Original Article

What drives the implied volatility of index options?

Received (in revised form): 3rd June 2009

Pradosh Simlai
is an assistant professor of Economics at the University of North Dakota, USA. He received his PhD in Economics and MS in Finance, both from the University of Illinois at Urbana-Champaign, USA. His teaching and research interest includes asset pricing, business statistics, corporate finance and financial econometrics. His recent publications are in the areas of empirical asset pricing, portfolio choice, volatility in financial markets, information and market linkages. He has the experience to work at various consultancy positions.

Correspondence: Pradosh Simlai, Department of Economics, College of Business and Public Administration, University of North Dakota, Grand Forks, ND 58202-8369, USA
E-mail: PSimlai@business.und.edu

ABSTRACT **The objective of this article is to identify and evaluate the correlates of volatility in index options during a speculative boom period. We investigate whether the persistence of changes in the CBOE volatility index is associated with Treasury-bill yield and common stock market index, such as S&P 500, during the period of tremendous stock market growth of the late 1990s. We examine if the positive information from the market index and the innovations in the predictive variables play any role in explaining the variability of index options. The implied volatility from our model shows that high volatility of index options is accompanying falling prices in the market index. Overall, the empirical findings suggest that there is more downside weakness than upside strength in the predictive power of the market index.**

Journal of Derivatives & Hedge Funds (2010) **16,** 85–99. doi:10.1057/jdhf.2009.20

Keywords: implied volatility; VIX; conditional variance; endogenous information; volatility forecasts

INTRODUCTION

One of the common tools that investors and analysts use to gauge the market movements and in risk management process is the Chicago Board Options Exchange (CBOE) Volatility Index (VIX). The VIX was introduced by CBOE in 1993 to measure market expectations of near-term volatility implied by standard and Poor's 500 stock index option prices. VIX provides a market-determined, forward-looking estimate of 1-month stock market volatility.[1] It also acts like a signal to anticipatory changes in the market.[2] According to Traub *et al*,[3] VIX 'is a good indicator of the level of fear or greed in US and global capital markets. When investors are fearful, the VIX level is significantly higher than normal' (p. 27). Simon[4] describes how markets tend to view extreme values of VIX as trading signals. Even though common wisdom treats VIX as a measure of 'fear gauge',[5] it can provide a barometer of market volatility and investor sentiment both in good and bad times. As a result, one can easily identify and relate the movement of the VIX to the overall market's confidence and fear.

Since September 2003, CBOE introduced a new version of the VIX. While the original VIX used only at-the money options, the new VIX (known as VXO) utilize a wide range of strike prices in order to incorporate information from the volatility skew. Despite the methodological difference, the fundamental features of VIX remain the same. VIX continues to provide investors consensus view of future expected stock market volatility, and within a short period of time has become the benchmark of stock market volatility forecasts.[6] Both theorists and practitioners utilize it to gauge the information content of implied market volatility. It has been widely cited in a large number of empirical works related to the information flow with implied volatility indices. Some representative references are Canina and Figlewski,[7] Christensen and Prabhala,[8] Fleming,[9] Blair *et al*,[10,11] Martens and Zein,[12] Koopman *et al*,[13] Bandi and Perron,[14] and Bekiros and Georgoutsos.[15]

In the empirical finance, there also exists an extensive literature on the ability of implied volatilities to predict the future volatility of speculative assets. Models using implied volatility of option prices are one prime example (for example, see Day and Lewis,[16] and Noh *et al*[17] for an early overview). In this article we develop a platform of such analysis that is explicitly directed towards the pattern and size of VIX fluctuations. Between the periods of second quarter of 1995 and the first quarter of 2001, stock markets representing technology sectors witnessed a phenomenal increase in their value in the western nations. For example, between 1995 and 2001, the technology-heavy National Association of Securities Dealers Automated (NASDAQ) composite index increased from 600 to 5000 points. Even though Dow Jones Industrial Average index also gained modestly during the same time period, the growth in NASDAQ was quite significant compared to non-tech stock prices. Many of the hot stocks were from the new technology firms that represented biotech, health care and most notably internet sectors and related fields. Our objective in this article is to identify and evaluate the correlates of volatility in index options during this speculative boom period. More specifically, we investigate whether the persistence of changes in the CBOE volatility index[18] is associated with Treasury Bill (T-bill) yield and common market index, such as S&P 500, during the period of tremendous stock market growth of the late 1990s. We examine if the positive information from the market index and the innovations in the predictive variables play any role in explaining the variability of index options. The implied volatility from our model shows that high volatility of index options accompanies falling prices in the market index. Overall, our empirical results demonstrate that there is more downside weakness than upside strength in the predictive power of the market index. This turns out to be very useful in the quantification of information flow.

The remainder of this article is organized as follows. In the next section we introduce various models of performance measurement to price options on a volatility index. The subsequent section contains our main empirical results. These include description of the data, summary statistics and estimation details. In the penultimate section we explore the implication of volatility persistence and outline performance evaluation, using various forecast measures. The final section concludes this article.

MODEL SPECIFICATION

Adoption of volatility options

In this subsection we briefly describe the specification issues in the adoption of volatility option models. Grünbichler and Longstaff[19] present one of the popular models to price options on a volatility index, given by

$$\mathrm{d}V = (\omega - \kappa V)\mathrm{d}t + \sigma\sqrt{V}\mathrm{d}Z \qquad (1)$$

where V is the volatility index, and ω, κ and σ are constants, and Z is a standard Wiener process. The Grünbichler and Longstaff specification (1) models the volatility index as a mean reverting process and keeps the possibility of conditional heteroskedasticity.

As it is well known that various forms of generalized autoregressive conditional heteroskedasticity (GARCH) specification provide a close and parsimonious approximation to the form of heteroskedasticity commonly encountered in financial time series data. In its simple form, the GARCH(1,1) model can be written as

$$h_t = \alpha + \delta h_{t-1} + \gamma\varepsilon_{t-1}^2 \qquad (2)$$

where

$$\varepsilon_t = \sqrt{h_t}z_t, \quad z_t\ i.i.d \sim N(0,1).$$

h_t is conditional volatility, and α, δ and γ are constant parameters. To avoid explosive behavior, it is a tradition to impose the following constraints: $0<\delta<1, 0<\gamma<1$ and $\delta+\gamma<1$ Nelson[20] shows that model (2) can be written as an approximation to a certain diffusion equation that has been postulated in the option pricing literature. Nelson developed conditions under which GARCH stochastic difference equations systems converge in distribution to Ito process and derives the following diffusion limits of the form

$$\mathrm{d}h = (\varphi - \varpi h)\mathrm{d}t + \vartheta h\mathrm{d}Z \qquad (3)$$

where φ, ϖ and ϑ are constant parameters, and Z is a standard Wiener process. The idea of Nelson is to investigate the convergence of stochastic difference equations for volatility models to stochastic differential equations as the length of the discrete time intervals between observations goes to zero. For example, one can rewrite the GARCH(1,1) model (2) as

$$h_t - h_{t-1} = [\alpha - (1-\delta-\gamma)h_{t-1}] + \delta h_{t-1}(Z_{t-1}^2 - 1)$$

which approximate the diffusion process (3) as

$$\mathrm{d}h = (\alpha^* - \varpi^* h)\mathrm{d}t + \delta^* h\mathrm{d}Z_h$$

when $E(\mathrm{d}Z)_h^2 = \mathrm{d}t$, $\alpha \to \alpha^*\mathrm{d}t$, $(1-\delta-\gamma) \to \varpi^*\mathrm{d}t$, and $\delta \to \delta^*\sqrt{\frac{\mathrm{d}t}{2}}$.

As mentioned by Engle and Mustafa,[21] when finer time intervals are used, the estimated GARCH parameters should exhibit small α, small δ and small ϖ. In addition, when $\varpi^*>0$, the diffusion exhibits mean reversion.

Specification of conditional volatility

Our main focus in this article is to identify and evaluate the impact of volatility on the variability of VIX changes, for which we use the following simplest specification:

$$r_t = b_0 + \sum_{i=1}^{k} b_i F_i + \varepsilon_t, \qquad \varepsilon_t = \sqrt{h_t}z_t,$$
$$Z_t\ i.i.d. \sim N(0,1) \qquad (4)$$

where r_t is the changes in VIX between periods $t-1$ and t, F_i are the risk factors, ε_t is the idiosyncratic error and $z_{t-1} = \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$. The potential candidates for F_t include lags of VIX change, level and lags of S&P 500 index returns, and changes in 10-year T-bill returns. Our objective is to use (4) in combination with popular variants of models of changing volatility. The focus is on various GARCH classes of models and their ability to give adequate forecasts. In these models, GARCH(1,1), given by equation (2), is the basic specification yielding time-varying volatility process.[22] In order to have better ability to deliver volatility forecasts, in addition to model (2), we use two other alternative specifications. They are Exponential GARCH (EGARCH) introduced by Nelson,[23] and GJR-GARCH proposed by Glosten *et al.*[24] These two models are very popular in the literature and given by the following:

$$\text{EGARCH}(1,1): \quad \ln(h_t) = \alpha + \delta \ln(h_{t-1}) + \theta[|z_{t-1}| - E(|z_{t-1}|)] + \gamma z_{t-1} \quad (5)$$

$$\text{GJR}-\text{GARCH}(1,1): \quad h_t = \alpha + \delta h_{t-1} + \theta S_{t-1}\varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 \quad (6)$$

where $S_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and 0 otherwise. For specification (5), if $\gamma < 0$, the negative shocks have a bigger impact on future volatility than positive shocks of the same magnitude. The advantage is that the specification allows h_t to respond asymmetrically to positive and negative shocks, and there are no non-negativity constraints of the GARCH parameters. In the specification (6), the dummy variable keeps track of whether the lagged residual error is positive or negative. If the coefficient θ is significant, the implication is that there is asymmetric information spillover effect. The coefficients of the time-varying conditional variance measure both the 'reactiveness' and persistence of volatility to shocks. Engle and Ng[25] find that compared to other asymmetric GARCH model, the GJR version is a better parametric model. Some recent references where the asymmetric relationship between return innovations and volatility is documented are Eraker *et al*,[26] and Bollerslev and Zhou.[27] To estimate all the GARCH models, we use the nonlinear optimization techniques based on Berndt–Hall–Hall–Hausman algorithm.

EMPIRICAL RESULTS

Data and summary statistics

For the empirical illustration we examine the behavior of VIX time series[28] over the 8-year period from September 1995 through August 2003. The data consist of weekly observations of the VIX series calculated by CBOE.[29] The total number of observations in the sample is 345. There is a general notion that aggregation reduces the amount of volatility in the model, the same can be true for VIX series. For example, Nelson[20,23] argues that estimation error owing to volatility model mis-specification decreases as observation frequency increases. Using weekly return data (instead of daily changes) in a modified GARCH framework gives us a platform to test that assertion. For the in-sample analysis we use the data from September 1995 to March 2001, and for out-of-sample analysis we employ the remaining sample observations between April 2001 and August 2003. Among other variables included in our empirical analysis, for S&P 500 series we use daily closing prices, and for risk-free rate we utilize the changes in the 10-year US Treasury yield during the same sample period.

Figures 1 and 2 illustrate the time evolution of VIX series and S&P 500 index level for the full sample period. Compared to the early 1990s (not considered in our sample), during the late 1990s VIX series becomes more volatile. More

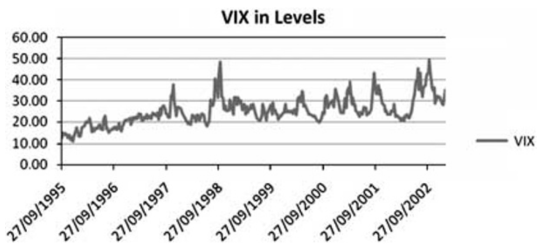

Figure 1: The Chicago Board Options Exchange (CBOE) volatility index, 27 September 1995 – 27 September 2003.

Figure 2: S&P 500 price index, 27 September 1995 – 27 September 2003.

specifically, between the last quarter of 1997 and first quarter of 2003, VIX oscillates in long swings. Between the periods of second quarter of 1995 and the first quarter of 2001, the dot-com boom was in full swing and it is evident that during these periods of excessive speculation the persistence of changes in VIX is somehow closely associated with S&P 500 index returns. Interestingly, spikes in the value of VIX in levels coincide with many economic and political crises (which is consistent with the idea of Whaley[5]) such as the Asian financial crisis of 1997, the Russian and LTCM crisis of 1998, the terrorist attack of 2001, accounting scandals of 2002, the Iraq war in 2003 and so on. It is important to note that the S&P 500 stocks trade on either of two largest American stock markets, the NYSE and NASDAQ. Out of all 500 large-cap common stocks, information technology sector represents roughly 15 per cent of total effective number of constituents. At the beginning of our sample period, there were 44 technology stocks trading in the market, and by early 2000, the trading number went up close to 150. Gradually, we see an increase in technology Initial Public Offerings (IPOs) over the entire span of our sample period, except between December 1998 and January 1999, when various technology stocks were taken off the list as they stopped trading on the market.

In Table 1, we provide some useful preliminary descriptive statistics for the changes in VIX, S&P 500 index and 10-year T-bill. In particular, it reports the sample mean, standard deviation, skewness and excess kurtosis for all the three time series. The overall sample skewness and kurtosis for both VIX and S&P 500 changes indicate high degree of non-normality and strong presence of autocorrelation. The Ljung–Box Q-statistics indicate significant temporal dependence and existence of conditional heteroskedasticity. The reported robust Q-statistics for lags of 10 and 20 support these findings. The Jarque–Bera (JB) normality test statistics indicate strong presence of non-normality as well. Additionally, we employ the augmented Dicky–Fuller and Phillips–Perron unit root test for non-stationarity, using all three time series, but do not find any evidence of non-stationarity (not reported). Finally, we also use the multivariate *t*-test to determine whether the measure of skewness or kurtosis is equal across VIX and S&P 500 series. The results indicate that the null hypothesis of equal skewness and equal kurtosis is rejected for both series. For both VIX and S&P 500 changes, almost all the time, over 50 per cent of the months realize a positive

Table 1: Descriptive statistics

Variable	*VIX change*	*S&P 500 change*	*10-year T-bill change*
Number of observations	345	345	345
% of observations>0	49	54	44
% of observations<0	51	46	56
Mean	0.07	0.19	0.01
t-statistic: mean=0	0.84	2.12	0.38
Standard deviation	3.33	3.02	1.89
Skewness	0.51	0.17	0.45
Excess kurtosis	6.09	5.42	3.58
Ljung–Box portmanteau			
Q(10)	24.7	22.6	15.6
Q(20)	30.3	29.7	13.1
Robust Q-statistic			
Q*(10)	16.2	17.1	13.2
Q*(20)	19.6	20.8	11.3
Autocorrelation			
$\rho(1)$	−0.21	−0.08	0.05
$\rho(10)$	−0.01	0.04	0.01
$\rho(20)$	0.001	0.01	0.02

payoff. The summary statistics also suggest relatively high variance of VIX and S&P 500 returns compared to 10-year T-bill returns.

In panel A of Table 2 we report the correlation matrix of three time series with and without one period lag. There exists relatively low correlation between one-period lagged VIX change and S&P 500 series without any lag. The contemporaneous correlation between VIX change and S&P 500 series is high and negative. In contrast, the contemporaneous correlation between T-bill and VIX change is negative but not very high. Overall, our preliminary investigation suggests a statistically robust role for both S&P 500 and 10-year T-bill. In addition, very low correlations across regressors in panel A suggest the absence of multicollinearity, and further influence their effectiveness as explanatory variable. Panel B of Table 2 reports two sample variance test statistics and their associated *P*-values. The *F* test statistics and corresponding *P*-value suggest that there is not enough evidence to infer that the variances of VIX change and S&P 500 returns (with or without one period lag) differ significantly. The same is not true for the variability between

Table 2: Correlation matrix and variance tests

Variable	*VIX change*		*S&P 500 change*		*10-year T-bill change*	
	No lag	*Lag1*	*No lag*	*Lag1*	*No lag*	*Lag1*
Panel A – Correlation matrix						
VIX no lag	1	—	—	—	—	—
VIX lag 1	−0.21	1	—	—	—	—
S&P 500 no lag	−0.80	0.10	1	—	—	—
S&P 500 lag 1	0.23	−0.80	−0.08	1	—	—
10-year T-bill no lag	−0.15	−0.04	0.11	0.08	1	—
10-year T-bill lag 1	0.08	−0.15	−0.06	0.11	0.06	1
Panel B – Two sample variance tests statistics (Null hypothesis (H_0): Equal variance, Alternative hypothesis (H_1): Unequal variance)						
10-year T-bill change	607.27	605.45	—	—	—	—
P-value	0.00	0.00	—	—	—	—
VIX change	—	—	1.21	0.81	—	—
P-value	—	—	0.07	0.97	—	—
S&P 500 Change	—	—	—	—	500.60	498.01
P-value	—	—	—	—	0.00	0.00

10-year T-bill returns and VIX change (or S&P 500 returns).

Estimation results and interpretation

Table 3 presents the results of various estimated regression models for changes in VIX series. The reported results suggest that the estimated intercept coefficient is significant, unless we include positive S&P 500 change as an explanatory variable in the mean model specification. The one period lag of VIX change is statistically significant only for models 1 and 2. The only variable that is statistically and economically significant for all five models is the return of the S&P 500 index. As expected from the descriptive statistics of Table 2, the slope coefficient for S&P 500 index returns is negative throughout. This indicates that a negative contemporaneous S&P 500 index return increases the VIX and a positive contemporaneous S&P 500 index return lowers the VIX. In contrast, the coefficient of one-period lagged S&P 500 index return is small but positive. As we can see for all the estimated models 1 through 5, no second-order lags turned out to be significant. The slope coefficient of changes in 10-year T-bill indicates a strong negative contemporaneous correlation with the VIX changes. The average value of adjusted R^2 is 0.58, with model 5 capturing the highest figure. The descriptive statistics for the estimated residuals do not change markedly across all the models. Only the skewness coefficient of the estimated residuals increases substantially for

Table 3: Estimation results of the simple linear regression models[a]

Independent variables	*Dependent variable: weekly % change in CBOE volatility index* *Models*				
	1	*2*	*3*	*4*	*5*
Constant	0.23*	0.22*	0.17*	0.16*	0.03
Lag 1 of VIX change	−0.13*	−0.13*	0.02	0.03	0.04
Lag 2 of VIX change	—	—	—	0.01	—
S&P 500 change	−0.87*	−0.86*	−0.86*	−0.86*	−1.11*
Lag 1 of S&P 500 change	—	—	0.21*	0.19*	0.26*
Lag 2 of S&P 500 change	—	—	—	0.05	—
Change in 10-year T-bill	—	−1.73*	−1.95*	−1.94*	−1.76*
Positive S&P 500 change	—	—	—	—	0.47*
Adj-R^2	0.55	0.58	0.59	0.59	0.61
Skewness	0.80	0.77	0.64	1.43	1.35
Kurtosis	6.73	6.71	5.91	5.94	5.54
JB tests	21.50	20.85	21.49	21.47	18.55
GARCH tests	65.41	45.08	47.25	39.10	36.79
Ljung–Box portmanteau					
Q(20)	26.94	30.45	18.59	16.79	21.84
$Q^2(20)$	32.05	33.10	26.96	22.00	46.73
Autocorrelation					
$\rho(1)$	−0.02	−0.01	−0.03	−0.03	−0.05
$\rho^2(1)$	−0.11	−0.12	−0.09	−0.07	−0.09

[a]Here, (*) means statistically significant at the 5% level. The *t*-statistics for all the models are based on Newey-West standard errors with three lags. JB test is the Jarque–Bera test statistic for non-normality. The null and alternative hypotheses for the tests for GARCH effects are no GARCH effects and GARCH(1,1) disturbance, respectively. Q(20) is the Ljung–Box statistic of residuals. $Q^2(20)$ is the Ljung–Box statistic of squared residuals. $\rho(1)$ is the first-order autocorrelation of residuals. $\rho^2(1)$ is the first-order autocorrelation of squared residuals.

models 4 and 5. Overall, in terms of diagnostic statistics, the residuals are somewhat right skewed, leptokurtic and highly non-normal. The Ljung–Box portmanteau statistics for both residuals and square residuals display serious temporal dependence. The test for GARCH effects supports the existence of conditional heteroskedasticity.

In Table 4, we re-evaluate our basic regression models of Table 3, by augmenting them with alternative GARCH specifications. Models 1–5 in Table 4 are based on the GARCH(1,1) specification (2), model 6 is based on the EGARCH(1,1) specification (5) and model 7 is based on the GJR-GARCH(1,1) specification (6).[30] The estimated results show that both GARCH(1,1) parameters, δ and γ, are significant for all models 1–5, indicating a significant presence of conditional variance and high degree of persistence.[31] The same is true when we introduce EGARCH and GJR-GARCH specification within the model. This outcome is not surprising. By simply fitting a mean model to weekly VIX return series (that is, models 1–5 in Table 3), we saw that the residuals have statistically significant unconditional skewness and kurtosis. The JB normality test statistics in Table 3 also suggest serious departures from normality. It is this remaining non-normality and time-varying

Table 4: Estimation results of the conditionally heteroskedastic regression models[a]

Independent variables	*Dependent variable: Weekly % change in CBOE volatility index Models*						
	1	*2*	*3*	*4*	*5*	*6*	*7*
Constant	0.20*	0.19*	0.18*	0.18*	−0.13	−0.10	−0.08
Lag 1 of VIX change	−0.13*	−0.14*	0.02	0.004	0.03	0.06	0.07
Lag 2 of VIX change	—	—	—	0.01	0.01	0.04	0.03
S&P 500 change	−0.85*	−0.84*	−0.84*	−0.85*	−1.12*	−1.14*	−1.09*
Lag 1 of S&P 500 change	—	—	0.21*	0.20*	0.26*	0.29*	0.31*
Lag 2 of S&P 500 change	—	—	—	0.04	0.07	0.08	0.07
Change in 10-year T-bill	—	−1.78*	−1.97*	−2.00*	−1.59*	−1.48*	−1.45*
Positive S&P 500 change	—	—	—	—	0.54*	0.66*	0.73*
α	0.90	0.88	0.79	0.79	0.66	0.23	1.02
γ	0.11	0.12	0.12	0.11	0.18	−0.04	0.16
δ	0.55	0.57	0.56	0.61	0.63	0.65	0.62
θ	—	—	—	—	—	0.17*	0.14*
Adj-R^2	0.66	0.67	0.68	0.70	0.71	0.73	0.74
Ljung–Box portmanteau							
Q(20)	11.55	12.24	11.90	13.08	13.21	13.72	14.40
$Q^2(20)$	10.21	11.52	9.20	10.59	11.03	10.07	10.56

[a]Here, (*) means statistically significant at the 5% level. The *t*-statistics for all the models are based on Newey-West standard errors with three lags. Q(20) is the Ljung–Box statistic of residuals. $Q^2(20)$ is the Ljung–Box statistic of squared residuals. Models 1–5 are based on the GARCH(1,1) specification: $h_t = \alpha + \gamma\varepsilon_{t-1}^2 + \delta h_{t-1}$. Model 6 is based on the EGARCH(1,1) specification: $\ln(h_t) = \alpha + \delta \ln(h_{t-1}) + \theta[|z_{t-1}| - E(|z_{t-1}|)] + \gamma z_{t-1}$. Model 7 is based on the GJR-GARCH(1,1) specification: $h_t = \alpha + \gamma\varepsilon_{t-1}^2 + \delta h_{t-1} + \theta S_{t-1}\varepsilon_{t-1}^2$, where S_{t-1}=1 if $\varepsilon_{t-1} < 0$, and 0 otherwise.

heteroskedasticty that we hope to correct for by augmenting various mean models with variants of GARCH specifications. The diagnostics statistics of the standardized residuals for various models in Table 4 suggest that we are indeed successful to do so. The Ljung–Box test statistics values for the first 20 residuals or squared residuals are not significant at conventional levels. On the flip side, the average explanatory power of the regressions increases to 0.70, with the GJR-GARCH model capturing the highest value of 0.74.

The coefficient estimate of lagged dependent variable in the mean equation is significant only for models 1 and 2, suggesting a strong presence of endogenous information. The estimated risk premium for S&P 500 is always negative and statistically significant at the conventional 5 per cent level. This outcome is not surprising and is in line with Fleming *et al*[32] and Bollen and Whaley.[33] As we include the lagged values of S&P 500 change, the endogenous effect becomes insignificant for VIX changes. The estimation result for models 3–7 suggests that one-period lagged S&P 500 returns always have a statistically significant positive slope coefficient. Interestingly, the positive information from S&P 500 change also has a statistically significant and positive effect on the variability of VIX change. It even makes the intercept coefficient insignificant, as all the regression intercepts for models 5–7 are consistently small. We also create an explanatory variable using the negative S&P 500 changes (not reported) but found out that, when used either jointly or separately, it is not statistically significant at all. This suggests an independent role of the information flow from positive component of S&P 500 returns. Similar to Table 3, in Table 4 we observe that all the estimated models are first-order models throughout. The estimation result of GJR-GARCH model parameters suggests that only the negative innovations have an impact on volatility. For EGARCH conditional variance specification (that is, model 6 in Table 4), the slope coefficient γ is significantly negative and θ is significantly positive. This suggests that EGARCH model does a good job in capturing asymmetric response of shocks to volatility. Therefore, our empirical results demonstrate that (1) there is a significant spillover effect from contemporaneous S&P 500 index returns to VIX changes; (2) even for weekly return data, complementing homoskedastic variance by time-varying conditional variance always improves the performance of risk valuation model; and (3) the significance of information flow from positive S&P 500 returns on VIX change is not linked to volatility persistence or the asymmetric response of shocks to volatility.

VOLATILITY PERSISTENCE AND PERFORMANCE EVALUATION USING MODEL FORECASTS

At this point, we are convinced about the pervasive role of endogenous information and contemporaneous effect of S&P 500 index returns on VIX changes. One important item that we have not talked about is the issue over parameterization. One of the easiest ways to check this issue is to use in-sample goodness-of-fit statistics and out-of-sample volatility forecasts corresponding to alternative models described in Table 4. As we mentioned before, our focus is to test the ability of various model specifications to deliver better volatility forecasts. In order to compare the in-sample and out-of-sample performance of various model specifications, following Marcucci,[34] we utilize four common statistical loss functions. They are two versions of mean squared error (MSE) and mean absolute deviation (MAD), given by the following:

$$\mathrm{MSE1} = \frac{1}{n}\sum_{t=1}^{n}\left(\hat{\sigma}_{t+1}^{\frac{1}{2}} - \hat{h}_{(t+1|t)}^{\frac{1}{2}}\right)^2,$$

$$\mathrm{MSE2} = \frac{1}{n}\sum_{t=1}^{n}\left(\hat{\sigma}_{t+1} - \hat{h}_{(t+1|t)}\right)^2$$

$$\mathrm{MAD1} = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{\sigma}_{t+1}^{\frac{1}{2}} - \hat{h}_{(t+1|t)}^{\frac{1}{2}}\right|,$$

$$\mathrm{MAD2} = \frac{1}{n}\sum_{t=1}^{n}\left|\hat{\sigma}_{t+1} - \hat{h}_{(t+1|t)}\right|$$

where $\hat{\sigma}_{t+1}$ is the true volatility and $\hat{h}_{(t+1|t)}$ is the one-step ahead volatility forecast made at time t. For in-sample results, the estimation period is from September 1995 through March 2001. For the evaluations of out-of-sample performance, we use a little more than the last 2 years of observation from April 2001 through August 2003.

Table 5 reports the results. Here, k refers to the number of the parameters being estimated in each model, logL is the log-likelihood value, AIC is the Akaike information criterion, SIC is the Schwartz information criterion and PER is the persistence of the shocks to the conditional variance. Models 1–7 are described in Table 4. We can see that, in terms of log-likelihood, the largest value corresponds to model 7 (that is, GJR-GARCH specification of Table 4). For the simple GARCH class of models 1–5, the largest log-likelihood value belongs to model 5 (that is, GARCH(1,1) specification of Table 4). In terms of model selection criterion, both AIC and SIC indicate that the best model is GJR-GARCH specification of Table 4. However, the persistence parameter of the GJR-GARCH model is not the highest, even though the estimate is higher compared to EGARCH model 1 (that is, model 6 of Table 4). In terms of MSE- and MAD-based performance, model 7 always performs at least as good as the other two. The only exception is MAD2, as model 5 has slightly lower value compared to model 7. Interestingly, the EGARCH specification always performs better than models 1–4, suggesting that these intermediate models are unable to capture entirely the variation in volatility. It also establishes the fact that shocks to systematic volatility is an important determinant of VIX changes.

In terms of out-of-sample results, performance of models 5 and 7 is consistent with our earlier conclusion. The results in Table 6 suggest that the pair of MSE and MAD figures corresponding to GJR-GARCH specification is always the lowest (except MSE2). For the EGARCH specification, only the MAD2 is slightly better than the MAD1 of models 1 through 5. In order to measure the accuracy of sign of volatility forecasts for various GARCH models, we also use a sign test, known as the so-called success ratio (SR). The SR test is simply based on the fraction of the volatility forecasts that has the same sign as the volatility realizations and is given by

$$\mathrm{SR} = \frac{1}{n}\sum_{j=0}^{n-1} I_{\{\bar{\sigma}_{t+j}\bar{h}_{t+j|t+j-1}>0\}}$$

where $\bar{\sigma}_{t+j}$ is the proxy for the actual volatility after subtracting its non-zero mean, $\bar{h}_{(t+j|t+j-1)}$ is the volatility forecasts after subtracting the corresponding mean and $I_{\{g>0\}}$ is an indicator function, that is, $I_{\{g>0\}} = 1$ if g is positive and 0 otherwise. From the results of the last column of Table 6, we observe that, overall, model 7 for

Table 5: In-sample performance of various regression models[a]

Models	Goodness-of-fit statistics					In-sample performance			
	k	*logL*	*AIC*	*SIC*	*PER*	*MSE1*	*MSE2*	*MAD1*	*MAD2*
1	6	−3620.1	−3631.5	−3622.3	0.66	5.24	1.66	1.06	1.05
2	7	−3455.7	−3487.6	−3497.1	0.69	5.10	1.57	1.04	0.99
3	8	−3448.1	−3454.7	−3455.5	0.68	5.08	1.53	1.01	0.98
4	10	−3225.6	−3213.0	−3216.8	0.72	5.12	1.56	1.03	0.99
5	11	−3188.4	−3194.1	−3196.8	0.81	5.01	1.51	0.97	0.91
6	12	−3170.2	−3172.3	−3174.6	0.61	5.00	1.52	0.98	0.93
7	12	−3152.1	−3164.3	−3166.1	0.79	4.98	1.49	0.96	0.90

[a]Here, *k* refers to the number of parameters being estimated in each model; logL is the log-likelihood value; AIC is the Akaike information criterion; SIC is the Schwartz information criterion; PER is the persistence of the shocks to the conditional variance; and MSE1, MSE2, MAD1 and MAD2 are the loss functions as defined in the text. Models 1–7 are detailed in Table 4.

GJR-GARCH specification does the best job in correctly predicting the sign of future volatility of VIX changes. Similarly, model 5 for GARCH (1,1) specification has the best ability to predict the correct sign of the volatility forecasts in its class.

We can clearly see the support of the above findings through the graphical representation of actual volatility and one-step-ahead volatility forecasts. In Figure 3, we have percentage change and conditional standard deviation of VIX returns. In Figures 4–6 we depict the one-step-ahead volatility forecasts for VIX returns with GARCH, EGARCH and GJR-GARCH specifications, respectively. It is clear that GJR-GARCH models do a great job in terms of their ability to predict correct forecasts, as the one-period-ahead forecasts from model 7 are pretty close to the true realizations. GARCH specification is better than EGARCH, but fails to disentangle the spikes in the VIX changes.

Table 6: Out-of-sample performance of various regression models[a]

Models	Out-of-sample performance				Sign test
	MSE1	*MSE2*	*MAD1*	*MAD2*	*SR*
1	22.07	7.65	1.96	2.46	0.65
2	22.01	7.41	1.93	2.40	0.67
3	21.28	7.32	1.91	2.37	0.71
4	20.66	6.98	1.89	2.24	0.74
5	19.04	5.77	1.59	2.07	0.80
6	19.52	5.91	1.63	2.01	0.79
7	18.55	5.80	1.52	2.00	0.82

[a]Here, MSE1, MSE2, MAD1 and MAD2 are the loss functions as defined in the text. Models 1–7 are detailed in Table 4.

CONCLUSIONS

As a measure of market risk, VIX is one of the popular and widely followed indexes. In this article, we study the persistence of changes in

  1753-9641 **Journal of Derivatives & Hedge Funds** Vol. 16, 2, 85–99

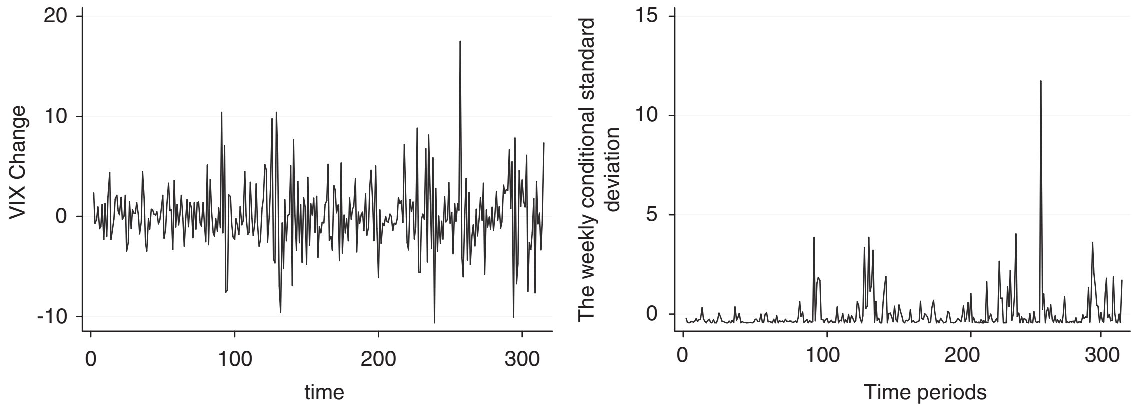

Figure 3: Percentage change and conditional standard deviation of VIX returns.

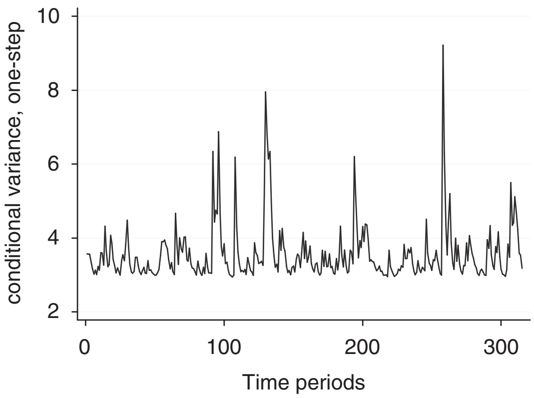

Figure 4: Predicted conditional variance of VIX returns, using GARCH specification.

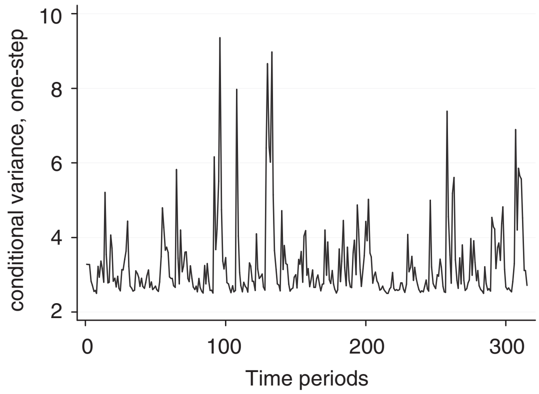

Figure 5: Predicted conditional variance of VIX returns, using EGARCH specification.

VIX series. Using various mean and variance specifications we investigate if there are any information spillovers between VIX changes and common market index such as S&P 500 returns. We evaluate the performance of various models to test the forecasting ability. We observe that GJR-GARCH specification seems to do a better job at forecasting the variation in actual volatility of VIX changes. Our result indicates a stronger presence of significant information flow from the market index to VIX change than has been previously shown.

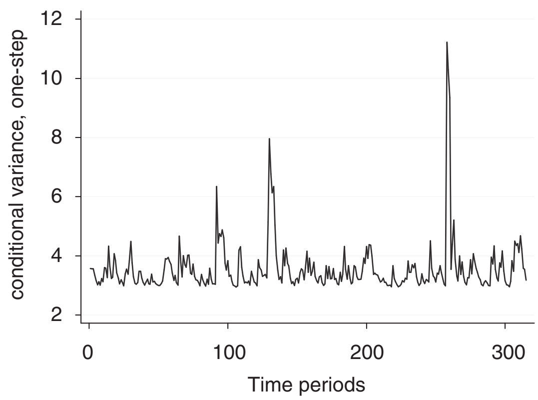

Figure 6: Predicted conditional variance of VIX returns, using GJR-GARCH specification.

REFERENCES AND NOTES

1 Hentschel, L. (2003) Errors in implied volatility estimation. *Journal of Financial and Quantitative Analysis* 38(4): 779–810.

2 Whaley, R.E. (1993) Derivatives on market volatility: Hedging tools long overdue. *Journal of Derivatives* 1(1): 71–84.

3 Traub, H., Ferreira, L., McArdle, M. and Antognelli, M. (2000) Fear and greed in global asset allocation. *Journal of Investing* 9(1): 27–32.

4 Simon, D.P. (2003) The NASDAQ volatility index during and after the bubble. *Journal of Derivatives* 11(2): 9–22.

5 Whaley, R.E. (2000) The investor fear gauge. *Journal of Portfolio Management* 26(3): 12–17.

6 Harvey, A. and Whaley, R.E. (1991) S&P 100 index option volatility. *Journal of Finance* 46(4): 1551–1561.

7 Canina, L. and Figlewski, S. (1993) The informational content of implied volatility. *Review of Financial Studies* 6(3): 659–681.

8 Christensen, B.J. and Prabhala, N.R. (1998) The relation between implied and realized volatility. *Journal of Financial Economics* 50(2): 125–150.

9 Fleming, J. (1998) The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance* 5(4): 317–345.

10 Blair, B.J., Poon, S.H. and Taylor, S.J. (2001a) Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency index returns. *Journal of Econometrics* 105(1): 5–26.

11 Blair, B.J., Poon, S.H. and Taylor, S.J. (2001b) Modeling S&P 100 volatility: The information content of stock returns. *Journal of Banking and Finance* 25(9): 1665–1679.

12 Martens, M. and Zein, J. (2004) Predicting financial volatility: High-frequency time-series forecasts vis-'a-vis implied volatility. *Journal of Futures Markets* 24(11): 1005–1028.

13 Koopman, S.J., Jungbacker, B. and Hol, E. (2005) Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements. *Journal of Empirical Finance* 12(3): 445–475.

14 Bandi, F.M. and Perron, B. (2006) Long memory and the relation between implied and realized volatility. *Journal of Financial Econometrics* 4(4): 636–670.

15 Bekiros, S.D. and Georgoutsos, D.A. (2008) Non-linear dynamics in financial asset returns: The predictive power of the CBOE volatility. *The European Journal of Finance* 14(5): 397–408.

16 Day, T.E. and Lewis, C.M. (1993) Forecasting futures market volatility using alternative models of conditional volatility. *Journal of Derivatives* 1(2): 33–50.

17 Noh, J., Engle, R. and Kane, A. (1994) Forecasting volatility and option prices of the S&P 500 index. *Journal of Derivatives* 2(1): 17–30.

18 See http://www.cboe.com/micro/vixfor more information.

19 Grünbichler, A. and Longstaff, F. (1996) Valuing futures and options on volatility. *Journal of Banking and Finance* 20(6): 985–1001.

20 Nelson, D. (1990) ARCH models as diffusion approximations. *Journal of Econometrics* 45(1–2): 7–38.

21 Engle, R.F. and Mustafa, C. (1992) Implied ARCH models from option prices. *Journal of Econometrics* 52(1–2): 289–311.

22 Engle, R.F. and Bollerslev, T.P. (1986) Modeling the persistence of conditional variances. *Econometric Review* 5(1): 1–50.

23 Nelson, D. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59(2): 347–370.

24 Glosten, L., Jagannathan, R. and Runkle, D. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48(5): 1779–1801.

25 Engle, R.F. and Ng, V.K. (1993) Measuring and testing the impact of news on volatility. *Journal of Finance* 48(5): 1749–1778.

26 Eraker, B., Johannes, M. and Polson, N. (2003) The impact of jumps in volatility and returns. *Journal of Finance* 58(3): 1269–1300.

27 Bollerslev, T.P. and Zhou, H. (2006) Volatility puzzles: A simple framework for gauging return-volatility regressions. *Journal of Econometrics* 131(1–2): 123–150.

28 The VIX is quoted on an annualized basis. For example, if the VIX is at 10, this implies that the expected annualized change is 10 per cent over the next 30 days.

29 Note that the calculation of VIX involves volatility of a variance swap (which can be replicated through vanilla puts and calls) but not that of a volatility swap (which requires dynamic hedging).

30 The intermediate regression results from EGARCH(1,1) and GJR-GARCH(1,1) model estimation are very similar to those reported in the paper using the GRACH(1,1) model and are omitted for sake of brevity. However, they are readily available upon request.

31 As we see from our discussion on equation (2), by volatility persistence we mean that shocks to the current volatility tend to stay for long periods well into the future, that is, if the estimated value of $\delta+\gamma \rightarrow 1$. One can visualize that any shocks to volatility will persist forever if the estimated value of $\delta+\gamma$ becomes one. Engle and Bollerslev[22] coined this type of conditional variance process as Integrated-GARCH. There are different ways to treat such non-stationary process. As

in our article most of the time the value of $\delta+\gamma$ is less than one, we restrict our discussion to the stationary case.

32 Fleming, J., Ostdiek, B. and Whaley, R.E. (1995) Predicting stock market volatility: A new measure. *Journal of Futures Markets* 15(3): 265–302.

33 Bollen, N.P.B. and Whaley, R.E. (2004) Does net buying pressure affect the shape of implied volatility functions? *Journal of Finance* 59(2): 711–753.

34 Marcucci, J. (2005) Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics* 9(4): 1–53.