and identically distributed. Therefore, it is asymptotically are normally distributed. w represents the intercept; σ_t^2 the conditional variance;α_i is the parameter of the model; ϵ_t^2 is the quadratic form of residuals; and, (L) is the lag operator:$\alpha(L) = \alpha_1(L) + \alpha_2(L^2) + +\alpha_i(L^q)$.

Bollerslev (1986) simplifies the ARCH framework, generalizing the specification of the model – analogously as the passage from an AR, to and ARMA model. Developing the Generalized ARCH (GARCH), the number of parameters are reduced; hence, the model parsimoniously explains conditional variance. Formally, the introduction of past volatility reduces the lag structure of ARCH models (DANÍELSSON, 2011).

$$\sigma^2 = w + \sum_{i=1}^{q} \alpha_i \epsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{11}$$

If $p > 0$, there is some order of conditional dependence in the model; α_i is the coefficient of the innovation term; and, β_j is the parameter of the past volatility. A restriction of positiveness in the parameters is done, whereas the sum of, $\alpha_i + \beta_j < 1$, is less than one [6].

Another important stylized fact about returns is the leverage effect. Volatility does not react at the same manner to positive and negative news. It tends to decline sharpen when negative shocks affect markets. The GARCH model captures the conditional heteroscedasticity

[6] Daníelsson (2011), highlights that using maximum likelihood as estimator can find out a global maximum where the second condition of $\alpha_i + \beta_j < 1$ cannot be attended. Thus, it can be flexible when the purpose is forecasting

behavior. However, the model does not consider the asymmetry in returns, what would drive the model to some misspecification in some situations.

In the literature of conditional heteroscedastic models there are two formulations seeking to model this fact. The first is proposed in Nelson (1991), the Exponential GARCH (EGARCH) model:

$$(\epsilon_t \mid \Phi_{t-1}) = \sigma_t u_t, u_t \sim i.i.d(0,1)$$

$$log\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i g(\epsilon_{t-i}) + \sum_{j=1}^{p} \beta j log\sigma_{t-j}^2$$

where $g(.)$ is the asymmetry parameter in the model, and ϵ_{t-i} is the innovation term. Given its logarithmic form, the restrictions on the signal of the parameters is not necessary[7]. Thus, θ captures the leverage effect and δ the magnitude of itself.

Positive news has the impact of

$$\theta\epsilon_t + \delta(\epsilon_t \check{} E\epsilon_t)$$

.

Whereas, positive news has its effects given by Eϵ_t) (FRANCQ; ZAKOIAN, 2011).

On the other hand, the Threshold GARCH model developed in Zakoian (1994) has its generalized form such:

$$\sigma_t^a = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^a + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^a + \sum_{k=1}^{r} {}_k(d_{\epsilon_{t-k\leq 0}}) \epsilon_{t-i}^a{}^a \tag{12}$$

Where,$d_{\epsilon_{t-k\leq 0}} \epsilon_{t-i}^a{}^a$ is a dummy variable. It splits the estimation in two phases. When the innovation term is positive the response of volatility would be $(\alpha_i)\epsilon_{t-i}^2$; and other when the impact is negative, the response of volatility will be $(\alpha_i +_i)\epsilon_{t-i}^2$.

2.3 Models Diagnostic

Though on the models described in this paper is possible to select the best fitted ones based on information criteria such Akaike(AIC), Schwarz (SBC) and Hann-Quen(HQ), or using a maximum local likelihood. Here is followed what Engle, Patton et al. (2001) points out: "A volatility model should be able to forecast volatility".

Hence, is followed the approach based on Cavaleri (2008) and Kuester, Mittnik e Paolella (2006) to measure the suitability of the forecasting methods used to Brazilian and USA stocks markets before the financial crisis. As stated by wheelwright1998forecasting, "To the consumer of forecasts, it is the accuracy of the future forecast that is most important"[8]

One of the measures used is the Mean Square Error (MSE), that takes the difference between the observed (y_t) and forecasted $(\hat{y}_t)$, intensifying the distance between both:

$$MSE = \sum_{t=1}^{n} \frac{(\hat{y}_t - y_t)^2}{n} \tag{13}$$

[7] Nelson (1991) assumes $g(.)$ as linear combination of its residuals: $g(\epsilon_t) = \omega + \theta\epsilon_t + \delta(\epsilon_t \check{} E\epsilon_t)$

[8] A further knowledge of statistical accuracy of forecasting methods is in Wheelwright, Makridakis e Hyndman (1998).

The next step to find the best variance forecasting method is the Mean Absolute Error (MAE). That is the distance between the observed (y_t) and the forecasting ($\hat{y}_t$) in module. It has the advantage to be more interpretable, the outcome does not depend upon the scale of the data. Mathematically:

$$MAE = \sum_{t=1}^{n} \frac{\mid \hat{y}_t - y_t \mid}{n} \tag{14}$$

2.4 Backtesting Methods

The Basel Accords in Committee (1996) requires that regulated financial institutions reserve capital to protect themselves from risk. The measure of risk used to enhance the exposure to risk is a VaR on a threshold of 1p.p., where as much violations are realized over this limit, as much capital this institutions have to set aside. However, these procedures might not be able to identify the model that provides the correct suitability to data.

There is no straight approach to choose the best model for forecasting risk than backtesting. It surpasses the approach of parameters test of significance or residuals analysis. Backtesting consists of a procedure that compare value-at-risk (VaR) forecast generated by a particular model, comparing them with post realized returns (DANÍELSSON, 2011). A complete discussion about backtesting methods can be found in Campbell (2005). This paper will adopt the procedures developed in Kupiec (1995) and Christoffersen (1998).

The former is concerned in the number of violations that a VaR faces given a α in a span of time. Kupiec (1995) constructs a sequence of zeros and ones following a Bernoulli distribution as a backtest. His proportion of failures techqniques represent on as a violation and zero as no violation. The following formalization is given in Daníelsson (2011) as:

$$H_0 : \eta \sim B(p) \tag{15}$$

Where B stands for the Bernoulli distribution. Probability (p) can be estimated by $\hat{p} = \frac{v_1}{W_T}$, being v the violations number and W_T the windows size of backtesting. Of course, part of the data sample (until W_E) will be used to estimate the model, and part is estimated in a rolling window, $W_E + 1$. Thus the restricted likelihood fuction is:

$$\mathcal{L}_R(p) = \prod_{t=W_E+1}^{T} (1 \breve{} p)^{1-\eta_t} = (1 \breve{} p)^{v_0}(p)^{v_1} \tag{16}$$

Whether can be used $\mathcal{L}_R = \mathcal{L}_U$,

$$LR = 2(log\mathcal{L}_U(p) \breve{} log\mathcal{L}_U(\hat{p})) \tag{17}$$

$$LR = 2log\frac{(1-\hat{p})^{v_0}(\hat{p})^{v_1}}{(1-p)^{v_0}(p)^{v_1}} \tag{18}$$

$$\sim^{asymptotic} \chi^2_1 \tag{19}$$

Although the Bernoulli coverage test does not assume distribution for the returns, it provides an intuitive benchmark to the VaR precision. (DANÍELSSON, 2011).As Campbell, Huisman e Koedijk (2001) stated two shortcomings deepen this test: i) it presents low power when sample size are small, systematically under reporting risk; ii) the test does not examine the occurrence of clusters in tails, what lead to an underestimation of risk.

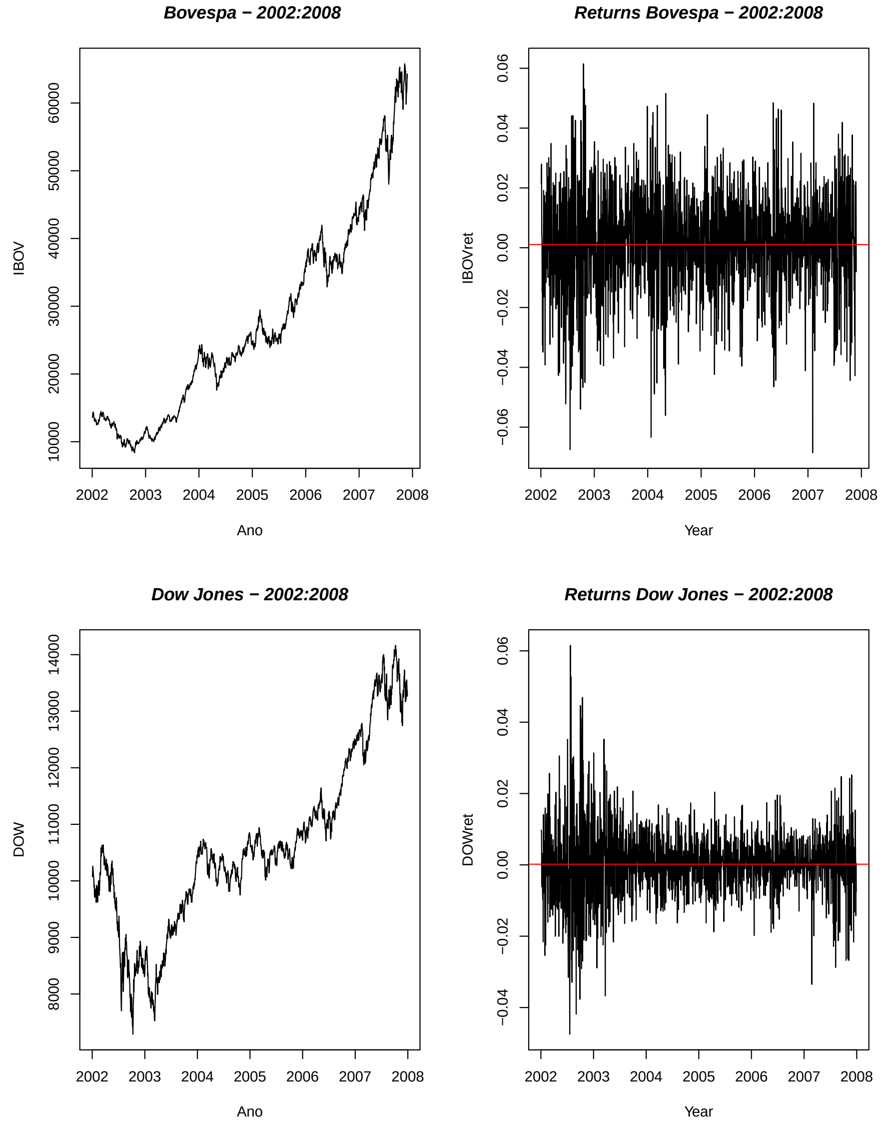
3 Results

3.1 Data

The time series used in this paper are Brazilian and USA indexes, respectively Bovespa and Dow Jones. They can be considered weighted portfolios, representing the overall performance of these countries stock markets. Criteria of negotiability and relevance are used to select the stocks participation in the index.

Using diary frequency, the sample covers the period that goes from 02/01/2002 to 31/12/2007 [1] of log return, collected in Yahoo Finances database, using the software Grapher OC.

Figure 1 – Bovespa and Dow Jones prices and log returns – Period: jan/2002 to dec/2008



In a preliminary inspection in the database in 2 is possible to observe the presence of clustering in the data, where maximum peaks are observed between 2002 and 2003. The perturbance in this period is marked by the burst of stocks .com bubble. Additionally, Brazil experienced the transition of presidency, when the Works Party candidate won the elections and raised concerns about fiscal and monetary policy. These events were preceded by the currency and Russian crisis, in 1999.

[1] The number of observations is diferent in each series. Due the difference in holidays and closed days of Stocks Exchange, Bovespa presented a length of 1488 price observations, while Dow Jones totalized 1510.

The period that goes from the end of 2003 to the beginning of 2008 is a relatively calm period in world financial markets. This can be evidenced by the USA stocks returns in ??. Even though, the Bovespa presents high variability in its returns, in this period the stock index behaves relatively less in the sample. In 2007, the series present a higher volatility, as reflect of the subprime crisis in USA. However, the peaks of these periods were lower than the ones of .com bubble.

In 1 is possible to identify the presence of a high excess kurtosis in both series, being 0.67 to Bovespa and 3.63 to Dow Jones indexes. Moreover, is possible to verify in ?? and ??, the presence of autocorrelation in the squared series[2]. Additionally, the presence of skewness is attest to Bovespa being negatively asymmetric distribution, while to Dow Jones it presents positive asymmetry in distribution.

Table 1 – Descriptive Statistics of Bovespa and Dow Jones - 02/01/2002 to 31/12/2007

Statistics	Bovespa	Dow Jones
Mean	0,0009	0,0001
SD	0,0172	0,0098
Max	0,0615	0,0615
Min	-0,0685	-0,0475
Kurtosis	0,6786	3,6310
Skewness	-0,2889	0,2124
Jarque Bera	48,04(0,00)	833,95(0,00)
ADF	-11,10(0,00)	-11,54(0,00)
Length	1487	1509

The non-normality is present in both series. However, the Augmented Dickey Fuller (ADF) test rejects the null hypothesis of unit root in the series. In a preliminary diagnosis, the time series used in this paper presents desirable features to be modelled by a GARCH family specification. The stylized facts of leptokurtosis, clustering and skewness are present in both markets.

3.2 Conditional Variance Accuracy

In ??, is possible to observe the results of the conditional variance forecasting precision to Bovespa. To filter the time series to the stylized facts presented in 3.1, are used a GARCH, EGARCH, TGARCH and iGARCH modelling. All these tests are run in both, normal and

[2] The squared autocorrelation of returns is a justification to use conditional variance modelling financial series

t-student distributions to allow fatter tails. Stating a moving window of 1000 observations, the estimation of the parameters is refitted every day, generating a variance forecasting of length 487 periods to Bovespa and 509 to Dow Jones.

Table 2 – Precision Measures of Time Varying Variances to Bovespa out-of-sample forecasting

				GARCH			
	MSE	MAE	DAC		MSE	MAE	DAC
GARCH(1,1) - Normal	0.2658	1.2286	0.5503	GARCH(1,1) - t-Student	0.2664	1.2282	0.5482
GARCH(1,2) - Normal	0.2658	1.2293	0.5462	GARCH(1,2) - t-Student	0.2662	1.2283	0.5482
GARCH(1,3) - Normal	0.2656	1.2284	0.5482	GARCH(1,3) - t-Student	0.2667	1.2335	0.5482
GARCH(2,1) - Normal	0.2659	1.2298	0.5523	GARCH(2,1) - t-Student	0.2661	1.2291	0.5462
GARCH(3,1) - Normal	0.2662	1.2312	0.5503	GARCH(3,1) - t-Student	0.2654	1.2277	0.5503
GARCH(2,2) - Normal	0.2658	1.2300	0.5523	GARCH(2,2) - t-Student	0.2659	1.2290	0.5462
GARCH(2,3) - Normal	0.2658	1.2298	0.5523	GARCH(2,3) - t-Student	0.2659	1.2296	0.5462
GARCH(3,2) - Normal	0.2657	1.2297	0.5523	GARCH(3,2) - t-Student	0.2657	1.2296	0.5462
GARCH(3,3) - Normal	0.2659	1.2310	0.5462	GARCH(3,3) - t-Student	0.2656	1.2295	0.5503
				EGARCH			
	MSE	MAE	DAC		MSE	MAE	DAC
EGARCH(1,1) - Normal	0.2669	1.2358	0.5236	EGARCH(1,1) - t-Student	0.2667	1.2335	0.5400
EGARCH(1,2) - Normal	0.2667	1.2350	0.5236	EGARCH(1,2) - t-Student	0.2658	1.2319	0.5441
EGARCH(1,3) - Normal	0.2663	1.2335	0.5379	EGARCH(1,3) - t-Student	0.2664	1.2326	0.5462
EGARCH(2,1) - Normal	0.2659	1.2324	0.5482	EGARCH(2,1) - t-Student	0.2658	1.2319	0.5441
EGARCH(3,1) - Normal	0.2661	1.2332	0.5400	EGARCH(3,1) - t-Student	0.2661	1.2325	0.5441
EGARCH(2,2) - Normal	0.2659	1.2324	0.5462	EGARCH(2,2) - t-Student	0.2658	1.2319	0.5441
EGARCH(2,3) - Normal	0.2659	1.2324	0.5462	EGARCH(2,3) - t-Student	0.2658	1.2319	0.5441
EGARCH(3,2) - Normal	0.2659	1.2326	0.5462	EGARCH(3,2) - t-Student	0.2658	1.2319	0.5441
EGARCH(3,3) - Normal	0.2659	1.2327	0.5462	EGARCH(3,3) - t-Student	0.2658	1.2319	0.5441
				TGARCH			
	MSE	MAE	DAC		MSE	MAE	DAC
TGARCH(1,1) - Normal	0.2663	1.2331	0.5359	TGARCH(1,1) - t-Student	0.2662	1.2306	0.5503
TGARCH(1,2) - Normal	0.2668	1.2344	0.5400	TGARCH(1,2) - t-Student	0.2666	1.2326	0.5462
TGARCH(1,3) - Normal	0.2658	1.2321	0.5441	TGARCH(1,3) - t-Student	0.2661	1.2311	0.5441
TGARCH(2,1) - Normal	0.2663	1.2332	0.5338	TGARCH(2,1) - t-Student	0.2662	1.2308	0.5462
TGARCH(3,1) - Normal	0.2664	1.2324	0.5441	TGARCH(3,1) - t-Student	0.2662	1.2321	0.5462
TGARCH(2,2) - Normal	0.2664	1.2339	0.5379	TGARCH(2,2) - t-Student	0.2662	1.2311	0.5462
TGARCH(2,3) - Normal	0.2661	1.2324	0.5462	TGARCH(2,3) - t-Student	0.2661	1.2319	0.5462
TGARCH(3,2) - Normal	0.2661	1.2308	0.5482	TGARCH(3,2) - t-Student	0.2658	1.2308	0.5503
TGARCH(3,3) - Normal	0.2659	1.2317	0.5462	TGARCH(3,3) - t-Student	0.2662	1.2312	0.5482
				iGARCH			
	MSE	MAE	DAC		MSE	MAE	DAC
iGARCH(1,1) - Normal	0.2657	1.2298	0.5462	iGARCH(1,1) - t-Student	0.2666	1.2293	0.5482
iGARCH(1,2) - Normal	0.2655	1.2288	0.5462	iGARCH(1,2) - t-Student	0.2663	1.2283	0.5503
iGARCH(1,3) - Normal	0.2655	1.2293	0.5462	iGARCH(1,3) - t-Student	NC	NC	NC
iGARCH(2,1) - Normal	0.2658	1.2304	0.5523	iGARCH(2,1) - t-Student	0.2663	1.2291	0.5482
iGARCH(3,1) - Normal	0.2660	1.2291	0.5523	iGARCH(3,1) - t-Student	0.2660	1.2291	0.5523
iGARCH(2,2) - Normal	0.2655	1.2297	0.5523	iGARCH(2,2) - t-Student	0.2656	1.2279	0.5482
iGARCH(2,3) - Normal	0.2663	1.2290	0.5503	iGARCH(2,3) - t-Student	0.2665	1.2295	0.5482
iGARCH(3,2) - Normal	0.2656	1.2303	0.5503	iGARCH(3,2) - t-Student	0.2665	1.2295	0.5482
iGARCH(3,3) - Normal	0.2658	1.2309	0.5462	iGARCH(3,3) - t-Student	0.2663	1.2290	0.5503 height

Note:The input of observations used goes from 2002 to 2008. Using daily returns frequency, is used a moving average window of 1000 days. The tests are run in out-of-sample observations totalling 487 to Dow Jones.*N.C. refers to Non Convergence.

Following Engle, Patton et al. (2001) our concern is to find the best model to forecasting. In other words, the least MSE and MAE measure. Whereas the concern of the DAC test is to attest the efficiency of the model to predict the signal of the excess return.

To find the best model, all the combinations of these model restricted to $(p \leq 3, q \leq 3)$. The purpose restricting the lags of the model is to maintain the parsimony at the same time pursuing the best forecasting specification.

Is possible to note in 2 that the best (p, q) combinations in a normal distribution based on MSE are $(1, 1)$ and $(3, 2)$ to GARCH; $(2, 1)$ to EGARCH; $(1, 3)$ to TGARCH; and, $(1, 2)$ to iGARCH. Considering a fatter tail, the best results are attained in GARCH $(3, 3)$, EGARCH$(1, 2)$, TGARCH$(3, 2)$, and iGARCH$(3, 1)$. Being the best model based on this measure to forecast financial risk to Brazil the iGARCH$(1, 2)$ - Normal. Using the MAE measure the best model is the GARCH$(1, 1)$ – t-Student.

3 present the results to Dow Jones Index. When considering the MSE results, all the models are almost tied. Most of them, when estimated in a t-Student distribution present Non Convergence (N.C.) in some forecasting windows, thus the outputs are not presented.The MAE test evidences GARCH$(1, 1)$, iGARCH$(1, 2)$ and iGARCH$(2, 1)$ as more precise.

3.3 VaR Backtesting

The best and most parsimonious specification attested in this work, with acceptability among practitioners is the GARCH$(1, 1)$. This specification will use different distributions assumptions, in order to averiguate the predictability of these models before the 2008 crisis. In 3.1, is observed that in the end of 2007, the volatility higher to both, Brazilian and USA stock market. However, the question that arises here: could the conditional variance models be well fitted in transition times?

To answer this question are used violation measures and independence measures such stated in Kupiec (1995) and Christoffersen (1998) in different thresholds. Stating a moving window of 1000 observations, the estimation of the parameters is refitted every day, generating a forecasting of length 487 periods to Bovespa and 509 to Dow Jones. The realized values are used to the model diagnostic.

In 4 is possible to observe that even though t-Student and normal distributions are extensively used in practice, the risk based on these distributions are underestimated when forecasted using an alpha of 5 or 1 p.p. When considered their skewed versions, the results improve but still are above the limit of violations. However, the results suggest that skewed versions of distributions presents suitability to the period covered in this paper. The best distribution to a GARCH$(1, 1)$ specification is the Jhonson's-SU (JSU) distribution, that exactly fits the number of expected violations to alpha of 1. Moreover, in a more rigorous test, only JSU and GED distributions does not reject the null hypothesis of excess of violations and, joint dependence and excess of violations.

The other models are in the region of rejection of the alternative hypothesis of excess of violations mostly when considered an alpha of 1% or in the dependence test when alpha is 5%. Hence, even though, the models are not underestimating risk to Brazil in a hit sequence, its violations are clustering.

On the other hand, to Dow Jones Index, using the Skewed GED distribution the model generated the most accurate forecasting. Although, none of the models respected the violation zone. More distributions presented suitability when considering the more rigorous backtesting tests: i) normal and skewness normal; ii) and GED and its skewness version. They presented significance rejecting the alternative hypothesis of excess of violations and dependence in the violations. Student distributions did not enable the convergence of the estimations, being excluded of the experiment.

Table 3 – Precision Measures of Time Varying Variances to Dow Jones out-of-sample forecasting

			GARCH				
	MSE	MAE	DAC		MSE	MAE	DAC
GARCH(1,1) - Normal	0.0060	0.5544	0.5343	GARCH(1,1) - t-Student	N.C	N.C	N.C
GARCH(1,2) - Normal	0.0060	0.5554	0.5206	GARCH(1,2) - t-Student	N.C	N.C	N.C
GARCH(1,3) - Normal	0.0061	0.5548	0.5265	GARCH(1,3) - t-Student	0.0060	0.5558	0.5147
GARCH(2,1) - Normal	0.0060	0.5547	0.5383	GARCH(2,1) - t-Student	N.C	N.C	N.C
GARCH(3,1) - Normal	0.0060	0.5547	0.5383	GARCH(3,1) - t-Student	N.C	N.C	N.C
GARCH(2,2) - Normal	0.0061	0.5548	0.5265	GARCH(2,2) - t-Student	N.C	N.C	N.C
GARCH(2,3) - Normal	0.0061	0.5551	0.5363	GARCH(2,3) - t-Student	N.C	N.C	N.C
GARCH(3,2) - Normal	0.0061	0.5551	0.5363	GARCH(3,2) - t-Student	N.C	N.C	N.C
GARCH(3,3) - Normal	0.0060	0.5551	0.5383	GARCH(3,3) - t-Student	N.C	N.C	N.C
			EGARCH				
	MSE	MAE	DAC		MSE	MAE	DAC
EGARCH(1,1) - Normal	0.0060	0.5569	0.5049	EGARCH(1,1) - t-Student	0.0060	0.5563	0.5049
EGARCH(1,2) - Normal	0.0060	0.5570	0.5049	EGARCH(1,2) - t-Student	0.0060	0.5564	0.5068
EGARCH(1,3) - Normal	0.0060	0.5570	0.5049	EGARCH(1,3) - t-Student	0.0060	0.5558	0.5147
EGARCH(2,1) - Normal	0.0060	0.5551	0.5206	EGARCH(2,1) - t-Student	0.0060	0.5550	0.5127
EGARCH(3,1) - Normal	0.0060	0.5574	0.5127	EGARCH(3,1) - t-Student	0.0060	0.5554	0.5166
EGARCH(2,2) - Normal	0.0060	0.5552	0.5166	EGARCH(2,2) - t-Student	N.C.	N.C.	N.C.
EGARCH(2,3) - Normal	0.0060	0.5557	0.5147	EGARCH(2,3) - t-Student	N.C.	N.C.	N.C.
EGARCH(3,2) - Normal	0.0060	0.5572	0.5147	EGARCH(3,2) - t-Student	N.C.	N.C.	N.C.
EGARCH(3,3) - Normal	0.0060	0.5578	0.5029	EGARCH(3,3) - t-Student	0.0060	0.5557	0.5225
			TGARCH				
	MSE	MAE	DAC		MSE	MAE	DAC
TGARCH(1,1) - Normal	0.0060	0.5558	0.5068	TGARCH(1,1) - t-Student	N.C.	N.C.	N.C.
TGARCH(1,2) - Normal	0.0060	0.5562	0.5049	TGARCH(1,2) - t-Student	N.C.	N.C.	N.C.
TGARCH(1,3) - Normal	0.0060	0.5551	0.5304	TGARCH(1,3) - t-Student	N.C.	N.C.	N.C.
TGARCH(2,1) - Normal				TGARCH(2,1) - t-Student	N.C.	N.C.	N.C.
TGARCH(3,1) - Normal	0.0060	0.5560	0.5147	TGARCH(3,1) - t-Student	N.C.	N.C.	N.C.
TGARCH(2,2) - Normal	0.0060	0.5539	0.5304	TGARCH(2,2) - t-Student	0.0060	0.5556	0.5127
TGARCH(2,3) - Normal	0.0060	0.5556	0.5166	TGARCH(2,3) - t-Student	0.0060	0.5553	0.5088
TGARCH(3,2) - Normal	0.0060	0.5552	0.5245	TGARCH(3,2) - t-Student	N.C.	N.C.	N.C.
TGARCH(3,3) - Normal	0.0060	0.5555	0.5029	TGARCH(3,3) - t-Student	0.0060	0.5555	0.5029
			iGARCH				
	MSE	MAE	DAC		MSE	MAE	DAC
iGARCH(1,1) - Normal	0.0060	0.5544	0.5363	iGARCH(1,1) - t-Student	N.C	N.C	N.C
iGARCH(1,2) - Normal	0.0060	0.5544	0.5383	iGARCH(1,2) - t-Student	N.C	N.C	N.C
iGARCH(1,3) - Normal	0.0060	0.5550	0.5206	iGARCH(1,3) - t-Student	N.C	N.C	N.C
iGARCH(2,1) - Normal	0.0060	0.5543	0.5422	iGARCH(2,1) - t-Student	N.C	N.C	N.C
iGARCH(3,1) - Normal	0.0060	0.5547	0.5422	iGARCH(3,1) - t-Student	N.C	N.C	N.C
iGARCH(2,2) - Normal	0.0060	0.5546	0.5343	iGARCH(2,2) - t-Student	N.C	N.C	N.C
iGARCH(2,3) - Normal	N.C.	N.C.	N.C.	iGARCH(2,3) - t-Student	N.C	N.C	N.C
iGARCH(3,2) - Normal	0.0060	0.5547	0.5324	iGARCH(3,2) - t-Student	N.C	N.C	N.C
iGARCH(3,3) - Normal	0.0060	0.5552	0.5402	iGARCH(3,3) - t-Student	N.C	N.C	N.C height

Note:The input of observations used goes from 2002 to 2008. Using daily returns frequency, is used a moving average window of 1000 days. The parameters are re estimated after every one-step-ahead forecasting. The tests are run in out-of-sample observations totalling 509 to Dow Jones.*N.C. refers to Non Convergence.

4 Conclusion

The objective of this paper was to contribute to the literature of risk forecasting presenting how conditional models based in different specifications and distributions behave in times of transition to Developed (USA) and Emerging Markets (Brazil). The methodology used follows Kuester, Mittnik e Paolella (2006) in the backtests used and Cavaleri (2008) in the estimation procedure.

The results suggest most of the models behave badly in a time of transition- the span between .com bubble and subprime crisis. The best models to predict one-step-ahead conditional variance were parsimonious iGARCH and standard GARCH models to both countries. The backtesting of one-step-ahead VaR models, pointed just $\text{GARCH}(1,1)$ – GED, $\text{GARCH}(1,1)$ –Skeweness GED and $\text{GARCH}(1,1)$ – JSU as the best models to Brazil. While to Dow Jones index, the allowance of fat tails using t-Student distributions presented problems of convergence in some windows of the forecasting. The best models to the index were $\text{GARCH}(1,1)$ – Normal, $\text{GARCH}(1,1)$ – Skewness Normal, $\text{GARCH}(1,1)$ – GED, $\text{GARCH}(1,1)$ –Skeweness GED.

A few lessons can be extracted from this experiment: i) restricted types of distributions (i.g. Normal) assumptions does not fits well modelling risk to all countries. Being more flexible ones (i.g. GED) desirable; and, ii) the financial manager whose considered the period between 2002 to 2008 to feed up its VaR models, consistently underestimated risk in this period considering more standard approaches.

Further applications should expand the methods of risk estimations. Focusing in tail modelling, such Extreme Value Theory(EVT) and algorithmic approaches. Moreover, different time span must be considered.

References

Bibliography

ALEXANDER, C. *Value-at-risk models*. [S.l.]: John Wiley & Sons, 2008.

ARTZNER, P. et al. Coherent measures of risk. *Mathematical finance*, Wiley Online Library, v. 9, n. 3, p. 203–228, 1999.

BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, Elsevier, v. 31, n. 3, p. 307–327, 1986.

CAMPBELL, R.; HUISMAN, R.; KOEDIJK, K. Optimal portfolio selection in a value-at-risk framework. *Journal of Banking & Finance*, Elsevier, v. 25, n. 9, p. 1789–1804, 2001.

CAMPBELL, S. *A Review of Backtesting and Backtesting Procedures. Federal Reserve Board, Washington, 2005*. 2005.

CAVALERI, R. Combinação de previsões aplicada à volatilidade. 2008.

CHRISTOFFERSEN, P. F. Evaluating interval forecasts. *International economic review*, JSTOR, p. 841–862, 1998.

CHRISTOFFERSEN, P. F.; DIEBOLD, F. X. How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics*, MIT Press, v. 82, n. 1, p. 12–22, 2000.

COMMITTEE on B. S. B. *Amendment to the capital accord to incorporate market risks*. [S.l.]: Bank for International Settlements, 1996.

DANIELSSON, J. The appropriate use of risk models: Part i. *CERP's Policy Portal*, VOX, 2011.

DANÍELSSON, J. *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*. [S.l.]: John Wiley & Sons, 2011. v. 588.

DANIELSSON, J.; VALENZUELA, M.; ZER, I. Learning from history: volatility and financial crises. 2016.

DRAKOS, A. A.; KOURETAS, G. P.; ZARANGAS, L. Predicting conditional autoregressive value-at-risk for stock markets during tranquil and turbulent periods. *Journal of Financial Risk Management*, Scientific Research Publishing, v. 4, n. 03, p. 168, 2015.

ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 987–1007, 1982.

ENGLE, R. F.; PATTON, A. J. et al. What good is a volatility model. *Quantitative finance*, Taylor & Francis, v. 1, n. 2, p. 237–245, 2001.

FRANCQ, C.; ZAKOIAN, J.-M. *GARCH models: structure, statistical inference and financial applications*. [S.l.]: John Wiley & Sons, 2011.

GAIO, L. E. et al. Value-at-risk in times of crisis: An analysis in the brazilian market. *African Journal of Business Management*, Academic Journals, v. 9, n. 5, p. 223, 2015.

HAUGOM, E.; ULLRICH, C. J. Forecasting spot price volatility using the short-term forward curve. *Energy Economics*, Elsevier, v. 34, n. 6, p. 1826–1833, 2012.

HULL, J. C. *Options, futures, and other derivatives*. [S.l.]: Pearson Education India, 2006.

JORION, P. *Value at risk*. [S.l.]: McGraw-Hill, New York, 1997.

KRISTJANPOLLER, W.; MINUTOLO, M. C. Gold price volatility: A forecasting approach using the artificial neural network–garch model. *Expert Systems with Applications*, Elsevier, v. 42, n. 20, p. 7245–7251, 2015.

KUESTER, K.; MITTNIK, S.; PAOLELLA, M. S. Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, Oxford Univ Press, v. 4, n. 1, p. 53–89, 2006.

KUPIEC, P. H. Techniques for verifying the accuracy of risk measurement models. *The journal of Derivatives*, Institutional Investor Journals, v. 3, n. 2, p. 73–84, 1995.

MABROUK, S. Forecasting daily conditional volatility and h-step-ahead short and long value-at-risk accuracy: Evidence from financial data. *The Journal of Finance and Data Science*, Elsevier, v. 2, n. 2, p. 136–151, 2016.

MANDELBROT, B. The variation of certain speculative prices. *The Journal of Business*, v. 36, n. 4, p. 394–419, 1963.

MARKOWITZ, H. Portfolio selection. *The journal of finance*, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.

NELSON, D. B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 347–370, 1991.

SERMPINIS, G. et al. Modeling, forecasting and trading the eur exchange rates with hybrid rolling genetic algorithms—support vector regression forecast combinations. *European Journal of Operational Research*, Elsevier, v. 247, n. 3, p. 831–846, 2015.

TSAY, R. S. *Analysis of financial time series*. [S.l.]: John Wiley & Sons, 2005. v. 543.

WHEELWRIGHT, S.; MAKRIDAKIS, S.; HYNDMAN, R. J. *Forecasting: methods and applications*. [S.l.]: John Wiley & Sons, 1998.

ZAKOIAN, J.-M. Threshold heteroskedastic models. *Journal of Economic Dynamics and control*, Elsevier, v. 18, n. 5, p. 931–955, 1994.

Appendix

Figure 2 – Bovespa and Dow Jones Correlograms – Period: jan/2002 to dec/2008

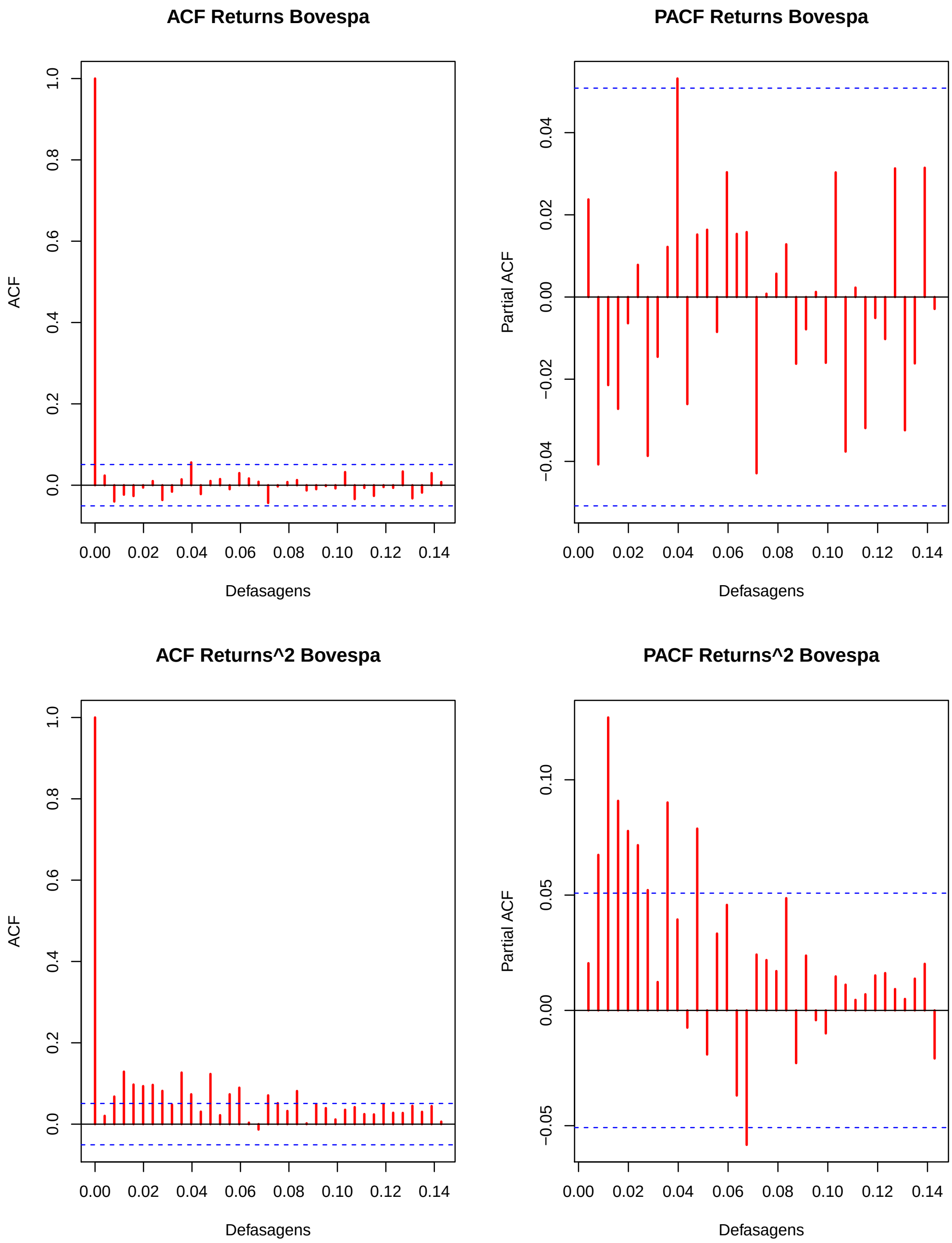


Table 4 – Backtesting of GARCH(1,1) - VaR models to Bovespa and Dow Jones

			Bovespa		
Model	Lenght For.	alpha	% Violations	LR_{uc}	LR_{cc}
Normal	487	0.05	7.2	4.344(0.03)	4.477(0.10)
		0.01	2.1	4.184(0.04)	NA
Skewed - Normal	487	0.05	7.2	4.344(0.03)	4.477(0.10)
		0.01	1.4	0.829(0.36)	NA
t-Student	487	0.05	7.4	5.147(0.0.02)	5.357(0.06)
		0.01	1.6	1.702(0.192)	NA
Skewed t-Student	487	0.05	7.2	4.344(0.03)	4.477(0.10)
		0.01	2.0	4.184(0.04)	NA
GED	487	0.05	7.2	4.344(0.03)	4.477(0.10)
		0.01	1.6	1.702(0.192)	NA
Skewed GED	487	0.05	7.0	3.603(0.05)	4.75(0.09)
		0.01	1.2	0.247(0.61)	NA
JSU	487	0.05	7.0	3.603(0.05)	4.75(0.09)
		0.01	1.0	0.003(0.953)	NA
			Dow Jones		
Model	Lenght For.	alpha	% Violations	LR_{uc}	LR_{cc}
Normal	509	0.05	6.9	3.394(0.06)	3.55(0.17)
		0.01	2.9	12.799(0.00)	NA
Skewed - Normal	509	0.05	6.3	1.646(0.19)	1.647(0.439)
		0.01	2.8	10.66(0.00)	NA
t-Student	509	0.05	N.C.	N.C.	N.C.
		0.01	N.C.	N.C.	N.C.
Skewed t-Student	509	0.05	N.C.	N.C.	N.C.
		0.01	N.C.	N.C.	N.C.
GED	509	0.05	6.7	2.748(0.09)	2.991(0.22)
		0.01	2.2	5.203(0.02)	N.A.
Skewed GED	509	0.05	6.3	1.646(0.19)	2.131(0.34)
		0.01	2.0	3.734(0.05)	NA
JSU	509	0.05	6.5	2.165(0.14)	2.518(0.28)
		0.01	2.6	8.684(0.00)	N.A.

Note:The time covered by this backtesting goes from 2002 to 2008. Using daily returns frequency, is used a moving average window of 1000 days. The parameters are re estimated after every one-step-ahead forecasting. The tests are run in out-of-sample observations totalling 487 to Bovespa and 509 to Dow Jones.