Cabedo, J. David; Moya Clemente, Ismael
Implied volatility as a predictor: the case of the Ibex 35 future contract
Estudios de Economía Aplicada, vol. 23, núm. 1, abril, 2005, pp. 67-78
Asociación Internacional de Economía Aplicada
Valladolid, España

Available in: http://www.redalyc.org/articulo.oa?id=30123104
Implied volatility as a predictor: the case of the Ibex 35 future contract

CABEDO J. DAVID(*) Y MOYA CLEMENTE ISMAEL(**)
(*)Departamento de Finanzas y Contabilidad. Universidad Jaume I.
(*)Campus Riu Sec; 12071 Castellón - Teléfono: 964 38 71 50. E-mail: cabedo@cofin.uji.es
(**)Camino de la Vera s/n; 46071 Valencia - Teléfono: 963 87 70 32 50. E-mail: imoya@esp.upv.es

ABSTRACT

In this paper we analyse the implied volatility (IMV) as a predictor of the realised volatility (REV) of the underlying asset. For this aim we use daily data from Ibex 35 future and option market (years 2000 to 2003). We obtain a series of daily IMV from the option prices and two additional series, one for historical and another for realised volatility, from the daily quotes of Ibex 35 future contracts. We test three hypotheses regarding the forecasting power of IMV on REV: IMV is an unbiased estimate of the future REV; IMV has more explanatory power than the historical volatility (HIV) when forecasting future REV; and HIV does not provide more information on REV than that provided by IMV. We also introduce a GARCH model for these tests. Our results indicate that, when forecasting, historical volatility does not provide additional information to that provided by implied volatility. However, IMV tends to overestimate realised volatility values.

Keywords: Volatility forecasting; Implied volatility; Ibex 35 future option market.

La volatilidad implícita como herramienta de predicción: una aplicación al contrato de futuro sobre Ibex 35

RESUMEN

En el presente trabajo se analiza la volatilidad implícita (IMV) como predictor de la volatilidad real (REV) del activo subyacente. Con este fin se utilizan datos diarios del mercado de futuros y opciones sobre Ibex 35 (años 2000 a 2003). A partir de las cotizaciones de los contratos de opción se elabora una serie diaria de IMV y a partir de las cotizaciones de los contratos de futuro se elaboran sendas series, una para la volatilidad real y otra para la histórica. En el trabajo se contrastan 3 hipótesis sobre el poder predictivo de la volatilidad implícita: Si IMV es un estimator insesgado de la futura REV; si IMV tiene un poder explicativo mayor que la volatilidad histórica (HIV) cuando se realizan predicciones sobre valores futuros de la REV; y si HIV proporciona información adicional a la aportada por la IMV cuando se realizan las mencionadas predicciones. Este contraste de hipótesis también se realiza introduciendo modelos GARCH. Los resultados muestran que, a la hora de predecir los valores de la volatilidad real, HIV no proporciona información adicional a la aportada por IMV. No obstante, las predicciones realizadas utilizando esta última volatilidad tienden a sobrestimar el valor de la real.

Palabras clave: Predicción de la volatilidad; volatilidad implícita; Mercado de futuros y opciones sobre Ibex 35

Clasificación JEL: G00,G23.
1. INTRODUCTION

The forecasting power of implied volatility on realised volatility has been a topic analysed in literature since seventies. Anyway, there has not been any generally accepted agreement on the magnitude of this power up to now.

From the option quotes, provided by market, we can get the market level of the volatility, by using appropriate option valuation models. This market level is known as implied volatility (IMV). If we demonstrate this IMV provides good forecasts for realised volatility (REV), implied volatility will become fundamental when designing investment and hedge strategies1.

Historical volatility (HIV) is the main alternative to IMV we can use when forecasting volatility. Therefore, an important question is determining the higher or lower forecasting power of IMV against HIV.

The first papers on the topic found that IMV fits the changes in REV better than HIV does. In this line we can mention some works on stock options, like the ones of Latane and Rendleman (1976), Chiras and Manaster (1978), Schmalensee and Trippi (1978) and Beckers (1981).

Later, some researchers have found that IMV does not provide an adequate forecasts for volatility, in the period till the option expiration date. In this line we can mention some works on stock options, like the ones of Canina and Figlewski (1993) do not find evidence about the relationship between IMV and REV in the remaining period till option maturity. However they find relationship between HIV and REV. Other authors, like Day and Lewis (1992) and Lamoureux and Lastrapes (1993) find forecasting power in the IMV. However they find a higher forecasting power in the HIV and the volatility calculated on a GARCH model.

Nevertheless other papers get opposite conclusions. Christiansen and Prabhala (1988) find the IMV as a good forecasting of REV.

Anyway, the main drawback of the above mentioned papers lies on the data they use: the quotes from option markets are used for estimating IMV, whilst for the estimation of REV the data come from other markets (those corresponding to the underlying assets). As data are not taken from the same market, quotes are not synchronised. Furthermore, transaction costs are different among different markets2. These facts can bias the conclusions achieved.

The objective of this paper is to analyse the forecasting power of implied volatility and compare it with the one provided by the HIV. For overcoming the above mentioned drawbacks we use the quotes of options and future contracts on the Ibex 35 index. Both contracts (options and futures) are simultaneously traded within the same market. We

---

1 Volatility forecasts are specially useful for risk management purposes. Specifically they can be used, for instance, in the value at risk calculation procedure.
2 See Fleming et al. (1996).
have structured the rest of the paper in the following way: section 2 shows the path followed for getting and calculating the data necessary for the analysis. Section 3 collects the hypothesis to test and the used methodology. In section 4 we show and analyse the results achieved. And the last section summarises the conclusions of the paper.

2. DATA

In this paper we use the quotes of future and option contract on Ibex 35, for the years 2000 - 2003. The Ibex 35 is a stock index, which collects the 35 most traded stocks of the Spanish market. The underlying asset for the future contract is the index whilst for option contract the underlying asset is the future contract.

We have estimated the IMV by using the Black and Scholes (1973) model. We have applied it on a daily series of option quotes. We have adjusted this series by implementing the procedure used by Szakmary et al. (2003). From an original set made up of all the daily settlement quotes, we have chosen those corresponding to the options on the future contract which maturity date is closest to the trading day. This guarantees the best fit between IMV and REV when we estimate the latter from the futures quotes. Anyway we have not used options with an expiration date lower than ten trading days, in order to overcome possible estimation errors. These errors stem when options with a small period to maturity are used. Moreover, this gives uniformity when estimating volatility.

Another topic we have considered when selecting the options for forming the daily series of prices is their money position. As stated by Rubinstein (1985) the IMV is lower for those options which are at the money (ATM). On the other hand, this volatility is higher for those out of the money (OTM) and in the money (ITM) options: the higher the distance from the ATM option, the higher the volatility. When returns distribution does not fit a normal one, the IMV calculated from ATM options is the most accurate one. Beckers (1981) and Canina and Figlewski (1993) demonstrate how IMV calculated from ATM options provides a better forecast on REV than that provided by the IMV calculated from options far from ATM position. That is why we use the closest to the ATM point option for every day.

For estimating the REV and the HIV we have used the settlement prices of the futures contracts for the analysis period (years 2000 to 2003).

3. METHODOLOGY

The main aim of this paper is analysing the forecasting power of IMV against HIV. For this aim we, firstly, study the relationship between IMV and REV and between HIV and REV. We define two regression models for this aim (models 1 and 2):
When estimating the regression coefficients if we obtain a statistically significant value for $\beta_1^{(1)}$, we will be able to conclude that IMV can be used for REV forecasting. We can interpret coefficient $\beta_1^{(2)}$ in a similar way when analysing the relationship between HIV and REV. Furthermore, we compare the R squared coefficients for both models (1 and 2) in order to determine which one of the independent variables (IMV and HIV) has a greater explanation power on REV (dependent variable).

Secondly, we define model (3). In this model we analyse if HIV provides additional information to that provided by IMV, when forecasting REV.

$$REV_t = \beta_0^{(3)} + \beta_1^{(3)} \cdot IMV_t + \beta_2^{(3)} \cdot HIV_t + \epsilon_t^{(3)}$$  \[3\]

If $\beta_2^{(3)}$ becomes statistically significant, we can conclude that HIV provides additional information. In the other case, when we can not reject the tested null hypothesis $\beta_2^{(3)} = 0$, we can conclude that HIV does not provide additional information to that provided by IMV, when forecasting REV.

For estimating models (1), (2) and (3) we have calculated the series corresponding to the dependent variable (REV) by using the procedure proposed by Canina and Figlewski (1993):

$$REV = \left( \frac{260}{T_M - 1} \sum_{c=1}^{T_M} \left( R_t - \bar{R} \right)^2 \right)^{0.5}$$  \[4\]

where $T_M$ denotes the number of trading days till the option expiration date; $R_t$ is the daily return of the underlying asset, estimated through the logarithmic approximation. And $\bar{R}$ denotes the average value of this return.

On the other hand, we have used the procedure described in the section of data of this paper for calculating the IMV series. Finally, the HIV has been calculated by estimating moving averages for different window lengths: 30, 60 and 90 days.

Additionally, we have estimated model (3) by using the series provided by an autoregressive conditionally heteroskedasticity (ARCH) model as HIV. We test if HIV, estimated through an ARCH model, provides additional information to that provided by IMV, when forecasting REV.

4. RESULTS

Table 1 shows the values estimated for the model (1) coefficients. As shown, the IMV coefficient is statistically significant at a 1% significance level. This indicates
that the IMV can be used for REV forecasting. Furthermore, as the value of the estimated coefficient is lower than 1, we can conclude the IMV overestimates REV values\(^3\).

**Table 1**: Model (1) estimation

| Parameter \( \beta_0^{(1)} \) | Estimate \(-0.0627917\) | Standard error \(5.04835\times10^{-3}\) | T value \(-12.4381\) | Prob level \(0.00000\) |
| Parameter \( \beta_1^{(1)} \) | Estimate \(0.486627\) | Standard error \(0.0182028\) | T value \(26.7337\) | Prob level \(0.00000\) |

R-squared = 41.92 percent

Regression model: \( REV_t = \beta_0^{(1)} + \beta_1^{(1)} \times IMV_t + \epsilon_t^{(1)} \)

\( IMV_t \): Implied volatility at time \( t \).

\( REV_t \): Realised volatility at time \( t \).

In tables 2 to 4 we show the results we have obtained estimating model (2) for different window lengths: 30, 60 and 90 days. As shown, the HIV coefficients are statistically significant at a 1\% significance level for all the lengths. This indicates that HIV can be used for REV forecasting purposes. Moreover, as the estimated coefficients are lower than 1, we can conclude that HIV also overestimates REV values.

**Table 2**: Model (2) estimation. Window length: 30 days

| Parameter \( \beta_0^{(2)} \) | Estimate \(-0.0275018\) | Standard error \(4.68426\times10^{-3}\) | T value \(-5.87111\) | Prob level \(0.00000\) |
| Parameter \( \beta_1^{(2)} \) | Estimate \(0.363879\) | Standard error \(0.0169664\) | T value \(21.447\) | Prob level \(0.00000\) |

R-squared = 32.37 percent

Regression model: \( REV_t = \beta_0^{(2)} + \beta_1^{(2)} \times HIV_t + \epsilon_t^{(2)} \)

\( HIV_t \): Historical volatility at time \( t \).

\( REV_t \): Realised volatility at time \( t \).

**Table 3**: Model (2) estimation. Window length: 60 days

| Parameter \( \beta_0^{(2)} \) | Estimate \(-0.0131808\) | Standard error \(6.11578\times10^{-3}\) | T value \(-2.15521\) | Prob level \(0.03140\) |
| Parameter \( \beta_1^{(2)} \) | Estimate \(0.30648\) | Standard error \(0.0220078\) | T value \(13.9259\) | Prob level \(0.00000\) |

R-squared = 17.24 percent

Regression model: \( REV_t = \beta_0^{(2)} + \beta_1^{(2)} \times HIV_t + \epsilon_t^{(2)} \)

\( HIV_t \): Historical volatility at time \( t \).

\( REV_t \): Realised volatility at time \( t \).

\(^3\) Biased forecasts are consistent with the results of Jorion (1995).
Another question to analyse in the models estimated up to this point is their explanation power. We analyse it through R squared coefficient values. As shown, the greatest R squared value equals 33.18% when we use HIV as the independent variable (see tables 2 to 4). It corresponds to a 90 days window length. However, this R squared value is lower than that provided by model (1): 41.92% (see table 1). This corresponds to the model where IMV is the independent variable. Therefore we can conclude that IMV has a greater REV forecasting power against HIV. IMV forecasting power is greater than HIV forecasting power for any of the windows length.

The following step of our analysis framework involves testing the information provided by HIV, additional to that provided by IMV, when forecasting REV values. We have estimated the model (3) coefficients and we show their values in table (5). As shown, the model (3) explanation power is the same than the model (1) explanation power. Their R squared coefficients are very close: 45% and 42% respectively. Furthermore, the model (3) IMV coefficient is statistically significant at a 1% significance level. However, the HIV coefficient in model (3) is not statistically significant at a 5% significance level. We can conclude that HIV does not provide additional information to that provided by IMV when forecasting REV.

Table 4: Model (2) estimation. Window length: 90 days

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>T value</th>
<th>Prob level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0^{(2)}$</td>
<td>-0.0283729</td>
<td>4.82189E-3</td>
<td>-5.8842</td>
<td>.00000</td>
</tr>
<tr>
<td>$\beta_1^{(2)}$</td>
<td>0.366432</td>
<td>0.0173239</td>
<td>21.1518</td>
<td>.00000</td>
</tr>
</tbody>
</table>

R-squared = 33.18 percent

Regression model: $REV_t = \beta_0^{(2)} + \beta_1^{(2)} \cdot HIV_t + \epsilon_t^{(2)}$

$HIV_t$: Historical volatility at time $t$.

$REV_t$: Realised volatility at time $t$.

Table 5: Model (3) estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>T value</th>
<th>Prob level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0^{(3)}$</td>
<td>-0.0654671</td>
<td>0.00510224</td>
<td>-12.8311</td>
<td>.0000</td>
</tr>
<tr>
<td>$\beta_1^{(3)}$</td>
<td>0.452519</td>
<td>0.0321659</td>
<td>14.0683</td>
<td>.0000</td>
</tr>
<tr>
<td>$\beta_2^{(3)}$</td>
<td>0.0508304</td>
<td>0.0273781</td>
<td>1.85661</td>
<td>.0634</td>
</tr>
</tbody>
</table>

R-squared = 45.10 percent

Regression model: $REV_t = \beta_0^{(3)} + \beta_1^{(3)} \cdot IMV_t + \beta_2^{(3)} \cdot HIV_t + \epsilon_t^{(3)}$

$IMV_t$: Implied volatility at time $t$.

$HIV_t$: Historical volatility at time $t$. Window length: 90 days.

$REV_t$: Realised volatility at time $t$. 
Obviously, these results depend on the HIV estimation procedure. We are assuming that IMV can be modelled through a moving average framework with different window lengths (30, 60 and 90 days). However, recent financial literature has been using a specific family of models when modelling volatility: autoregressive conditional heteroskedasticity models. Several works have shown that these models adequately fit the volatility of the main financial series. In this paper we use an ARCH model for calculating a series of HIV alternative to that calculated through a moving average framework.

Autoregressive conditional heteroskedasticity models (ARCH) were introduced by Engle (1982). They start out from the error distribution of a dynamic linear regression model (5):

\[ y_t = X_t \beta + \varepsilon_t \]  

where \( y_t \) represents the dependent or endogenous variable; \( X_t \) is the vector of explanatory exogenous variables of the model, amongst which may be included lagged values of the dependent variable; \( \beta \) is the parameter vector; and \( \varepsilon_t \) represents the error term.

The stochastic error distribution \( \varepsilon_t \), conditioned to the set of information available in the immediately preceding moment (\( \Psi_{t-1} \)) is fitted to a normal distribution with mean equal to 0 and variance equal to \( h_t \) (6).

\[ \varepsilon_t / \Psi_{t-1} \sim N(0, h_t) \]  

where the set of information available in moment \( t \) 1 gathers the value of the exogenous and lagged endogenous variables in one or more time periods (7):

\[ \Psi_{t-1} = \{ x_{t-1}, y_{t-1}, x_{t-2}, y_{t-2}, \ldots \} \]  

and where the variance, \( h_t \), is a linear function of the lagged squared errors (8):

\[ h_t = \alpha_0 + \alpha_1 \cdot \varepsilon^2_{t-1} + \ldots + \alpha_q \cdot \varepsilon^2_{t-q} \]  

\[ \alpha_0 > 0 \ \text{y} \ \alpha_i > 0 \ \forall \ i = 1, \ldots, q \]  

The order of the ARCH process is determined precisely by the number of lags considered in the errors when calculating the variance (\( q \)).

In the ARCH(\( q \)) model exposed, a major shock is characterized by a long deviation of \( y_t \) in relation to its conditioned mean (\( X_t \beta \)), or, put in a different way, for one \( \varepsilon_t \), of any sign, of a high magnitude. If we take into account that the conditioned variance of these errors (\( h_t \)) is an increasing function with respect to the magnitude of the lagged errors (squared), the errors of large (or small) magnitude tend to be followed by errors of large (or small) magnitude.

\[ \text{For this analysis we use the HIV values calculated by using a 90 days window length. This is the one which has provided the highest R squared coefficient.} \]
The empirical application of the ARCH models showed that on many occasions, the use of a high number of lags was necessary to specify the $h_t$ variance. In this vein, various solutions have been proposed for the problem of the excessive number of parameters to be estimated, which resulted in Bollerslev’s (1986) definitive proposal in the form of the generalization of ARCH models. The generalized ARCH model (GARCH) put forward by this author can be defined in similar terms to that of Engle (1982) (expressions (5) and (6)). However, the nuance introduced by Bollerslev is that the behaviour of the conditioned variance can be modelled, not only as a linear function of the lagged values of the squared errors, but also as a linear function of the lagged values of the conditioned variance (9):

$$h_t = \alpha_0 + \alpha_1 \cdot \varepsilon^2_{t-1} + \ldots + \alpha_q \cdot \varepsilon^2_{t-q} + \beta_1 \cdot h_{t-1} + \ldots + \beta_p \cdot h_{t-p}$$  \[9\]

The model defined by expressions (5), (6) and (9) is known as GARCH (p,q), where $p$ and $q$ are the lags considered for the conditioned variance and for the square of the past errors respectively.

The daily variations in the future prices suggest an autoregressive conditional heteroscedasticity (ARCH) behaviour for these variations since, as can be seen in figure 1, large variations are followed by large variations, and small variations by small variations.

**Figure 1:** Daily returns in future prices

---

5 The main developments in these models in the field of financial markets can be referred to in Bollerslev et al. (1992).
These indices must be confirmed through the use of suitable econometric tools. In order to confirm that it follows an ARCH model, two tests were carried out: the analysis of the autocorrelation of the squared series, as proposed by Enders (1996), and Lagrange multiplier test, as proposed by Engle (1982). The results obtained are detailed in Table 6.

This table shows that the Lagrange multiplier test rejects the null hypothesis tested, thus confirming that an autoregressive conditional heteroscedasticity scheme is adequate to model the behaviour of the series under study.

Within the family of autoregressive conditional heteroscedasticity models, the GARCH (1,1) is considered to be the most suitable to model the variance in financial series. This framework has been chosen to model the variance. The results of the estimation of the parameters of the model are detailed in table 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>T value</th>
<th>Prob level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.000014208</td>
<td>0.000004051</td>
<td>3.50723</td>
<td>.00045279</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.095498913</td>
<td>0.044190909</td>
<td>2.16105</td>
<td>.03069122</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.848999307</td>
<td>0.142992079</td>
<td>5.93739</td>
<td>.00000000</td>
</tr>
</tbody>
</table>

Regression model: $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$, $\epsilon_{t}$: Squared error lagged 1 period. (See model 9).

As shown in table 7, both regression coefficients (the one corresponding to the lagged variance and the one corresponding to the lagged squared error) are statistically significant. By using these coefficients and the series of daily returns in future prices.
we have estimated the daily variance series. We have introduced this series (called from here on HGV) in model (3), replacing HIV data. The resulting equation is (10):

\[ REV_t = \beta_0^{(10)} + \beta_1^{(10)} \cdot IMV_t + \beta_2^{(10)} \cdot HGV_t + \epsilon_t^{(10)} \]  \[10\]

Table 8 shows the results we have obtained when estimating model (10). As shown, the model forecasting power does not increase when we introduce in the model the volatility estimated through a GARCH framework. Regardless using HIV or HGV as independent variable, the R squared coefficient is almost the same: 42% (see tables 1 and 8). Furthermore, as shown in table 8, the HGV coefficient is not statistically significant. Therefore HGV does not provide additional information to that provided by IMV when forecasting REV.

Table 8: Model (10) estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>T value</th>
<th>Prob level</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0^{(10)} )</td>
<td>-0.062657</td>
<td>0.005052</td>
<td>-12.4028</td>
<td>.0000</td>
</tr>
<tr>
<td>( \beta_1^{(10)} )</td>
<td>0.468037</td>
<td>0.024263</td>
<td>19.2898</td>
<td>.0000</td>
</tr>
<tr>
<td>( \beta_2^{(10)} )</td>
<td>0.110708</td>
<td>0.094986</td>
<td>1.1655</td>
<td>.2441</td>
</tr>
</tbody>
</table>

R-squared = 41.89 percent
Regression model: \( REV_t = \beta_0^{(10)} + \beta_1^{(10)} \cdot IMV_t + \beta_2^{(10)} \cdot HGV_t + \epsilon_t^{(10)} \)

IMVt: Implied volatility at time t.
HGVt: Historical volatility at time t, estimated using a GARCH (1,1) model. Window length: 90 days.

Moreover, if we compare models (10) and (3) we can see that the explanation power is higher for the latter: the R squared equals 45% (see table 5) against a level of 42% for model (10) (see table 8).

Summarising we can conclude that the explanation power of the historical volatility when forecasting the REV is higher when we estimate this historical volatility through a moving average scheme than when we estimate it through an autoregressive conditional heteroskedasticity model.

5. CONCLUDING REMARKS

In this paper we have analysed the forecasting power of implied volatility on realised volatility when the underlying asset is the future contract on Ibex 35 index. The Ibex 35 is the most representative index of the Spanish stock market, and we have chosen it because the options on Ibex 35 have the future contracts as underlying assets and both, future and option, are simultaneously traded within the same market.
We have taken synchronised data for both markets (options and futures): those corresponding to the daily settlement prices. By this way we overcome the problems and bias stemming when the volatility is estimated by using series referred to different times.

The data set is compounded of the daily settlement prices for Ibex 35 future and option contracts in the years 2000 – 2003. With these data we have analysed the forecasting power of implied volatility and historical volatility on realised volatility. The results indicates that implied volatility can be used for forecasting realised volatility. However, implied volatility overestimates realised volatility values.

Our results also indicate that historical volatility can be used for forecasting realised volatility values. Nevertheless, the explanation power of historical volatility is lower than the one of implied volatility. Moreover, our results show that the historical volatility does not provide additional information to that provided by implied volatility, when forecasting realised volatility values.

Finally we have tested the forecasting power of historical volatility when we estimate it through an autoregressive conditional heteroskedasticity model. In accordance with the achieved results, even when we use a GARCH model for estimating the volatility series, this volatility does not provide additional information to that provided by implied volatility, when forecasting realised volatility values.

The behaviour of the implied volatility we have observed in this paper can be tested for other derivative contracts. In future research, this and other paths will be explored.

REFERENCES


