MODELLING AND FORECASTING THE VOLATILITY OF BRAZILIAN ASSET RETURNS: A REALIZED VARIANCE APPROACH

BY M. R. C. CARVALHO, M. A. S. FREIRE, M. C. MEDEIROS, AND L. R. SOUZA

ABSTRACT. The goal of this paper is twofold. First, using five of the most actively traded stocks in the Brazilian financial market, this paper shows that the normality assumption commonly used in the risk management area to describe the distributions of returns standardized by volatilities is not compatible with volatilities estimated by EWMA or GARCH models. In sharp contrast, when the information contained in high frequency data is used to construct the realized volatilies measures, we attain the normality of the standardized returns, giving promise of improvements in Value at Risk statistics. We also describe the distributions of volatilities of the Brazilian stocks, showing that the distributions of volatilities are nearly lognormal. Second, we estimate a simple linear model to the log of realized volatilities that differs from the ones in other studies. The main difference is that we do not find evidence of long memory. The estimated model is compared with commonly used alternatives in an out-of-sample experiment.

KEYWORDS. Realized volatility, high frequency data, risk analysis, volatility forecasting, GARCH models.

VERY PRELIMINARY AND INCOMPLETE.

1. INTRODUCTION

Given the fast growth of financial markets and the development of new and more complex financial instruments, there is an ever-growing need for theoretical and empirical knowledge of the volatility of financial time series. It is widely known that daily returns of financial assets, especially of stocks, are hard to predict, if not impossible, although the volatility of the returns seems to be relatively easier to forecast. Therefore, the volatility has played a central role in modern pricing and risk-management theories. There is, however, an inherent problem to the use of models that have the volatility measure taking a central role, as the conditional variance is not directly observable. The conditional variance can be estimated, among other approaches, by the (Generalized) Autoregressive Conditional Heteroskedastic – (G)ARCH – family of models proposed by Engle (1982) and Bollerslev (1986), stochastic volatility models (Taylor 1986), or the exponentially weighted moving averages (EWMA) as advocated by the Riskmetrics methodology (Morgan 1996). These approaches are heavily based on the assumption that the conditional returns of financial time series are approximately Gaussian. However, as pointed out by Bollerslev (1987), Teräsvirta (1996), and Carnero, Peña, and Ruiz (2001), among others, this is not a compatible assumption with the estimated volatility from the above mentioned models, since the standardized returns still have excess of kurtosis.

Date: May, 24 2004.

The search for an adequate framework for the estimation and prediction of the conditional variance of financial assets returns has led us to the analysis of high-frequency intraday data. Merton (1980) already noted that the variance over a fixed interval can be estimated arbitrarily accurately by the sum of squared realizations, provided the data are available at a sufficiently high sampling frequency. More recently, Andersen and Bollerslev (1998), showed that ex-post daily foreign exchange volatility is best measured by aggregating 288 squared five-minute returns. The five-minute frequency is a trade-off between accuracy, which is theoretically optimized using the highest possible frequency, and noise due to, for example, micro-structure frictions. Ignoring the small remaining measurement error the ex-post volatility essentially becomes "observable". Andersen and Bollerslev (1998) used this new volatility measure to evaluate the out-of-sample forecasting performance of GARCH models. This same approach was adopted by Mota and Fernandes (2004) to compare different volatility models to the index of the São Paulo stock market.

As volatility becomes "observable", it can be modeled directly, rather than being treated as a latent variable. Recent studies, based on the theoretical results of Andersen, Bollerslev, Diebold, and Labys (2001a), Andersen, Bollerslev, Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2002a,b), and Meddahi (2002), documented the properties of realized volatilities constructed form high-frequency data. For example, Andersen, Bollerslev, Diebold, and Labys (2001a) study the bilateral exchange rates between the Japanese yen (¥), the Deutsche Mark (DM), and the U.S. Dollar ($), Ebens (1999) the Dow Jones index, Andersen, Bollerslev, Diebold, and Labys (2001b) the 30 stocks underlying the Dow Jones index, and Areal and Taylor (2002) the FTSE 100 index. Pong, Shackleton, Taylor, and Xu (2002) analyzed the £/$ (£ is the British Pound), Li (2002) the ¥/$, DM/$, and £/$ exchange rates, Hol and Koopman (2002) the S&P 100 index, and Martens and Zein (2002) the ¥/$, S& 500 and Light, Sweet, and Crude Oil.

Several important characteristics of the realized volatilities came out from these studies. First, the unconditional distribution of daily returns is not skewed, but it does exhibit excess kurtosis. Daily returns are not autocorrelated (except for the first order in some cases). Second, daily returns standardized by the realized variance measure are Gaussian. Third, the unconditional distributions of realized variance and volatility are distinctly non-normal and extremely right skewed. On the other hand, the natural logarithm of the volatility is close to normality. Third, the log of the realized volatility displays a high degree of (positive) autocorrelation which dies out very slowly. Fourth, realized volatility does not seem to have a unit root, but there is clear evidence of fractional integration, roughly of order 0.40.

The main goal of this paper is twofold. First, using five of the most actively traded stocks in Bovespa, this paper shows that the normality assumption commonly used in the risk management area to describe the distributions of returns standardized by volatilities is not compatible with volatilities estimated by EWMA or GARCH models. In sharp contrast, when we use the information contained in high frequency data to construct the realized volatilities measures, we attain the normality of the standardized returns, giving promise of improvements in Value at Risk statistics. We also describe the distributions of volatilities of the Brazilian stocks, showing that the distributions of volatilities are nearly lognormal. Second, we estimate a simple linear model to the log of realized volatilities that differs from the ones in other studies. The main difference is that we do not find evidence of long memory. The estimated model is compared with commonly used alternatives in an out-of-sample experiment.

The paper proceeds as follows. In Section 2, we briefly describe the calculation of the realized volatility. Section 3 describes the data used in the paper and carefully analyze the distribution of the standardized returns and realized volatility. In Section 4 we estimate a simple linear model to the realized volatility and an out-of-sample experiment is conducted to evaluate the forecasting performance of the estimated models. Finally, Section 5 concludes.

2. Realized Variance and Realized Volatility

The present section is strongly based on Oomen (2001). The term "realized variance" refers to the sum of squared intra-day returns and "realized volatility" is the squared root of the realized variance. The realized variance is an estimator for the average or integral of instantaneous variance over the interval of interest. In fact, in a continuous time framework, it has been shown by Andersen, Bollerslev, Diebold, and Labys (2001a) and Andersen, Bollerslev, Diebold, and Labys (2003) that when the return process is assumed to follow a special semi-martingale the realized variance measure can be made arbitrarily close to the integral of instantaneous variance, provided that the intra-period returns are sampled at a sufficiently high frequency. In the present context, however, the focus will be on a discrete time model.

Let $p_{t,j}$ denote the jth intra day-t logarithmic price of the security under consideration and $\mathcal{I}_{t,j}$ be the σ-algebra generated by $\{p_{a,b}\}_{a=-\infty,b=0}^{a=t,b=j}$. Under the assumption of N equally time-spaced intradaily observations of p $(j = 1,\ldots,N)$, the daily return is defined as:

$$r_t = p_{t,N} - p_{t-1,N}, \quad t = 1,\ldots,T.$$

At sampling frequency f, we can construct $N_f = \frac{N}{f}$ intradaily returns:

$$r_{t,i} = p_{t,if} - p_{t,(i-1)f}, \quad i = 1,\ldots,N,$$

where $p_{t,0} = p_{t-1,N}$.

In the following, it is assumed that the asset's (excess) return at the daily frequency can be characterized as:

$$r_t = h_t^{1/2}\varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}_{t=1}^{T}$ is a sequence of independent and normally distributed random variables with zero mean and unit variance, $\varepsilon_t \sim \text{NID}(0,1)$, and h_t is the daily variance. Note that $\mathbb{E}[r_t^2|\mathcal{I}_{t,0}] = h_t$ and that $\mathbb{V}[r_t^2|\mathcal{I}_{t,0}] = 2h_t^2$. Now consider the situation in which intradaily returns, at sampling frequency f, are uncorrelated and can be characterized as:

$$r_{t,i} = h_{t,i}^{1/2}\varepsilon_{t,i}, \tag{2}$$

where $\varepsilon_{t,i} \sim \text{NID}(0, N_f^{-1})$.

From (2) it is clear that $r_t = \sum_{i=1}^{N_f} r_{t,i}$. Then,

$$r_t^2 = \left[\sum_{i=1}^{N_f} r_{t,i}\right]^2 = \sum_{i=1}^{N_f} r_{t,i}^2 + 2\sum_{i=1}^{N_f-1}\sum_{j=i+1}^{N_f} r_{t,i}r_{t,j}, \tag{3}$$

and

$$\mathbb{E}\left[r_t^2|\mathcal{I}_{t,0}\right] = \mathbb{E}\left[\sum_{i=1}^{N_f} r_{t,i}^2\middle|\mathcal{I}_{t,0}\right] + 2\mathbb{E}\left[\sum_{i=1}^{N_f-1}\sum_{j=i+1}^{N_f} r_{t,i}r_{t,j}\middle|\mathcal{I}_{t,0}\right]. \tag{4}$$

Under the assumption that the intradaily returns are uncorrelated, it directly follows that

$$\mathbb{E}\left[\sum_{i=1}^{N_f} r_{t,i}^2\middle|\mathcal{I}_{t,0}\right] = \mathbb{E}\left[r_t^2|\mathcal{I}_{t,0}\right] = h_t.$$

As a result, two unbiased estimators for the average day-t return variance exist, namely the squared day-t return and the sum of squared intra day-t returns. However, it can be shown that

$$\mathbb{V}\left[\sum_{i=1}^{N_f} r_{t,i}^2\middle|\mathcal{I}_{t,0}\right] = \frac{2}{N_f}\sum_{i=1}^{N_f}\frac{h_{t,i}^2}{N_f} < \frac{2}{N_f}\left[\sum_{i=1}^{N_f}\frac{h_{t,i}}{\sqrt{N_f}}\right] = \mathbb{V}[r_t^2|\mathcal{I}_{t,0}], \tag{5}$$

since

$$\mathbb{E}\left[\left(\sum_{i=1}^{N_f} h_{t,i}\varepsilon_{t,i}^2\right)^2\middle|\mathcal{I}_{t,0}\right] = \frac{3}{N_f^2}\sum_{i=1}^{N_f} h_{t,i}^2 + \frac{2}{N_f^2}\sum_{i=1}^{N_f-1}\sum_{j=i+1}^{N_f} h_{t,i}h_{t,j},$$

and

$$h_t = \frac{1}{N_f}\sum_{i=1}^{N_f} h_{t,i}.$$

In words, the average daily return variance can be estimated more accurately by summing up squared intradaily returns rather than calculating the squared daily return. In addition, when returns are observed (and uncorrelated) at any arbitrary sampling frequency, it is possible to estimate the average daily variance free of measurement error as

$$\lim_{N_f\to\infty}\mathbb{V}\left[\sum_{i=1}^{N_f} r_{t,i}^2\middle|\mathcal{I}_{t,0}\right] = 0.$$

The only (weak) requirement on the dynamics of the intradaily return variance for the above to hold is that

$$\sum_{i=1}^{N_f} h_{t,i}^2 \propto N_f^{1+c},$$

where $0 \leq c < 1$. Finally, note that although the daily realized variance measure employs intradaily return data, there is no need to take the (well documented) pronounced intra-day variance pattern of the return process into account. This feature of the realized variance measure contrasts sharply with popular parametric variance models which generally require the explicit modeling on intradaily regularities in return variance.

However, when the returns are correlated, the realized volatility will be a biased estimator of the daily volatility. Although, in the context of efficient markets, the finding of correlated intradaily returns may at first sight appear puzzling, it has a sensible explanation in the context of the market micro-structure literature; see Campbell, Lo, and Mackinlay (1997, Chapter 3). When the returns are sampled at higher frequencies, market microstructure may introduce some autocorrelation in the intra-day returns, thus, driving the realized variance to be a biased estimator of the daily variance. On the other hand, lower frequencies may lead to an estimator with a higher variance. The effects of micro-structure and the optimal sampling of intradaily returns have been discussed in several papers, such as, for example, Oomen (2001), Andersen, Bollerslev, Diebold, and Labys (2003), and Bandi and Russel (2003), among others.

3. The Data

In this paper we use data of five out of the ten major stocks from the São Paulo Stock Market (BOVESPA), namely: Bradesco (BBDC4), Embratel (EBTP4), Petrobrás (PETR4), Telemar (TNLP4), and Vale do Rio Doce (VALE5). The data set consists of intra-day prices observed every 15-minute from 10/01/2001 to 11/30/2003 (539 daily observations). We use data from 10/01/2001 to 04/11/2003 (379 daily observations) for in-sample evaluation and the remaining for out-of-sample analysis.

One important point to mention is the choice of the sampling frequency. We heuristically tested the bias-efficiency trade-off involved for three different frequencies: 15 minutes, 30 minutes, and 45 minutes. Based on Andersen, Bollerslev, Diebold, and Labys (2000) and Barndoff-Nielsen and Shephard (2002a), we use a simple method to choose the sampling frequency. First, we estimate the realized volatility using three different frequencies as mentioned above and average them over the sample. Table 1 shows the average of the daily realized volatility. As pointed out by Andersen, Bollerslev, Diebold, and Labys (2000), if microstructure effects are present the average of the realized volatility may be differ according to the sampling frequency. As we can see by inspection of 1, the mean is rather stable.

TABLE 1. Mean daily realized volatility.

Asset	15-minute window	30-minute window	45-minute window
Bradesco	0.0215	0.0199	0.0200
Embratel	0.0434	0.0399	0.0404
Petrobrás	0.0200	0.0188	0.0190
Telemar	0.0228	0.0218	0.0221
Vale	0.0172	0.0159	0.0159

Notes: The table shows the average of the daily realized volatility estimated using different sampling frequencies. The estimation period is 10/01/2001 – 04/11/2003.

On the other hand, to estimate the precision of the estimator we make use of the result of Barndoff-Nielsen and Shephard (2002a)

$$\frac{\log\left(\sum_{i=1}^{N_f} r_{t,i}^2\right) - \log\left(h_t\right)}{\sqrt{\frac{2\sum_{i=1}^{N_f} r_{t,i}^4}{3\left(\sum_{i=1}^{N_f} r_{t,i}^2\right)^2}}} \xrightarrow{\mathcal{D}} \mathrm{N}(0,1). \tag{6}$$

Table 2 shows the average size of the 95% confidence interval for the realized volatility calculated from (6). As can be observed it seems that a 15-minute frequency is the “optimal” frequency, when a bias-efficiency trade-off is considered. Thus, this will be the chosen frequency in the remaining of this paper.

TABLE 2. Mean of the confidence intervals of the daily realized volatility.

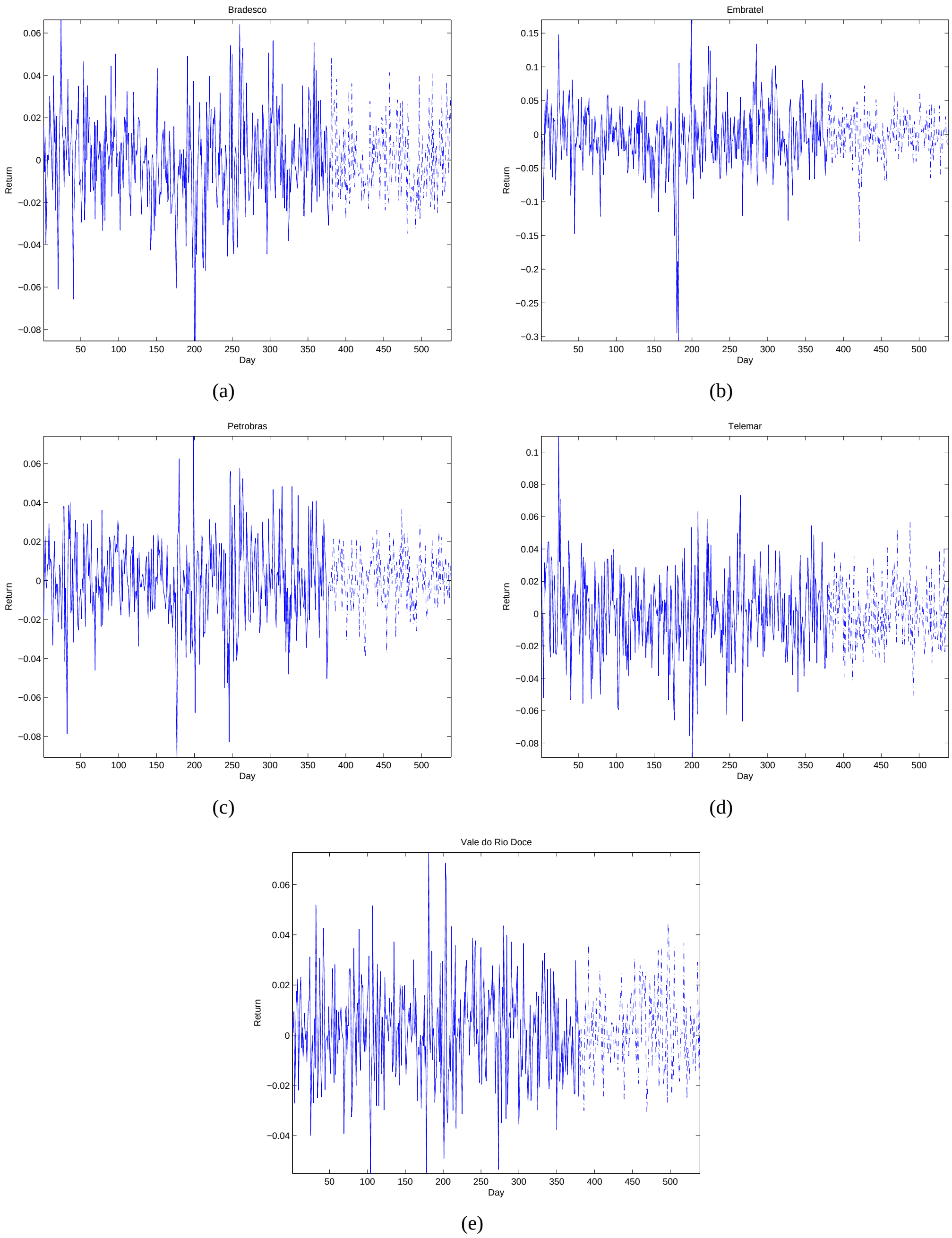
Asset	15-minute window	30-minute window	45-minute window
Bradesco	0.000871	0.001000	0.001100
Embratel	0.004800	0.005100	0.005200
Petrobrás	0.000867	0.000967	0.001100
Telemar	0.000930	0.001100	0.001300
Vale	0.000670	0.000724	0.000787

Notes: The table shows the average of the confidence interval of the daily realized volatility estimated using different sampling frequencies. The estimation period is 10/01/2001 – 04/11/2003.

Figure 1 shows the daily returns. The dashed lines represent the out-of-sample period.

3.1. **The Distribution of Standardized Returns and Realized Volatility.** Table 3 shows, for each of the daily returns of the five stocks considered in this paper, the mean, the standard deviation, the skewness, the kurtosis, and the p-value of the Jarque-Bera normality test. As can be observed, as expected, all the five series have excess of kurtosis, specially Embratel. One interesting fact is that four of the series are negatively skewed, whereas Vale do Rio Doce is positive skewed. The Jarque-Bera test strongly rejects the null hypothesis of normality for all the five series.

Table 4 shows descriptive statistics for the standardized returns. To compare the realized volatility approach with other methods to compute the daily volatility, we estimate the following models: a GARCH(1,1), a EGARCH(1,1) (Nelson 1991), and a GJR-GARCH(1,1) (Glosten, Jagannanthan, and Runkle 1993). In addition we also compute the



(a) (b) (c) (d) (e)

FIGURE 1. Daily returns. The dashed lines represent the out-of-sample period. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

volatility with the Riskmetrics methodology that is based on a exponentially weighted moving average of the squared returns (EWMA) with a decay factor $\lambda = -0.94$ as suggested in Morgan (1996). For each of the daily standardized returns of the five stocks considered in this paper, Table 4 shows the mean, the standard deviation, the skewness, the kurtosis, and the p-value of the Jarque-Bera normality test. It seems that the realized volatility methodology produces (nearly) Gaussian standardized returns for all the five series. The same result does not hold for the other models. The

TABLE 3. Daily returns: Descriptive statistics.

Asset	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera
Bradesco	1.64×10^{-4}	0.0225	-0.1669	4.0555	9.00×10^{-5}
Embratel	-4.90×10^{-3}	0.0511	-0.9743	8.9420	0
Petrobrás	-1.15×10^{-4}	0.0224	-0.2611	4.4836	5.80×10^{-9}
Telemar	3.91×10^{-4}	0.0256	-0.0107	4.0867	1.28×10^{-4}
Vale	1.60×10^{-3}	0.0192	0.1847	3.8186	2.20×10^{-3}

Notes: The table shows the mean, the standard deviation, the skewness, and the kurtosis of the daily returns, and the p-value of the Jarque-Bera test.

only exceptions are the GARCH(1,1), the EGARCH(1,1), and the GJR-GARCH(1,1) models estimated for Bradesco and Vale do Rio Doce and the EGARCH(1,1) and the GJR-GARCH(1,1) for Petrobrás. Figure 2 shows the histograms of the returns and standardized returns when the daily variance is estimated by the realized volatility approach.

Table 5 shows descriptive statistics for the realized volatility. It is clear that, for all the five series, the realized volatility is strongly positively skewed and non-Gaussian. However, in accordance with the international literature, the natural logarithm of the realized volatilities are nearly Gaussian as shown in Table 6. Figures 3 and 4 show the evolution and the histogram of the realized volatility and the log realized volatility.

4. MODELLING AND FORECASTING REALIZED VOLATILITY

4.1. **In-sample Analysis.** In order to compare the performance of different methods/models to extract the daily volatility, we estimate 95% confidence intervals for the daily returns and check the number of observations of the absolute daily returns that are greater than the interval. Table 7 shows the number of exceptions of the 95% interval and Table 8 shows the p-values of the tests of unconditional coverage, independence, and conditional coverage (Christoffersen 1998). All the methods/models considered in the paper seems to produce "correct" intervals.

It seems, by inspection of Figure 3 that the natural logarithm of the realized volatilities, on the contrary of the international empirical evidence, is not very persistent. Figure 5 shows the autocorrelation and partial correlation functions for the log realized volatilities. Table 9 presents the statistics and the respective p-values of the Augmented-Dickey-Fuller (ADF) and Philipps-Perron (PP) tests for the null hypothesis of a unit-root. The unit-root hypothesis is strongly rejected for all the five series. Furthermore, there is no evidence of long-memory in the series.

Based on the evidence of no long memory in the log realized volatility series, we proceed by estimating a simple linear model for each series defined as

$$\log(h_t) = \alpha + \beta r_{t-1}^2 + \phi \log(h_{t-1}) + \delta \log(h_{t-1}) \times (r_{t-1} < 0) + \theta \varepsilon_{t-1} + u_t, \tag{7}$$

where $\{u_t\}_{t=1}^{T}$ is a sequence of independent and identically distributed random variables with zero mean and variance σ^2, $u_t \sim \text{IID}\left(0, \sigma^2\right)$. The details of the estimated models are described in Table 10.

TABLE 4. Daily standardized returns: Descriptive statistics.

Asset	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera
		Panel I: Realized Volatility			
Bradesco	0.0152	0.9956	0.0831	2.7107	0.3890
Embratel	-0.0963	0.9976	0.1003	2.4461	0.0577
Petrobrás	-0.0076	1.0088	0.0122	2.4885	0.1133
Telemar	0.0236	1.0370	0.0672	2.5952	0.2177
Vale	0.1056	1.0748	0.0387	2.7161	0.4728
		Panel II: EWMA ($\lambda = 0.94$)			
Bradesco	$-7.37 \times 10{-}4$	1.0334	-0.1169	3.7898	0.0061
Embratel	-0.1089	1.0309	-0.4971	4.8002	$6.88 \times 10{-}15$
Petrobrás	-0.0171	1.0483	-0.4817	4.7595	$1.85 \times 10{-}14$
Telemar	0.0032	1.0357	-0.1474	3.7688	0.0060
Vale	0.0811	1.0312	-0.0796	4.2522	$8.95 \times 10{-}6$
		Panel II: GARCH(1,1)			
Bradesco	0.0030	1.0005	-0.0757	3.5296	0.1063
Embratel	0.0067	1.0010	-0.3006	4.0334	$1.80 \times 10{-}5$
Petrobrás	-0.0014	0.9976	-0.1800	3.5370	0.0434
Telemar	-0.0064	1.0068	-0.0598	3.9023	0.0019
Vale	0.0021	0.9995	0.0180	3.5813	0.0815
		Panel III: EGARCH(1,1)			
Bradesco	0.0088	1.0006	-0.0365	3.5296	0.1407
Embratel	-0.0008	1.0010	-0.3467	4.0334	$1.43 \times 10{-}7$
Petrobrás	0.0054	0.9989	-0.0989	3.5370	0.5153
Telemar	-0.0069	1.0089	-0.0971	3.9023	0.0164
Vale	0.0009	0.9996	0.0292	3.5813	0.0930
		Panel IV: GJR-GARCH(1,1)			
Bradesco	0.0098	1.0007	-0.0343	3.4954	0.1599
Embratel	0.0053	1.0010	-0.3212	4.0521	$8.98 \times 10{-}6$
Petrobrás	0.0045	0.9992	-0.1131	3.2742	0.3978
Telemar	-0.0059	1.0074	-0.0527	3.9142	0.0017
Vale	-0.0013	0.9994	-0.0250	3.4940	0.1644

Notes: The table shows the mean, the standard deviation, the skewness, the kurtosis, and the p-value of the Jarque-Bera test of the daily standardized returns.

4.2. **Out-of-sample Analysis.** To evaluate the forecasting performance of the models estimated before, we conduct an out-of-sample experiment. Figure 6 shows the daily returns and the 95% confidence interval computed with the forecasted volatilities. The dashed lines represent the out-of-sample period. Table 11 shows the frequency of observations of the absolute returns that are greater than the 95% confidence interval over the out-of-sample period.

TABLE 5. Realized volatility: Descriptive statistics.

Asset	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera
Bradesco	0.0215	0.0079	1.7447	9.6061	0
Embratel	0.0434	0.0225	5.2217	53.4585	0
Petrobrás	0.0200	0.0091	2.0659	9.5716	0
Telemar	0.0228	0.0079	0.7927	3.5341	3.44×10^{-10}
Vale	0.0172	0.0084	2.4204	12.0763	0

Notes: The table shows the mean, the standard deviation, the skewness, the kurtosis, and the p-value of the Jarque-Bera test of the daily realized volatilities.

TABLE 6. Daily log realized volatilities: Descriptive statistics.

Asset	Mean	Standard deviation	Skewness	Kurtosis	Jarque-Bera
Bradesco	-3.9002	0.3451	0.0753	3.5299	0.1062
Embratel	-3.2200	0.3806	0.8128	5.1035	0
Petrobrás	-3.9939	0.3929	0.4693	3.4192	2.82×10^{-4}
Telemar	-3.8394	0.3473	-0.1016	2.7485	0.414
Vale	-4.1530	0.4081	0.5341	3.7153	2.87×10^{-6}

Notes: The table shows the mean, the standard deviation, the skewness, the kurtosis, and the p-value of the Jarque-Bera test of the daily log realized volatilities.

TABLE 7. In-sample analysis: Frequency of observations of the absolute returns that are greater than a given confidence interval.

Asset	Realized Volatility	EWMA ($\lambda = 0.94$)	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
			Panel I: 99% Confidence Interval		
Bradesco	0.0026	0.0264	0.0185	0.0211	0.0211
Embratel	0.0026	0.0264	0.0211	0.0290	0.0211
Petrobras	0.0026	0.0158	0.0211	0.0158	0.0185
Telemar	0.0053	0.0237	0.0132	0.0185	0.0132
Vale	0.0158	0.0185	0.0185	0.0132	0.0158
			Panel II: 95% Confidence Interval		
Bradesco	0.0449	0.0686	0.0660	0.0765	0.0712
Embratel	0.0422	0.0660	0.0554	0.0528	0.0554
Petrobras	0.0422	0.0501	0.0554	0.0528	0.0501
Telemar	0.0528	0.0686	0.0607	0.0686	0.0660
Vale	0.0554	0.0554	0.0475	0.0501	0.0475
			Panel III: 90% Confidence Interval		
Bradesco	0.1029	0.0976	0.1055	0.1108	0.1029
Embratel	0.1003	0.1003	0.0923	0.0844	0.0923
Petrobras	0.1108	0.1082	0.1082	0.1055	0.1108
Telemar	0.1161	0.1108	0.0950	0.0976	0.1003
Vale	0.1266	0.1055	0.1029	0.1082	0.1029

TABLE 8. In-sample analysis: p-value of the test of the null hypothesis of correct unconditional coverage, independence, and correct conditional coverage, at nominal significance level 0.05.

Asset	Realized Volatility	EWMA	GARCH(1,1)	EGARCH(1,1)	GJR-GARCH(1,1)
			Panel I: Unconditional Coverage		
			99% Confidence Interval		
Bradesco	0.0866	0.0078	0.1383	0.0585	0.0585
Embratel	0.0866	0.0078	0.0585	0.0025	0.0585
Petrobrás	0.0866	0.2930	0.0585	0.2930	0.1383
Telemar	0.3098	0.0223	0.5515	0.1383	0.5515
Vale	0.2930	0.1383	0.1383	0.5515	0.2930
			95% Confidence Interval		
Bradesco	0.6402	0.1149	0.1731	0.0275	0.0737
Embratel	0.4755	0.1731	0.6346	0.8062	0.6346
Petrobrás	0.4755	0.9906	0.6346	0.8062	0.9906
Telemar	0.8062	0.1149	0.3550	0.1149	0.1731
Vale	0.6346	0.6346	0.8214	0.9906	0.8214
			Panel II: Independence		
			99% Confidence Interval		
Bradesco	0.9419	0.2535	0.6073	0.5564	0.5564
Embratel	0.9419	0.4610	0.5564	0.4167	0.5564
Petrobrás	0.9419	0.0742	0.1497	0.0742	0.1084
Telemar	0.8840	0.5076	0.7143	0.6073	0.7143
Vale	0.6600	0.6073	0.6073	0.7143	0.6600
			95% Confidence Interval		
Bradesco	0.2057	0.8214	0.7790	0.8684	0.9561
Embratel	0.2343	0.0510	0.7852	0.1348	0.1159
Petrobrás	0.7005	0.8062	0.8673	0.9520	0.9616
Telemar	0.3832	0.6402	0.0490	0.1176	0.0897
Vale	0.1159	0.8214	0.9616	0.9616	0.8743
			Panel III: Conditional Coverage		
			99% Confidence Interval		
Bradesco	0.2298	0.0151	0.2921	0.1404	0.1404
Embratel	0.2298	0.0220	0.1404	0.0074	0.1404
Petrobrás	0.2298	0.1168	0.0591	0.1168	0.0919
Telemar	0.5907	0.0590	0.7832	0.2921	0.7832
Vale	0.5222	0.2921	0.2921	0.7832	0.5222
			95% Confidence Interval		
Bradesco	0.4025	0.1003	0.3801	0.0868	0.2017
Embratel	0.3821	0.0841	0.7533	0.3173	0.2595
Petrobrás	0.7200	0.1947	0.8808	0.9689	0.9988
Telemar	0.6634	0.6704	0.0940	0.0848	0.0937
Vale	0.2595	0.8514	0.9988	0.9988	0.9627

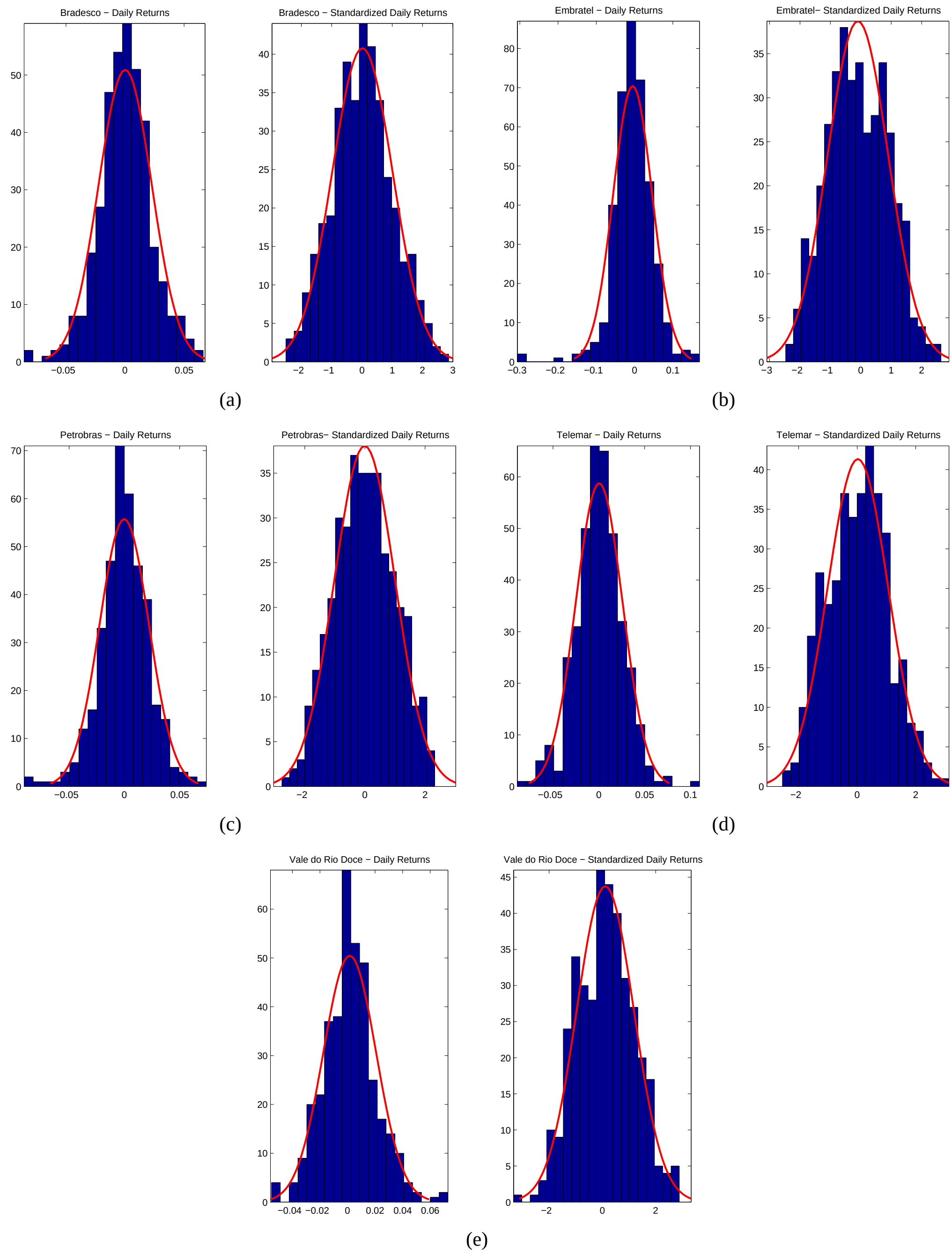


FIGURE 2. Histograms of the daily returns and standardized daily returns. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

5. CONCLUSIONS

The goal of this paper was twofold. First, by using the realized variance estimated by summing up intraday squared returns of five major Brazilian stock assets we estimated the distribution of the standardized returns. The main finding,

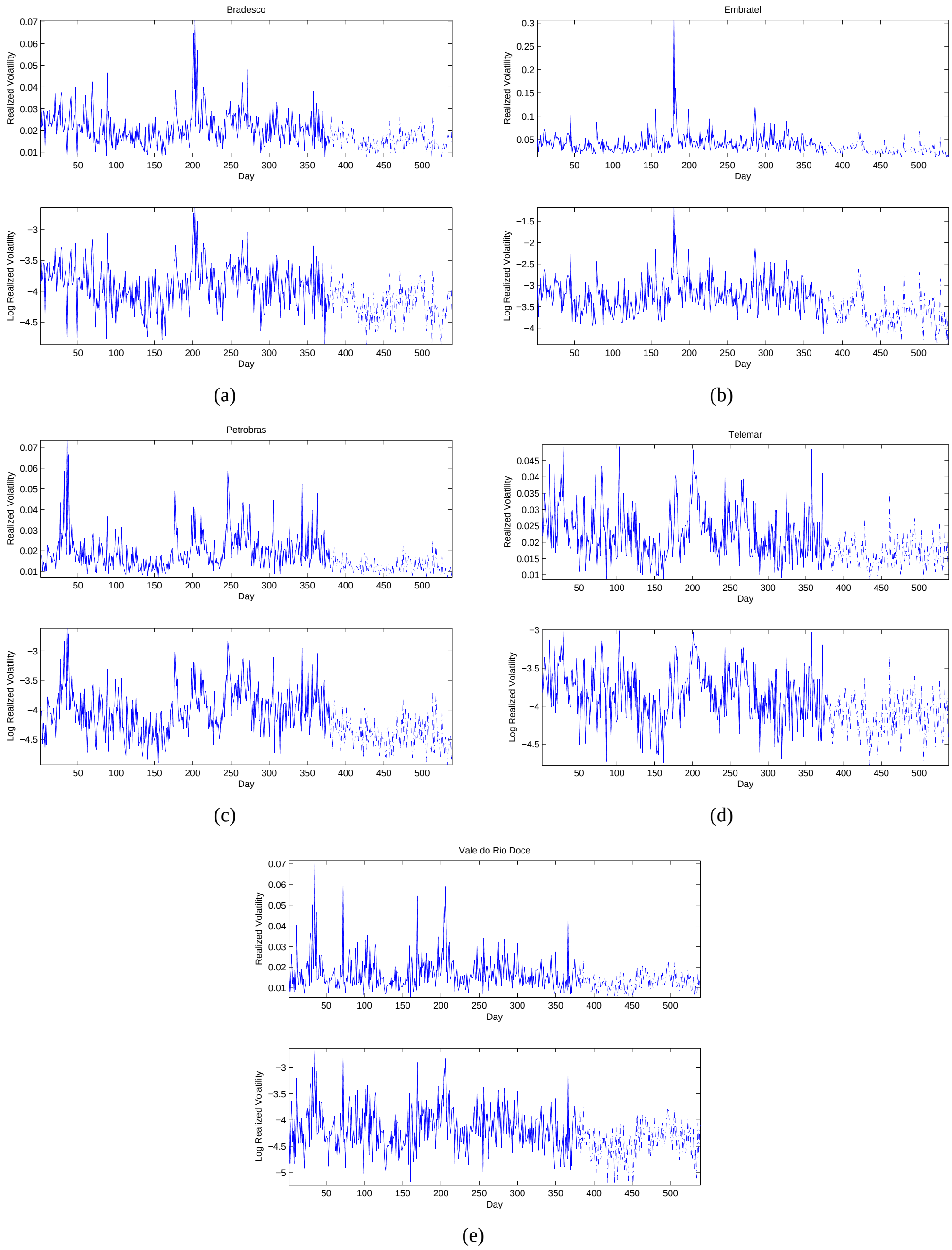


FIGURE 3. Daily realized volatilities. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

in accordance with the international literature, is that the distribution of the standardized returns is Gaussian. Furthermore, the distribution of the realized volatility (the squared root of the realized variance) is strongly skewed and non-Gaussian. However, the log realized volatility is nearly Gaussian. On the other hand, when the returns are standardized with the volatility given by models of the ARCH family, its distribution still has excess of kurtosis. Second,

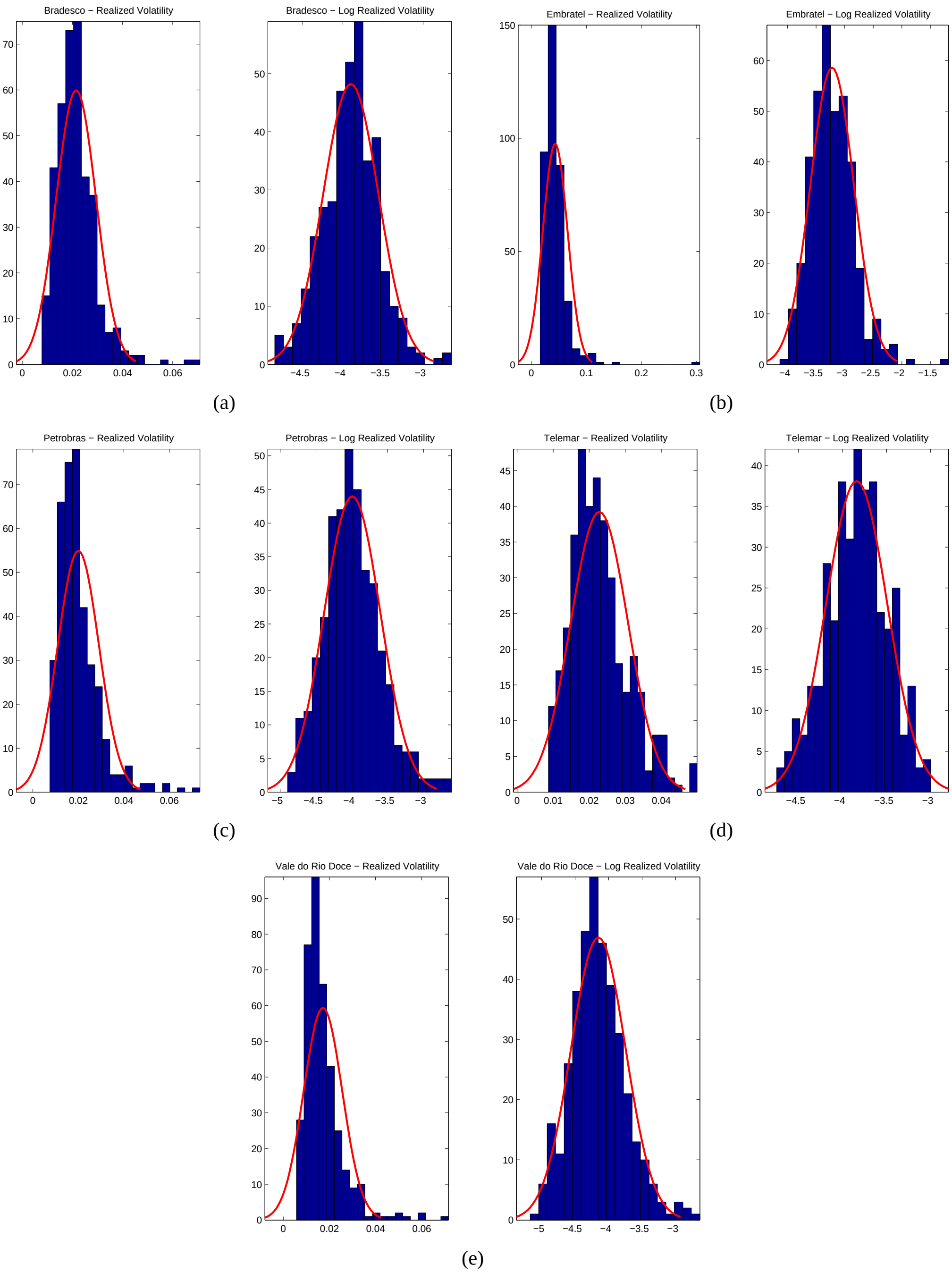


FIGURE 4. Histograms of the daily realized volatilities and log daily realized volatilities. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

by considering the log realized volatility as an observed variable, instead of latent as in the ARCH approach, we estimated a simple linear model to forecast out-of-sample values. When standard methods to evaluate volatility measures were used to compare different methods, it is difficult to discriminate the performance of the different alternatives.

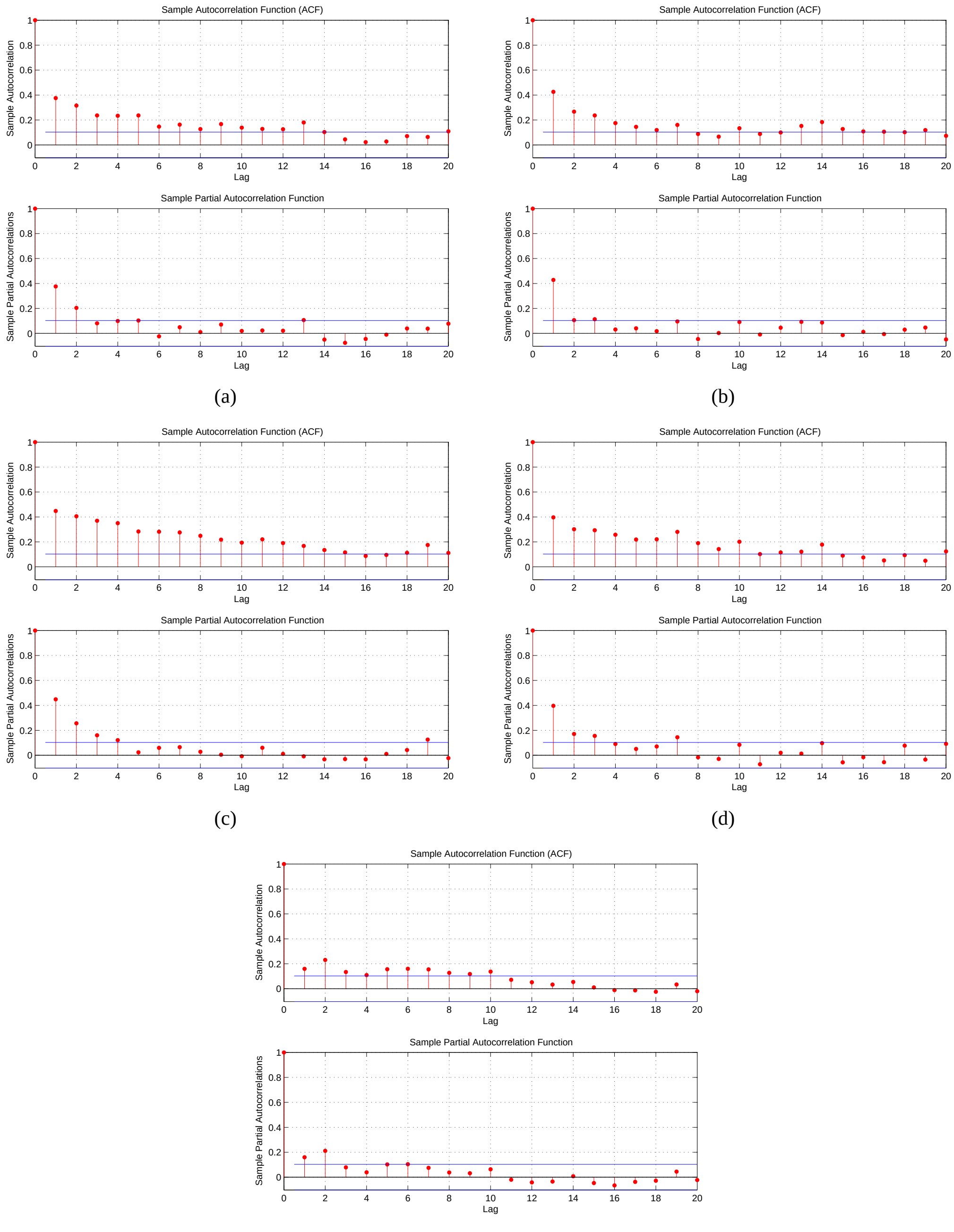


FIGURE 5. Autocorrelation and partial autocorrelation functions of the log realized volatility. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

REFERENCES

ANDERSEN, T., AND T. BOLLERSLEV (1998): "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39, 885–905.

ANDERSEN, T., T. BOLLERSLEV, F. X. DIEBOLD, AND P. LABYS (2000): "Market Microstructure Effects and the Estimation of Integrated Volatility," Work in progress, Duke University and University of Pennsylvania.

TABLE 9. In-sample analysis: Unit-root tests.

Asset	Dickey-Fuller	Philipps-Perron
Bradesco	-8.77	-14.38
	(0)	(0)
Embratel	-12.22	-12.97
	(0)	(0)
Petrobras	-5.19	-13.97
	(0)	(0)
Telemar	-6.92	-14.29
	(0)	(0)
Vale	-10.16	-17.69
	(0)	(0)

Notes: The table shows the p-value of several unit-roots test applied to the log of the realized volatilities.

TABLE 10. In-sample analysis: Estimated models.

$$\log(h_t) = \alpha + \beta r_{t-1}^2 + \phi\log(h_{t-1}) + \delta\log(h_{t-1}) \times (r_{t-1} < 0) + \theta\varepsilon_{t-1} + \varepsilon_t$$

Parameters	Bradesco	Embratel	Petrobrás	Telemar	Vale
α	-1.02^{**}	-2.69^{**}	-1.01^{**}	-1.35^{**}	-1.50^{**}
	(0.19)	(0.56)	(0.26)	(0.23)	(0.49)
β	130.14^{**}	17.56^{**}	87.99^{**}	123.89^{**}	151.66^{**}
	(19.83)	(4.82)	(27.16)	(20.72)	(49.71)
ϕ	0.88^{**}	0.59^{**}	0.89^{**}	$0\ \ .84^{**}$	0.83^{**}
	(0.02)	(0.09)	(0.03)	(0.03)	(0.06)
δ	—-	—-	$-\ 0.01\ ^{*}$	$-\ 0.02\ ^{**}$	—-
			(0.006)	(0.005)	
θ	-0.82^{**}	-0.33^{**}	-0.70^{**}	-0.77^{**}	-0.75^{**}
	(0.04)	(0.10)	(0.05)	(0.05)	(0.07)
$R^2_{adj.}$	0.27	0.23	0.32	0.30	0.12
JB	0.18	0	0	0.33	0
$LM_{SC}(1)$	0.12	0.48	0.88	0.32	0.08
$LM_{SC}(4)$	0.45	0.87	0.84	0.85	0.06
$LM_{ARCH}(1)$	0.05	0.70	0.90	0.55	0.68
$LM_{ARCH}(4)$	0.42	0.70	0.98	0.93	0.007

——— (2001a): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

——— (2001b): "Exchange rate returns standardized by realized volatility are (nearly) Gaussian," *Multinational Finance Journal*, forthcoming.

——— (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

AREAL, N. M. P. C., AND S. J. TAYLOR (2002): "The Realized Volatility of the FTSE-100 Future Prices," *Journal of Futures Markets*, 22, 627–648.

BANDI, F., AND J. R. RUSSEL (2003): "Microstructure noise, realized volatility, and optimal sampling," Working paper, Graduate School of Business, The University of Chicago.

BARNDOFF-NIELSEN, O., AND N. SHEPHARD (2002a): "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253–280.

——— (2002b): "Estimating Quadratic Variation Using Realized Volatility," *Journal of Applied Econometrics*, 17, 457–477.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 21, 307–328.

——— (1987): "A Conditional Heteroskedasticity Time Series Model for Speculative Prices and Rates of Return," *The Review of Economic and Statistics*, 69, 542–547.

CAMPBELL, J. Y., A. W. LO, AND A. C. MACKINLAY (1997): *The Econometrics of Financial Markets*. Princeton University Press.

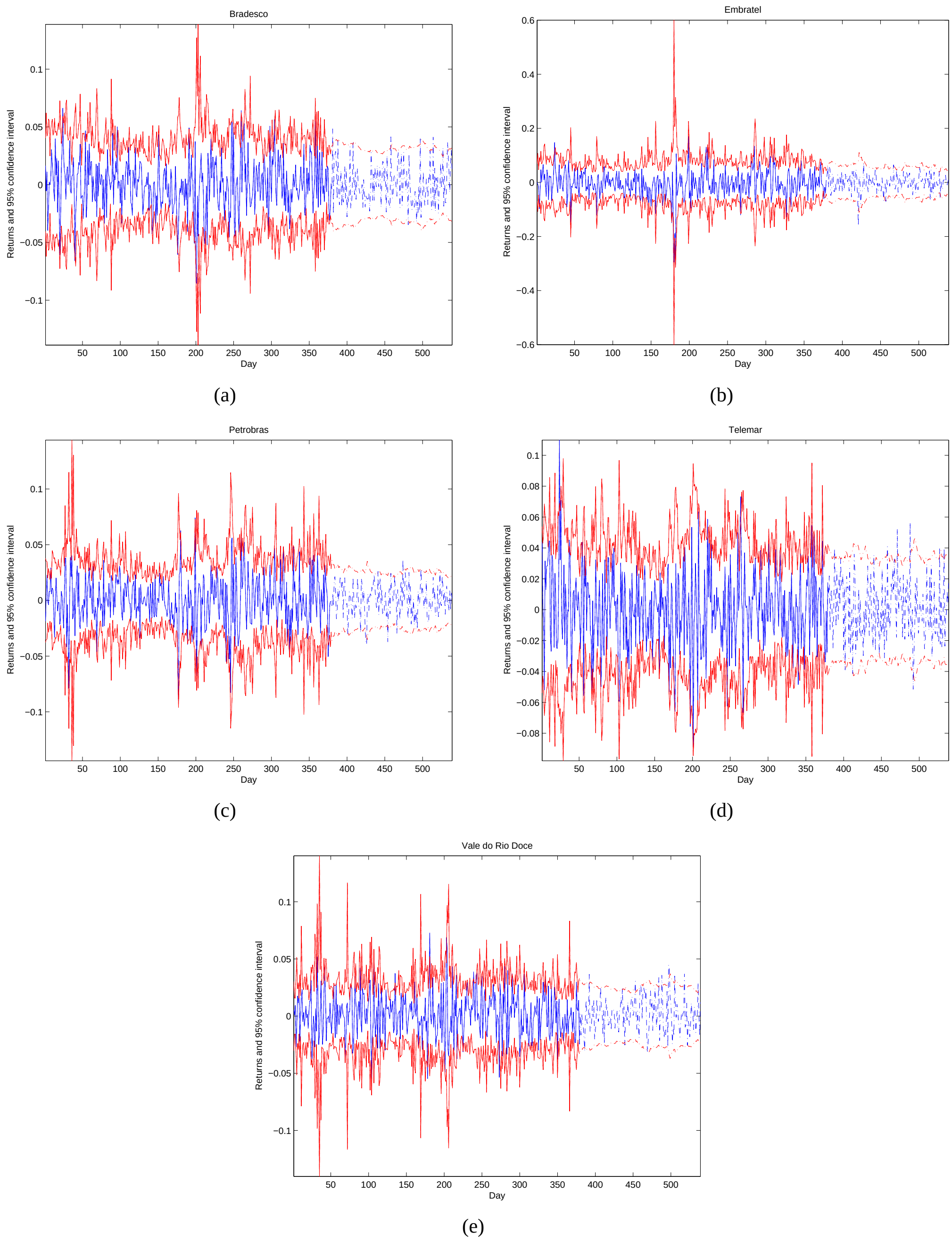


FIGURE 6. Daily returns and a 95% confidence interval computed with estimated and forecasted realized volatilities. The dashed lines represent the out-of-sample period. Panel (a): Bradesco. Panel (b): Embratel. Panel (c): Petrobrás. Panel (d): Telemar. Panel (e): Vale do Rio Doce.

CARNERO, M. A., D. PEÑA, AND E. RUIZ (2001): "Is Stochastic Volatility More Flexible Than GARCH?," Working Paper Series in Statistics and Econometrics 01-08, Universidad Carlos III de Madrid.

CHRISTOFFERSEN, P. F. (1998): "Evaluating interval forecasts," *International Economic Review*, 39, 841–862.

EBENS, H. (1999): "Realized Stock Volatility," Unpublished manuscript, Johns Hopkins University.

TABLE 11. Out-of-sample analysis: Frequency of observations of the daily absolute returns are greater than a 95% confidence interval.

Asset	RV	EWMA ($\lambda = 0.94$)	GARCH	EGARCH	GJR-GARCH	RV + GARCH	RV + EGARCH	RV + GJR-GARCH
Panel I: 99% Confidence Interval								
Bradesco	0.0187	0.0125	0	0	0	0.0063	0.0063	0.0063
Embratel	0.0187	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063	0.0063
Petrobrás	0.0250	0.0063	0	0	0	0.0063	0.0063	0.0063
Telemar	0.0187	0.0125	0.0063	0.0063	0.0063	0.0125	0.0125	0.0125
Vale	0.0437	0.0125	0	0	0	0.0125	0.0063	0.0125
Panel II: 95% Confidence Interval								
Bradesco	0.0563	0.0437	0.0250	0.0125	0.0250	0.0437	0.0437	0.0375
Embratel	0.0750	0.0437	0.0187	0.0250	0.0187	0.0313	0.0313	0.0437
Petrobrás	0.0750	0.0375	0.0313	0.0313	0.0250	0.0437	0.0375	0.0375
Telemar	0.0813	0.0437	0.0187	0.0187	0.0187	0.0563	0.0563	0.0563
Vale	0.0938	0.0813	0.0313	0.0313	0.0313	0.0563	0.0563	0.0563
Panel III: 90% Confidence Interval								
Bradesco	0.1125	0.1063	0.0750	0.0563	0.0688	0.0938	0.0875	0.0875
Embratel	0.0875	0.0813	0.0500	0.0688	0.0500	0.0750	0.0680	0.0750
Petrobrás	0.1250	0.0813	0.0563	0.0500	0.0500	0.0875	0.0813	0.0750
Telemar	0.1375	0.0938	0.0625	0.0563	0.0625	0.0875	0.0813	0.0875
Vale	0.1375	0.1187	0.0813	0.0688	0.0750	0.1250	0.1125	0.1313

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflations," *Econometrica*, 50, 987–1007.

GLOSTEN, L., R. JAGANNANTHAN, AND R. RUNKLE (1993): "On The Relationship Between The Expected Value and The Volatility of The Nominal Excess Returns on Stocks," *Journal of Finance*, 48, 1779–1801.

HOL, E., AND S. J. KOOPMAN (2002): "Stock Index Volatility Forecasting with High Frequency Data," Discussion Paper 2002-068/4, Tinbergen Institute.

LI, K. (2002): "Long Memory Versus Option-Implied Volatility Predictions," *Journal of Derivatives*, 9, 9–25.

MARTENS, M., AND J. ZEIN (2002): "Forecasting Financial Volatility: High-Frequency Time-Series Forecasts Vis-a-Vis Implied Volatility," Working paper, Erasmus University.

MEDDAHI, N. (2002): "A theoretical Comparison Between Integrated and Realized Volatility," *Journal of Applied Econometrics*, 17, 479–508.

MERTON, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

MORGAN, J. P. (1996): *J. P. Morgan/Reuters Riskmetrics – Technical Document*. J. P. Morgan, New York.

MOTA, B., AND M. FERNANDES (2004): "Desempenho de Estimadores de Volatilidade na Bolsa de Valores de São Paulo," *Revista Brasileira de Economia*, forthcoming.

NELSON, D. B. (1991): "Conditinal Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

OOMEN, R. C. A. (2001): "Using High Frequency Stock Market Index Data to Calculate, Model, and Forecast Realized Return Variance," Working Paper 2001/6, European University Institute.

PONG, S., M. B. SHACKLETON, S. J. TAYLOR, AND X. XU (2002): "Forecasting Sterling/Dollar Volatility: Implied Volatility Versus Long Memory Intraday Models," Working paper, Lancaster University.

TAYLOR, S. J. (1986): *Modelling Financial Time Series*. John Wiley.

TERÄSVIRTA, T. (1996): "Two Stylized Facts anf the GARCH(1,1) Model," Working Paper Series in Economics and Finance 96, Stockholm School of Economics.

(M. R. C. Carvalho) DEPARTMENT OF ECONOMICS, PONTIFICAL CATHOLIC UNIVERSITY OF RIO DE JANEIRO, RIO DE JANEIRO, RJ, BRAZIL.

E-mail address: marcelorcc@globo.com

(M. A. S. Freire) DEPARTMENT OF ECONOMICS, PONTIFICAL CATHOLIC UNIVERSITY OF RIO DE JANEIRO, RIO DE JANEIRO, RJ, BRAZIL.

E-mail address: mfreire@econ.puc-rio.br

(M. C. Medeiros – Corresponding author) DEPARTMENT OF ECONOMICS, PONTIFICAL CATHOLIC UNIVERSITY OF RIO DE JANEIRO, RIO DE JANEIRO, RJ, BRAZIL.

E-mail address: mcm@econ.puc-rio.br

(L. R. Souza) MINISTÉRIO DO PLANEJAMENTO, BRASÍLIA, DF, BRAZIL.

E-mail address: leonardo.r.souza@planejamento.gov.br