

## Value at Risk (VaR) Using Volatility Forecasting Models: EWMA, GARCH and Stochastic Volatility

**Fernando Caio Galdi\***  
*University of São Paulo*

**Leonel Molero Pereira\*\***  
*University of São Paulo*

**ABSTRACT:** This paper explores three models to estimate volatility: exponential weighted moving average (EWMA), generalized autoregressive conditional heteroskedasticity (GARCH) and stochastic volatility (SV). The volatility estimated by these models can be used to measure the market risk of a portfolio of assets, called Value at Risk (VaR). VaR depends on the volatility, time horizon and confidence interval for the continuous returns under analysis. For empirical assessment of these models, we used a sample based on Petrobras stock prices to specify the GARCH and SV models. Additionally, we adjusted these models by violation backtesting for one-day VaR, to compare the efficiency of the SV, GARCH and EWMA volatility models (suggested by RiskMetrics). The results suggest that VaR calculated considering EWMA was less violated than when considering SV and GARCH for a 1500-observation window. Hence, for our sample, the model suggested by RiskMetrics (1999), which uses exponential smoothing and is easier to implement, did not produce inferior violation test results when compared to more sophisticated tests such as SV and GARCH.

**Keywords:** VaR, Stochastic Volatility, GARCH, EWMA

---

*Received in 12/15/2006; revised in 01/03/2007; accept in 02/05/2007.*

**Corresponding authors:**

\* E-mail: [fecaio@uol.com.br](mailto:fecaio@uol.com.br)

\*\* E-mail: [leonel@usp.br](mailto:leonel@usp.br)

**Editor's note:** This paper was accepted by Alessandro Broedel Lopes.

## 1. INTRODUCTION

**F**The main objective of volatility models is to provide a measure that can be used in managing financial risks, helping in the selection of portfolio assets and in derivatives pricing.

The value-at-risk (VaR) models used in risk management by financial institutions, as a measure of the risk of financial loss for a determined confidence interval and time horizon, need a volatility estimate for their formulation. Volatility forecasting models, such as GARCH and stochastic volatility, are proposed as alternatives for this estimation.

In this sense, the present paper suggests the use of autoregressive conditional heteroskedasticity and stochastic volatility models to predict the volatility used in VaR measures.

We specify the models and their estimated parameters using an extended sample of continuous returns of preferred Petrobras shares. Additionally, we use a violation test to compare the VaR limits of the models obtained by GARCH, stochastic volatility (SV) and that suggested by RiskMetrics (1999) for the marked-to-the-market returns our portfolio of Petrobras shares.

The paper is organized in eight sections plus appendixes. The second section presents the VaR measure and the adjustment of the volatility models to this measure. The third section introduces the main concepts of the GARCH model, and the fourth section briefly presents the stochastic volatility model. The fifth and sixth sections present the data used and specify the volatility models. The seventh section then details the backtesting to compare the efficiency of the volatility prediction models used in calculating the VaR. The eighth section presents some final conclusions, and the appendixes contain more detailed results of each model's estimation.

## 2. THE VALUE AT RISK (VaR) OF A PORTFOLIO

Value at Risk (VaR) seeks to measure the market risks in terms of asset price volatility. VaR, as defined by Jorion (2001, p.19), synthesizes the greatest (or worst) loss expected from a portfolio, within determined time periods and confidence intervals.

Formally, VaR is defined for a long position in an asset S over a time horizon  $j$ , with probability  $p$  ( $0 < p < 1$ ):

$$p = P(\Delta P_j \leq VaR) = F_j(VaR)$$

(0.1)

where  $\Delta P_j$  represents the gain or loss of position  $P$ , given by  $\Delta P_j = P_{t+j} - P_t$  and  $F_j(.)$  is the accumulated distribution function (a.d.f.) of the random variable  $\Delta P_j$ .

The VaR is given in monetary units and represents the  $p$ -quantile of the distribution  $F_j(.)$ . According to Moretin (2004, p.179), this quantile is estimated from an empirical distribution of the returns. The VaR calculated in (2.1) has a negative value, because someone with a long position suffers a loss if  $\Delta P_j < 0$ . The amount in monetary units in calculating the VaR is obtained by multiplying the value of the financial position by the VaR of the return.

Calculation of the VaR can be simplified if it is possible to suppose a normal distribution of the continuous returns ( $y_t$ ), or log-returns, of the assets composing the portfolio. Starting from the estimates of the distribution parameters, such as the standard deviation of the returns, the expected portfolio loss can be determined as follows:

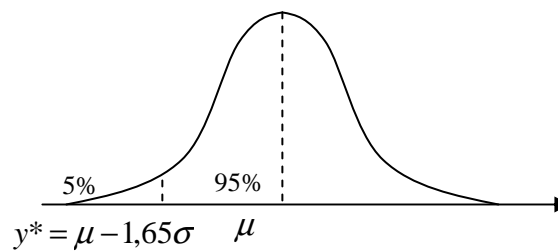
$$p = \int_{-\infty}^{y^*} f(y)dy, f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y^*-\mu)^2}{2\sigma^2}}$$

(0.2)

where  $y^*$  is the critical return value for calculating the VaR for time horizon  $j$ .

The following figure illustrates the VaR calculation at a 5% confidence interval for a supposedly normal distribution of returns, with mean  $\mu$  and standard deviation  $\sigma$  :

**Figure 1: VaR with normal distribution**



Source: Prepared by the authors.

In this way it is possible to calculate the VaR from the accumulated probability density function of a standard normal distribution. However, care must be taken to convert the log-return of the VaR into a discrete percentage variation  $i$ , in the following manner:

$$i = e^{y^*} - 1 \quad (0.3)$$

The absolute VaR suggested by RiskMetrics, in Longestae and More (1995), starts from the premise that the conditional distribution of the returns is normal and has mean zero and variance  $\sigma_{t+1}^2$ , with  $y_{t+j} | I_t \sim N(0, (j-t)\sigma_{t+1}^2)$  and  $\sigma_{t+j}^2 = (j-t)\sigma_{t+1}^2$ .

To calculate the VaR it is necessary to have an estimate of the volatility of the asset's log-returns for the analysis horizon. In this study, we evaluate three different approaches of estimating the volatility to calculate the VaR. The first approach, which is based on the model proposed by RiskMetrics, the most common method among VaR users, utilizes exponential smoothing with a decay factor  $\lambda$  of 0.94 and assumes the returns are normally distributed. This approach can be considered a particular case of the generalized autoregressive conditional heteroskedasticity (GARCH) model, and according to Jorion (2001, p. 175), it is represented by the following equation:

$$h_t = \lambda h_{t-1} + (1-\lambda)y_{t-1}^2 \quad (0.4)$$

The second method analyzed uses the concept of conditional volatility, modeled through a combination of the autoregressive moving average (ARMA) plus Gaussian GARCH, and the third performs the volatility prediction through the stochastic volatility (SV) model.

Berkowitz and O'Brien (2002) evaluated VaR models for a sample of six banks using historic series. They compared the models employed by the banks with the VaR calculated based on an ARMA(1,1) plus GARCH(1,1) model, assuming a normal distribution. They found by backtesting that the banks' VaR, although more conservative, did not follow the profit and loss (P&L) volatility of their portfolios and was outperformed by the GARCH model in terms of violation of the VaR limits. Jorion (2001, p. 170) states that the models for calculating VaR that use GARCH are more precise, principally in cases where there are volatility clusters.

In this study, we use the conditional volatility model in its reduced form to compute the VaR, as done in the work of Berkowitz and O'Brien (2002), as originally proposed by Zangari (1997). The model we use is composed of an autoregressive component of the returns, represented by an AR(1) model<sup>1</sup>:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \quad (0.5)$$

combined with a GARCH(1,1) conditional volatility model

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}. \quad (0.6)$$

To calculate the VaR with the AR(1) plus GARCH(1,1) conditional volatility model, we use the conditional mean and variance one step ahead, estimated by the model:

$$y_{t+1} | I_t \sim N(\hat{y}_t(1), \hat{\sigma}_t^2(1)) \quad (0.7)$$

In this context, supposing a VaR of 5% for one day (with  $p = 5\%$  and  $z = 1.65$ ), it should be calculated in the following way:

$$VaR_{5\%} = \hat{y}_t(1) - 1.65 \hat{\sigma}_t^2(1) \quad (0.8)$$

In following sections we describe the models used to predict volatility for calculating the VaR.

### 3. GARCH MODEL

Engle (1982) shows that it is possible to model the mean and variance simultaneously. For this, he uses the concept of conditional variance, which can be modeled as an autoregressive term:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_p \hat{\varepsilon}_{t+1-q}^2 + v_t, \quad (0.1)$$

where  $\hat{\varepsilon}_t$  is the estimated residual of the model  $y_t = a_0 + \Phi y_{t-1} + \varepsilon_t$  and  $v_t$  is the white noise.

The representation of the above equation is the base for the autoregressive conditional heteroskedasticity (ARCH) model. Nevertheless, in terms of estimation of  $\varepsilon_t$ , it is not the most suitable, given that to carry out the joint estimation of  $\{y_t\}$  and the conditional variance, the maximum likelihood technique is used. Hence, a more suitable specification is to treat  $v_t$  as a multiplicative term. Hence, the equation can be written as follows:

$$\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{i=1}^q \alpha_i \hat{\varepsilon}_{t-i}^2}, \quad (0.2)$$

where  $\alpha_0$  and  $\alpha_i$  are constant parameters such that  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  and  $0 \leq \sum \alpha_i \leq 1$ , for the variance given by  $\sigma_\varepsilon^2 = \alpha_0 / 1 - \sum_{i=1}^q \alpha_i$  not to be negative and/or explosive.

Bollerslev (1986) expanded the model given by (3.2) to permit the conditional variance to be modeled as an autoregressive moving average (ARMA) model. The generalized autoregressive conditional heteroskedasticity (GARCH) model is "a generalized ARCH model in which the conditional variance of  $n$  at instant  $t$  depends not only on the past squared perturbations, but also the past conditional variances." (Gujarati, 2005, p.440). The most common typification is the GARCH(1,1) model, where the first number refers to the lag of the autoregressive terms and the second refers to the number of lags in the model's moving average component. GARCH ( $p, q$ ) models are specified as:

$$\varepsilon_t = v_t \sqrt{h_t} \quad (0.3)$$

<sup>1</sup> The identification of the ARMA ( $p, q$ ) model used in this study followed the method presented by Box and Jenkins (1970).

where  $\sigma_v^2 = 1$  and

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}. \quad (0.4)$$

The constraints of this model are:  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_i \geq 0$  and  $0 \leq \sum \alpha_i + \sum \beta_i \leq 1$ .

It is interesting to note that GARCH models are conditionally heteroskedastic, but have a constant unconditional variance.

To specify GARCH models, it is necessary to assume the conditional distribution of the error terms  $\varepsilon_t$ . The literature usually employs the following distributions: i) normal; ii) Student's  $t$ ; and/or iii) generalized errors.

For a given distribution, the model is estimated by the maximum likelihood method. For example, for a GARCH (1,1) model, with  $T$  observations, assuming a normal distribution of the error terms, the log-likelihood is given by:

$$\log L = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(h_t) - \frac{1}{2} \sum_{t=1}^T (y_t - a_0 - y_{t-1}\phi)^2 / h_t \quad (0.5)$$

Assuming a Student's  $t$  distribution implies:

$$\log L = -\frac{1}{2} \log \left( \frac{\pi(v-2)\Gamma(v/2)^2}{\Gamma((v+1)/2)^2} \right) - \frac{1}{2} \sum_{t=1}^T \log h_t - \frac{(v+1)}{2} \log \left( 1 + \frac{(y_t - a_0 - y_{t-1}\phi)^2}{h_t(v-2)} \right), \quad (0.6)$$

where  $v$  is the number of degrees of freedom and  $\Gamma(x)$  is the usual Gamma function, i.e.,

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy.$$

In this paper we use the estimations of the GARCH model considering the maximum likelihood based on the normal distribution<sup>2</sup>.

#### 4. STOCHASTIC VOLATILITY (SV) MODEL

According to Morettin (2004, p.164) "the models of the ARCH family suppose that the conditional variance depends on the past returns." The stochastic volatility (MV) model, first proposed by Taylor (1986), does not make this assumption. This model's premise is that the present volatility depends on its past values, but that it is independent of the past returns. Considering the price of the financial asset at  $t$  ( $S_t$ ), the discrete time stochastic volatility model, presented by Harvey, Ruiz and Shephard (1994), can be written as:

$$y_t = \sigma_t \varepsilon_t, \quad t = 1, \dots, T, \quad (0.1)$$

where  $y_t$  represents the continuous return of the asset in period  $t$ , calculated by  $y_t = \ln(S_t / S_{t-1})$ , and  $\log \sigma^2$  follows an AR(1) process. It is assumed that  $\varepsilon_t$  is a series of independent and identically distributed (iid) random terms. Usually  $\varepsilon_t$  is specified to have a standard distribution, so its variance  $\sigma_\varepsilon^2$  is unknown. Thus, for a normal distribution  $\sigma_\varepsilon^2$  is equal to one, while for a  $t$  distribution with  $v$  degrees of freedom, it is  $v/(v-2)$ . According to the convention in the literature, one can write:

$$y_t = \sigma \varepsilon_t e^{0,5ht} \quad (0.2)$$

<sup>2</sup> We also tested the models based on the Student  $t$  distribution. The differences in the results were not significant.

where  $h_t$  is the logarithmic volatility at  $t$  and  $\sigma$  is a constant scale factor, for which reason there is no need for a constant in the first-order stationary autoregressive term, according to the following equation:

$$h_{t+1} = \Phi h_t + \eta_t, \eta_t \sim iid(0, \sigma_\eta^2), |\Phi| < 1 \quad (0.3)$$

If  $\varepsilon_t$  has finite variance, the variance of  $y_t$  is given by:

$$Var(y_t) = \sigma^2 \sigma_\varepsilon^2 e^{(\sigma_\varepsilon^2/2)}, \quad (0.4)$$

where  $\sigma_h^2$  is the variance of  $h_t$ .

One of the advantages of the discrete time stochastic volatility model is that it is analogous to the continuous time models utilized in articles on options pricing, such as in Hull and White (1987). The basic econometric properties of stochastic volatility models are discussed in Taylor (1986, 1994), Shephard (1996), Ghysels, Harvey and Renault (1996) and Jacquier, Polson and Rossi (1994). One of the key characteristics of this model is that it can be linearized by applying the logarithm squared of the observations in (4.2):

$$\log y_t^2 = h_t + \log \varepsilon_t^2 + \log \sigma^2 \quad (0.5)$$

Afterward, the term  $E(\log \varepsilon_t^2)$  is added to and subtracted from expression (4.5), to obtain:

$$\log y_t^2 = h_t + \log \varepsilon_t^2 - E(\log \varepsilon_t^2) + \log \sigma^2 + E(\log \varepsilon_t^2) \quad (0.6)$$

The representation of this expression can be written as:

$$\log y_t^2 = k + h_t + \xi_t, \quad (0.7)$$

where  $k = \log \sigma^2 + E(\log \varepsilon_t^2)$  and  $\xi_t = \log \varepsilon_t^2 - E(\log \varepsilon_t^2)$ .

As shown in Harvey, Ruiz and Shephard (1994), the state space form given by equations (4.3) and (4.7) supply the basis for estimating the model's parameters by applying a Kalman filter. Harvey, Ruiz and Shephard (1994) estimate the parameters  $\theta = (\Phi, \sigma_\eta^2) \in (-1, 1)$  by maximizing the following quasi-likelihood function:

$$\log L_Q(y | \theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^n \log F_t - \frac{1}{2} \sum_{t=1}^n v_t^2 / F_t, \quad (0.8)$$

where  $y = (y_1, y_2, \dots)$ ,  $v_t$  is the projection of the error one step ahead for the best estimator of  $\log y_t^2$  and  $F_t$  is the corresponding quadratic error.

The estimation carried out by the quasi-maximum likelihood method is consistent and asymptotically follows a normal distribution. In the next section, we detail the characteristics of the sample, with market data on the shares, used to apply the models presented.

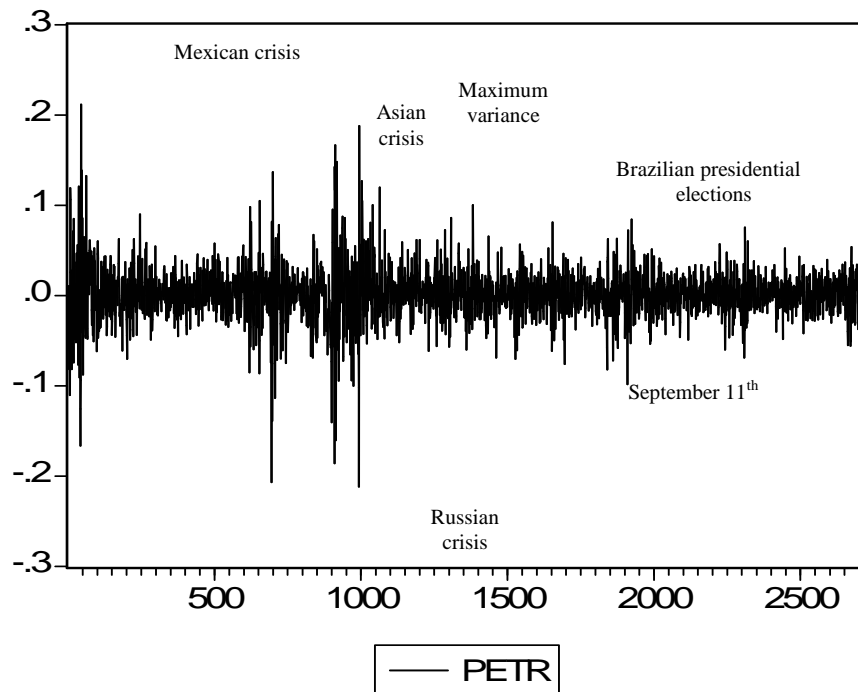
## 5. DESCRIPTION OF THE DATA

To apply the volatility models to the calculation of the value at risk (VaR), we chose the preferred shares of Petrobras (Petr4), given their liquidity and the number of trading days with an ample window of data available. The data are daily (trading days) and cover the period from January 2, 1995 to January 12, 2006, for a total of 2729 observations. Alexander (2005, p. 90) reports that "in the GARCH model there is a dichotomy between whether to have sufficient data for the estimates of the parameters to be stable, according to the moving data window, or to have excessive data, so that the predictions do not appropriately reflect current market conditions." Our sample represents a window of eleven years for applying the models, because we prioritized the stability of the parameters. Figure 2

shows the behavior of the selected series of returns, highlighting events that caused high-volatility clusters.

[Mexican crisis / Asian crisis / Russian crisis / Maximum variance / September 11<sup>th</sup> / Brazilian presidential elections]

**Figure 2: Continuous Returns of Petrobras PN**



We initially ran tests to identify the existence or not of a unit root and conditional heteroskedasticity in the series, for adequate application of the models. According to the tests conducted (augmented Dickey-Fuller, Phillips-Perron, KPSS and correlogram of squared returns), it can be said that the series is stationary and heteroskedastic, which qualifies it for application of the models analyzed (see the appendixes).

#### SPECIFICATION OF THE MODELS

In this section we specify the GARCH and SV models using the data sample mentioned above. We specified the combined AR(1) and GARCH (1,1) model with the help of the EViews program. The AR(1) specified was:

$$y_t = 0,132674 y_{t-1} + \varepsilon_t \quad (0,01992) \quad (0.1)$$

We followed the procedures below to estimate the complete GARCH model:

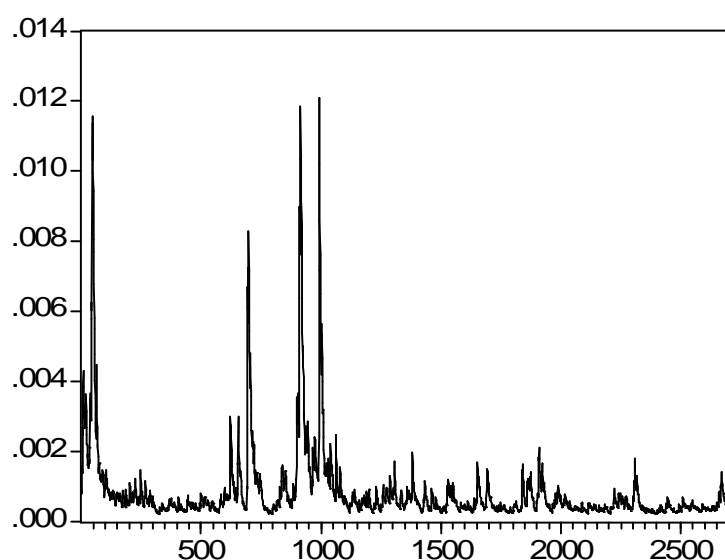
- We estimated the AR(1) model following the method suggested by Box and Jenkins (1970);
- We verified the lag of the GARCH model by analyzing the autocorrelation and partial autocorrelation functions of the squared residuals, resulting in specification of a GARCH (1,1) model due to its parsimony and because models with more lags did not converge satisfactorily;
- We applied Student's  $t$  test to the estimated parameters, rejecting the null hypothesis of equality to zero;

$$h_t = 0,0000247 + 0,12972\varepsilon_{t-1}^2 + 0,83784h_{t-1} \quad (0,0000038) \quad (0,00813) \quad (0,01115) \quad (0.2)$$

- The model's constraints were satisfied, since the estimates of the parameters were positive and the sum less than 1;
- We verified the absence of autocorrelation in the autocorrelation function by the Ljung Box test;
- In the squared residuals tests, we verified the existence of autocorrelation (Ljung-Box), and by the LM test we rejected the null hypothesis of the absence of autocorrelation in the squared residuals;
- Using the Jacque-Bera normality test of the standardized residuals, we could verify in the histogram that the distribution of the residuals is leptokurtic, rejecting the hypothesis of normality.

According to these tests, the GARCH (1,1) model is suitable to estimate the conditional volatility, and is thus used to calculate the VaR. The figure below contains the static prediction one step ahead of the conditional variance based on the model specified.

**Figure 3: Prediction of the conditional variance one step ahead**



Source: Prepared by the authors.

We specified the stochastic volatility (SV) model with the help of the STAMP program. A practical problem arises in estimating this model, namely the existence of zeros in the data series. Since the calculations are carried out in logarithms, the values of the returns cannot be nil. We employed the following transformation, suggested by Breidt and Carriquiry (1994), to get around this problem:

$$\log y_t^2 \cong \log(y_t^2 + cs_y^2) - cs_y^2 / (y_t^2 + cs_y^2), \quad t = 1, 2, \dots, T, \quad (0.3)$$

where  $s_y^2$  is the sample variance of  $y$  and  $c$  is a small number (in STAMP it is 0.02).

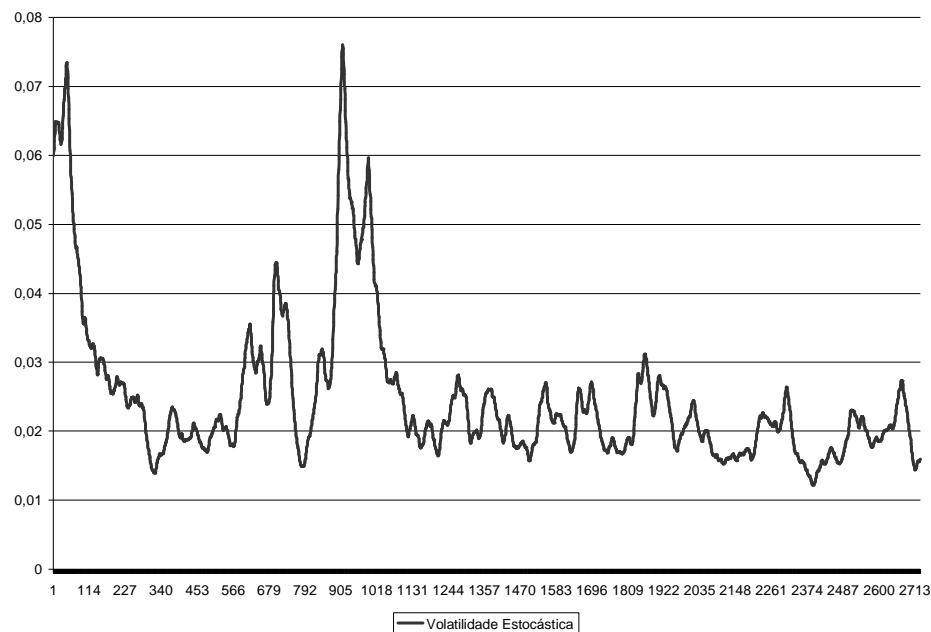
After carrying out this transformation, the model is estimated from the quasi-maximum likelihood method via a Kalman filter, resulting in the following equations:



$$\left\{ \begin{array}{l} \log y_t^2 = \kappa + h_t + \zeta_t \rightarrow \log y_t^2 = -8,5769 + h_t + \zeta_t \\ h_{t+1} = \Phi h_t + \eta_t \rightarrow h_{t+1} = 0,986323h_t + \eta_t \end{array} \right.$$

The model converges very strongly in 14 iterations, with the estimated standard deviation of  $\zeta_t$  equal to 1.6888 and that of  $\eta_t$  equal to 0.13382. It is important to stress that the Ljung-Box test applied to the residuals estimated by the model suggests there is no autocorrelation. The value obtained in the autoregressive component, of 0.9863, is very high and suggests there is an adjustment equivalence between the GARCH(1,1) and SV models. Figure 4 shows the SV model for the Petrobras stock returns. Just as in the GARCH model, there are volatility clusters coinciding with certain events, but the general volatility level of the SV model appears more stable.

**Figure 4: SV model for the Petrobras returns**



Source: Prepared by the authors.

Thus, for the case under analysis, either the GARCH (1,1) or the SV model are adequate to estimate the VaR. In this sense, in the next section we apply backtesting to compare the VaR based on the RiskMetrics (1999) methodology, which is widely used in the market, with the VaR based on the GARCH and SV methods.

## 6. BACKTESTING

According to the RiskMetrics (1999) manual, backtesting compares the results obtained with the measures generated by the model, to measure the efficiency of the model used by financial institutions.

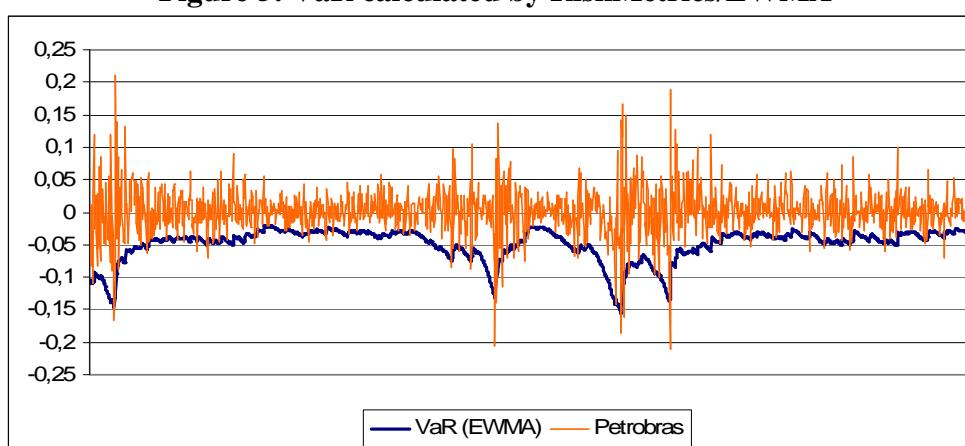
One of the methods used to evaluate the efficiency of models through backtesting is to test for violations of the VaR limits, given by the number of excesses outside the confidence interval.

This test uses the portfolio's value marked to the market, counting the number of times the portfolio's returns exceeded the confidence interval stipulated for the VaR. The number

violations can be differentiated into: upper limits, when the return exceeds the confidence interval on the right side of the tail; and lower limits, when the return is more negative than the critical return determined by the VaR. In this work we apply the violation test for the lower limits, using the marked-to-the-market returns of the Petrobras shares for a window of 1500 observations. In this form, it is possible to compare the efficiency of the GARCH and RiskMetrics models.

The next figure shows the Petrobras returns and the VaR calculated by the RiskMetrics model, with volatility estimated by the exponential weighted moving average (EWMA), with a decay factor  $\lambda=0.94$ . It can be seen that at the moments of greatest volatility, the Petrobras returns exceeded the lower VaR limits.

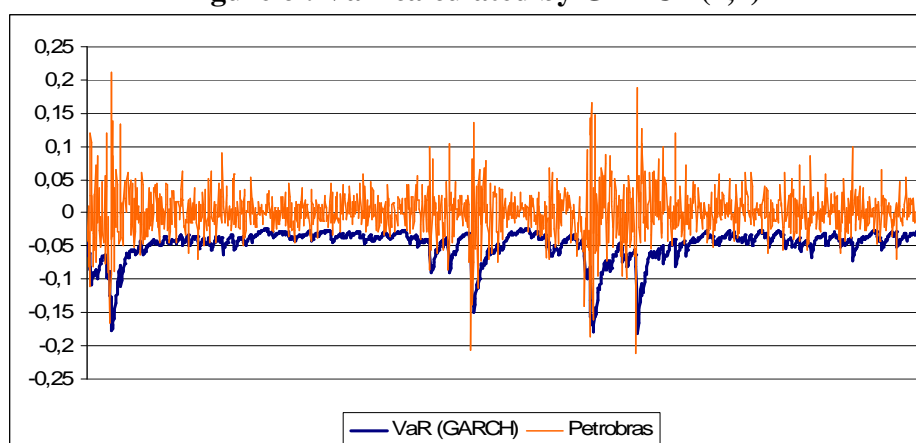
**Figure 5: VaR calculated by RiskMetrics/EWMA**



Source: Prepared by the authors.

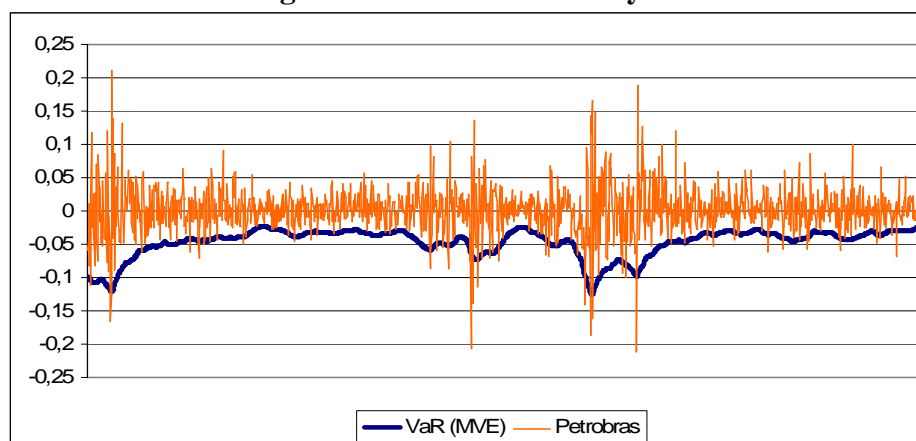
The next figure presents the VaR calculated by the GARCH (1,1) model and the limit violations.

**Figure 6 : VaR calculated by GARCH(1,1)**



Source: Prepared by the authors.

Figure 7 presents the VaR calculated by the SV model and its respective limit violations.

**Figure 7: VaR calculated by SV**

Source: Prepared by the authors.

The results of the limits-violation testing show that in 4.54% of the observations, the Petrobras returns exceed the VaR limits calculated with the GARCH model, while the corresponding percentages for the SV model were 3.87% and for the EWMA model only 3.20%. It is important to point out that these results are only indicative for a sample, and by comparing them it is not possible to conclude which is more efficient. But it is possible to infer that the VaR calculated by the EWMA method, through the model proposed by RiskMetrics, suffered fewer violations of the limits than the VaR calculated with the volatility forecast by the GARCH (1,1) and SV methods. Nevertheless, it should be remembered that all the models tested remained within the 5% significance level used in the VaR.

## 7. FINAL CONSIDERATIONS

This article analyzed three models used to estimate volatility: exponential weighted moving average (EWMA), generalized autoregressive conditional heteroskedasticity (GARCH) and stochastic volatility (SV). The volatility estimated by these models is the basis for calculating the VaR, a metric widely used by financial institutions and companies with exposures, to evaluate the risk of probable losses in their portfolios caused by asset price variations. The VaR measure depends on the volatility, the time horizon and the confidence interval for the continuous returns calculated through the logarithmic differences of the asset prices.

For empirical analysis, we used a sample of prices of preferred Petrobras shares to specify the generalized autoregressive conditional heteroskedasticity and the stochastic volatility models. Both the GARCH and the SV models proved adequate to model the volatility. Additionally, we carried out limits-violation backtesting for a VaR of 5% calculated one step ahead, to compare the efficiency of the GARCH and SV models with that proposed by RiskMetrics (EWMA). The results of these tests were not conclusive, but we verified that the VaR calculated by EWMA suffered fewer violations than those calculated by the GARCH and SV models, for a window of 1500 observations. The model suggested by RiskMetrics (1999), which uses the volatility calculated by exponential smoothing, besides being favored by the simplicity of its implementation, did not provide inferior results in the violation test in comparison with the more sophisticated volatility estimation models.

For subsequent works, we suggest the use of portfolios with more than one asset, or verification of the models for longer projection horizons than one day.

**REFERENCES**

- ALEXANDER, C. *Modelos de Mercado: Um guia para a análise de informações financeiras*. São Paulo: Editora Saraiva, 2005.
- BOLLERSLEV, T. *Generalized autoregressive conditional heteroskedasticity*. *Journal of Econometrics* 31. 1986, pp. 307-327.
- BOX, G.E, JENKINS, G.M. *Time series analysis: forecasting and control*. San Francisco: Holden Day, 1970.
- BREIDT, F.J., CARRIQUIRY, A.L. *Improved quasi-maximum likelihood estimation for stochastic volatility models*. In LEE, J.C., Zellner, A. (Eds.), *Modeling and prediction: honoring Seymour Geisser*, pp. 228-247. New York: Springer, 1994.
- ENDERS, W. *Applied econometric time series*. New York: John Wiley & Sons Inc, 2004.
- ENGLE, R.F. *Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation*. *Econometrica* 50. 1982, pp. 987-1007.
- GHYSELS, E., HARVEY, A. C., RENAULT, E. *Stochastic volatility*. In RAO, C. R., MADDALA G. S. (Eds.), *Statistical Methods in Finance*, pp.119-191. Amsterdam: North-Holland, 1996.
- GUJARATI, D.N. *Econometria básica*. São Paulo: Pearson Education do Brasil, 2004.
- HARVEY, A.C., RUIZ, E., SHEPHARD, N. *Multivariate stochastic variance models*. *Review of Economic Studies* 61. 1994, pp. 247-264.
- HULL, J., WHITE, A. *The pricing of options on assets with stochastic volatilities*. *Journal of Finance* 42. 1987, pp. 281-300.
- JACQUIER, E., POLSON, N. G., ROSSI, P. E. *Bayesian analysis of stochastic volatility models (with discussion)*. *Journal of Business and Economic Statistics* 12. 1994, pp. 371-417.
- JORION, P. *Value at Risk*. New York: McGraw Hill, 2001.
- MORETTIN, P.A., TOLOI, C.M.C. *Análise de séries temporais*. São Paulo: Editora Edgard Blücher do Brasil, 2004.
- MORETTIN, P.A. *Econometria Financeira* (available at [www.ime.usp.br/pam](http://www.ime.usp.br/pam)). São Paulo, 2004.
- RISKMETRICS. *Risk management: a practical guide*. 1<sup>st</sup> Edition. New York: RMG, 1999
- SHEPHARD, N. *Statistical aspects of ARCH and stochastic volatility*. In COX, D. R., BARNDORFF-NIELSON, O.E., HINKLEY, D.V. (Eds.), *Time Series Models in Econometrics, Finance and other Fields*, pp. 1-67. London: Chapman & Hall, 1996.
- SPANOS, A. *Statistical foundations of econometric models*. Cambridge: University Press, 1986.
- TAYLOR, S. J. *Modelling financial time series*. Chichester: John Wiley, 1986.
- TAYLOR, S. J. *Modelling stochastic volatility*. *Mathematical Finance* 4. 1994, pp. 183-204.

## APPENDIXES

### i) Autocorrelation of the series

The correlogram of the stock's returns is presented below:

Included observations: 2729

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.113	0.113	35.014	0.000
		2	-0.038	-0.051	38.869	0.000
		3	-0.069	-0.060	52.003	0.000
		4	-0.042	-0.029	56.823	0.000
		5	-0.033	-0.031	59.878	0.000
		6	-0.037	-0.037	63.532	0.000
		7	0.020	0.021	64.581	0.000
		8	0.051	0.039	71.604	0.000
		9	0.044	0.030	77.015	0.000
		10	0.020	0.015	78.134	0.000
		11	-0.015	-0.011	78.749	0.000
		12	0.002	0.013	78.756	0.000
		13	-0.013	-0.008	79.254	0.000
		14	-0.011	-0.004	79.572	0.000
		15	0.038	0.041	83.463	0.000
		16	-0.006	-0.020	83.574	0.000
		17	0.028	0.029	85.800	0.000
		18	-0.004	-0.010	85.847	0.000
		19	0.018	0.021	86.691	0.000
		20	0.003	0.003	86.714	0.000
		21	0.008	0.013	86.893	0.000
		22	-0.038	-0.040	90.925	0.000
		23	-0.040	-0.030	95.431	0.000
		24	0.016	0.020	96.120	0.000
		25	0.025	0.013	97.812	0.000
		26	0.027	0.019	99.890	0.000
		27	0.012	0.004	100.29	0.000
		28	-0.009	-0.010	100.49	0.000
		29	-0.023	-0.019	101.90	0.000
		30	-0.009	0.002	102.13	0.000
		31	-0.031	-0.025	104.72	0.000
		32	-0.028	-0.024	106.89	0.000
		33	-0.026	-0.027	108.75	0.000
		34	-0.019	-0.028	109.72	0.000
		35	-0.042	-0.048	114.53	0.000
		36	-0.015	-0.018	115.19	0.000

From the analysis of the ACF and PACF, it is not very clear which model the behavior of these functions represents. However, a significant reduction can be perceived both in the ACF and PACF in the series' first lag. In this context, the AR(1), MA(1) and ARMA(1,1) are estimated and the respective information criteria are analyzed to select the model to be estimated with GARCH:

#### ARMA(1,1) Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.058242	0.153047	-0.380551	0.7036
MA(1)	0.179051	0.150903	1.186528	0.2355
R-squared	0.012326	Mean dependent var		0.001218
Adjusted R-squared	0.011964	S.D. dependent var		0.028971
S.E. of regression	0.028797	Akaike info criterion		-4.256375
Sum squared resid	2.260544	Schwarz criterion		-4.252041
Log likelihood	5807.696	Durbin-Watson stat		2.003717
Inverted AR Roots	-.06			
Inverted MA Roots	-.18			

#### AR (1) Model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.114847	0.019009	6.041861	0.0000
R-squared	0.011464	Mean dependent var		0.001218
Adjusted R-squared	0.011464	S.D. dependent var		0.028971
S.E. of regression	0.028804	Akaike info criterion		-4.256236
Sum squared resid	2.262518	Schwarz criterion		-4.254069
Log likelihood	5806.505	Durbin-Watson stat		1.989847
Inverted AR Roots	.11			

MA (1) Model				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	0.122887	0.019000	6.467663	0.0000
R-squared	0.012546	Mean dependent var		0.001196
Adjusted R-squared	0.012546	S.D. dependent var		0.028988
S.E. of regression	0.028806	Akaike info criterion		-4.256124
Sum squared resid	2.263601	Schwarz criterion		-4.253957
Log likelihood	5808.481	Durbin-Watson stat		2.004299
Inverted MA Roots	-.12			

Based on the significance of the estimated coefficients, as well as the information criteria of the models, we chose to use the AR(1) specification for the series.

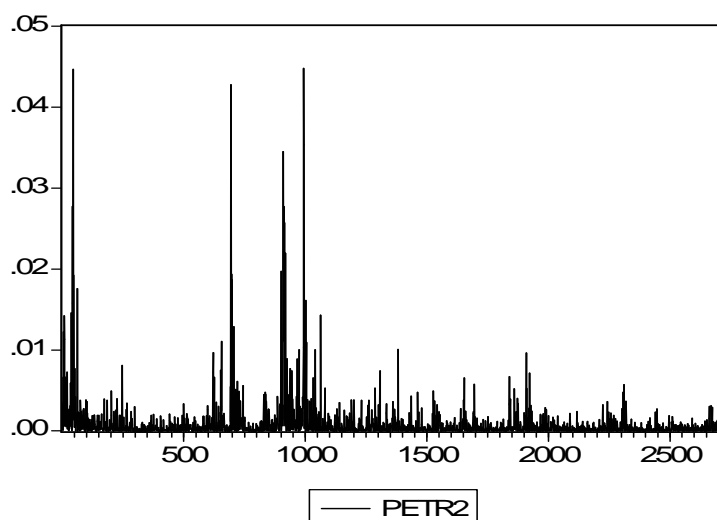
ii) Analysis of the series for adjustment of a GARCH model

The correlogram of the stock's squared returns is presented below:

Included observations: 2729

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.341	0.341	317.46	0.000
		2	0.335	0.248	624.20	0.000
		3	0.297	0.153	865.26	0.000
		4	0.205	0.022	980.12	0.000
		5	0.227	0.080	1120.7	0.000
		6	0.163	0.008	1193.1	0.000
		7	0.139	0.005	1245.9	0.000
		8	0.164	0.057	1319.6	0.000
		9	0.205	0.114	1434.3	0.000
		10	0.159	0.019	1503.4	0.000
		11	0.126	-0.023	1547.2	0.000
		12	0.163	0.051	1619.8	0.000
		13	0.078	-0.053	1636.3	0.000
		14	0.114	0.018	1671.9	0.000
		15	0.120	0.042	1711.1	0.000
		16	0.053	-0.036	1718.8	0.000
		17	0.096	0.015	1744.0	0.000
		18	0.097	0.030	1769.7	0.000
		19	0.084	0.011	1789.1	0.000
		20	0.102	0.022	1817.7	0.000
		21	0.047	-0.043	1823.7	0.000
		22	0.065	0.011	1835.3	0.000
		23	0.100	0.053	1862.6	0.000
		24	0.070	-0.002	1876.2	0.000
		25	0.083	0.027	1895.0	0.000
		26	0.097	0.034	1921.0	0.000
		27	0.108	0.033	1953.3	0.000
		28	0.099	0.012	1980.5	0.000
		29	0.075	-0.025	1996.0	0.000
		30	0.092	0.028	2019.4	0.000
		31	0.075	0.006	2034.8	0.000
		32	0.065	-0.022	2046.4	0.000
		33	0.084	0.034	2066.0	0.000
		34	0.081	0.020	2084.2	0.000
		35	0.076	-0.008	2100.3	0.000
		36	0.102	0.042	2129.3	0.000

The squared return series presents strong autocorrelation, which provides indications that the generalized autoregressive conditional heteroskedasticity model can be used for the best modeling of the series. The graph of the squared return series shows characteristic clusters of the GARCH model:



iii) Unit root tests of the series:

#### ADF Test – Model with intercept and trend

Null Hypothesis: PETR has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 2 (Automatic based on SIC, MAXLAG=27)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-31.79295	0.0000
Test critical values:		
1% level	-3.961409	
5% level	-3.411456	
10% level	-3.127584	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(PETR)

Method: Least Squares

Sample (adjusted): 4 2729

Included observations: 2726 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PETR(-1)	-0.988080	0.031079	-31.79295	0.0000
D(PETR(-1))	0.103309	0.025471	4.055912	0.0001
D(PETR(-2))	0.059656	0.019120	3.120107	0.0018
C	0.000799	0.001102	0.725248	0.4684
@TREND(1)	3.02E-07	6.99E-07	0.432184	0.6656
R-squared	0.447128	Mean dependent var		-8.18E-07

Adjusted R-squared	0.446315	S.D. dependent var	0.038593
S.E. of regression	0.028717	Akaike info criterion	-4.260837
Sum squared resid	2.243884	Schwarz criterion	-4.249996
Log likelihood	5812.521	F-statistic	550.1436
Durbin-Watson stat	1.999082	Prob(F-statistic)	0.000000

**Conclusion: By the augmented Dickey-Fuller (ADF) unit root test, the null hypothesis that there is a unit root in the stock's return series cannot be accepted at 1%, 5% and 10% significance. The conclusions are the same for the model without intercept and trend and with intercept.**

#### PP Test – Model with intercept and trend

Null Hypothesis: PETR has a unit root  
 Exogenous: Constant, Linear Trend  
 Bandwidth: 11 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-46.33250	0.0000
Test critical values:		
1% level	-3.961407	
5% level	-3.411454	
10% level	-3.127583	

\*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.000828
HAC corrected variance (Bartlett kernel)	0.000664

Phillips-Perron Test Equation  
 Dependent Variable: D(PETR)  
 Method: Least Squares  
 Sample (adjusted): 2 2729  
 Included observations: 2728 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PETR(-1)	-0.886767	0.019019	-46.62638	0.0000
C	0.000693	0.001103	0.628217	0.5299
@TREND(1)	2.86E-07	7.00E-07	0.408488	0.6829

R-squared	0.443767	Mean dependent var	2.46E-05
Adjusted R-squared	0.443358	S.D. dependent var	0.038593
S.E. of regression	0.028793	Akaike info criterion	-4.256243
Sum squared resid	2.259185	Schwarz criterion	-4.249743
Log likelihood	5808.516	F-statistic	1087.012
Durbin-Watson stat	1.989712	Prob(F-statistic)	0.000000

**Conclusion: By the Phillips-Perron (PP) unit root test, the null hypothesis that there is a unit root in the stock's return series cannot be accepted at 1%, 5% and 10% significance. The conclusions are the same for the model without intercept and trend and with intercept.**



## iv) Stationarity test

In order to confirm the stationarity of the series, identified in the unit root tests, we applied the KPSS test, reported below:

**KPSS Test– Model with intercept and trend**

Null Hypothesis: PETR is stationary

Exogenous: Constant, Linear Trend

Bandwidth: 8 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.040889
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)	
Residual variance (no correction)	0.000840
HAC corrected variance (Bartlett kernel)	0.000815

## KPSS Test Equation

Dependent Variable: PETR

Method: Least Squares

Sample: 1 2729

Included observations: 2729

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000683	0.001110	0.615912	0.5380
@TREND(1)	3.76E-07	7.04E-07	0.533665	0.5936
R-squared	0.000104	Mean dependent var		0.001196
Adjusted R-squared	-0.000262	S.D. dependent var		0.028988
S.E. of regression	0.028992	Akaike info criterion		-4.242870
Sum squared resid	2.292122	Schwarz criterion		-4.238537
Log likelihood	5791.396	F-statistic		0.284798
Durbin-Watson stat	1.771974	Prob(F-statistic)		0.593617

**Conclusion: By the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) stationarity test, the null hypothesis that there is a unit root in the stock's return series cannot be rejected at 1%, 5% and 10% significance. The conclusion is the same for the model with intercept.**

## v) GARCH model

Dependent Variable: PETR  
 Method: ML - ARCH (Marquardt)  
 Sample(adjusted): 2 2729  
 Included observations: 2728 after adjusting endpoints  
 Convergence achieved after 15 iterations  
 Variance backcast: OFF

	Coefficient	Std. Error	z-Statistic	Prob.
AR(1)	0.132674	0.019922	6.659689	0.0000
Variance Equation				
C	2.47E-05	3.80E-06	6.517232	0.0000
ARCH(1)	0.129720	0.008130	15.95504	0.0000
GARCH(1)	0.837840	0.011150	75.13943	0.0000
R-squared	0.011145	Mean dependent var		0.001218
Adjusted R-squared	0.010056	S.D. dependent var		0.028971
S.E. of regression	0.028825	Akaike info criterion		-4.600617
Sum squared resid	2.263248	Schwarz criterion		-4.591949
Log likelihood	6279.241	Durbin-Watson stat		2.023680
Inverted AR Roots	.13			

LM Test of the residuals

ARCH Test:

F-statistic	5.419314	Probability	0.004478
Obs*R-squared	10.80755	Probability	0.004500

Test Equation:

Dependent Variable: STD\_RESID^2

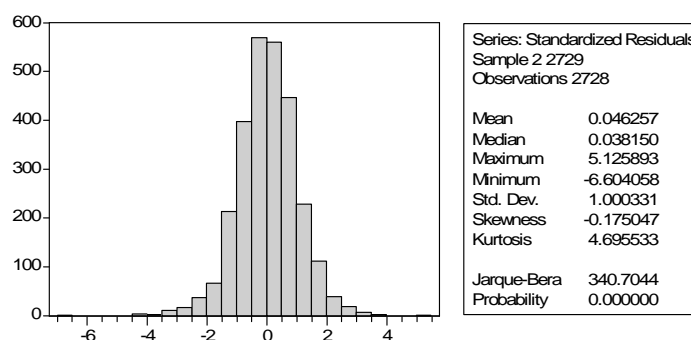
Method: Least Squares

Sample(adjusted): 4 2729

Included observations: 2726 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.938109	0.045594	20.57535	0.0000
STD_RESID^2(-1)	0.001367	0.019126	0.071459	0.9430
STD_RESID^2(-2)	0.062959	0.019129	3.291311	0.0010
R-squared	0.003965	Mean dependent var		1.002573
Adjusted R-squared	0.003233	S.D. dependent var		1.917730
S.E. of regression	1.914627	Akaike info criterion		4.138023
Sum squared resid	9981.968	Schwarz criterion		4.144528
Log likelihood	-5637.125	F-statistic		5.419314
Durbin-Watson stat	1.995807	Prob(F-statistic)		0.004478

Normality Test of the Residuals



vi) Stochastic volatility (SV) model

Complete results:

Method of estimation is maximum likelihood

The present sample is: 0 to 2728

MaxLik initialising...

it 1 f= -0.61823 e0= 0.00000 step= 1.00000

MaxLik iterating...

it 4 f= -0.56087 df= 0.00480 e1= 0.00347 e2= 0.11477 step= 1.00000

it 9 f= -0.55836 df= 0.00009 e1= 0.00021 e2= 0.03684 step= 1.00000

it 14 f= -0.55835 df= 0.00000 e1= 0.00000 e2= 0.00000 step= 0.00010

Equation 1.

SVpetr = Level + AR(1) + Irregular

Estimation report

Model with 3 parameters ( 1 restrictions).

Parameter estimation sample is 0. 1 - 2728. 1. (T = 2729).

Log-likelihood kernel is -0.5583489.

Very strong convergence in 14 iterations.

( likelihood cvg 1.789563e-015

gradient cvg 1.219025e-008

parameter cvg 1.623527e-010 )

Eq 1 : Diagnostic summary report.

Estimation sample is 0. 1 - 2728. 1. (T = 2729, n = 2728).

Log-Likelihood is -1523.73 (-2 LogL = 3047.47).

Prediction error variance is 3.04931

Summary statistics

	SVpetr
Std.Error	1.7462
Normality	114.33
H(909)	0.93420
r(1)	-0.0018508
r(39)	0.011165
DW	1.9980
Q(39,37)	38.960
R <sup>2</sup>	0.11195

Eq 1 : Estimated variances of disturbances.

Component	SVpetr (q-ratio)
Irr	2.8520 ( 1.0000)
Ar1	0.017907 ( 0.0063)

Eq 1 : Estimated standard deviations of disturbances.

Component	SVpetr (q-ratio)
Irr	1.6888 ( 1.0000)
Ar1	0.13382 ( 0.0792)

Eq 1 : Estimated autoregressive coefficient. *BBR, Braz.*

*Bus. Rev. (Engl. ed., Online),*

[www.bbronline.com.br](http://www.bbronline.com.br)

*Vitória, v. 4, n. 1, Art. 5, p. 74 - 94, jan.-apr. 2007*

The AR(1) rho coefficient is 0.986323.

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value
Lvl	-8.5769	0.18600	-46.113 [ 0.0000]
Ar1	-0.76884	0.45764	

Normality test for Residual SVpetr

Sample Size	2728	
Mean	-0.053013	
Std.Devn.	0.998594	
Skewness	-0.257566	
Excess Kurtosis	-0.580191	
Minimum	-3.280796	
Maximum	3.113678	
Skewness Chi <sup>2</sup> (1)	30.163	[0.0000]
Kurtosis Chi <sup>2</sup> (1)	38.263	[0.0000]
Normal-BS Chi <sup>2</sup> (2)	68.425	[0.0000]
Normal-DH Chi <sup>2</sup> (2)	114.33	[0.0000]

Goodness-of-fit results for Residual SVpetr

Prediction error variance (p.e.v)	3.049310
Prediction error mean deviation (m.d)	2.511845
Ratio p.e.v. / m.d in squares	0.938204
Coefficient of determination R <sup>2</sup>	0.111953
... based on differences RD <sup>2</sup>	0.466639
Information criterion of Akaike AIC	1.117114
... of Schwartz (Bayes) BIC	1.123613

Serial correlation statistics for Residual SVpetr.

Durbin-Watson test is 1.99805.

Asymptotic deviation for correlation is 0.019146.

Lag	dF	SerCorr	BoxLjung	ProbChi2(dF)					
1	0	-0.0019			27	24	-0.0182	25.9750	[0.3544]
2	0	0.0213			28	25	-0.0177	26.8430	[0.3638]
3	0	0.0197			29	26	-0.0319	29.6463	[0.2825]
4	1	-0.0394	6.5441	[0.0105]	30	27	0.0433	34.8198	[0.1435]
5	2	0.0128	6.9890	[0.0304]	31	28	-0.0167	35.5886	[0.1534]
6	3	-0.0149	7.5947	[0.0552]	32	29	0.0074	35.7379	[0.1813]
7	4	0.0012	7.5987	[0.1074]	33	30	0.0176	36.5987	[0.1891]
8	5	-0.0104	7.8932	[0.1622]	34	31	-0.0230	38.0570	[0.1790]
9	6	0.0103	8.1829	[0.2250]	35	32	0.0089	38.2765	[0.2059]
10	7	0.0271	10.1904	[0.1780]	36	33	-0.0097	38.5380	[0.2333]
11	8	0.0027	10.2104	[0.2506]	37	34	-0.0008	38.5398	[0.2716]
12	9	-0.0161	10.9238	[0.2810]	38	35	0.0052	38.6145	[0.3096]
13	10	0.0263	12.8205	[0.2339]	39	36	0.0112	38.9597	[0.3380]
14	11	-0.0097	13.0774	[0.2883]	40	37	0.0160	39.6695	[0.3519]
15	12	-0.0182	13.9897	[0.3014]	41	38	0.0167	40.4403	[0.3631]
16	13	-0.0045	14.0456	[0.3706]	42	39	0.0055	40.5251	[0.4029]
17	14	-0.0191	15.0495	[0.3748]	43	40	-0.0124	40.9534	[0.4285]
18	15	-0.0135	15.5496	[0.4126]	44	41	-0.0001	40.9535	[0.4727]
19	16	-0.0163	16.2809	[0.4335]	45	42	-0.0336	44.0926	[0.3832]
20	17	-0.0024	16.2973	[0.5029]	46	43	-0.0027	44.1124	[0.4244]
21	18	-0.0274	18.3652	[0.4319]	47	44	-0.0005	44.1131	[0.4669]
22	19	0.0151	18.9955	[0.4571]	48	45	-0.0148	44.7190	[0.4838]
23	20	0.0163	19.7278	[0.4751]	49	46	0.0259	46.5891	[0.4480]
24	21	-0.0229	21.1722	[0.4485]	50	47	0.0122	47.0029	[0.4724]
25	22	-0.0102	21.4600	[0.4925]					
26	23	-0.0361	25.0574	[0.3473]					

vii) Programming routine in VBA for calculation of the conditional variance of the exponential decay model:

```

Function ewma(lambda, As Range returns)
i = 1
lambda2 = 1
ewma = 0
Do While lambda2 > 0.00001
ewma = ewma + returns (i, 1) ^ 2 * lambda2
lambda2 = lambda2 * lambda
i = i + 1
Loop
ewma = ewma * (1 - lambda)
End Function

```

