BRAZILIAN BUSINESS REVIEW
v. 7, n.1
Vitória-ES, Jan - Apr .2010
p.42-58 ISSN 1808-2386 DOI: http://dx.doi.org/10.15728/bbr.2010.7.1.3

Does implied volatility contain information about future volatility? Evidence from the Petrobras options market

José Valentim Machado Vicente †
Ibmec-RJ Business School

Tiago de Sousa Guedes Ω
Ibmec-RJ Business School

ABSTRACT: The aim of this work is to study the relationship between implied and realized volatilities. For this purpose, we analyze the markets of Petrobras stocks and calls between January 2006 and December 2008. Regression analysis with no overlapping monthly data of in-the-money, at-the-money and out-of-the-money calls indicates that the implied volatility of out-of-the-money options contains more information about future volatility than does historical volatility. On the other hand, the implied volatility of the in-the-money and at-the-money calls has poor explanatory power about future volatility.

Keywords: Options, volatility, market efficiency.

Received in 07/09//2010; revised in 03/11/2010; accept in 08/04/2010

Corresponding authors*:
† Ph.D. in Mathematical Economics from the Institute of Pure and Applied Mathematics - IMPA.
Professor at Ibmec Business School.
Address: Av. Gastão Senges 55 ap. 803 Barra da Tijuca – Rio de Janeiro - RJ – Brazil
CEP: 22631-280.
E-mail: jvalent@terra.com.br
Telephone: (21) 2572-0692.

Ω Master's in Economics from Ibmec Business School
Address: Rua Gomes Carneiro, no.80, 302, Ipanema, Rio de Janeiro – RJ – Brazil
CEP 22071-110.
E-mail: tguedes84@hotmail.com
Telephone: (21) 9801-0654

Editor's note: *This paper was accepted by Antonio Lopo Martinez*

1. INTRODUCTION

Implied volatility is a measure of the expectations of market participants plus a risk premium. If this premium is small or at least constant over time, implied volatility will be a good forward-looking estimator of future volatility. Thus, some economists believe that the use of implied volatility as a forecast of future volatility shows more promising results than that offered by models based on historical data. The aim of this paper is to examine whether implied volatility contains information about future volatility. For this purpose, we examine the relationship between the implied volatility of Petrobras options (calls) and the realized volatility in a subsequent period.

If the rational expectations hypothesis is correct and markets are efficient, then implied volatility should contain significant information about future volatility. Various studies have tried to determine the best estimator of future volatility. Different results have been obtained, with arguments in favor of both the use of implied volatility and historical volatility.

Day & Lewis (1992) and Lamoureux & Lastrapes (1993) concluded that implied volatility is a biased and inefficient forecast of future volatility and that historical volatility contains more information about future volatility than does implied volatility.

Day & Lewis (1992) studied the S&P 100 index between 1985 and 1989, with an option maturity period of 36 business days. Lamoureux & Lastrapes (1993) focused on the options of the ten most liquid stocks of the S&P 100, between 1982 and 1984, with a time to maturity of 119 business days. Canina & Figlewski (1993) also utilized data on the S&P 100 index before 1987. They observed a more radical result, indicating that the implied volatility has no predictive power regarding future volatility.

Christensen & Prabhala (CP) (1998) studied at-the-money (ATM) call options with one-month maturity on the S&P 100 index using data between 1983 and 1995. Their conclusion is contrary to the results of the above authors. They found that the implied volatility is a better estimator of future volatility than historical volatility, being efficient and less biased in relation to the previous studies, particularly in the period after the 1987 New York Stock Exchange crash.

This event represented a structural break in the stock market and can explain why implied volatility was a biased estimator in the previous studies. CP also indicated some failings in the works of Day & Lewis (1992) and Lamoureux & Lastrapes (1993).

Among them is the use of small and overlapping samples of options with maturities longer than one month.

They also argued that the earlier works contained a problem of maturity miss-match, namely the forecasting power of implied volatility was tested only one day ahead by Lamoureux & Lastrapes (1993) and one week ahead by Day & Lewis (1992). In line with CP, Gwilym & Buckle (1999), analyzing data on ATM options on the main United Kingdom stock index, the UK FTSE 100, concluded that implied volatility contains more information on future volatility than the historical volatility.

For the Brazilian market, Gabe & Portugal (2004) used overlapping data from Telemar options in the period from October 2, 2000 to October 10, 2002. They found that the volatility estimated by the GARCH and EGARCH models was an efficient and unbiased estimator, better than implied volatility. On the other hand, Tabak & Chang (2006) showed that the implied volatility of foreign exchange options is a better estimator of future volatility than that obtained via GARCH models.

In this study we revisit this question, based on the Petrobras options market. Since these options did not have good liquidity before 2006, the study examines, besides ATM options, the series corresponding to ITM (in-the-money) and OTM (out-of-the-money) options, in order to increase the database. Additionally, this allows testing separately the information contained in ITM, ATM and OTM options.

Unlike other studies of the Brazilian market, our goal is not to determine the best future volatility forecast, but rather to assess the explanatory power of implied and historical volatilities with respect to future volatility[i].

While the works of Gabe & Portugal (2004) and Tabak & Chang (2006) compared implied volatility with volatility predicted via traditional econometric models, we analyze the relationship between implied and realized volatilities. Therefore, while a bias in implied volatility represents a negative point for these authors, in our approach this is interpreted as a risk premium.

For ITM options, we found no relation between their implied volatility and future volatility, both in level and in log. The implied volatility of ATM options was only significant in the equation in level, but was strongly biased.

The implied volatility of OTM options contained information about realized volatility, both in level and in log, and was efficient and less biased than the implied volatility of ATM options. In the case of the log of the implied volatility of OTM options, the result obtained was similar to those of Christensen & Prabhala (1998) and Gwilym & Buckle (1999), indicating that a forward-looking approach in Brazil can be successfully employed.

The rest of this paper is organized as follows. In Section 2 we present the concepts of implied and realized volatilities. In Section 3 we present the database. In Section 4 we discuss the empirical results. In Section 5 we compare our results with those of other studies. Section 6 concludes.

2. IMPLIED AND REALIZED VOLATILITIES

In finance, volatility corresponds to the standard deviation of an asset's returns. While it is impossible today to know the exact prices of securities in the future, statistical methods and regressions can be utilized to estimate the future volatility of the respective underlying asset. These volatility estimates shed light on the expected future price movements of the asset in question.

An alternative to econometric methods to obtain information about an asset's volatility consists of analyzing the options market. The option price is a direct function of the volatility of its underlying asset. Since this price is observable, one can extract the volatility from it. This requires the use of some pricing model. The most famous of them is undoubtedly the Black & Scholes (BS) (1973) model.

The BS option pricing model consists of an equation that provides the fair price of options through no-arbitrage arguments. The price at *t* for a European option maturing at *T* is calculated according to the following expression:

$$c_t = S_t N(d_1) - K e^{-r_f \tau} N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S_t}{K}\right) + \left(r_f + \frac{1}{2}\sigma^2\right)\tau}{\sigma\sqrt{\tau}}$$

$$d_2 = d_1 - \sigma\sqrt{\tau},$$

Where c_t is the theoretical value of a call option, S_t is the price of the underlying asset, K is the strike price of the option, $\tau = T - t$ is the time to maturity, σ is the volatility, r_f is the risk-free rate and N is the standard normal cumulative function [ii].

The implied volatility (IV) at *t* is simply the value of the constant σ which makes the theoretical price of the call option equal to the market price. Although the calls traded on the São Paulo Stock Exchange (Bovespa) are American options (they can be exercised at any time before the expiry date), they are protected against dividends (the amounts received as dividends are deducted from the strike price).

Hence, there is no advantage to early exercise and they can be characterized as European options (see Hull, 1997)[iii]. However, although the pricing model has an analytic form, there is no closed solution expressing the volatility as a function of the asset price. To solve this problem we used Newton's method.

The realized volatility (RV) between dates t and T is defined as:

$$RV_t = \sqrt{\frac{1}{\tau}\sum_{i=1}^{n}\left(r_{t+i} - \overline{r_{tT}}\right)^2} ,$$

Where n is the number of days between t and T, r_{t+i} is the daily return on day $t + i$ and $\overline{r}_{tT}$ is the average of the daily returns between t and T. The historical volatility is defined analogously as in the previous equation, but only considering one period before t (in the empirical exercise presented in Section 4, this period, as well as the maturity of the options, is one month).

3. DATABASE

Our empirical study is based on data on the preferred shares of Petrobras (PETR4) and their call options, for the period covering January 2006 to December 2008. The data were obtained from the Bovespa site.

Since this period includes a two-for-one stock split on April 28, 2008, we divided all the stock and strike prices before this date by two. The use of Petrobras shares is due to their high liquidity and substantial participation (around 14%) in the main Brazilian stock index (Ibovespa), allowing their use as a proxy for this index.

Figure 1 presents the monthly returns of preferred Petrobras shares and the Ibovespa (IBOV) between January 2006 and December 2008. Note the high volatility for both series in the period. In October 2008 the American economy suffered the worst of the subprime crisis, causing the failure of large banks and triggering a credit crunch. Thus the global economy was severely affected, leading to a strong capital outflow from emerging markets, negatively influencing the main stock indexes and increasing the risk of investments, primarily in equities.

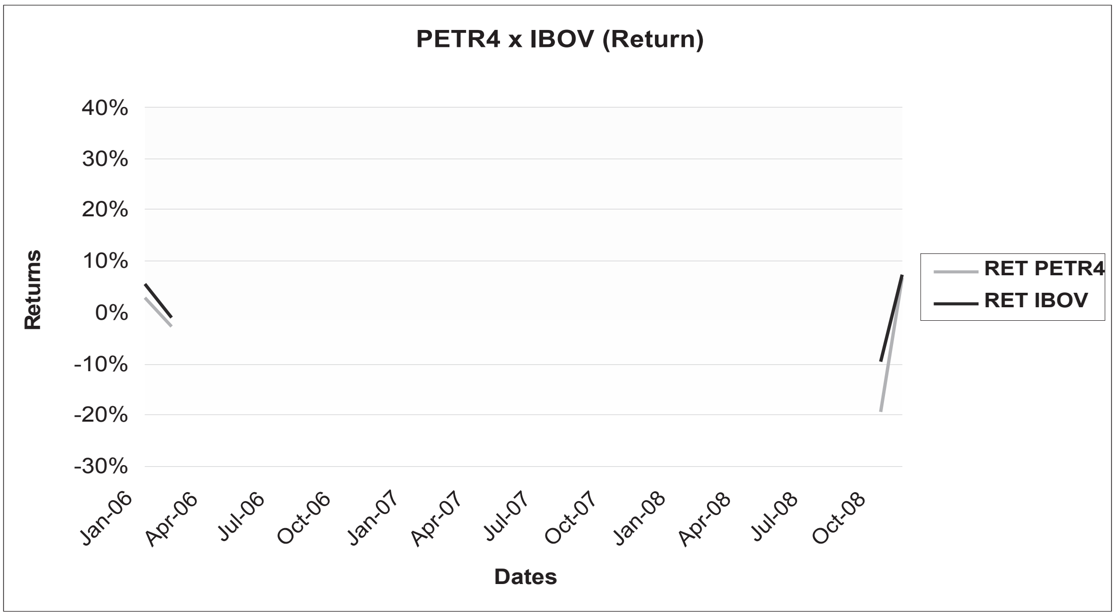

Figure 1 - Returns of Petrobras Stocks and the Ibovespa.
Nota: compares the returns of Petrobras stocks and the Ibovespa between January 2006 and December 2008.
Source: elaborate by the authors

In Brazil, there is only liquidity in the market for call options. The options expire monthly on the third Monday of each month.

This market was dominated by Telemar calls until the middle of 2006, after which the shares of Petrobras held by the public (the company is controlled by the Brazilian government) became the most frequently traded shares on the Bovespa and those with the greatest weight in the Ibovespa. Since then the market for Petrobras calls has been the largest in Brazil.

We divided the database into 35 sub-periods (monthly data from January 2006 through December 2008), each one starting on the maturity of the previous series and ending on the last business day before the maturity of the next series, with no overlap. In other words, the options have a term of one month, or approximately 21 business days.

We used the closing prices. To expand the scope of this work, besides ATM options, we also studied OTM and ITM options. In each sub-period we also calculated the realized volatility.

Options can be classified according to their moneyness into three types: in-the-money (ITM), at-the-money (ATM) and out-of-the-money (OTM). There is no fixed definition of these three categories of options. In this study, we defined an ATM option as one whose strike price is nearest the price of the underlying asset, and the ITM and

OTM options, respectively, as the options whose strike prices are below and above the ATM.

For example, on January 2, 2007, Petrobras preferred shares closed at R$ 46.74. The option with a strike price nearest this was series PETRA48, with strike equal to R$ 46.77. We thus considered this as ATM.

The first option with a strike price below R$ 46.77 was series PETRA46, with a strike price of R$ 44.77, so we defined this as the ITM option. Likewise, the first option with strike price above R$ 46.77 was PETRA50, with strike equal to R$ 48.19, so we defined it as the OTM option[iv].

In each sub-period we calculated the historical volatility of Petrobras shares and the implied volatility of the ITM, ATM and OTM options on the first day of each sub-period. The realized volatility (RV) was calculated as the standard deviation of the daily returns of the asset over the periods corresponding to the life of each option.

Figure 2 shows the evolution of the implied and realized volatilities in the period studied. Both these are expressed in annual terms. Note that the only period in which the realized volatility was higher than the implied was after July 2008, when the rumors about the financial crisis in the United States started, leading to increased nervousness in global markets.

This fact provides evidence that the options market was not able to anticipate the crisis. According to Hull (1997), after 1987 (after the NYSE crash), the graph of the implied volatilities of a series of equity options tended to form a volatility skew. That is, the higher the strike price of an option, the lower its implied volatility tends to be. One of the reasons given for this phenomenon is leverage.

When the prices of a company's shares decline, its leverage increases as the stock price decreases, and for this reason ITM options tend to have greater implied volatility than ATM and OTM options.

The implied volatilities of the options in our sample behaved with volatility skew until April 2008, when Brazil received an investment grade rating from Standard & Poor's. After this period, the Petrobras options started to have very similar implied volatility.

Table 1 contains the descriptive statistics of the volatility series. The means of the implied volatility series are higher than those of the realized volatility, both in level and in log (indicated by the letter L). Besides this, the difference between the series in level (implied volatility minus realized volatility) is lower than for the series in log.

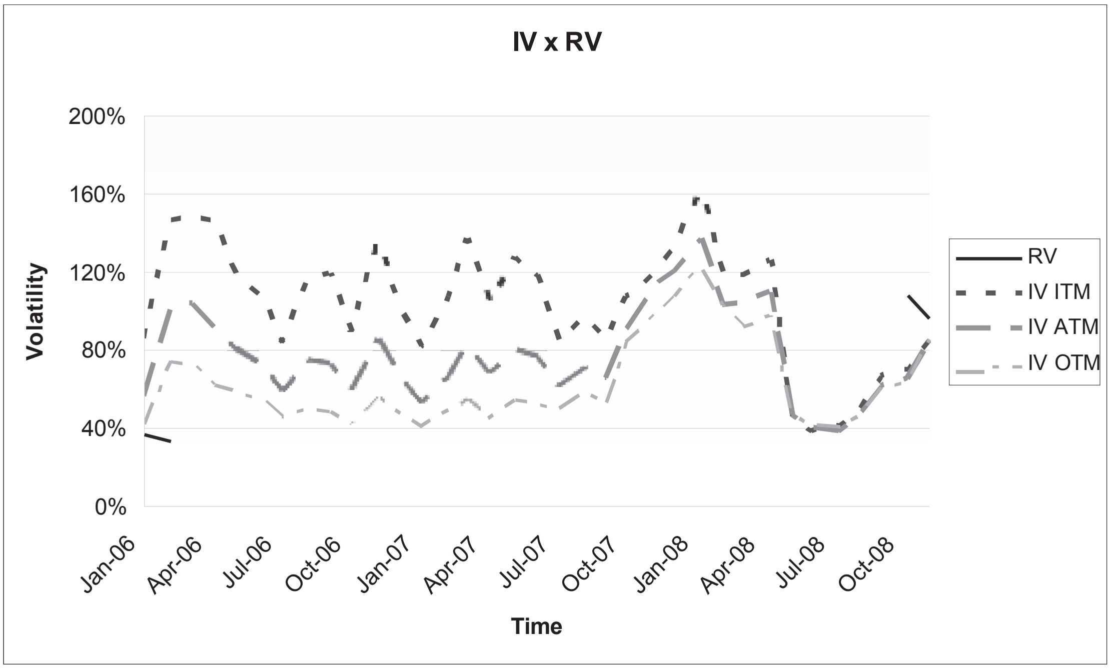

Figure 2 – Comparison of the volatilities series.
Nota: This graph compares the volatility series during the 35 sub-periods into which the sample was divided. Abbreviations: RV: realized volatilty; IV: implied volatility; ITM: in-the-money; ATM: at-the-money; OTM: out-of-the-money.
Source: elaborate by the authors

The maximum and minimum values of the ITM series are much higher than those of the other options, particularly the minimum value of the ITM IV, which is over four times the minimum of the RV. Note also that the standard deviation of the series in level is higher for the IV series, while in the series in log the realized volatilities have a larger standard deviation.

The series also show problems of asymmetry and kurtosis. All of them have positive asymmetry except the ITM LIV, which is negative, indicating all the series except the ITM LIV series have a long right-hand tail. With respect to kurtosis, the ATM IV, OTM IV, RV and OTM LIV series are leptokurtic (elongated in relation to the normal distribution).

The other series present platykurtic characteristics (flattened in relation to the normal distribution), meaning they are less affected by large oscillations. Also note that the values of the asymmetry and kurtosis statistics are similar in the ATM IV, OTM IV and RV series and very different in the ITM IV series.

We will show later that this fact is in line with the greater explanatory power found in the OTM options. All the volatility series are stationary, according to the results of the augmented Dickey-Fuller test (Dickey & Fuller, 1979), at 1% significance.

Table 1 – Descriptive statistics of the volatilities series.

Statistics	IV ITM	IV ATM	IV OTM	RV	LIV ITM	LIV ATM	LIV OTM	LRV
Mean	1.1412	0.8035	0.6054	0.3251	0.1122	-0.2485	-0.5488	-1.1792
Median	1.1583	0.7470	0.5435	0.3042	0.1469	-0.2917	-0.6097	-1.1899
Maximum	1.6120	1.3681	1.2230	1.0844	0.4775	0.3135	0.2013	-0.4094
Minimum	0.8090	0.5366	0.4032	0.1620	-0.2120	-0.6225	-0.9084	-1.8204
SD	0.2316	0.2100	0.2116	0.1167	0.2044	0.2434	0.2965	0.3345
Asymmetry	0.2346	1.0516	1.6220	1.1874	-0.0505	0.5728	1.1302	0.3169
Kurtosis	2.0580	3.5236	4.7752	4.2834	1.9218	2.6419	3.4345	2.8904

Nota: This table presents the descriptive statistics of the volatility series. Abbreviations: IV: implied volatility; ITM: in-the-money; ATM: at-the-money; OTM: out-of-the-money; RV: realized volatility; LIV: log of the implied volatility; LRV: log of the realized volatility; SD: standard deviation.
Source: elaborate by the authors

To analyze the dynamic properties of the series, we estimated ARIMA (p,d,q) models, in the following form:

$$\Phi(B)\left(\Delta^{d} x_t - \mu\right) = \Theta(B)\varepsilon_t ,$$

Where x_t represents a series in log of the volatilities (LRV, ITM LIV, ATM LIV and OTM LIV), the parameter μ is the mean, ε_t is a white noise term, Φ and Θ are polynomials of order p and q in B, the lag operator, defined by $Bx_t = x_{t-1}$, and $\Delta = 1 - B$ is the first difference operator.

Table 2 presents the results of fitting the ARIMA model. Note that the non-integrated series fit the volatilities better than the integrated series. With the exception of the LRV series, the coefficients of the ARIMA (1,1,1) model are all non-significant at 5%.

Analysis of the Box-Pierce statistic (Box & Pierce, 1970) with 12 lags (Q_{12}) and the Akaike (AIC) and Bayesian (BIC) information criteria[v] showed that the implied volatility series are better adjusted by an AR(2) process, while the realized volatility is better adjusted by an ARMA(1,1) process.

Table 2 – ARIMA (*p,d,q*) model for implied and realized volatilities.

Model	μ	ϕ_1	ϕ_2	θ_1	Q_{12}	AIC	BIC
ITM LIV							
ARMA(1,1)	-0.01	**0.56**		0.23	6.47	0.35	0.48
AR(1)	-0.01	**0.69**			6.26	0.30	0.39
AR(2)	-0.02	**0.75**	-0.07		8.59	0.28	0.42
ARIMA(1,1,1)	-0.02	0.09		-0.23	11.21	0.44	0.57
ATM LIV							
ARMA(1,1)	**-0.22**	0.22		**0.70**	11.56	-0.04	0.10
AR(1)	-0.10	**0.64**			10.22	0.00	0.09
AR(2)	**-0.11**	**0.80**	-0.17		11.31	-0.08	0.05
ARIMA(1,1,1)	-0.01	0.04		-0.06	9.99	0.13	0.27
OTM LIV							
ARMA(1,1)	-0.19	**0.59**		0.25	12.47	-0.11	0.03
AR(1)	-0.13	**0.72**			12.13	-0.14	-0.05
AR(2)	-0.12	**0.83**	-0.08		15.86	-0.20	-0.07
ARIMA(1,1,1)	0.00	-0.05		0.02	17.73	-0.08	0.05
LRV							
ARMA(1,1)	**0.17**	**1.12**		**-1.00**	14.72	0.60	0.73
AR(1)	-0.27	**0.72**			13.87	0.79	0.88
AR(2)	-0.11	**0.56**	0.29		7.65	0.82	0.95
ARIMA(1,1,1)	**0.02**	**0.45**		**-0.97**	12.29	0.67	0.81

Nota: This table presents the estimates of the ARIMA model for the logs of volatility series (LRV, ITM LIV, ATM LIV and OTM LIV). Values in boldface indicate that the parameters are significant at 5%. The AIC and BIC columns represent the Akaike and Bayesian information criteria, while Q_{12} is the Box-Pierce statistic. Abbreviations: ITM: in-the-money; ATM: at-the-money; OTM: out-of-the-money; LIV: log of the implied volatility; LRV: log of the realized volatility.
Source: elaborate by the authors

4. EMPIRICAL RESULTS

In this section we first examine whether implied volatility provides information about future volatility. Then we test whether the lagged variables, both for HV and IV, have explanatory power about RV. We should point out that other variables can also influence RV. However, the scope of this paper is to verify whether there is any correlation between IV and RV.

The information content of implied volatility is typically assessed in the literature by estimating a regression of the form:

$$RV_t = \alpha + \beta IV_t + \varepsilon_t, \qquad (1)$$

Where RV is the realized volatility and IV is the implied volatility.

From this equation, we can draw some conclusions. First of all, if implied volatility has some predictive power about the realized volatility, then β must be statistically different from zero. Second, if implied volatility is an unbiased forecast of realized volatility, we should find that $\alpha = 0$ and $\beta = 1$. Finally, if implied volatility is efficient, then the residuals must be white noise and not correlated with any variable in the market's information set.

Table 3 presents the results of Equation 1. The implied volatility of the ITM options, both in level and in log, are not relevant to explain the realized volatility. The p-values of the coefficients of IV are very high (0.7650 in level and 0.9728 in log), so the null hypothesis that they are statistically different from zero is rejected.

Table 3 – Results of Equation 1.

	Variable	Coefficient	Standard error	P-value	Adjusted R^2	DW
Level ITM	C	0.2889	0.1220	**0.0267**	-0.0393	1.2990
	ITM_IV	0.0317	0.1048	0.7650		
Level ATM	C	0.1338	0.0868	0.1370	0.1479	1.8065
	ATM_IV	0.2380	0.1047	**0.0327**		
Level OTM	C	0.1156	0.0573	0.0554	0.3674	2.1264
	OTM_IV	0.3460	0.0895	**0.0008**		
Log ITM	C	-1.1779	0.0783	0.0000	-0.0344	1.2515
	ITM_LIV	-0.0118	0.3412	0.9728		
Log ATM	C	-1.0690	0.0934	**0.0000**	0.0651	1.6440
	ATM_LIV	0.4434	0.2713	0.1157		
Log OTM	C	-0.8433	0.1227	**0.0000**	0.2637	1.9768
	OTM_LIV	0.6121	0.1976	**0.0051**		

Nota: This table shows the results of Equation 1. The coefficients in boldface are statistically significant at 5%. Abbreviations: IV: implied volatility; ITM: in-the-money; ATM: at-the-money; OTM: out-of-the-money; RV: realized volatility; LIV: log of the implied volatility; DW: Durbin-Watson statistic.
Source: elaborate by the authors

Regarding the IV of the ATM options, in the regression in level, the intercept is not statistically significant at 5% and β is 0.2380. This value of β shows that the implied volatility was nearly four times greater than the realized volatility in the period observed. Hence, the implied volatility of the ATM options contains information about the future volatility, although this is highly biased (β much smaller than 1). This large difference between implied and realized volatilities can be explained by the high risk premium demanded by investors, showing that in risky environments, options tend to have a high extrinsic value[vi]. With the equation in log, the implied volatility coefficient of the ATM options is no longer statistically significant (p-value of 0.1157). In other words, according to the equation, we cannot guarantee that variations in the implied volatility of ATM options have any explanatory power over their realized volatility.

In the regression in level for the OTM options, the value of β is 0.3460, which is statistically significant at 5%. This indicates that the implied volatility represents, on average for the study period, around one-third of the realized volatility. It is noteworthy that the constant, α, is statistically zero. The OTM regressions in log show that a variation of one percentage point in implied volatility on average causes an increase of 0.6121% in realized volatility. Furthermore, the intercept is statistically significant and negative, showing a possible bias in the estimation.

To sum up, the results indicate that the implied volatility of both ATM and OTM options contains information about future volatility. However, the estimators are biased, because $\alpha = 0$ and $\beta = 1$ do not hold in any situation. These conclusions are confirmed by analyzing the adjusted R^2 coefficient. Adjusted R^2 denotes the relation between the variation explained by the multiple regression equation and the total variation of the dependent variable considering the number of variables. It is higher for the OTM options and lower for the ITM ones, both in level and in log. The estimators generated by the implied volatilities of the OTM and ATM options (the latter only in level) are efficient, given that their Durbin-Watson statistics are not statistically different from two[vii].

To ascertain whether implied or historical volatility has more explanatory power of future volatility, we first estimated the following regression:

$$RV_t = \alpha + \beta RV_{t\text{-}1} + \varepsilon_t \qquad (2)$$

Table 4 – Results of Equation 2

	Variable	Coefficient	Standard error	P-value	Adjusted R^2	DW
Level RV(-1)	C	-0.7669	0.2668	**0.0088**	0.1070	2.0059
	RV(-1)	0.3485	0.2147	0.1187		
Log RV(-1)	C	0.2155	0.0733	**0.0076**	0.0989	2.0337
	LRV(-1)	0.3419	0.2199	0.1343		

Nota: This table shows the results of Equation 2. The coefficients in boldface are statistically significant. Abbreviations: RV (-1): lag of the realized volatility; LRV(-1): log of the lag of the realized volatility DW: Durbin-Watson statistic.
Source: elaborate by the authors

Table 4 shows the results of Equation 2. It can be seen that the past volatility appears not to contain information about future volatility. The coefficients of the variables RV(-1) and LRV(-1) are not statistically significant at 5%, indicating that past volatility does not help in predicting future volatility.

This result is plausible. Even if today's volatility is strongly influenced by yesterday's, in a highly volatile market like Brazil's, the monthly volatilities are not good predictors of future volatility.

By adding the IV series in Equation 2, we have:

$$RV_t = \alpha + \beta_1 IV_t + \beta_2 RV_{t\text{-}1} + \varepsilon_t \quad (3)$$

Table 5 – Results of Equation 3

	Variable	Coefficient	Standard error	P-value	Adjusted R^2	DW
Level ITM	C	0.2227	0.1312	0.1043	0.0991	2.0374
	ITM_IV	-0.0076	0.1133	0.9471		
	RV(-1)	0.3468	0.2366	0.1576		
Level ATM	C	0.1064	0.0939	0.2696	0.2134	1.9274
	ATM_IV	0.2446	0.1400	0.0951		
	RV(-1)	0.0569	0.2661	0.8328		
Level OTM	C	0.1240	0.0634	0.0640	0.4543	1.7908
	OTM_IV	0.4488	0.1214	**0.0013**		
	RV(-1)	-0.2402	0.2355	0.3193		
Log ITM	C	-0.7036	0.3056	**0.0317**	0.1156	2.0261
	ITM_LIV	-0.1672	0.3690	0.6550		
	LRV(-1)	0.3840	0.2322	0.1131		
Log ATM	C	-0.8791	0.2844	**0.0055**	0.1557	1.9844
	ATM_LIV	0.3876	0.3520	0.2833		
	LRV(-1)	0.1797	0.2629	0.5017		
Log OTM	C	-0.9023	0.2383	**0.0011**	0.3481	1.9400
	OTM_LIV	0.7418	0.2661	**0.0110**		
	LRV(-1)	-0.0938	0.2458	0.7065		

Nota: This table shows the results of Equation 3. The statistically significant coefficients are in boldface. Abbreviations: IV: implied volatility; ITM: in-the-money; ATM: at-the-money; OTM: out-of-the-money; RV: realized volatility; LIV: log of the implied volatility RV(-1): lag of the realized volatility; LRV(-1): log of the lag of the realized volatility; DW: Durbin-Watson statistic.
Source: elaborate by the authors

Table 5 shows the results of Equation 3. It can be seen that none of the past volatility equations contain any type of information about future volatility, because the coefficients are statistically nil in all of them.

Just as in the article by CP, the use of lagged HV increased the coefficients of the implied volatility series. In the case of the OTM series, the coefficients continued being significant and the adjusted R^2 increased, indicating greater predictive power of that equation.

Finally, it is important to verify whether the lagged variables of IV help explain future volatility. To do this, we estimated the following regression:

$$RV_t = \alpha + \beta IV_{t\text{-}1} + \varepsilon_t \quad (4)$$

Table 6 shows the results of Equation 4. It can be seen that only OTM IV series in level lagged by one period has explanatory power about the realized volatility. In the other cases, the coefficients are not statistically significant at 5%.

Table 6 – Results of Equation 4

	Variable	Coefficient	Standard error	P-value	Adjusted R^2	DW
Level ITM	C	0.3984	0.1340	**0.0070**	0.0144	1.2654
	ITM_IV(-1)	-0.0668	0.1174	0.5753		
Level ATM	C	0.1821	0.1098	0.1113	0.0733	1.5753
	ATM_IV(-1)	0.1812	0.1374	0.2006		
Level OTM	C	0.1472	0.0801	0.0797	0.1926	2.0419
	OTM_IV(-1)	0.3041	0.1327	**0.0319**		
Log ITM	C	-1.1642	0.0790	**0.0000**	0.0168	1.2713
	ITM_LIV(-1)	-0.2275	0.3707	0.5457		
Log ATM	C	-1.0929	0.1121	**0.0000**	0.0485	1.4845
	ATM_LIV(-1)	0.3434	0.3242	0.3010		
Log OTM	C	-0.8796	0.1631	**0.0000**	0.1603	1.9268
	OTM_LIV(-1)	0.5287	0.2579	0.0525		

Nota:This table shows the results of Equation 4. The statistically significant coefficients are in boldface. Abbreviations: IV: implied volatility; ITM: *in-the-money*; ATM: *at-the-money*; OTM: *out-the-money*; RV: realized volatility; (-1): lag; LIV: log of the implied volatility; RV(-1): lag of the realized volatility; LRV(-1): log of the lag of the realized volatility; DW: Durbin-Watson statistic.
Source: elaborate by the authors

5. COMPARISON WITH PREVIOUS WORKS

Previous articles have expressed various opinions about the information content of implied volatility in relation to future volatility. Day & Lewis (1992), studying options on the S&P 100 index, and Lamoureux & Lastrapes (1993), working with the ten most liquid stocks in the S&P 100, obtained different results than those found here.

They concluded that historical volatility is a better forecast of future volatility than is implied volatility, which was highly biased and inefficient. Canina & Figlewski (1993), studying options on the S&P 100 index before 1987, also obtained an opposite result to that found in this work: only historical volatility contained information on future volatility.

In contrast, in the Brazilian case historical, volatility appears to have no predictive power about future volatility when dealing with non-overlapping monthly data.

More recently, Christensen & Prabhala (1998) repeated the study of options on the S&P 100 index, utilizing ATM options with monthly maturities, without overlapping data and based on a much larger sample (139 observations). The results found for the period after the 1987 NYSE crash are very similar to those reported here in Equation 1 in relation to the estimator of the IV of OTM options.

In our case, implied volatility outperformed past volatility in forecasting future volatility, being, however, more biased than found by CP. There are other differences with respect to historical volatility and the intercept. CP showed that historical volatility has predictive power, but the intercept was not statistically significant, results opposite to those found in this study. Our results agree with those of Gwilym & Buckle (1999), according to which implied volatility contains more information about the realized volatility than does historical volatility when using call options with one-month maturity.

The main difference between our study and the others cited above is that the IV series of the options with highest predictive power were the OTM instead of ATM options, as in the previous works.

In Brazil, Gabe & Portugal (2004) concluded that the volatility estimated by different statistical models of the GARCH family produce better forecast of future volatility than implied volatility. It should be noted, however, that GARCH models specify a different regression relation than those studied here.

For the comparison between IV and HV to be more balanced, it is necessary to analyze the information content of volatility through identical models. However, GARCH models consider not only the past volatility, but also other information, such as past returns.

Moreover, the database of Gabe & Portugal (2004) was smaller than ours and contained overlapping data. The reason is that their focus was different than ours. While they aimed to forecast (so that balance mattered little), we focus on analysis of the information content.

6. CONCLUSION

In this paper we compared the explanatory power of implied and historical volatilities in relation to future volatility, using data from the Petrobras call options market. Our results indicate that the implied volatility of OTM options has a higher correlation with future volatility than does historical volatility.

The weak explanatory power of ATM and ITM options reveals either that the volatility risk premium of these options is high or the market has inefficiencies. We found no evidence in the period studied (January 2006 to December 2008) that past volatility has any correlation with future volatility in monthly terms. Therefore, the use of a forward-looking approach, as is the case of the implied volatilities of options, appears to be an alternative to the use of past data.

REFERENCES

BLACK, F.; SCHOLES, M. The Pricing of options and corporate liabilities. **Journal of Political Economy**, 1973.

BOX, G. P.; PIERCE, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, 65, p. 1509 – 1526, 1970.

CANINA, L.; FIGLEWSKI, S. The Informational content of implied volatility. **Review of Financial Studies**, v. 6, p. 659-681, 1993.

CHRISTENSEN, B. J.; PRABHALA, N. R. The Relation between implied and realized volatility. **Journal of Financial Economics**, v. 50, p. 125-150, 1998.

DAY, T.; LEWIS, C. Stock market volatility and the information content of stock index options. **Journal of Econometrics**, 52, 1992

DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, 74, p. 427–431.

FRENCH, K.; SCHWERT, G.W.; STAMBAUGH, R. Expected stock returns and volatility. **Journal of Financial Economics**, 19, p. 3 - 30, 1987.

GABE, J.; PORTUGAL, M. S. Volatilidade implícita versus volatilidade estatística: um exercício utilizando opções e ações da Telemar S.A. **Revista Brasileira de Finanças**, v. 2, n. 1, p. 47-73, 2004.

GWILYM, O.A.; BUCKLE, M. Volatility forecasting in the framework of the option expiry circle. **The European Journal of Finance**, 5, 73–94, 1999.

HULL, John C. **Options, futures and other derivatives**. Upper Saddle River: Prentice-Hall Inc., 1997.

LAMOUREUX, C. G.; LASTRAPES, W. Forecasting stock return variance: towards understanding stochastic implied volatility. **Review of Financial Studies**, 6. 1993.

TABAK, B. M; CHANG, E. J. Are implied volatilities more informative? The Brazilian real exchange rate case. **Applied Financial Economics**, 17, 569-576, 2006.

---

[i] Thus, we observed the precautions indicated by CP (no sample overlap, fixed-maturity options).

[ii] In this study we used the CDI (overnight) rate as the risk-free rate, expressed in a base year of 252 business days.

[iii] The options on the S&P 100 index used in CP are also American options, but are not protected for dividends. This introduces a bias into the calculation of the implied volatility via the BS model that does

not happen in the case of Petrobras options. Therefore, correction of the errors of the variables used in the regressions, as presented in Section 4 of the CP article, is not necessary here.

[iv] Other more usual moneyness criteria, such as that defined by the option's Delta, require constructing a volatility surface for choosing the ITM, ATM and OTM options. Because of the low number of available options series in our database, the interpolation error on the volatility surface could impair the results.

[v] The information criteria are defined as: AIC = - $2(l/T) + 2(k/T)$ and BIC = $-2(l/T) + k\log(T)/T$, where l is the value of the likelihood function, k is the number of parameters and T is the sample size.

[vi] An option's value can be divided into an intrinsic and an extrinsic value, or its value in time. The intrinsic value is the difference between the spot price of the underlying asset and the strike price of the call option. The extrinsic price reflects the opportunity cost and market expectations.

[vii] The Durbin-Watson statistic (DW) measures the serial correlation of the residuals. A result near two indicates there is no first-order serial correlation.