Journal of Banking & Finance 32 (2008) 2401–2411

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices ☆

Eirini Konstantinidi [a], George Skiadopoulos [a,b,*], Emilia Tzagkaraki [a]

[a] Department of Banking and Financial Management, University of Piraeus, Greece
[b] Financial Options Research Centre, Warwick Business School, University of Warwick, UK

ARTICLE INFO

Article history:
Received 12 February 2008
Accepted 24 February 2008
Available online 8 March 2008

JEL classification:
C53
G10
G12
G13
G14

Keywords:
Implied volatility
Implied volatility indices
Interval forecasts
Market efficiency
Predictability
Volatility derivatives

ABSTRACT

We address the question whether the evolution of implied volatility can be forecasted by studying a number of European and US implied volatility indices. Both point and interval forecasts are formed by alternative model specifications. The statistical and economic significance of these forecasts is examined. The latter is assessed by trading strategies in the recently inaugurated CBOE volatility futures markets. Predictable patterns are detected from a statistical point of view. However, these are not economically significant since no abnormal profits can be attained. Hence, the hypothesis that the volatility futures markets are efficient cannot be rejected.

1. Introduction

The question whether the dynamics of implied volatility per se can be forecasted is of paramount importance to both academics and practitioners.[1] Given that the implied volatility is a reparameterisation of the market option price, this question falls within the vast literature on the predictability of asset prices. In addition, implied volatility is often used as a measure of the market risk and hence it can be used in many asset pricing models. Therefore, understanding whether the variation in implied volatility is predictable can help us understand how expected returns change over time (see e.g., Corrado and Miller, 2006). From a practitioner's point of view, in the case where market participants can predict changes in implied volatility, then they can possibly form profitable option trading strategies. This will also have implications about the efficiency of the option markets.

Among others, David and Veronesi (2002) and Guidolin and Timmerman (2003) have developed asset pricing models that explain theoretically why implied volatility may change in a predictable fashion. The main idea is that investors' uncertainty about the economic fundamentals (e.g., dividends) affects implied volatility. This uncertainty evolves over time. In the case where it is persistent, the models induce predictable patterns in implied volatility.

The empirical evidence on the predictability of implied volatility is mixed. Dumas et al. (1998) and Gonçalves and Guidolin (2006) have investigated whether the dynamics of the S&P 500 implied volatilities across option strike prices and expiry dates (implied volatility surface) can be predicted over different time periods. The first study finds that the specifications under scrutiny are unstable over time for the purposes of option pricing and hedging. The second finds a statistically predictable pattern. This pattern cannot be

☆ This paper was originally submitted to Professor Giorgio Szegö on 2 May, 2007 and was revised once prior to submission through EES.

* Corresponding author. Address: Department of Banking and Financial Management, University of Piraeus, Greece. Tel.: +30 210 4142363; fax: +30 210 4142341.
E-mail addresses: ekonst@webmail.unipi.gr (E. Konstantinidi), gskiado@unipi.gr (G. Skiadopoulos), etzagk@unipi.gr (E. Tzagkaraki).

[1] This question is distinct from the question whether implied volatility can forecast the future realised volatility. There is also some distinct literature that has investigated the dynamics of implied volatilities across options with different strike prices and maturities by means of Principal Components Analysis solely for the purposes of option pricing and hedging (see e.g., Skiadopoulos et al., 1999).

0378-4266/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2008.02.003

exploited in an economically significant way since no abnormal profits can be obtained in the case where sufficiently high transaction costs are injected. There is also some literature that has explored whether the evolution of short-term at-the-money implied volatility, rather than the entire implied volatility surface, can be forecasted over time in various markets. Harvey and Whaley (1992), Guo (2000) and Brooks and Oozeer (2002) have addressed this question in the S&P 100, Philadelphia Stock Exchange currency, and LIFFE long gilt futures options markets, respectively. To this end, they used sets of economic variables as predictors. They found that changes in implied volatility are partially statistically predictable. However, their results are not economically significant just as in Gonçalves and Guidolin (2006). In a related study, Gemmill and Kamiyama (2000) have found that the changes in the implied volatilities of index options in a specific market are driven by the previous period changes of implied volatilities in another market (lagged spillover effects); the FTSE 100 (UK), NK225 (Japan), and S&P 500 (US) options are employed. However, the economic significance of their results is not examined. On the other hand, Goyal and Saretto (2007) have found that there is both a statistically and economically significant predictable pattern in the dynamics of implied volatility by using information from the cross-section of implied volatilities across various stock options.

This paper makes at least four contributions to the ongoing discussion about the predictability of implied volatility in equity markets. First, it employs an extensive data set of European and US implied volatility indices. Implied volatility indices have mushroomed over the last 15 years in the European and US markets and have particularly attractive characteristics for the purposes of our analysis as will be discussed below. In addition, the nature of the data set will shed light on whether the results may differ across countries and industry sectors. Second, both point and interval forecasts are formed and evaluated; the previously mentioned papers have only considered point forecasts. Interval forecasts are particularly useful for trading purposes (see e.g., Poon and Pope, 2000 for an application to option markets). Third, we perform a horse race among alternative model specifications so as to check the robustness of the obtained results; tests for predictability form a joint hypothesis test of the question under scrutiny and the assumed model. Finally, the economic significance of the statistical evidence is assessed by means of trading strategies in the newly introduced and fast growing Chicago Board Options Exchange (CBOE) volatility futures markets. The results will have implications about the efficiency of these markets that has not yet been investigated, as far as we are concerned.

To fix ideas, an implied volatility index tracks the implied volatility of a synthetic option that has constant time-to-maturity. The data on the implied volatility indices are the natural choice to study whether implied volatility is predictable. This is because the various methods to construct the index eliminate measurement errors in the calculated implied volatilities (see Hentschel, 2003), and take into account the traded option prices (or implied volatilities). Moreover, the possible presence of a predictable pattern in the evolution of implied volatility indices is of particular importance because these can be used in a number of applications. They serve as the underlying asset to implied volatility derivatives and they can be interpreted as variance and volatility swap rates.[2] Furthermore, the implied volatility index can also be used for Value-at-Risk purposes (Giot, 2005), to identify profitable opportunities in the stock market (see e.g., Banerjee et al., 2007), and to forecast the future market volatility (see e.g., Moraux et al., 1999; Giot, 2005; Becker et al., 2007 among others).

There are a number of papers that have studied the dynamics of implied volatility indices for the purposes of pricing implied volatility derivatives (see e.g., Dotsis et al., 2007, and the references therein). However, the question whether the dynamics of implied volatility indices can be predicted has received little attention. To the best of our knowledge, Aboura (2003), Ahoniemi (2006) and Fernandes et al. (2007) are the only related studies. All three studies differ in the time period they consider, focus on a limited number of indices and forecasting models, and provide only point forecasts. They all find that the evolution of implied volatility indices is statistically predictable. Only the second paper examines the economic significance of the obtained forecasts and finds that a trading strategy with the S&P 500 options cannot attain abnormal profits. Our research approach is more general; a range of European and US implied volatility indices is employed over a common time period, point and interval forecasts are formed by a number of alternative model specifications, and both their statistical and economic significance is assessed.

The remainder of the paper is structured as follows. In the next Section, the data sets are described. Section 3 presents the models to be used for forecasting. The in-sample performance of each model is examined in Section 4. The out-of-sample predictive performance of the models and the economic significance of the generated forecasts are evaluated in Sections 5 and 6, respectively. The last Section concludes.

2. The data set

Daily data on seven implied volatility indices, a set of economic variables (closing prices), and the CBOE volatility futures (settlement prices) are used. The various implied volatility indices have been listed on different dates. Hence, we consider the period from February 2, 2001 to September 28, 2007, so as to study the seven indices over a common time period. The subset from February 2, 2001 to March 17, 2005 will be used for the in-sample evaluation and the remaining data will be used for the out-of-sample one. This choice is dictated by the sample period (March 18, 2005 up to September 28, 2007) spanned by the volatility futures data; these will be used to assess the economic significance of the out-of-sample results.

In particular, four major American (VIX, VXO, VXN, VXD) and three European (VDAX-New, VCAC, and VSTOXX) implied volatility indices are examined. All indices but VXO are constructed by the VIX algorithm (see the CBOE VIX white paper, and Carr and Wu, 2006, for a description of the VXO algorithm).[3] VXO is constructed from the implied volatilities of options on the S&P 100. VIX, VXN, and VXD are based on the market prices of options on the S&P 500, Nasdaq 100, and Dow Jones Industrial Average (DJIA) index, respectively. VDAX-New, VCAC and VSTOXX are constructed from the market prices of options on DAX (Germany), CAC 40 (France) and the DJ EURO STOXX 50 index, respectively. The data for VDAX-New and VCAC are obtained from Bloomberg while for the other indices are obtained from the websites of the corresponding exchanges. VXO represents the implied volatility of an at-the-money synthetic option with constant time-to-maturity (thirty calendar days) at any point in time. We study the adjusted VXO, $\mathrm{VXOA} = \sqrt{\frac{22}{30}} \times \mathrm{VXO}$ rather than VXO itself. This adjustment allows interpreting VXOA as the volatility swap rate under general assumptions; the remaining indices represent the 30-day variance swap rate once they are squared (see Carr and Wu, 2006, and the references therein).

[2] A variance swap is actually a forward contract where the buyer (seller) receives the difference between the realised variance of the returns of a stated index and a fixed variance rate, termed variance swap rate, if the difference is positive (negative). The volatility swap is defined similarly; a volatility rather than a variance index serves as the underlying asset.

[3] The CBOE white paper can be retrieved from http://www.cboe.com/micro/vix/vixwhite.pdf.

Table 1
Summary statistics

	VIX	VXOA	VXN	VXD	VDAX_NEW	VCAC	VSTOXX
Panel A: Summary statistics for implied volatility indices (levels): February 2, 2001 to March 17, 2005							
Mean	0.22	0.21	0.37	0.21	0.29	0.26	0.28
Standard deviation	0.07	0.07	0.14	0.07	0.12	0.10	0.11
Skewness	0.75	0.67	0.30	0.69	0.82	1.04	0.91
Kurtosis	2.93	2.74	1.90	2.68	2.75	3.47	2.97
ρ_1	0.95*	0.96*	0.96*	0.96*	0.98*	0.98*	0.97*
ADF	−3.18	−2.91	−2.30	−2.34	−2.12	−2.14	−2.32
Panel B: Summary statistics for implied volatility indices (daily differences): February 2, 2001 to March 17, 2005							
Mean	−0.0004	−0.0004	−0.0008	−0.0003	−0.0002	−0.0001	−0.0002
Standard deviation	0.01	0.01	0.01	0.01	0.02	0.02	0.02
Skewness	0.05	0.17	−0.24	0.33	0.82	1.79	1.4
Kurtosis	5.29	6.02	6.02	6.92	10.47	16.44	18.09
ρ_1	0.01	−0.03	0.05	0.03	−0.02	−0.03	−0.03
ADF	−15.37*	−32.01*	−29.26*	−30.40*	−32.26*	−32.79*	−24.51*

	Levels			Daily differences		
	Shortest	2nd Shortest	3rd Shortest	Shortest	2nd Shortest	3rd Shortest
Panel C: Summary statistics for VIX futures: March 18, 2005 to September 28, 2007						
# Observations	630	608	590			
Mean	142.28	148.90	154.79	0.00	0.00	0.00
Standard deviation	29.30	23.86	20.00	0.04	0.03	0.02
Skewness	2.37	2.15	1.90	0.83	0.99	0.56
Kurtosis	9.23	8.01	6.92	14.25	8.30	8.21
ρ_1	0.99*	0.98*	0.95*	−0.01	−0.02	−0.06
Average volume	699.57	367.06	333.14			
(min–max)	(5–9,139)	(5–4,683)	(5–5,072)			
Panel D: Summary statistics for VXD futures: March 18, 2005 to September 28, 2007						
# Observations	490	370	290			
Mean	136.78	144.52	151.59	0.00	0.00	0.00
Standard deviation	30.24	26.58	22.93	0.05	0.03	0.03
Skewness	2.01	1.58	1.28	0.78	0.77	0.25
Kurtosis	7.12	5.04	3.92	11.32	7.99	8.22
ρ_1	0.91*	0.84*	0.79*	−0.03	0.03	−0.06
Average volume	63.75	38.83	38.4			
(min–max)	(5–328)	(5–308)	(5–336)			

Entries report the summary statistics of each one of the implied volatility indices in the levels and the first daily differences. The first order autocorrelation ρ_1, the Jarque-Bera and the Augmented Dickey Fuller (ADF) (an intercept has been included in the test equation) test values are also reported. One asterisk denotes rejection of the null hypothesis at the 1% level. The null hypothesis for the Jarque-Bera and the ADF tests is that the series is normally distributed and has a unit root, respectively. Summary statistics for the VIX and VXD futures in levels and changes are also provided.

The set of economic variables consists of the returns of the stock indices that serve as an underling asset to the options that are used to construct the corresponding volatility indices, the USD Libor and Euribor one-month interbank interest rates, the Euro/USD exchange rate, the WTI and Brent crude oil prices, the slope of the yield curve calculated as the difference between the prices of the 10-year government bond and the one-month interbank interest rate, and the volume of the futures contract of the underlying stock index. The time series of the economic variables were downloaded from Datastream.[4]

The CBOE VIX and VXD volatility futures were listed in March 2004 and April 2005, respectively. The liquidity of these markets keeps increasing. Measured on January 3, 2007, the open interest for the VIX (VXD) futures had increased by 95% (133%). The contract size of the volatility futures is $1000.[5] On any day, up to six near-term serial months and five months on the February quarterly cycle contracts are traded. The contracts are cash settled on the Wednesday that is thirty days prior to the third Friday of the calendar month immediately following the month in which the contract expires. Three time series of futures prices were constructed by ranking the data according to their expiry date: the shortest, second shortest and third shortest maturity series. To minimise the impact of noisy data, we roll to the second shortest series in the case where the shortest contract has less than five days to maturity. Prices that correspond to a volume of less than five contracts were discarded.

Table 1 shows the summary statistics of the implied volatility indices (in levels and first differences, Panels A and B, respectively), and volatility futures in levels and first differences (for VIX and VXD, Panels C and D, respectively). Information on the volume in the volatility futures markets is also provided. The augmented Dickey-Fuller (ADF) test for unit roots is also reported. We can see that none of the indices exhibit strong autocorrelation in the daily changes. The values of the ADF test also show that implied volatility indices are non-stationary in the levels, stationary in the first differences though; the same result holds for most of the economic variables (not reported here due to space limitations). The VIX futures are more liquid than the VXD ones, as expected.

3. The forecasting models

3.1. The economic variables model

The economic variables model employs certain economic variables as predictors to forecast the evolution of each implied

[4] Data on the volume of the S&P 100 futures contract are not available since this contract is not traded.

[5] Prior to March 26, 2007, the underlying asset of the VIX (VXD) futures contract was an "Increased-Value index" termed VBI (DVB) that was 10 times the value of VIX (VXD) at any point in time. The contract size of the volatility futures was $100 times the value of the underlying index. We have rescaled our series accordingly.

volatility index (see also Ahoniemi, 2006, for a similar approach). In particular, the following general forecasting specification is employed

$$\Delta IV_t = c_1 + a_1^+ R_{t-1}^+ + a_1^- R_{t-1}^- + \beta_1 i_{t-1} + \gamma_1 fx_{t-1} + \delta_1 \mathrm{oil}_{t-1} + \zeta_1 \Delta HV_{t-1} + \rho_1 \Delta IV_{t-1} + \kappa_1 \Delta ys_{t-1} + \xi_1 \mathrm{vol}_{t-1} + \varepsilon_t, \quad (1)$$

where ΔIV_t denotes the daily changes of the given implied volatility index, c_1 is a constant, and R_t^+, R_t^- denote the corresponding underlying stock index positive and negative log-returns (e.g., R_t^+ is filled with the positive returns and zeroes elsewhere), respectively so as to capture the possible presence of the asymmetric effect of index returns on implied volatility. i_t denotes the one-month US interbank (Euribor) interest rate for the European (US) market, fx_t the Euro/USD exchange rate, oil_t the WTI (Brent Crude Oil) price for the American (European) market; all three variables are measured in log-differences. ΔHV_t denotes the changes of the 30-days historical volatility, Δys_t the changes of the slope of the yield curve calculated as the difference between the yield of the ten year government bond and the one-month interbank interest rate, and vol_t the volume in log-differences of the futures contract of the underlying index. The choice of these variables is supported by the large literature on the predictability of asset returns (see e.g., Goyal and Welch, accepted for publication). This is because the implied volatility index is related to the expected return of the underlying stock index and therefore it may be forecasted by these variables (see Harvey and Whaley, 1992, for an explanation). The historical volatility is calculated as a 30-days moving average of equally weighted past squared returns. Furthermore, following Harvey and Whaley (1992) and Guo (2000), we augment the above mentioned set of economic variables by adding the changes of historical volatility and the term ΔIV_{t-1} as explanatory variables.

3.2. Univariate autoregressive and VAR models

Univariate autoregressive and VAR models are employed in order to examine whether the evolution of any given implied volatility index can be forecasted using its previous values, as well as the information from the evolution of implied volatility indices in the other option markets (see also Aboura, 2003, for a similar approach). First, for each implied volatility index an AR(1) model is employed, i.e.:

$$\Delta IV_t = c_1 + \lambda_1 \Delta IV_{t-1} + \varepsilon_t. \quad (2)$$

One lag is used since this is found to minimise the BIC criterion (within a range up to ten lags). The VAR specification is given by

$$Y_t = C + \Phi_1 Y_{t-1} + \varepsilon_t, \quad (3)$$

where Y_t is the vector of the seven implied volatility indices in their first differences that are assumed to be endogenously (jointly) determined. C is a (7×1) vector of constants, Φ_1, is the (7×7) matrix of coefficients to be estimated, and ε_t is the (7×1) vector of the VAR residuals.

3.3. The principal components model

Principal Components Analysis (PCA) is a non-parametric technique that summarises the dynamics of a set of variables by means of a smaller number of variables (principal components-PCs). Stock and Watson (2002) have shown that PCA can be employed for forecasting purposes. In particular, the PCs are used as predictors in a linear regression equation since they are proven to be consistent estimators of the true latent factors under quite general conditions. Moreover, the forecast constructed from the PCs is shown to converge to the forecast that would be obtained in the case where the latent factors were known. These properties make PCA a very powerful technique for forecasting purposes since it lets the data decide on the predictors to be used. This is in contrast to the approach taken in Eqs. (1)–(3) where the set of forecasting variables was chosen a priori.

First, we apply PCA to the daily changes of implied volatility indices. The first four PCs are retained. These explain 94% of the total variance of the changes of implied volatility indices. To identify any possible economic interpretation of the retained PCs, their pairwise correlations with the economic variables employed in Eq. (1) are calculated (see also Mixon, 2002, for a similar approach). Strong correlations appear only in the case of the first two PCs with the returns of the underlying stock indices. Interestingly, the first PC moves all implied volatility indices to the same direction, and hence it can be interpreted as a global factor (results are available upon request). Next, the changes of each volatility index are regressed on the previous day values of the first four PCs (PCA model),

$$\Delta IV_t = c_1 + r_1 PC1_{t-1} + r_2 PC2_{t-1} + r_3 PC3_{t-1} + r_4 PC4_{t-1} + \varepsilon_t, \quad (4)$$

where $r_i, i = 1, \ldots, 4$ are coefficients to be estimated.

3.4. ARIMA and ARFIMA models

ARIMA(p,d,q) and ARFIMA(p,d,q) models are employed to take into account the possible presence of short and long memory characteristics in the dynamics of implied volatility, respectively (see Fernandes et al., 2007, for a similar approach). The ARIMA(p,d,q) specification is given by

$$\Phi(L)\Delta^d IV_t = c + \Theta(L)\varepsilon_t, \quad (5)$$

where d is an integer that dictates the order of integration needed to produce a stationary and invertible process (in our case $d = 1$), L is the lag operator, $\Phi(L) = 1 + \phi_1 L + \ldots + \phi_p L^p$ is the autoregressive polynomial, $\Theta(L) = 1 + \theta_1 L + \ldots + \theta_p L^p$ is the moving average polynomial, μ is the mean of $\Delta^d IV_t$, $c = -\mu(1 + \phi_1 + \ldots + \phi_p)$, and ε_t is a Gaussian white noise process with zero mean and variance σ_ε^2. The ARFIMA(p,d,q) model is defined by

$$\Phi(L)(1-L)^d(\Delta IV_t - \mu) = \Theta(L)\varepsilon_t, \quad (6)$$

where now d denotes the non-integer order of fractional integration, $(1 - L)^d$ is the fractional difference operator, and μ denotes the expected value of ΔIV_t. In the case where $|d| < 0.5$, the ARFIMA(p,d,q) process is invertible and second-order stationary. In particular, if $0 < d < 0.5$ ($-0.5 < d < 0$) the process is said to exhibit long-memory (antipersistent) in the sense that the sum of the autocorrelation functions diverges to infinity (a constant) (see Baillie, 1996, for a review on fractional integration).

We choose $p = q = 1$ based on the BIC criterion and to avoid over-fitting the data (the differences in the BIC values are miniscule across a range of values for p and q). We follow Pong et al. (2004) to estimate the ARFIMA($1,d,1$) model and subsequently form the forecasts. In particular, maximum likelihood estimation is performed in the frequency domain by using the Whittle approximation of the Gaussian log-likelihood. Next, forecasts are obtained by taking the infinite autoregressive expansion of the ARFIMA $(1,d,1)$ process. Thus, one-step ahead forecasts are formed by

$$E(IV_{t+1} \mid I_t) = IV_t + \mu - \sum_{j=1}^{\infty} \pi_j(\Delta IV_{t-j+1} - \mu), \quad (7)$$

where $\pi_j = \sum_{i=0}^{j}(b_i + \phi b_{i-1})(-\theta)^{j-i}$, $b_i = \frac{\Gamma(-d+i)}{\Gamma(-d)\Gamma(i+1)}$ and $\Gamma(\cdot)$ denotes the gamma function. To implement Eq. (7), the infinite summation is truncated at $k = 150$.

4. In-sample evidence

Tables 2–5 show the in-sample performance of the economic variables, AR(1)/VAR, PCA, and ARIMA(1, 1, 1)/ARFIMA(1, d, 1) models, respectively. The estimated coefficients, the t-statistics within parentheses and the adjusted R^2 are reported for each one of the implied volatility indices, respectively. One and two asterisks indicate that the estimated parameters are statistically significant at 1% and 5% level, respectively. In the case of the economic variables model Table 2 we can see that the adjusted R^2 is nearly zero for all indices and takes the largest value (2.5%) for VCAC. The statistically significant variables for VCAC are CAC's positive return, the lagged changes in historical volatility and the lagged VCAC changes. In the remaining indices, almost all economic variables are insignificant. This comes at no surprise (see e.g., Harvey and Whaley, 1992, for similar results on predicting the evolution of the implied volatility of the S&P 100 options). Interestingly, our results do not depend on the degree of capitalisation of the underlying stock index. This is in contrast to the evidence provided by the literature on the predictability of stock returns where the small size stocks manifest greater predictability compared with big size stocks (see e.g., Fama and French, 1988). Finally, it should be noticed that the reported results are not subject to problems in statistical inference that arise due to the fact that the predictors may be nearly integrated (see e.g., Ferson et al., 2003). This is because the first order autocorrelation coefficient of the changes of each one of the economic variables is well far from unity (the maximum is 0.3 for the interest rate variable).

Table 3 (Panel A) shows the results from the AR(1) model Eq. (2). We can see that the adjusted R^2 are zero for all implied volatility indices. The fact that there is no mean-reversion in dynamics of the changes of the implied volatility indices is in contrast to the results found in Dotsis et al. (2007); their results were obtained for a different time period though. Table 3 (Panel B) shows the results from the estimation of the VAR model by ordinary least squares (OLS). For each one of the seven equations in the VAR, the estimated coefficients are reported. The greatest value of the adjusted R^2 is obtained for VCAC (11.7%), while the lowest is obtained for VIX (1.2%).

Table 4 shows the results from the PCA model Eq. (4). We can see that the model fits poorly most volatility indices; the only exception occurs for VCAC and VSTOXX ($R^2 = 11.2\%$, $R^2 = 6.8\%$, respectively). Table 5 shows the results for the ARIMA(1, 1, 1) and the ARFIMA(1, d, 1) models (Panel A and B, respectively). We can see that the adjusted R^2's are zero for all implied volatility indices. Moreover, the fractional integration parameter is statistically significant in most cases and lies within the range $-0.5 < d < 0$. Therefore, the changes in the implied volatility index do not exhibit long memory. Overall, within sample, the VAR and PCA models perform best. In general, they fit better the European than the US indices. This implies that each European index manifests a certain predictable pattern in its dynamics that could be exploited by the information extracted from the other volatility indices. For instance, VCAC is affected by VXD and VSTOXX, and it affects the other three US indices and VSTOXX.

5. Out-of-sample forecasting performance

We assess the out-of-sample performance of each model specification that we considered in Section 4. The out-of-sample exercise is performed from March 18, 2005 to September 28, 2007 by increasing the sample size by one observation and re-estimating each model as time goes by. Point and interval forecasts are formed for each one of the seven implied volatility indices. Every day,

Table 2
Forecasting with the economic variables model

	Dependent variable: ΔVIX_t	Dependent variable: $\Delta VXOA_t$	Dependent variable: ΔVXN_t	Dependent variable: ΔVXD_t	Dependent variable: $\Delta VDAX\_New_t$	Dependent variable: $\Delta VCAC_t$	Dependent variable: $\Delta VSTOXX_t$
Included Obs.	954	955	953	950	1015	1017	1015
	Coeff. (t-Statistic)	Coeff. (t-Statistic)	Coeff. (t-Statistic)	Coeff. (t-Statistic)	Coeff. (t-Statistic)	Coeff. (t-Statistic)	Coeff. (t-Statistic)
C_1	0.000	0.000	−0.001	0.000	−0.001	0.000	0.000
	(0.556)	(0.041)	(−0.822)	(−0.220)	(−0.772)	(−0.327)	(0.261)
R_{t-1}^{+}	−0.020	−0.027	−0.064	−0.105	0.009	*−0.164***	0.002
	(−0.215)	(−0.596)	(−1.452)	(−1.612)	(0.117)	(−2.417)	(0.022)
R_{t-1}^{-}	0.147	0.050	−0.018	−0.046	−0.054	−0.184	0.054
	(1.099)	(0.407)	(−0.281)	(−0.437)	(−0.528)	(−1.814)	(0.433)
i_{t-1}	−0.020	−0.020	−0.054	−0.056	−0.075	0.012	−0.078
	(−0.375)	(−0.366)	(−1.054)	(−1.378)	(−0.599)	(0.121)	(−0.697)
fx_{t-1}	−0.084	−0.074	−0.018	−0.055	0.124	0.088	0.185
	(−1.184)	(−1.009)	(−0.195)	(−0.894)	(1.230)	(0.854)	(1.796)
oil_{t-1}	0.020	−0.007	0.022	−0.009	−0.005	−0.024	−0.017
	(1.227)	(−0.475)	(1.09)	(−0.537)	(−0.230)	(−1.401)	(−0.770)
ΔHV_{t-1}	0.107	0.025	*0.092***	0.086	0.034	*0.131***	0.049
	(1.407)	(0.366)	(2.151)	(1.464)	(0.517)	(1.984)	(0.477)
ΔIV_{t-1}	0.072	−0.019	0.014	−0.036	−0.043	*−0.144**	−0.004
	(0.753)	(−0.201)	(0.269)	(−0.516)	(−0.516)	(−3.151)	(−0.037)
Δys_{t-1}	0.009	0.007	0.004	0.004	−0.018	−0.009	−0.013
	(1.428)	(1.135)	(0.512)	(0.699)	(−1.160)	(−0.694)	(−0.779)
vol_{t-1}	−0.001	–	0.001	0.000	−0.001	0.000	0.000
	(−0.927)	–	(0.443)	(0.476)	(−0.366)	(0.027)	(−0.185)
Adj. R^2	*0.002*	*−0.004*	*0.006*	*0.004*	*−0.004*	*0.025*	*−0.003*

The entries report results from the regression of each implied volatility index on a set of lagged economic variables, augmented by an AR(1) term. The following specification is estimated $\Delta IV_t = c_1 + a_1^{+}R_{t-1}^{+} + a_1^{-}R_{t-1}^{-} + \beta_1 i_{t-1} + \gamma_1 fx_{t-1} + \delta_1 \mathrm{oil}_{t-1} + \zeta_1 \Delta HV_{t-1} + \rho_1 \Delta IV_{t-1} + \kappa_1 \Delta ys_{t-1} + \xi_1 \mathrm{vol}_{t-1} + \varepsilon_t$ where ΔIV: the changes of the implied volatility index, R^{+}: the underlying positive stock index return, R^{-}: the underlying negative stock index return, i: the one-month interbank/Euribor interest rate for the US/European market, log-differenced, f_x: the EUR/USD exchange rate log-differenced, oil: WTI/Brent crude oil price for the American/European market, in log-differences, HV: the 30-days historical volatility, Δys: the changes of the yield spread calculated as the difference between the yield of the 10-year government bond and the one-month interbank interest rate, and vol: the volume in log-differences of the futures contract of the underlying index. The estimated coefficients, Newey-West t-statistics in parentheses, and the adjusted R^2 are reported. One and two asterisks denote rejection of the null hypothesis of a zero coefficient at the 1% and 5% level, respectively. The model has been estimated for the period February 2, 2001 to March 17, 2005.

Table 3
Forecasting with the univariate autoregressive and VAR models

Included Obs.	Dependent variable: ΔVIX$_t$ 956 Coeff. (t-Statistic)	Dependent variable: ΔVXOA$_t$ 955 Coeff. (t-Statistic)	Dependent variable: ΔVXN$_t$ 953 Coeff. (t-Statistic)	Dependent variable: ΔVXD$_t$ 956 Coeff. (t-Statistic)	Dependent variable: ΔVDAX_New$_t$ 1015 Coeff. (t-Statistic)	Dependent variable: ΔVCAC$_t$ 1017 Coeff. (t-Statistic)	Dependent variable: ΔVSTOXX$_t$ 1015 Coeff. (t-Statistic)
Panel A: AR(1) model							
c_1	0.000	0.000	*−0.001***	0.000	0.000	0.000	0.000
	(−1.104)	(−1.203)	(−2.082)	(−1.127)	(−0.315)	(−0.178)	(−0.233)
ΔIV_{t-1}	0.008	−0.026	0.052	0.025	−0.016	−0.029	−0.029
	(0.169)	(−0.545)	(1.386)	(0.590)	(−0.435)	(−0.777)	(−0.539)
Adj. R^2	*−0.001*	*0.000*	*0.002*	*0.000*	*−0.001*	*0.000*	*0.000*
	Dependent variable: ΔVIX$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVXOA$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVXN$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVXD$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVDAX_New$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVCAC$_t$ Coeff. (t-Statistic)	Dependent variable: ΔVSTOXX$_t$ Coeff. (t-Statistic)
Panel B: VAR model							
ΔVIX$_{t-1}$	0.158	*0.478**	0.206	*0.316**	*0.459**	0.211	*0.383**
	(1.694)	(5.203)	(1.890)	(3.953)	(3.900)	(1.903)	(3.119)
ΔVXOA$_{t-1}$	−0.038	*−0.394**	−0.085	0.051	−0.014	0.112	0.055
	(−0.462)	(−4.881)	(−0.891)	(0.729)	(−0.138)	(1.147)	(0.508)
ΔVXN$_{t-1}$	−0.049	−0.045	−0.063	−0.028	*−0.113***	−0.044	−0.079
	(−1.235)	(−1.148)	(−1.374)	(−0.826)	(−2.275)	(−0.936)	(−1.514)
ΔVXD$_{t-1}$	−0.070	−0.038	0.118	*−0.283**	−0.017	*−0.199***	0.099
	(−0.925)	(−0.507)	(1.342)	(−4.361)	(−0.176)	(−2.225)	(0.994)
ΔVDAX_New$_{t-1}$	−0.007	0.005	−0.009	−0.042	*−0.298**	0.066	0.033
	(−0.128)	(0.101)	(−0.143)	(−0.919)	(−4.431)	(1.039)	(0.466)
ΔVCAC$_{t-1}$	*−0.108**	*−0.116**	*−0.098**	−0.050	−0.064	*−0.237**	*−0.113**
	(−3.349)	(−3.669)	(−2.630)	(−1.833)	(−1.587)	(−6.216)	(−2.683)
ΔVSTOXX$_{t-1}$	0.028	0.027	0.032	0.043	*0.196**	*0.259**	*−0.130***
	(0.566)	(0.560)	(0.572)	(1.035)	(3.187)	(4.482)	(−2.037)
C	0.000	0.000	−0.001	0.000	0.000	0.000	0.000
	(−0.903)	(−0.957)	(−1.867)	(−0.776)	(−0.609)	(−0.066)	(−0.640)
Adj. R^2	*0.012*	*0.037*	*0.021*	*0.043*	*0.063*	*0.117*	*0.084*

Panel A: The entries report results from the estimation of a univariate AR(1) specification for the daily changes ΔIV of each implied volatility index, i.e. $\Delta IV_t = c_1 + \lambda_1\, \Delta IV_{t-1} + \varepsilon_t$. Panel B: The entries report the estimated coefficients of a VAR, for the set of the seven implied volatility (IV) indices: $Y_t = C + \Phi_1 Y_{t-1} + \varepsilon_t$, where Y_t is the (7×1) vector of IV indices (in differences), C is a (7×1) vector of constants, Φ_1 is the (7×7) matrix of coefficients to be estimated, and u_t is a (7×1) vector of errors. The estimated coefficients, Newey-West t-statistics in parentheses and the adjusted R^2 are reported. One and two asterisks denote rejection of the null hypothesis of a zero coefficient at the 1% and 5% level, respectively. The models have been estimated for the period February 2, 2001 to March 17, 2005.

Table 4
Forecasting with the principal components analysis model

Included Obs.	Dependent variable: ΔVIX$_t$ 932 Coeff. (t-Statistic)	Dependent variable: ΔVXOA$_t$ 931 Coeff. (t-Statistic)	Dependent variable: ΔVXN$_t$ 931 Coeff. (t-Statistic)	Dependent variable: ΔVXD$_t$ 932 Coeff. (t-Statistic)	Dependent variable: ΔVDAX_New$_t$ 950 Coeff. (t-Statistic)	Dependent variable: ΔVCAC$_t$ 953 Coeff. (t-Statistic)	Dependent variable: ΔVSTOXX$_t$ 949 Coeff. (t-Statistic)
c	0.000	0.000	*−0.001***	0.000	0.000	0.000	0.000
	(−0.972)	(−1.106)	(−2.068)	(−1.084)	(−0.668)	(−0.299)	(−0.522)
PC1$_{t-1}$	0.000	0.000	*−0.001***	−0.001	−0.002	*−0.004**	*−0.003**
	(0.810)	(0.462)	(−2.460)	(−1.454)	(−2.820)	(−5.567)	(−3.992)
PC2$_{t-1}$	*0.001***	0.001	*0.001***	*0.001**	0.002	0.001	*0.004**
	(2.070)	(1.773)	(2.374)	(2.719)	(3.610)	(1.082)	(5.164)
PC3$_{t-1}$	0.001	0.001	0.001	0.001	0.001	*0.004**	0.001
	(1.880)	(1.516)	(1.683)	(1.126)	(0.848)	(6.167)	(1.037)
PC4$_{t-1}$	0.000	0.000	−0.001	0.000	−0.002	0.001	*−0.001***
	(−0.231)	(0.191)	(−0.782)	(−0.750)	(−2.554)	(1.298)	(−2.159)
Adj. R^2	*0.012*	*0.007*	*0.020*	*0.015*	*0.036*	*0.112*	*0.068*

The entries report results from the regression $\Delta IV_t = c_1 + r_{1j}\, PC1_{t-1} + r_{2j} PC2_{t-1} + r_{3j} PC3_{t-1} + r_{4j} PC4_{t-1} + \varepsilon_t$ of the changes ΔIV of each implied volatility index on the lagged first four principal components $PC1$, $PC2$, $PC3$ and $PC4$ derived from the set of the seven IV indices. The estimated coefficients, Newey-West t-statistics in parentheses, and the adjusted R^2 are reported. One and two asterisks denote rejection of the null hypothesis of a zero coefficient at the 1% and 5% level, respectively. The model has been estimated for the period February 2, 2001 to March 17, 2005.

10,000 simulation runs have been generated to construct the interval forecasts.

5.1. Point forecasts

In line with Gonçalves and Guidolin (2006), we use three metrics to assess the out-of-sample performance of the employed models in a statistical setting: the root mean squared prediction error (RMSE), the mean absolute prediction error (MAE) and the mean correct prediction (MCP) of the direction of change in the value of the implied volatility index (see Gonçalves and Guidolin, 2006, for the definition of each metric). The models are compared to the random walk model that is used as a benchmark. The modified Diebold–Mariano test of Harvey et al. (1997) and a ratio test

Table 5
Forecasting with the ARIMA(1, 1, 1) and the ARFIMA(1, d, 1) models

	Dependent variable: ΔVIX_t	Dependent variable: $\Delta VXOA_t$	Dependent variable: ΔVXN_t	Dependent variable: ΔVXD_t	Dependent variable: $\Delta VDAX\_New_t$	Dependent variable: $\Delta VCAC_t$	Dependent variable: $\Delta VSTOXX_t$
Included Obs.	994	993	992	994	1030	1033	1030
	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)
Panel A: ARIMA(1, 1, 1) model							
c	0.000	0.000	*−0.001***	0.000	0.000	0.000	0.000
	(−1.209)	(−1.612)	(−1.963)	(*−0.911*)	(−0.474)	(−0.209)	(−0.399)
ϕ	*0.735**	*−0.773**	*0.703**	0.534	*−0.879**	0.629	*−0.856**
	(2.590)	(−5.989)	(4.921)	(1.014)	(−8.333)	(1.323)	(−6.790)
θ	*0.774**	*−0.848**	*0.773**	0.574	*−0.909**	0.588	*−0.898**
	(2.935)	(−7.352)	(6.076)	(1.133)	(−9.448)	(1.204)	(−8.251)
Adj. R^2	*0.001*	*0.012*	*0.009*	*0.000*	*0.002*	*0.001*	*0.004*
	Dependent variable: ΔVIX_t	Dependent variable: $\Delta VXOA_t$	Dependent variable: ΔVXN_t	Dependent variable: ΔVXD_t	Dependent variable: $\Delta VDAX\_New_t$	Dependent variable: $\Delta VCAC_t$	Dependent variable: $\Delta VSTOXX_t$
Included Obs.	995	994	993	995	1031	1034	1031
	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)	Coeff. (*t*-Statistic)
Panel B: ARFIMA (1, d, 1) model							
d	*−0.210**	*−0.178**	*−0.071***	*−0.169**	*−0.078***	−0.031	*−0.091**
	(−3.725)	(−3.459)	(−2.112)	(−2.996)	(−2.269)	(−1.002)	(*−2.974*)
ϕ	−0.172	−0.121	*0.622**	−0.177	0.366	*0.667**	*0.627**
	(−0.831)	(−0.511)	(5.009)	(−0.822)	(1.326)	(3.317)	(3.152)
θ	0.033	0.021	*0.723**	0.011	0.431	*0.641**	*0.688**
	(0.190)	(0.099)	(6.993)	(0.058)	(1.677)	(3.080)	(3.818)
Adj. R^2	*0.016*	*0.012*	*0.010*	*0.011*	*0.003*	*0.002*	*0.008*

Panel A: The entries report results from the estimation of an ARIMA(1, 1, 1) model. The specification $(1+\phi L)\Delta IV_t = c + (1+\theta L)\varepsilon_t$ is used. Panel B: The entries report the results from the estimation of an ARFIMA(1, *d*, 1) model. The specification $(1+\phi L)(1-L)^d(\Delta IV_t - \mu) = (1+\theta L)\varepsilon_t$ is used. The estimated coefficients, *t*-statistics in parentheses, and the adjusted R^2 are reported. One and two asterisks denote rejection of the null hypothesis of a zero coefficient at the 1% and 5% level, respectively. The models have been estimated for the period February 2, 2001 to March 17, 2005.

are used to assess whether any model under consideration outperforms the random walk model in a statistically significant sense under the RMSE/MAE and the MCP metrics, respectively. The null hypothesis is that the random walk model and the model under consideration perform equally well.[6]

Table 6 shows the results on the out-of-sample performance of the alternative model specifications for each one of the seven implied volatility indices. One and two asterisks denote rejection of the null hypothesis at 1% and 5% significance levels, respectively. There are 35 combinations of implied volatility indices and predictability measures (out of possible total of 126) in which one of the six models has outperformed the random walk. Therefore, in 28% of the cases one of the models performs better than the random walk. This indicates that a statistically predictable pattern exists in the dynamics of implied volatility indices (by assuming independence at a level of significance 5%).

Consistently with the in-sample evidence, the predictable pattern is stronger in the case of the European indices where in 41% (22/54) of the cases, the models under consideration outperform the random walk; in the case of the US indices, only in 18% (13/72) of the cases one of the models outperforms the random walk. Regarding the question which model performs best, the VAR and PCA models outperform all competing models in the case of the European indices since they beat the random walk under all metrics. The ARIMA(1, 1, 1) and ARFIMA(1, *d*, 1) models perform best in the case of the US indices. The results imply that there are implied volatility spillovers between the markets; the information contained in all implied volatility indices can be used to predict each European index separately. This is not the case for the US indices where instead their autocorrelation structure should be taken into account in order to predict their evolution.

5.2. Interval forecasts

To evaluate the goodness of the out-of-sample interval forecasts, Christoffersen's (1998) likelihood ratio test of unconditional coverage (LR_{unc}) is used. Let an observed sample path $\{IV_t\}_{t=1}^{T}$ of the time series of the implied volatility index and a series of constructed interval forecasts $\{(L_{t/t-1}(\alpha), U_{t/t-1}(\alpha))\}_{t=1}^{T}$ at a significance level α%. $L_{t/t-1}(\alpha)$ and $U_{t/t-1}(\alpha)$ denote the constructed at time $(t-1)$ lower and upper bound of the α%-interval forecast for time *t*, respectively. The null hypothesis is that the percentage of times that the realised index value at time *t* falls outside the constructed at time $(t-1)$ intervals is *a*%. Given that the power of the test may be sensitive to the sample size, we base the accept/reject decisions of the null hypothesis on MC simulated *p*-values.

Table 7 shows the percentage of observations that fall outside the constructed 5% intervals, and the values of Christoffersen's (1998) test obtained by the economic variables, AR(1), VAR, PCA, ARIMA(1, 1, 1), and ARFIMA(1, *d*, 1) models (Panels A, B, C, D, E, and F, respectively) for each one of the seven implied volatility indices. One and two asterisks denote rejection of the null hypothesis at 1% and 5% significance levels, respectively. We can see that there is no single model that yields accurate forecasts for all indices just as was the case with the point forecasts; the VAR model performs best in the horse race among models. Overall, the null hypothesis is accepted in 48% of the cases (20 cases out of a possible total of 42). Interestingly, 17 out of these 20 cases pertain to the US indices. These results imply that there is also a predictable pattern in an interval forecast sense that is stronger for the US indices. This is in contrast to the point forecasts case where predictability was stronger for the European indices. On the other hand, the presence of volatility spillovers is useful for forecasting purposes just as was the case in the point forecasts.

[6] Strictly speaking, the MCP cannot be calculated under the random walk model. Hence, in the ratio test, we treat the random walk model as a naïve model that would yield MCP = 50%.

Table 6
Out-of-sample performance of the model specifications for each one of the implied volatility indices

	VIX	VXOA	VXN	VXD	VDAX New	VCAC	VSTOXX
Panel A: Random walk							
RMSE	1.07	1.01	1.03	1.01	0.94	1.04	1.00
MAE	0.68	0.64	0.70	0.64	0.67	0.71	0.70
Panel B: Regression model based on economic variables							
RMSE	1.08	1.02	1.04	1.00	0.94	1.05	1.01
MAE	0.67	0.63	0.70	0.63	0.67**	0.72	0.70
MCP	54.71%**	50.67%	53.70%	49.50%	55.31%*	47.05%	49.84%
Panel C: AR(1) Model							
RMSE	1.07	1.01	1.04	1.01	0.94	1.04	1.01
MAE	0.67	0.63**	0.70	0.63	0.68	0.71	0.70
MCP	52.86%	53.20%	56.06%*	52.36%	50.24%	52.47%	49.21%
Panel D: VAR Model							
RMSE	1.07	1.01	1.06	0.99	0.85*	0.99*	0.89*
MAE	0.68	0.63	0.70	0.63	0.62*	0.68*	0.65*
MCP	51.54%	55.65%*	52.74%	52.06%	61.20%*	58.22%*	60.43%*
Panel E: PCA Model							
RMSE	1.08	1.02	1.05	1.01	0.85*	0.99*	0.90*
MAE	0.68	0.64	0.70	0.64	0.62*	0.69*	0.65*
MCP	52.91%	53.25%	50.34%	50.00%	59.87%*	58.56%*	58.43%*
Panel F: ARIMA(1,1,1) Model							
RMSE	1.07	1.00**	1.04	1.00	0.94	1.04	1.03
MAE	0.67	0.63**	0.70	0.63	0.67	0.71	0.71
MCP	53.20%	56.23%*	49.83%	49.66%	53.09%	53.27%	52.05%
Panel G: ARFIMA(1,d,1) Model							
RMSE	1.06	1.00	1.03	1.01	0.94	1.04	1.00
MAE	0.67	0.63*	0.68**	0.64	0.67	0.70	0.69
MCP	53.41%**	55.36%*	56.49%*	53.90%**	54.38%**	53.53%**	52.96%

The root mean squared prediction error (RMSE), the mean absolute prediction error (MAE), and the mean correct prediction (MCP) of the direction of change in the value of the implied volatility index are reported. The random walk model (Panel A), economic variables model (Panel B), AR(1) model (Panel C), VAR model (Panel D), PCA model (Panel E), ARIMA(1,1,1) model (Panel F) and the ARFIMA(1,d,1) model (Panel G) have been implemented. The null hypothesis is that the random walk and the model under consideration perform equally well, against the alternative that the model under consideration performs better, have been tested via the Modified Diebold–Mariano test (for RMSE and MAE) and the ratio test (for MCP). One and two asterisks denote rejection of the null hypothesis at 1% and 5% significance levels, respectively. The models have been estimated recursively for the period March 18, 2005 to September 28, 2007.

Table 7
Statistical accuracy of the interval forecasts

	VIX	VXOA	VXN	VXD	VDAX_New	VCAC	VSTOXX
Panel A: Economic variables model interval forecasts							
# Violations	6.23%	5.22%	2.69%	6.40%	1.43%	3.51%	3.34%
LRunc	1.76	0.06	7.94*	2.25	23.36*	3.26	4.12**
Panel B: AR(1) interval forecasts							
# Violations	6.06%	5.22%	2.86%	6.90%	1.43%	2.87%	1.26%
LR^{unc}	1.32	0.06	6.71*	4.07**	23.36*	7.02*	26.29*
Panel C: VAR interval forecasts							
# Violations	5.99%	5.65%	3.77%	6.34%	1.17%	3.52%	1.17%
LR_{unc}	1.14	0.50	2.04	2.03	26.38*	3.04	26.46*
Panel D: PCA interval forecasts							
# Violations	6.16%	5.48%	3.42%	7.02%	1.00%	3.36%	1.00%
LR^{unc}	1.56	0.27	3.41	4.48**	29.52*	3.82**	29.60*
Panel E: ARIMA(1,1,1) interval forecasts							
# Violations	7.24%	6.56%	4.38%	8.59%	1.74%	3.51%	1.89%
LR^{unc}	5.54**	2.80	0.51	13.36*	18.62*	3.26	16.72*
Panel F: ARFIMA(1,d,1) interval forecasts							
# Violations	5.52%	5.36%	2.92%	6.49%	1.41%	2.83%	1.56%
LR^{unc}	0.34	0.16	6.54*	2.65	24.03*	7.47*	21.67*

Entries report the percentage of the observations that fall outside the constructed intervals, and the values of Christoffersen (1998) **'slikelihood ratio test of unconditional coverage (LR_{unc}) for each implied volatility index. The null hypothesis is that the percentage of times that the actually realised index value falls outside the constructed α%-intervals is a %. One and two asterisks denote rejection of the null hypothesis at 1% and 5% significance levels, respectively. The results are reported for daily 5%-interval forecasts over the period March 18, 2005 to September 28, 2007 generated by the economic variables model (Panel A), AR(1) model (Panel B), VAR model (Panel C), PCA model (Panel D), ARIMA(1,1,1) model (Panel E) and ARFIMA (1,d,1) model (Panel F).

6. Economic significance

To assess the economic significance of the point and interval forecasts formed by each one of the six employed models, trading strategies with VIX (VXD) futures are constructed. The strategies employ each one of the three shortest VIX (VXD) futures series. The strategies are implemented for each model separately, despite the fact that some of the models do not generate statistically

Table 8
Trading strategy with VIX/VXD futures based on point forecasts from March 18, 2005 to September 28, 2007

	VIX			VXD		
	Shortest	2nd Shortest	3rd Shortest	Shortest	2nd Shortest	3rd Shortest
Panel A: Economic variables model point forecasts						
Sharpe Ratio	0.0306	0.0175	0.0062	-0.0176	−0.0787	−0.1081
95% CI	(−0.05, 0.11)	(−0.06, 0.10)	(−0.08, 0.09)	(−0.11, 0.08)	(−0.19, 0.04)	(−0.26, 0.03)
A_p	0.2487	0.1026	0.0347	−0.3606	−0.7400	−0.7774
95% CI	(−0.51, 1.00)	(−0.45, 0.67)	(−0.49, 0.53)	(−1.41, 0.66)	(− 1.82, 0.28)	(−1.90, 0.28)
Panel B: AR(1) point forecasts						
Sharpe Ratio	−0.0190	−0.0392	−0.0366	−0.0680	−0.0654	−0.1234
95% CI	(−0.10, 0.06)	(−0.12, 0.04)	(−0.12, 0.05)	(−0.15, 0.03)	(−0.18, 0.06)	(−0.25, 0.02)
A_p	−0.3375	−0.3367	−0.2385	*−1.1944**	−0.6825	−0.8191
95% CI	(−0.96, 0.25)	(−0.81, 0.13)	(−0.67, 0.19)	(−2.13, −0.34)	(−1.55, 0.14)	(−1.84, 0.11)
Panel C: VAR point forecasts						
Sharpe Ratio	−0.0140	−0.0186	−0.0369	0.0812	−0.0209	0.0071
95% CI	(−0.09, 0.06)	(−0.10, 0.06)	(−0.12, 0.05)	(−0.02, 0.17)	(−0.15, 0.10)	(−0.15, 0.14)
A_p	−0.2098	−0.1626	−0.2377	0.8104	−0.2192	0.0825
95% CI	(−0.93, 0.52)	(−0.71, 0.40)	(−0.74, 0.26)	(−0.27, 1.95)	(−1.30, 0.94)	(−0.96, 1.20)
Panel D: PCA point forecasts						
Sharpe Ratio	−0.0664	−0.0596	−0.0828	*0.1137**	0.0746	0.0773
95% CI	(−0.14,0.01)	(−0.14, 0.02)	(−0.16, 0.00)	(0.02, 0.21)	(−0.05, 0.19)	(−0.06, 0.22)
A_p	−0.6747	−0.4272	−0.5023	*1.1268**	0.6274	0.5529
95% CI	(−1.42, 0.06)	(−1.00, 0.14)	(−0.99, 0.00)	(0.06, 2.25)	(−0.41, 1.77)	(−0.48, 1.64)
Panel E: ARIMA(1, 1, 1) point forecasts						
Sharpe Ratio	0.0101	0.0202	0.0274	0.0571	0.0373	−0.0214
95% CI	(−0.07, 0.09)	(−0.06, 0.10)	(−0.06, 0.11)	(−0.04, 0.15)	(−0.09, 0.16)	(−0.16, 0.12)
A_p	0.0612	0.1211	0.1527	0.6021	0.3201	−0.1464
95% CI	(−0.66, 0.81)	(−0.44, 0.69)	(−0.35, 0.65)	(−0.48, 1.71)	(−0.76, 1.43)	(−1.22, 0.88)
Panel F: ARFIMA(1, d, 1) point forecasts						
Sharpe Ratio	−0.0127	−0.0286	−0.0268	0.0494	−0.0133	0.1220
95% CI	(−0.09, 0.07)	(−0.11, 0.05)	(−0.11, 0.06)	(−0.05, 0.15)	(−0.13, 0.11)	(−0.02, 0.26)
A_p	−0.2471	−0.2528	−0.1752	0.3651	−0.1755	0.9140
95% CI	(−0.91,0.40)	(−0.75, 0.24)	(−0.63, 0.27)	(−0.65, 1.37)	(−1.20, 0.82)	(−0.09, 1.92)

The entries show the annualised Sharpe ratio and Leland's Alpha (A_p) and their respective bootstrapped 95% confidence intervals (CI). The strategy is based on point forecasts obtained from the economic variables model (Panel A), AR(1) model (Panel B), VAR model (Panel C), PCA model (Panel D), ARIMA(1,1,1) model (Panel E), and ARFIMA(1,d,1) model (Panel F). The Sharpe ratio for the S&P 500 and the Dow Jones Industrial Average is 0.0265 [95% CI = (−0.05, 0.10)] and 0.0319 [95% CI = (−0.04, 0.11)], respectively. One asterisk denotes rejection of the null hypothesis of a zero Sharpe ratio (A_p) at a 5% level of significance.

significant forecasts. This is because the statistical evidence does not always corroborate a financial criterion (see also Ferson et al., 2003). The CBOE transaction costs are taken into account ($0.5 per contract).

6.1. Trading strategy based on point forecasts

To assess the economic significance of the point forecasts, the following trading rule is employed. The investor goes long (short) in the volatility futures in the case where the forecasted value of the implied volatility index is greater (smaller) than its current value. Table 8 shows the annualised Sharpe ratio (SR) and Leland's (1999) alpha (A_p) obtained for each one of the three shortest VIX and VXD futures.[7] Results are reported for the trading strategy based on the point forecasts formed by the economic variables (Panel A), AR(1) (Panel B), VAR (Panel C), PCA (Panel D), ARIMA(1,1,1) (Panel E), and ARFIMA(1,d,1) (Panel F) models. To evaluate the statistical significance of SR and A_p, 95% confidence intervals have been bootstrapped and reported within parentheses. One asterisk denotes rejection of the null hypothesis of a zero SR (A_p) at a 5% level of significance. We can see that SR and A_p are statistically insignificant in almost all cases. Therefore, the statistically predictable pattern found in Section 5.1 is not economically significant in that no abnormal profits can be attained. A naive buy and hold strategy did not yield an economically significant performance, either.

6.2. Trading strategy based on interval forecasts

To evaluate the economic significance of the constructed interval forecasts, the following trading rule is used

$$\text{If } IV_{t-1} < (>) \frac{U_{t/t-1}(\alpha) + L_{t/t-1}(\alpha)}{2}, \quad \text{then go long(short)}.$$

$$\text{If } IV_{t-1} = \frac{U_{t/t-1}(\alpha) + L_{t/t-1}(\alpha)}{2}, \quad \text{then do nothing}.$$

The rational is that in the case where the value of the volatility index is closer to the lower (upper) bound of the next day's forecast interval, the index price is expected to increase and a long (short) position is taken in the volatility futures. Notice that the criterion requires a contemporaneous comparison of the volatility index value and the constructed intervals at time ($t-1$); this is in contrast to Christoffersen's test.[8]

Table 9 shows the annualised SR and A_p, and their corresponding bootstrapped 95% confidence intervals obtained for each one of the three shortest VIX and VXD futures series. Results are reported

[7] A_p is used since the distribution of the returns of the futures trading strategy is found to be non-normal. It is calculated by using the S&P 500 and the DJIA indices to proxy the benchmark (market) portfolio in the VIX and VXD futures strategies, respectively. To check the sensitivity of the results on A_p to the choice of the benchmark portfolio, the VIX and VXD indices were also used to proxy the market portfolio; the results did not change.

[8] We have also considered implementing an alternative trading strategy where trades would be triggered only when the implied volatility index crosses the limits of the constructed interval forecast. Again, a contemporaneous comparison of the volatility index value and the constructed interval forecast is required. However, this rule did not trigger any trades since the value of the volatility index did not cross the bounds of the interval forecast through our sample.

Table 9
Trading strategy with VIX/VXD futures based on interval forecasts from March 18, 2005 to September 28, 2007

	VIX			VXD		
	Shortest	2nd Shortest	3rd Shortest	Shortest	2nd Shortest	3rd Shortest
Panel A: Economic variables model interval forecasts						
Sharpe Ratio	0.0029	−0.0098	−0.0284	0.0144	−0.0409	−0.0872
95% CI	(−0.08, 0.08)	(−0.09, 0.07)	(−0.11, 0.06)	(−0.08, 0.11)	(−0.16, 0.08)	(−0.23, 0.05)
A_p	−0.0092	−0.0827	−0.1705	−0.0154	−0.4002	−0.6389
95% CI	(−0.75, 0.74)	(−0.65, 0.47)	(−0.69, 0.34)	(−1.07, 1.05)	(−1.47, 0.63)	(−1.77, 0.41)
Panel B: AR(1) interval forecasts						
Sharpe Ratio	−0.0450	−0.0751	−0.0714	−0.0411	−0.0269	−0.0345
95% CI	(−0.12, 0.03)	(−0.15, 0.01)	(−0.15, 0.01)	(−0.13, 0.06)	(−0.14, 0.10)	(−0.18, 0.11)
A_p	−0.5678	−0.5776	−0.4402	−0.8593	−0.3295	−0.1765
95% CI	(−1.21,0.04)	(−1.06, −0.10)	(−0.89, 0.00)	(−1.84, 0.06)	(−1.25, 0.55)	(−1.15, 0.84)
Panel C: VAR interval forecasts						
Sharpe Ratio	−0.0324	−0.0394	−0.0675	0.0316	0.0018	0.0258
95% CI	(−0.11, 0.04)	(−0.12, 0.04)	(−0.15, 0.02)	(−0.07, 0.13)	(−0.12, 0.12)	(−0.12,0.16)
A_p	−0.4023	−0.3095	−0.4172	0.2681	−0.0048	0.2070
95% CI	(−1.13, 0.31)	(−0.86, 0.24)	(−0.92, 0.09)	(−0.82, 1.41)	(−1.07, 1.13)	(−0.85, 1.33)
Panel D: PCA interval forecasts						
Sharpe Ratio	−0.0375	−0.0385	−0.0791	0.0644	0.0817	0.0821
95% CI	(−0.11, 0.04)	(−0.12, 0.04)	(−0.16,0.01)	(−0.03, 0.16)	(−0.04, 0.20)	(−0.06, 0.23)
A_p	−0.3954	−0.2837	−0.4830	0.5768	0.7058	0.5770
95% CI	(−1.15, 0.35)	(−0.85,0.27)	(−0.99, 0.03)	(−0.50, 1.67)	(−0.38, 1.18)	(−0.48, 1.61)
Panel E: ARIMA(1,1,1) interval forecasts						
Sharpe Ratio	0.0177	0.0491	0.0439	0.0659	0.0495	−0.0323
95% CI	(−0.07, 0.09)	(−0.03, 0.13)	(−0.04, 0.13)	(−0.03, 0.16)	(−0.07, 0.17)	(−0.18, 0.11)
A_p	0.1823	0.5480	0.2522	0.7467	0.4476	−0.2393
95% CI	(−0.54, 0.96)	(−0.22, 0.90)	(−0.24, 0.76)	(−0.32, 1.85)	(−0.64, 1.55)	(−1.35, 0.80)
Panel F: ARFIMA(1,d,1) interval forecasts						
Sharpe Ratio	−0.0230	−0.0386	−0.0354	0.0209	−0.0405	0.1107
95% CI	(−0.10, 0.06)	(−0.12,0.04)	(−0.12,0.05)	(−0.07, 0.12)	(−0.16, 0.08)	(−0.03, 0.25)
A_p	−0.3425	−0.3228	−0.2251	−0.0366	−0.4248	0.8353
95% CI	(−1.01, 0.31)	(−0.82, 0.17)	(−0.68, 0.21)	(−1.05, 0.96)	(−1.41, 0.57)	(−0.17, 1.85)

The entries show the annualised Sharpe ratio, Leland (1999) **'sAlpha (A_p) and their respective bootstrapped 95% confidence intervals (CI). The trading game is based on interval forecasts obtained from the economic variables model (Panel A), AR(1) model (Panel B), VAR model (Panel C), PCA model (Panel D), ARIMA(1,1,1) model (Panel E) and ARFIMA(1,d,1) model (Panel F). The Sharpe ratio for the S&P 500 and the Dow Jones Industrial Average is 0.0265 [95%CI = (−0.05, 0.10)] and 0.0319 [95% CI = (−0.04, 0.11)], respectively. One asterisk denotes rejection of the null hypothesis of a zero Sharpe ratio (A_p) at a 5% level of significance.

for the interval forecasts derived by the economic variables (Panel A), AR(1) (Panel B), VAR (Panel C), PCA (Panel D), ARIMA(1,1,1) (Panel E) and the ARFIMA(1,d,1) (Panel F) models. We can see that the obtained SR and A_p are statistically insignificant for all VIX and VXD futures series and for all six models; the same results hold for a naive buy and hold strategy. Therefore, no economically significant profits can be obtained just as was the case with the trading strategy based on point forecasts.[9]

7. Conclusions

This paper has contributed to the literature on whether the evolution of implied volatility can be forecasted in the equity markets by using a number of European and US implied volatility indices. To this end, six alternative model specifications (economic variables, AR(1), VAR, PCA, ARIMA and ARFIMA models) have been employed to generate point as well as interval forecasts. The accuracy of the generated out-of-sample forecasts was evaluated both in a statistical and economic setting. The economic significance was assessed by employing for the first time trading strategies with the VIX and VXD volatility futures.

We found that both the point and interval forecasts are statistically significant. The evidence on the predictability of the point forecasts is stronger for the European indices where the VAR and PCA models perform best among the competing models. In the case of the interval forecasts, the predictable pattern is stronger for the US indices; the VAR model performs best. However, the generated point and interval forecasts are not economically significant.

These results have at least three implications. First, the previous literature that had considered only point forecasts is extended in that it is found that implied volatility can be statistically predicted in both a point and interval forecast setting. Second, the presence of implied volatility spillover effects between the various markets is also confirmed. Finally, the results indicate that the newly CBOE volatility futures markets are informational efficient just as other derivative markets. Given that the answer on the predictability question always depends on the assumed specification of the predictive regression, alternative model specifications should be considered (see e.g., Goyal and Welch, accepted for publication). Also longer horizons can be examined. In the interests of brevity, these topics are best left for future research.

Acknowledgement

We are grateful to Ike Mathur (the Editor) and two anonymous referees for their stimulating, thorough, and constructive

[9] The robustness of the reported results (statistical and economic significance of point and interval forecasts) across various sub-periods was assessed by a recursive "pseudo" out-of-sample scheme (see also Gonçalves and Guidolin, 2006, for a similar approach). First, the sample from February 2, 2001–March 17, 2005 was used to form forecasts for the observations over the next 100 observations (first out-of-sample period). Then, we added these 100 observations to the initial sample and generated forecasts for the next 100 observations (second out-of-sample period) and so forth. Overall, six out-of-sample periods were formed. All models are re-estimated at each time step (i.e. daily). We found that the reported results were not sensitive to the period under consideration. Another robustness check was conducted by implementing the trading strategies without taking into account the CBOE transaction costs. Again, the reported results were not affected.

comments. We would like also to thank Luc Bauwens, Marcelo Fernandes, Jim Gatheral, Gordon Gemmill, Massimo Guidolin, Daniel Giamouridis, Stewart Hodges, George Kapetanios, Alexandros Kostakis, Dimitris Malliaropoulos, Joshua Rosenberg, Lucio Sarno, Alessio Saretto, Dimitris Thomakos, and the participants at the Cass Business School, Federal Reserve Bank of N. York, University of Louvain-La-Neuve (CORE), University of Piraeus, Queen Mary University, Warwick Business School seminar series and the 2006 Europe RISK Quant Congress (London), 2006 Hellenic Finance and Accounting Association (Thessaloniki), 2006 USA RISK Quant Congress (N. York), 2007 Conference on Research in Economic Theory and Econometrics (Naxos), 2007 International Congress on Insurance: Mathematics and Economics (Piraeus), 2007 International Workshop in Economics and Finance (Tripoli), for useful discussions and comments. Financial support from the Research Centre of the University of Piraeus is gratefully acknowledged. Any remaining errors are our responsibility alone. George Skiadopoulos dedicates this paper to the memory of his beloved grandmother Katerina Lykou.

References

Aboura, S., 2003. International transmission of volatility: A study on the volatility indexes VX1, VDAX, VIX. Working paper, ESSEC Business School.

Ahoniemi, K., 2006. Modeling and forecasting implied volatility: An econometric analysis of the VIX index. Working paper, Helsinki School of Economics.

Baillie, R.T., 1996. Long memory processes and fractional integration in econometrics. Journal of Econometrics 73, 5–59.

Banerjee, P.J., Doran, J.S., Peterson, D.R., 2007. Implied volatility and future portfolio returns. Journal of Banking and Finance 31, 3183–3199.

Becker, R., Clements, A.E., White, S.I., 2007. Does implied volatility provide any information beyond that captured in model-based volatility forecasts? Journal of Banking and Finance 31, 2535–2549.

Brooks, C., Oozeer, M.C., 2002. Modeling the implied volatility of options on long gilt futures. Journal of Business Finance and Accounting 29, 111–137.

Carr, P., Wu, L., 2006. A tale of two indices. Journal of Derivatives 13, 13–29.

Christoffersen, P.F., 1998. Evaluating interval forecasts. International Economic Review 39, 841–862.

Corrado, C.J., Miller, T.W., 2006. Estimating expected excess returns using historical and option-implied volatility. Journal of Financial Research 29, 95–112.

David, A., Veronesi, P., 2002. Option prices with uncertain fundamentals: Theory and evidence on the dynamics of implied volatilities. Working paper, University of Chicago.

Dotsis, G., Psychoyios, D., Skiadopoulos, G., 2007. An empirical comparison of continuous-time models of implied volatility indices. Journal of Banking and Finance 31, 3584–3603.

Dumas, B., Fleming, J., Whaley, R.E., 1998. Implied volatility functions: Empirical tests. Journal of Finance 53, 2016–2059.

Fama, E.F., French, K.R., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.

Fernandes, M., Medeiros, M.C., Scharth, M., 2007. Modeling and predicting the CBOE market volatility index. Working paper, Queen Mary University.

Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics. Journal of Finance 58, 1393–1413.

Gemmill, G., Kamiyama, N., 2000. International transmission of option volatility and skewness: When you're smiling, does the whole world smile? Working paper, City University Cass Business School.

Giot, P., 2005. Implied volatility indexes and daily Value-at-Risk models. Journal of Derivatives 12, 54–64.

Gonçalves, S., Guidolin, M., 2006. Predictable dynamics in the S&P 500 index options implied volatility surface. Journal of Business 79, 1591–1635.

Goyal, A., Saretto, A., 2007. Option returns and volatility mispricing. Working paper, Purdue University.

Goyal, A., Welch, I., accepted for publication. A comprehensive look at the empirical performance of equity premium prediction. Review of Financial Studies, doi:10.1093/rfs/hhm014.

Guidolin, M., Timmerman, A., 2003. Option prices under Bayesian learning: Implied volatility dynamics and predictive densities. Journal of Economic Dynamics and Control 27, 717–769.

Guo, D., 2000. Dynamic volatility trading strategies in the currency option market. Review of Derivatives Research 4, 133–154.

Harvey, C.R., Whaley, R.E., 1992. Market volatility prediction and the efficiency of the S&P 100 index option market. Journal of Financial Economics 31, 43–73.

Harvey, D.I., Leybourne, S.J., Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting 13, 281–291.

Hentschel, L., 2003. Errors in implied volatility estimation. Journal of Financial and Quantitative Analysis 38, 779–810.

Leland, H.E., 1999. Beyond mean-variance: Performance measurement in a nonsymmetrical world. Financial Analysts Journal 55, 27–35.

Mixon, S., 2002. Factors explaining movements in the implied volatility surface. Journal of Futures Markets 22, 915–937.

Moraux, F., Navatte, P., Villa, C., 1999. The predictive power of the French market volatility index: A multi horizons study. European Finance Review 2, 303–320.

Pong, S., Shackleton, M.B., Taylor, S.J., Xu, X., 2004. Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. Journal of Banking and Finance 28, 2541–2563.

Poon, S.H., Pope, P.F., 2000. Trading volatility spreads: A test of index option market efficiency. European Financial Management Journal 6, 235–260.

Skiadopoulos, G., Hodges, S.D., Clewlow, L., 1999. The dynamics of the S&P 500 implied volatility surface. Review of Derivatives Research 3, 263–282.

Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American Statistical Association 97, 1167–1179.