The Implied Volatility Bias: A No-Arbitrage Approach for Short-Dated Options*

João Pedro Ruas

ISCTE - IUL Business School

José Dias Curto

BRU-UNIDE, Lisbon University Institute (ISCTE-IUL)

João Pedro Vidal Nunes†

BRU-UNIDE and ISCTE - IUL Business School

January, 2012

*Financial support by FCT's grant number PTDC/EGE-ECO/099255/2008 is gratefully acknowledged.

†Corresponding author. ISCTE - IUL Business School, Complexo INDEG/ISCTE, Av. Prof. Aníbal Bettencourt, 1600-189 Lisboa, Portugal. Tel: +351 21 790 3932, e-mail: joao.nunes@iscte.pt

The Implied Volatility Bias: A No-Arbitrage Approach for Short-Dated Options

Abstract

Under the efficient market hypothesis, the implied volatility associated to an option price should be the best possible forecast of the future realized volatility of the underlying asset. However, a vast number of studies in the financial literature have found that implied volatility is a biased estimator of the future realized volatility. In this paper we introduce a new *ex-post* measure of implied volatility (the *Market Neutral Volatility*) that is simply based on no-arbitrage and preference-free arguments. We demonstrate that the implied volatility retrieved from options prices is an *unbiased* estimator of the *Market Neutral Volatility* for options close to maturity. This result is empirically tested and validated using S&P 500 monthly option prices for the period between January 2005 and June 2011.

Keywords: Market Neutral Volatility; Implied Volatility; Realized Volatility; Straddles

JEL Classification: C53; G13.

1. Introduction

The implied volatility (hereafter IV) associated to an option price can be seen as the market's forecast of the future volatility for the underlying asset over the option remaining life. Under the rational expectations assumption that the market uses all the available information and under the market's efficiency hypothesis, the IV should be the best possible forecast of the future volatility given the currently available information. If this is not true, i.e. if an option fails to embody optimal forecasts of the underlying asset's future volatility, the option pricing theory implies that a profitable trading strategy is available and its implementation should push the option price towards the best possible forecast of future volatility.

Most of the studies in the literature find evidence that IV forecasts of future volatility are informative, superior to other measures of volatility, but upward-biased. The bias problem is not new and has been documented and studied for decades—see, for instance, Canina and Figlewski (1993), Day and Lewis (1992), Jorion (1995), and Lamoureux and Lastrapes (1993). For a more in depth comparison of 93 studies that deal with the forecast of future realized volatility see Poon and Granger (2003).

Poon and Granger (2003) argue that the existence of an implied volatility bias carries a severe problem, since it should only persist if it is difficult to perform arbitrage trades that may remove the mispricing. In the case of stocks indexes, both options and futures contracts are highly liquid which should mitigate this problem, but the bias still persists in the majority of studies. Doran and Ronn (2008) argue that the extra premium resulting from the bias explains why options traders tend to be short, since they can write options at the higher implied volatility and reap the extra premium by delta-hedging their price exposures.

Due to this problem, several explanations for the implied volatility bias problem have been put forward in the literature. A first explanation is focused on model risk: options implied volatilities are model dependent. Given an observed option market value, its implied volatility is estimated by using a pricing model, and the one commonly used is the Black and Scholes (1973) model. However, the underlying model may be miss-specified. Facing this problem, the investors may require an extra premium for incurring in model risk.

A second explanation is related with econometric issues, namely the variables' measurement error and the empirical methodologies used to estimate the realized volatility. Related with the first issue, Christensen and Prabhala (1998) argue that implied volatility is biased because of errors in the measurement of variables, such as the (neglected) American-style feature of the option contract or the use of overlapping observations. Ederington and Guan (1999, 2002) study the implied volatility of futures options on the S&P 500 index and also find that the apparent bias and inefficiency of implied volatility detected in early studies are due to variables' measurement errors. Concerning the second econometric issue, Poteshman (2000) finds that half of the bias in the S&P 500 futures options market can be removed when actual volatility is estimated with a more efficient volatility estimator based on intra-day 5-minutes returns.

A third explanation is the existence of a volatility risk premium. This explanation is based on the notion that volatility risk is non-diversifiable, which leads investors to require a premium for bearing this risk—see, among many others, Chernov (2007), Poteshman (2000), Doran and Ronn (2005, 2008), Coval and Shumway (2001), and Bakshi and Kapadia (2003).[1]

A fourth explanation arises from general equilibrium models. Guidolin and Timmermann (2003) propose an equilibrium model for asset prices under Bayesian learning. The introduction of Bayesian learning generates biases similar to those observed in option prices, offering a new economic explanation for Black and Scholes (1973) biases. Garcia *et al.* (2003) propose an utility based option pricing model with stochastic volatility and jump features. In this model, preferences matter for option pricing and this can help to explain the bias in IV as well as other empirical features of IV (e.g asymmetric IV curves).

In restricting ourselves to no-arbitrage arguments within the Black and Scholes (1973) framework, we introduce the concept of Market Neutral Volatility (hereafter MNV) and derive a theoretical model that captures the relationship between this new concept and the IV. In this setting, we only address the bias problem; i.e. within our framework, and in opposition with the above mentioned general equilibrium models, no explanation is provided for other IV features (such as the IV smile). Nevertheless, we demonstrate that IV is directly forecasting the Market Neutral Volatility, being an unbiased estimator of this new

[1] This is the case, for instance, of the Heston (1993) model that allows volatility to be stochastic and its market price of risk to be non-zero.

no-arbitrage measure. Hence, we are able to provide a simpler explanation for the IV bias associated to option contracts that are close to maturity, which is only based on no-arbitrage arguments and is, therefore, valid for any system of preferences.

The outline of the paper is as follows. In the next section we discuss the MNV concept and we derive the theoretical model that captures the relationship between this new concept and the IV. In section 3, our theoretical result is tested on the S&P 500 options market. The final section summarizes our conclusions.

2. The market neutral volatility model

This section presents the theoretical framework proposed to address the relationship between MNV and IV.

2.1. Model assumptions

In what follows, uncertainty will be represented by a complete probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where $\mathbb{Q}$ denotes the martingale probability measure obtained when the money market account is taken as the numéraire of the economy. We also assume that the market trades a stock index, whose market price at the close of day t will be represented by S_t. Hence, the probability space is endowed with a right continuous and complete filtration $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$, where $\mathcal{F}_t = \sigma(S_u, 0 \leq u \leq t)$ is the sigma-algebra generated by the observation of the trading prices for the stock index.

Additionally, the market also trades futures contracts and European-style futures options (with "stock-style margining") on the same stock index, and for the same expiry date. Hereafter, $F_{t,T}$ will denote the time-t price of a futures contract on the stock index S, and for delivery at time T $(\geq t)$. Similarly, $c_t(F_{t,T}, K, T; \sigma)$ and $p_t(F_{t,T}, K, T; \sigma)$ will represent the time-t price of a European-style call or put option, respectively, on the futures price $F_{t,T}$, with strike $K \in \mathbb{R}_+$, with maturity at time T, and computed through the estimate σ of the realized volatility of the underlying futures contract during the time-interval $[t, T]$. As usual, both option prices are taken to be adapted to the filtration $\mathbb{F}$. Note also that the well known put-call parity for European-style futures options implies that at-the-money (ATM) forward calls and puts must have the same price (for any volatility level σ, and for any expiry date T):

$$c_t(F_{t,T}, F_{t,T}, T; \sigma) = p_t(F_{t,T}, F_{t,T}, T; \sigma). \tag{1}$$

The expiry date of the future and option contracts will be taken as the next trading day, i.e. $T - t = \frac{1}{\Delta}$, where Δ is an annualizing factor. Moreover, the market is assumed to be *complete* in the sense that tradable European-style and ATM-forward calls and puts (on the stock index futures contract) with a time-to-maturity of just one trading day are always available before the close of the market at each day t. This assumption is not restrictive as it seems: in today's financial markets, daily and even intra-day options for stock indexes, currencies and commodities are readily available.[2]

To define the novel concept of Market Neutral Volatility, it is useful to introduce first the notion of a (long) *ATM-forward straddle*. This is simply a portfolio of options contracts designed to bet on a high volatility scenario, which involves a long position on both ATM-forward call and put options with the same expiry date (time $t + \frac{1}{\Delta}$), and on the same underlying futures contract. The time-t value of such portfolio will be denoted by $g_t(\sigma)$, and the absence of arbitrage opportunities requires that

$$g_t(\sigma) = c_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t + \frac{1}{\Delta}; \sigma\right) + p_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t + \frac{1}{\Delta}; \sigma\right). \tag{2}$$

2.2. Market neutral volatility

The time-t Market Neutral Volatility ($MNV_{t,\frac{1}{\Delta}}$), for the time to maturity $\frac{1}{\Delta}$, will be defined in terms of the time-t fair value of an ATM-forward straddle. For this purpose, next proposition assumes, for the moment, that $\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right|$ is $\mathcal{F}_t$-measurable.

Proposition 1 *Under the strong assumption that the information about the absolute variation between the close of day t and the close of day $t + \frac{1}{\Delta}$ is available at the end of day t,*

[2] Daily options on a wide range of stock indexes, currencies and commodities can be traded, for example, on the IG Index (a regulated Financial Services Authority company). Moreover, daily options on the AEX index are also available on NYSE Euronext.

then the time-t no-arbitrage price of an ATM-forward straddle would be equal to

$$g_t^e = e^{-\frac{r}{\Delta}} \left| F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}} \right|, \tag{3}$$

for any volatility level σ, and where r represents the constant risk-free (and continuously compounded) interest rate.

Proof. Using equation (2), and since Geman *et al.* (1995, Corollary 2) imply that the relative price of any financial asset (in terms of the money market account numéraire) must be a $\mathbb{Q}$-martingale, then

$$\begin{aligned} g_t(\sigma) &= e^{-\frac{r}{\Delta}} \mathbb{E}_{\mathbb{Q}} \left[\left(F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}} \right)^+ \middle| \mathcal{F}_t \right] \\ &\quad + e^{-\frac{r}{\Delta}} \mathbb{E}_{\mathbb{Q}} \left[\left(F_{t,t+\frac{1}{\Delta}} - F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} \right)^+ \middle| \mathcal{F}_t \right], \end{aligned}$$

where $(x)^+ = \max(x, 0)$, for any $x \in \mathbb{R}$. Furthermore, imposing that $\left| F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}} \right|$ is $\mathcal{F}_t$-measurable, the previous equation can be rewritten as

$$g_t^e = e^{-\frac{r}{\Delta}} \left[\left(F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}} \right)^+ + \left(F_{t,t+\frac{1}{\Delta}} - F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} \right)^+ \right],$$

which is exactly equivalent to equation (3). ■

In practice, at the end of day t no agent is endowed with the information about the absolute variation between the close of day t and the close of day $t + \frac{1}{\Delta}$. Therefore, equation (3) is, in practice, simply the time $t + \frac{1}{\Delta}$ discounted payoff generated by the ATM-forward straddle strategy. Nevertheless, such *no-arbitrage* price of the ATM-forward straddle enables the definition of the Market Neutral Volatility concept.

Definition 1 *The time-t Market Neutral Volatility ($MNV_{t,\frac{1}{\Delta}}$), for the time to maturity $\frac{1}{\Delta}$, is the time-t implied volatility level that would price an ATM-forward straddle at its no-arbitrage value, i.e.*

$$\begin{aligned} g_t\left(MNV_{t,\frac{1}{\Delta}} \right) &= g_t^e \\ &= e^{-\frac{r}{\Delta}} \left| F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}} \right|. \end{aligned} \tag{4}$$

The MNV is the time-t implied volatility that ensures, for any time-t bet on a straddle strategy, that the payout will be zero. In practice, no agent at time-t knows the value of $MNV_{t,\frac{1}{\Delta}}$. Nevertheless, we will argue that the IV retrieved from options prices at time t is forecasting this *no-arbitrage* implied volatility $MNV_{t,\frac{1}{\Delta}}$.

2.3. The relationship between IV and MNV

Following, for instance, Jackwerth and Rubinstein (1996), it is well known that the usual geometric Brownian motion assumption is not able to accommodate the negative skewness and the high kurtosis that are typically implicit in the empirical asset return distributions. Nevertheless, it is common market practice to compute implied volatilities through the Black and Scholes (1973) model, for options on the spot, or through the Black (1976) model, for futures options, respectively. For a European-style call option on a futures contract this procedure corresponds to finding the root $IV_{t,\tau}$ of the following non-linear equation:

$$c_t\left(F_{t,T}, K, T; IV_{t,\tau}\right) = e^{-r\tau}\left[F_{t,T}\Phi\left(d_+\right) - K\Phi\left(d_-\right)\right], \tag{5}$$

where

$$d_\pm := \frac{\ln\left(\frac{F_{t,T}}{K}\right) \pm \frac{IV_{t,\tau}^2}{2}\tau}{IV_{t,\tau}\sqrt{\tau}},$$

$\tau := T - t$ is the time-to-maturity (in years) of the futures and option contracts, $\Phi\left(\cdot\right)$ is the distribution function of the standard univariate normal probability law, and the left-hand side of equation (5) is read as the market price of the European-style call.

Even though no closed-form solution (in $IV_{t,\tau}$) exists for equation (5), the following result arises as a straightforward extension of Brenner and Subrahmanyam (1988, Equation 3) to European-style futures options, and provides an accurate approximation for the time-t implied volatility associated to an ATM-forward call option:

$$IV_{t,\tau} \approx \sqrt{\frac{2\pi}{\tau}}\frac{c_t\left(F_{t,T}, F_{t,T}, T; IV_{t,\tau}\right)}{F_{t,T}}e^{r\tau}. \tag{6}$$

This (first order) proportional relationship between volatility and the option price is obtained through a power series expansion of the distribution function $\Phi(\cdot)$—along the lines of Abramowitz and Stegun (1972, Equation 26.2.10)—and is only valid for the ATM-forward (i.e. $K = F_{t,T}$) option contracts to be considered in the forthcoming empirical analysis. Moreover, the accuracy of such first order approximation is also higher for smaller values of $d_{\pm}$, which are implicit in the one day ahead option contracts to be priced.

Based on approximation (6), next proposition offers our main theoretical result: The time-t Market Neutral Volatility is approximately equal to 125% of the one day ahead annualized absolute return.

Proposition 2 *Under the non-restrictive assumptions described in Subsection 2.1,*

$$\begin{aligned} MNV_{t,\frac{1}{\Delta}} &\approx \frac{\sqrt{2\pi}}{2}\sqrt{\Delta}\frac{\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right|}{F_{t,t+\frac{1}{\Delta}}} \\ &\cong 1.25\sqrt{\Delta}\frac{\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right|}{F_{t,t+\frac{1}{\Delta}}}. \end{aligned} \tag{7}$$

Proof. Replacing $IV_{t,\tau}$ by $MNV_{t,\frac{1}{\Delta}}$ in approximation (6), then

$$MNV_{t,\frac{1}{\Delta}} \approx \sqrt{2\pi\Delta}\frac{c_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t+\frac{1}{\Delta}; MNV_{t,\frac{1}{\Delta}}\right)}{F_{t,t+\frac{1}{\Delta}}}e^{\frac{r}{\Delta}}. \tag{8}$$

The call price on the right-hand side of equation (8) can be rewritten in terms of the underlying futures price through the combination of equations (1) and (2),

$$\begin{aligned} g_t\left(MNV_{t,\frac{1}{\Delta}}\right) &= c_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t+\frac{1}{\Delta}; MNV_{t,\frac{1}{\Delta}}\right) \\ &\quad + c_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t+\frac{1}{\Delta}; MNV_{t,\frac{1}{\Delta}}\right), \end{aligned}$$

and using Definition 1:

$$e^{-\frac{r}{\Delta}}\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right| = 2c_t\left(F_{t,t+\frac{1}{\Delta}}, F_{t,t+\frac{1}{\Delta}}, t+\frac{1}{\Delta}; MNV_{t,\frac{1}{\Delta}}\right). \tag{9}$$

Combining equations (8) and (9), it follows that

$$MNV_{t,\frac{1}{\Delta}} \approx \frac{\sqrt{2\pi}}{2}\sqrt{\Delta}\frac{e^{-\frac{r}{\Delta}}\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right|}{F_{t,t+\frac{1}{\Delta}}}e^{\frac{r}{\Delta}}, \tag{10}$$

which is equivalent to equation (7). ■

Note that if we define the time $t + \frac{1}{\Delta}$ annualized volatility as

$$\sigma^r_{t,\frac{1}{\Delta}} := \frac{\left|F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} - F_{t,t+\frac{1}{\Delta}}\right|}{F_{t,t+\frac{1}{\Delta}}}\sqrt{\Delta}, \tag{11}$$

then

$$MNV_{t,\frac{1}{\Delta}} \approx 1.25\sigma^r_{t,\frac{1}{\Delta}},$$

i.e. the time-t Market Neutral Volatility can be understood as corresponding to approximately 125% of the one day ahead annualized realized volatility. The definition of the realized market volatility through equation (11) arises from the notion that the payoff of a European-style option at expiration only depends on the difference between initial and terminal underlying asset prices, and not on the prices observed in the market between those two dates. This way of measuring volatility is also inspired on former works found in the literature. Figlewski (1997) argues that taking deviations from zero instead of the sample mean typically increases volatility forecasting accuracy. Ding *et al.* (1993) suggest measuring volatility directly from absolute returns. Andersen and Bollerslev (1998) refer to the daily realized volatility as the sum of intra-day squared returns.

Having derived, in Proposition 2, the theoretical value of the MNV, and being the MNV an ex-post *no-arbitrage* measure of implied volatility, we argue that the IV retrieved from options prices should be forecasting this ex-post value. This hypothesis is tested in the following section.

Table 1: Average Daily Volume of all SPX and S&P 500 Futures Options per year

	2007	2008	2009	2010
SPX	629,819	707,688	614,562	695,601
S&P Futures	76,913	64,737	41,073	42,369

SPX and S&P Futures stand for options on the SPX and the S&P 500 futures, respectively. Data retrieved from CBOE for SPX and from CME for S&P 500 futures.

3. Empirical analysis

3.1. Data

This paper uses data from the closing values of the S&P 500 index (SPX) and S&P 500 monthly options from January 2005 to June 2011. We choose the S&P index options market and not the S&P futures options market to mitigate possible liquidity-related biases. The average daily volume of the options on the SPX is more than 10 times the average daily volume of the options on the S&P futures, as can be seen in Table 1. The SPX options market is widely used in the literature and, as pointed by Rubinstein (1994), in this market the conditions required by the Black and Scholes (1973) model are well approximated.

The IV and the index price data are retrieved from the Thinkorswim trading platform, a division of TD Ameritrade. The implied volatility is retrieved from the nearest ATM straddle (nearest ATM call, plus nearest ATM put). As pointed by Hentschel (2003), by using ATM straddles we avoid the additional errors in IV usually associated to options that are far from the ATM level.

The SPX options are European-style contracts and may be exercised only on the last business day before expiration. The trading in SPX options will ordinarily cease on the business day (usually a Thursday) preceding the day on which the exercise-settlement value is calculated.

Since our model only considers one day ahead forecasts, for each month we get the data for the SPX and SPX 500 options when just one business day is left to trade the option (usually a Wednesday). At the close of that day we retrieve the implied volatility of the nearest ATM straddle and the value of the SPX. At the close of the following day we retrieve

Table 2: Average Daily Volumes and Open Interest per year of the S&P 500 options used

	Total	2005	2006	2007	2008	2009	2010	2011
	Panel A: Average Daily Volume							
Calls	7,898	4,378	9,621	14,999	4,022	6,964	8,388	6,520
Puts	7,238	3,668	9,145	17,292	6,033	2,939	4,349	6,648
	Panel B: Average Open Interest							
Calls	27,323	24,768	25,148	60,481	11,298	20,493	27,142	21,639
Puts	23,988	21,787	17,746	55,437	20,293	14,017	15,547	21,821

The calls and puts considered in Table 2 are selected monthly, from January 2005 to June 2011, and correspond to the nearest ATM contracts with just one business day left to maturity. Data retrieved from Thinkorswim.

the value of the SPX and compute the MNV value from equation (7).[3]

All option contracts used satisfy the usual lower / upper no-arbitrage boundaries, and have trading volumes that are higher than the cutoff point used by Gonçalves and Guidolin (2006) of 100 traded contracts per day. Table 2 reports the average daily volume (in Panel A) and the average open interest (in Panel B) of the contracts used.

3.2. Descriptive statistics

Our data includes 154 observations for the S&P 500 index that allow us to compute 77 values for the MNV—since we need an observation from time t and another one from time $t + \frac{1}{\Delta}$ to calculate the MNV at time t. The same number of 77 monthly observations is also collected for the IV associated to the S&P 500 monthly options prices. One observation is excluded from the sample (November 20, 2008). The maximum value of 160.72% attained for the MNV in this day is only comparable with the values obtained in the 87 crash and, therefore, this extreme outlier is removed—similarly, Bakshi and Kapadia (2003) exclude all options with IV above 100% and Guo and Whitelaw (2006) exclude the trading days between October 17, 1987, and October 28, 1987.

Panel A of Table 3 provides the descriptive statistics for both the implied volatility and

[3]For this purpose, we simply take $F_{t+\frac{1}{\Delta},t+\frac{1}{\Delta}} = S_{t+\frac{1}{\Delta}}$ (by obvious no-arbitrage arguments) and $F_{t,t+\frac{1}{\Delta}} \approx S_t$. The latter identity is justified because the effect of interest rates and dividends is negligible for the single trading day to maturity option contracts considered.

Table 3: Descriptive Statistics for IV and MNV

Variable	*Obs*	*Mean*	*Std.Dev.*	*Min*	*Max*	*t*-sig	*l*-sig
	Panel A: Options with just one trading day to maturity						
IV	77	20.48%	12.77%	8.90%	85.17%	0.6882	4.06e-6
MNV	77	19.31%	22.07%	0.20%	103.78%		
	Panel B: Options with 15 days to maturity						
IV	77	18.75%	10.29%	8.82%	62.43%	0.9462	0.0110
MNV	77	18.62%	13.81%	5.43%	95.16%		

IV and MNV stand for Implied Volatility and Market Neutral Volatility, respectively. Obs is the number of observations, Mean is the arithmetic mean, Std. Dev. is the standard deviation, while Min and Max represent the minimum and the maximum value for each series. t-sig is the probability associated with the two-independent samples t-test where in the null it is assumed the equality of IV and MNV means. l-sig is the probability associated with the Levene's test for the equality of IV and MNV variances.

the MNV. The means' difference between the IV and the MNV is not statistically significant, which can be justified by the high probability (in the seventh column) associated with the two independent samples t-test, where in the null it is assumed that the population means are equal. However, the low probability associated to the Levene's test result (in the last column of Table 3 shows that the dispersion (measured by the variance) is statistically higher in the MNV distribution when compared to the one of the implied volatility law. Therefore, the implied volatility seems to be a smoothed expectation of the MNV.

The range between the maximum and the minimum observations is considerably large, which is consistent with the relative low values of volatility prevailing between 2005-2007 and the extremely high values observed in 2008.

3.3. Testing for unbiasedness

To test the hypothesis that implied volatility is an *unbiased* predictor of MNV we estimated the following two linear regressions, that are in line with the common practice in the literature—see for instance Christensen and Prabhala (1998) when studying the relationship between IV and realized volatility:

$$MNV_{t,\tau} = \alpha + \beta IV_{t,\tau} + \varepsilon_{t+\tau}, \tag{12}$$

and

$$MNV_{t,\tau} = \alpha + \beta IV_{t,\tau} + \gamma MNV_{t^-,\tau} + \varepsilon_{t+\tau}, \tag{13}$$

where $MNV_{t,\tau}$ denotes the MNV at period t and for a time to maturity of τ years (i.e. calculated using the time $t+\tau$ and the time t closing values of the SPX, for $\tau > 0$), $IV_{t,\tau}$ denotes the implied volatility at period t and for maturity after τ years, $MNV_{t^-,\tau}$ denotes the MNV calculated in the previous month, and $\varepsilon_{t+\tau}$ is a zero mean homoskedastic error that is uncorrelated with the explanatory variables.

Three hypotheses are typically tested in the literature concerning this regression. First, IV is *informative* only if the estimate for β is significantly greater than zero. Second, IV is an *unbiased* estimator if the joint hypotheses of α equal to zero and β equal to 1 is not rejected. Third, IV is informational efficient if no other variable (e.g. past values) is statistically significant, that is, if γ is equal to zero.

For implied volatility to be an unbiased estimator of the MNV, we must have the case that $\alpha = 0$ and $\beta = 1$. Thus, the two restrictions must be tested simultaneously and the decision must point for the nonrejection of the null, where it is assumed that $\alpha = 0$ and $\beta = 1$ in regressions (12) and (13). For implied volatility to be informational efficient the decision must point for the nonrejection of the null, where it is assumed that $\gamma = 0$ in regression (13).

Panel A of Table 4 reports the results from the two regressions, where α and β are the estimated coefficients from the regressions (12) and (13), γ is the estimated coefficient from regression (13), and $\tau = \frac{1}{\Delta}$. The t-test values are presented in brackets and F gives the probability associated with the F-test of *unbiasedness*, where in the null it is assumed that $H_0 : \alpha = 0$ and $\beta = 1$.

Based on the t-tests presented in Panel A of Table 4, and for a 5% significance level, only the estimates for β are significantly different from zero in both regressions; the estimates for α in both regressions are not statistically significant and the estimate for γ in regression (13) is also not statistically significant. The joint hypothesis of $\alpha = 0$ and $\beta = 1$ is not rejected in both regressions due to the probability associated with the F-test shown in the last column of Table 4. This result allows us to confirm the hypothesis that implied volatility is directly forecasting the one day ahead Market Neutral Volatility.

Table 4: MNV Regression over Implied Volatility

	α	β	γ	R^2	Obs	F
	Panel A: Options with just one trading day to maturity					
Regression (12)	-4.005	1.138	-	0.426	77	0.548
	(-1.100)	(7.534)				
Regression (13)	-4.083	1.278	-0.150	0.437	76	0.292
	(-1.116)	(7.255)	(-1.478)			
	Panel B: Options with 15 days to maturity					
Regression (12)	-1.928	1.095	-	0.661	77	0.561
	(-1.008)	(12.24)				
Regression (13)	-2.469	1.279	-0.158	0.664	76	0.247
	(-1.257)	(7.595)	(-1.271)			

MNV stands for Market Neutral Volatility. α, β and γ are the estimates for the coefficients of equations (12) and (13); the t-statistics are presented in brackets below the coefficients' estimates. R^2 is the coefficient of determination from equations (12) and (13). Obs is the number of observations and F is the probability associated with the F-test of unbiasedness where in the null it is assumed that $H_0 : \alpha = 0$ and $\beta = 1$.

3.4. Checking for robustness

To access the robustness of our results we rerun the same analysis done in the previous section, but now considering options that have 15 days to maturity (descriptive statistics are presented in Panel B of Table 4). This analysis is outside the main purpose of this paper, and the results should be taken with care since to perform this analysis one has to assume that it would be possible to trade a series of consecutive forward start options with one day to maturity each. We carry this analysis to access if possible measurement errors, when dealing with options with just one day to maturity, have a significant impact on the obtained results.

Panel B of Table 4 reports the estimation results from the regressions (12) and (13), but now considering the 15 day period until maturity. The estimates for α and γ are still not significantly different from zero, and the joint hypothesis of $\alpha = 0$ and $\beta = 1$ is again not rejected in both regressions. These results allow us to confirm the robustness of the conclusions drawn from Panel A of Table 4.

4. Conclusions

The finding, in numerous studies, that the IV is a biased estimator of future realized volatility has puzzled researchers for decades since it contradicts the efficient market hypothesis.

Several explanations for this anomaly have been put forward in the literature, namely the existence of model risk, the use of inadequate econometric methodologies, the presence of a volatility risk premium, and idiosyncrasies concerning investor's attitudes and preferences.

By introducing the new concept of Market Neutral Volatility, we follow a different path to access the IV bias problem. Based on the benchmark Black (1976) model, we show that the MNV is an ex-post *no-arbitrage* measure of IV for the options market. We argue that the IV is directly forecasting the MNV, and we are able to explain the upward bias of IV for option contracts that are close to maturity by simply using no-arbitrage and preference-free arguments.

The relationship between MNV and IV was tested using actual market data retrieved from the S&P 500 options market, and the results obtained support our hypothesis that IV is an *unbiased* forecast of the MNV, which corresponds to approximately 125% of the one day ahead annualized realized volatility.

References

Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions* 1972 (Dover Publications: New York).

Andersen, T. and Bollerslev, T., Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *International Economic Review*, 1998, **39**, 885–905.

Bakshi, G. and Kapadia, N., Delta-Hedged Gains and the Negative Market Volatility Risk Premium. *Review of Financial Studies*, 2003, **16**, 527–566.

Black, F., The Pricing of Commodity Contracts. *Journal of Financial Economics*, 1976, **3**, 167–179.

Black, F. and Scholes, M., The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 1973, **81**, 637–654.

Brenner, M. and Subrahmanyam, M., A Simple Formula to Compute the Implied Standard Deviation. *Financial Analysts Journal*, 1988, **44**, 80–83.

Canina, L. and Figlewski, S., The Informational Content of Implied Volatility. *Review of Financial Studies*, 1993, **6**, 659–681.

Chernov, M., On the Role of Risk Premia in Volatility Forecasting. *Journal of Business and Economic Statistics*, 2007, **25**, 411–426.

Christensen, B. and Prabhala, N., The Relation between Implied and Realized Volatility. *Journal of Financial Economics*, 1998, **50**, 125–150.

Coval, J. and Shumway, T., Expected Option Returns. *Journal of Finance*, 2001, **56**, 983–1009.

Day, T. and Lewis, C., Stock Market Volatility and the Information Content of Stock Index Options. *Journal of Econometrics*, 1992, **52**, 267–287.

Ding, Z., Granger, C. and Engle, R., A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, 1993, **1**, 83–106.

Doran, J. and Ronn, E., The Bias in Black-Scholes/Black Implied Volatility: An Analysis of Equity and Energy Markets. *Review of Derivatives Research*, 2005, **8**, 177–198.

Doran, J. and Ronn, E., Computing the Market Price of Volatility Risk in the Energy Commodity Markets. *Journal of Banking and Finance*, 2008, **32**, 2541–2552.

Ederington, L. and Guan, W., The Information Frown in Option Prices. *Working Paper, University of Oklahoma*, 1999.

Ederington, L. and Guan, W., Is Implied Volatility an Informationally Efficient and Effective Predictor of Future Volatility?. *Journal of Risk*, 2002, **4**, 29–46.

Figlewski, S., Forecasting Volatility. *Financial Markets, Institutions and Instruments*, 1997, **6**, 1–88.

Garcia, R., Luger, R. and Renault, E., Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables. *Journal of Econometrics*, 2003, **116**, 49–83.

Geman, H., Karoui, N.E. and Rochet, J.C., Changes of Numéraire, Changes of Probability Measure and Option Pricing. *Journal of Applied Probability*, 1995, **32**, 443–458.

Gonçalves, S. and Guidolin, M., Predictable Dynamics in the S&P 500 Index Options Implied Volatility Surface. *Journal of Business*, 2006, **79**, 1591–1635.

Guidolin, M. and Timmermann, A., Option Prices under Bayesian Learning: Implied Volatility Dynamics and Predictive Densities. *Journal of Economic Dynamics and Control*, 2003, **27**, 717–769.

Guo, H. and Whitelaw, R., Uncovering the Risk-Return Relation in the Stock Market. *Journal of Finance*, 2006, **61**, 1433–1463.

Hentschel, L., Errors in Implied Volatility Estimation. *Journal of Financial and Quantitative Analysis*, 2003, **38**, 779–810.

Heston, S., A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 1993, **6**, 327–343.

Jackwerth, J. and Rubinstein, M., Recovering Probability Distributions from Option Prices. *Journal of Finance*, 1996, **51**, 1611–1631.

Jorion, P., Volatility in the Foreign Exchange Market. *Journal of Finance*, 1995, **50**, 507–528.

Lamoureux, C. and Lastrapes, W., Forescasting Stock-Return Variance: Towards an Understanding of Stochastic Implied Volatilities. *The Review of Financial Studies*, 1993, **6**, 293–326.

Poon, S. and Granger, C., Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature*, 2003, **41**, 478–539.

Poteshman, A., Forecasting Future Volatility from Option Prices. *Working Paper, University of Illinois at Urbana-Champaign*, 2000.

Rubinstein, M., Implied Binomial Trees. *Journal of Finance*, 1994, **69**, 771–818.