The Information Content of the Implied Volatility Term Structure on Future Returns

Yaw-Huei Wang & Kuang-Chieh Yen*

ABSTRACT

We adopt the Heston (1993) stochastic volatility (SV) model framework to examine the theoretical relationship between the term structure of implied volatility and the expected excess returns of underlying assets. Three alternative approaches are adopted for our compilation of the variables representing the information on the squared VIX level and term structure in support of our empirical investigation of the information content of the level and term structure variables on future excess returns in the S&P 500 index. Our empirical results provide support for the important role of the term structure in the determination of future excess returns, with such predictive power being discernible for various horizons. Overall, the information content of the term structure variable is found to be significant, and indeed, a strong complement to that of the level variable. In particular, due to the mean-reversion behavior of volatility, the information in the term structure of implied volatility is found to be very effective in the prediction of shorter-term excess returns.

Keywords: VIX term structure, Predictability, S&P 500 index returns.

JEL Classification: G13, G14

1. INTRODUCTION

According to the methodology proposed by the CBOE for the implementation of the model-free implied volatility formula (developed by Jiang and Tian, 2005),[1] the computation of the implied volatility index (VIX) is based upon consideration of all of the available market prices of the S&P 500 index options. Such an approach facilitates the approximation of the expected aggregate volatility of the S&P 500 index during the subsequent 30 calendar-day period,[2] and indeed, this method has been used not only as a measure of sentiment, but also as an instrument for timing the market, particularly in the aftermath of the sub-prime mortgage crisis.

Following Whaley (2000), who proposed the use of the VIX as an effective fear indicator, Giot (2005) subsequently identified a strongly negative correlation between contemporaneous changes and future market index returns, along with a positive correlation between such future returns and current levels of the implied volatility indices. Both Guo and Whitelaw (2006) and Banerjee, Doran and Peterson (2007) went on to report similar findings; for example, Banerjee et al. (2007) derived the theoretical relationship between the level and innovation of the VIX and

---

[1]   The CBOE's revision of the methodology of the VIX formula was based largely upon the results of Carr and Madan (1998) and Demeterfi et al. (1999). As regards theoretical fundamentals, Britten-Jones and Neuberger (2000) derived model-free implied volatility under the diffusion assumption, with Jiang and Tian (2005) subsequently extending this to the jump diffusion assumption.
[2]   When it was first introduced in 1993, the VIX was originally compiled from the implied volatility of eight S&P 100 index options, comprising of near at-the-money, nearby and second nearby calls and puts; however, ever since 2003, the VIX has been calculated from the prices of S&P 500 index options using a model-free formula with almost all of the available contracts; that is, with a wide range of strike prices.

future returns by adopting the stochastic volatility (SV) model of Heston (1993), with their results providing empirical support for the predictive ability of the VIX-related variables on future portfolio returns.

Since the current version of the VIX is compiled for each maturity period with the incorporation of almost all of the available contracts, it should prove to be more informative than the old version when used to investigate its predictive ability on future equity returns; however, the VIX remains dependent upon maturity periods, with the 30-day version being the most frequently used. This therefore gives rise to the interesting and important question of whether any useful information is contained in the VIX term structure for potential use in the forecasting of returns.[3]

Furthermore, since it is well known that volatility has some special stylized facts, such as clustering and mean-reversion, the relative positions of the VIX levels for different maturity periods may imply the expectations of market participants on market volatility, and thus, on changes in the S&P 500 index due to the mean-variance relationship in the conventional theory of the 'capital asset pricing model' (CAPM). Therefore, our aim in the present study is to contribute to the extant related literature by comprehensively investigating whether the VIX term structure contains any useful information on future returns in the S&P 500 index.

---

[3]　See for example, Egloff, Leippold and Wu (2010), Bakshi, Panayotov and Skoulakis (2011), Duan and Yeh (2011), Johnson (2012), Luo and Zhang (2012b), Feunou, Fontaine, Taamouti, and Tedongap (2014).

Based upon the SV model framework of Heston (1993), we refine the theoretical work of Banerjee et al. (2007) by going on to derive a theoretical model which reveals positive relationships between expected excess returns in the S&P 500 index and both the squared VIX levels and the difference between forward and current squared VIX levels. Since the forward squared VIX level can be computed from two squared VIX values with different horizons, it can therefore be regarded as a proxy for the VIX term structure. Hence, this model provides theoretical fundamentals for the potential of the VIX term structure with regard to the prediction of excess returns in the S&P 500 index.

We propose three alternative empirical methods for compiling the variables representing the information in both the squared VIX level and its term structure in order to investigate whether the information content of the term structure provides an additional contribution to the forecasting of future excess returns in the S&P 500 index.

Firstly, adhering closely to the theoretical model, we use the 30-day squared current VIX level and the difference between the 30-day forward squared VIX and the 30-day squared current VIX level as the respective level and term structure variables. Secondly, we run a 'principal component analysis' (PCA) of the squared VIX levels under various time horizons to generate the first and second components for our investigation; the first and second components respectively represent the

level and slope of the VIX term structure. As for our third approach, we employ the two-factor SV model proposed by Egloff et al. (2010) as the means of transforming the maturity-dependent squared VIX values into maturity-independent instantaneous variance and stochastic central tendency, and then use the former along with the difference between the latter and the former to serve as the respective proxies for the squared VIX level and term structure.

Our findings reveal that the information content of the term structure plays an important role in the prediction of excess returns, regardless of which approach is used for the compilation of the variables; furthermore, this predictive power is also discernible for the excess returns under various horizons. When comparing the incremental contribution provided by the level and term structure variables with regard to the effective prediction of returns, we find that the information content of the squared VIX term structure is significant, and indeed, a strong complement to that of the squared VIX level. In particular, as a result of the mean-reversion behavior of volatility, the term structure variable is found to be more informative over shorter-term prediction horizons. Our empirical results are also found to be insensitive to the use of either overlapping or non-overlapping data.

Attempts have been made in recent studies to estimate specific price dynamics by

incorporating the VIX term structure under a two-factor volatility framework;[4] however, the information content of the VIX term structure on future excess returns in the S&P 500 index has yet to attract any significant attention. Cochrane and Piazzesi (2005) provided an effective approach for extracting the information content of the interest rate term structure for the prediction of excess bond returns, with Bakshi, Panayotov and Skoulakis (2011) and Luo and Zhang (2012b) subsequently applying the approach to their empirical investigations of the relationship between forward variances and future excess returns; however, neither study offered any theoretical fundamentals for the prediction of these relationships.

Furthermore, although Johnson (2012) and Feunou et al. (2014) derived a state-dependent relationship between S&P 500 excess returns and the VIX for different maturity periods, the direction of the impact remained somewhat ambiguous. Our study therefore contributes to the literature by not only proposing a theoretical model but also providing empirical evidence in support of the important role of the information in the VIX term structure on the prediction of future excess returns in the S&P 500 index.

The remainder of this paper is organized as follows. The theoretical fundamentals adopted for our study are provided in Section 2, followed in Section 3

---

[4]   Examples include Egloff et al. (2010), Duan and Yeh (2011) and Luo and Zhang (2012a).

by a description of the data and empirical methodologies used for our empirical analyses. The empirical results and robustness tests are subsequently presented in Section 4. Finally, the conclusions drawn from this study are presented in Section 5.

2.   THEORETICAL FUNDAMENTALS

Giot (2005) provided support for the predictive power of current implied volatility on the returns of the underlying asset, whilst Banerjee et al. (2007) demonstrated that the changes in implied volatility from the previous period were also associated with the future returns of the underlying asset. Subsequently, Bakshi et al. (2011) empirically showed that forward variance could predict excess returns in the S&P 500 index. Moreover, Feunou et al. (2014) show that the term structure of option-implied variance can reveal two predictors of the equity premium.

In the present study, we propose a model for examining the relationship between future excess returns and the term structure of implied volatility by referring to the Banerjee et al. (2007) model. Following the stochastic volatility (SV) setting outlined in Heston (1993), we specify the dynamics of the asset prices, S_t, and the variance in the asset prices, V_t, under the real-world P measure as:

$$dS_t = \mu S_t dt + \sqrt{V_t} S_t dW_t^P, \tag{1}$$

$$dV_t = \kappa(\theta - V_t)dt + \lambda_V V_t + \sigma_V \sqrt{V_t} dZ_t^P, \tag{2}$$

and $cov(dW_t^P, dZ_t^P) = \rho dt$, where $\mu = r_f + \lambda V_t$ is the expected return; r_f is the risk-free rate;

κ is the volatility mean-reversion speed; θ is the long-run variance level; σ_V is the volatility of the volatility; ρ is the correlation between the price and variance innovations; and W_t^P and Z_t^P, and λ and λ_v are the respective market prices of the price and variance risks. Using the Girsanov theorem to transform the real-world processes into their risk-neutral Q-measure equivalents, we have the following Q-measure dynamics:

$$dS_t = r_f S_t dt + \sqrt{V_t} S_t dW_t^Q, \tag{3}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t} dZ_t^Q, \tag{4}$$

and $cov(dW_t^Q, dZ_t^Q) = \rho dt$, where W_t^Q and Z_t^Q are the respective Q-measure price and variance innovations.

By discretizing the P-measure dynamics in Equations (1) and (2) and then replacing V_t with the realized variance, $\sigma_{RV,t}^2$, due to the P-measure, we obtain:

$$S_{t+1} - S_t = \left(r_f + \lambda \sigma_{RV,t}^2\right) S_t + \sigma_{RV,t} S_t \epsilon_t \tag{5}$$

and $\sigma_{RV,t+1}^2 - \sigma_{RV,t}^2 = \kappa\left(\theta - \sigma_{RV,t}^2\right) + \lambda_V \sigma_{RV,t}^2 + \sigma_V \sigma_{RV,t}\left(\rho \epsilon_t + \sqrt{1 - \rho^2}\, \epsilon_t^V\right), \tag{6}$

where ε_t and ε_t^V are the discrete-version price variance innovations under the P-measure and ε_t and ε_t^V are uncorrelated (i.e. $Cov(\epsilon_t, \epsilon_t^V) = 0$).

Similarly, by discretizing the Q-measure volatility dynamic in Equation (4), or applying the Girsanov theorem to Equation (6), and then replacing V_t with the implied variance, $\sigma_{IV,t}^2$, due to the Q-measure, we obtain:

$$\sigma_{IV,t+1}^2 - \sigma_{IV,t}^2 = \kappa\left(\theta - \sigma_{IV,t}^2\right) + \sigma_V \sigma_{IV,t}\left(\rho\epsilon_t^* + \sqrt{1-\rho^2}\,\epsilon_t^{V*}\right), \tag{7}$$

where ε_t^{V*} and ε_t^{V*} the discrete-version price variance innovations under the Q-measure

and they are uncorrelated (i.e. $Cov(\epsilon_t^*, \epsilon_t^{V*}) = 0$).

In order to connect the price risk with the variance risk, we assume that:

$$\lambda = -\delta\lambda_V, \tag{8}$$

where δ is restricted to be positive.

Whilst the price risk premiums, λ, are widely known to be positive, in some of the related studies, such as Carr and Wu (2009), the variance risk premiums, λ_V, have been found to be strongly negative; our finding of a proportional relationship between the two risk premiums would therefore appear to be consistent with the findings in the extant literature.[5]

By applying Equation (8) to Equation (6), and then taking the difference with Equation (7), we obtain:

$$\lambda\sigma_{RV,t}^2 = \delta\left(\Delta\sigma_{IV,t}^2 - \Delta\sigma_{RV,t}^2\right) + \kappa\delta\left(\sigma_{IV,t}^2 - \sigma_{RV,t}^2\right) + \delta g_1(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}) \tag{9}$$

where $\quad g_1(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}) = \sigma\left(\sigma_{RV,t}\left(\rho\epsilon_t + \sqrt{1-\rho^2}\epsilon_t^V\right) - \sigma_{IV,t}\left(\rho\epsilon_t^* + \sqrt{1-\rho^2}\epsilon_t^{V*}\right)\right)$

and $\Delta\sigma_t^2 = \sigma_{t+1}^2 - \sigma_t^2$.

Equations (5) and (9) then jointly result in the following relationship between future excess returns and realized and implied volatility levels:

---

[5] Our assumption on the relationship between the price and volatility risks differs from that of Banerjee et al. (2007), in which $\lambda = \lambda_v + \delta$, where δ is restricted to be greater than the absolute value of the variance risk premium. Our assumption is more general, and since it imposes less restriction on the δ parameter, it always results in positive price and negative variance risk premiums.

$$r_{t+1} - r_f = \delta\left(\Delta\sigma_{IV,t}^2 - \Delta\sigma_{RV,t}^2\right) + \kappa\delta\left(\sigma_{IV,t}^2 - \sigma_{RV,t}^2\right) + g_2(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}), \quad (10)$$

where $\quad r_{t+1} = \frac{S_{t+1} - S_t}{S_t}$ and $g_2(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}) = \sigma_{RV,t}S_t\epsilon_t - g_1(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*})$.

Since it is already well documented that the implied volatility computed from option prices is an efficient, albeit upwardly-biased, forecast of realized volatility[6] and highly self-correlated, we assume that:

$$\sigma_{RV,t}^2 = \alpha + \Psi\sigma_{IV,t}^2, \quad (11)$$

where Ψ should be less than 1, due to the upwardly-biased prediction.

By applying Equation (11) to Equation (10), we obtain the relationship between future excess returns and implied volatility levels across time points, as follows:

$$r_{t+1} - r_f = \delta(1 - \Psi)\Delta\sigma_{IV,t}^2 + \kappa\delta(1 - \Psi)\sigma_{IV,t}^2 + g_3(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}), \quad (12)$$

where $g_3(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}) = g_2(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*}) - \alpha\kappa\delta$.

Taking the conditional expectation for Equation (12), we derive the relationship between the expected future excess returns and the expected future and current implied volatility levels as:

$$E_t(r_{t+1} - r_f) = A_1\left(E_t(\sigma_{IV,t+1}^2) - \sigma_{IV,t}^2\right) + A_2\sigma_{IV,t}^2 + K, \quad (13)$$

where $\quad A_1 = \delta(1 - \Psi), A_2 = \kappa\delta(1 - \Psi)$, and $K = E_t\left(g_3(\epsilon_t, \epsilon_t^V, \epsilon_t^*, \epsilon_t^{V*})\right)$.

Since both δ and κ are positive, and Ψ is less than 1, then both A_1 and A_2 are also positive.

Equation (13) provides a theoretical fundamental for the potential predictive

---

[6] See, for example, Christensen and Prabhala (1998), Fleming (1998), Blair, Poon and Taylor (2001), and Jiang and Tian (2005).

ability on excess returns arising not only from the current implied volatility level, but also from its term structure which contains the expectation on future implied volatility. The positive sign on A_2 indicates a positive association between expected excess returns and risk-neutral expected variance, which is consistent with the positive return-risk relationship stated in the conventional asset pricing models. The positive sign of A_1 indicates a positive linkage between expected excess returns and the expectation on changes in expected variance, which implies that the expectation of an upward change in implied variance may drive future excess returns in the same direction.

3. DATA AND METHODOLOGY

3.1 Data

The daily S&P 500 index levels and the prices of the options written on the index are obtained from OptionMetrics, with our sample period (2 January 1998 to 31 August 2012) providing a rather 'rich' data period since it includes both bull and bear regimes. The VIX levels for all of the available time horizons are calculated based upon the prices of S&P 500 index options and their corresponding time to maturity. The bond data for analysis in this study were obtained from the Federal Reserve Bank of St. Louis[7], whilst the high-frequency index levels were obtained from OlsenData.

---

[7] See the website (http://research.stlouisfed.org/) for more details.

3.2 Methodology

We construct the squared VIX for several standardized horizons, comprising of 30, 60, 90, 180, 270 and 360 calendar days, following the method and the interpolation process adopted by the CBOE[8]. The squared VIX value serves as the measure for the risk-neutral expected variance. As shown in Table 1, the excess returns of the S&P 500 index are found to be more correlated with $DVIX^2_{t,30}$ than $VIX^2_{t,30}$, which may be a signal of the potential predictive ability of the VIX term structure with regard to returns. The highly negative correlation (–0.62) between $DVIX^2_{t,30}$ and $VIX^2_{t,30}$ arises as a result of the mean reversion property of volatility.

<Table 1 is inserted about here>

The correlation matrix of the squared VIX values across various maturity periods is presented in Table 2. Since it is clear that all of the values are highly correlated, it is inappropriate to include them simultaneously within a regression model, which lends additional support to the appropriateness of our approach of extracting the term structure information from the slope or through PCA.[9]

<Table 2 is inserted about here>

As suggested in many of the prior studies, several option-implied variables (such as the variance risk premium, model-free skewness and kurtosis) and

---

[8] See the VIX white paper (http://www.cboe.com/micro/vix/vixwhite.pdf) for more details.
[9] Johnson (2012) and Feunou et al. (2014) also used PCA to extract the information content of the VIX term structure.

macroeconomic variables (such as the interest rate term spread and default spread) are also informative with regard to future stock returns;[10] thus, we follow Bollerslev et al. (2009) to select three factors as the control variables in our regression analysis, these are: (i) the variance risk premium (*VRP*), which is defined as the difference between the VIX level and the realized variance compiled from the high-frequency five-minute transaction prices;[11] (ii) the default spread (*DFS*), which is defined as the difference between Moody's BAA and AAA bond yield indices; and (iii) the term spread (*TMS*), which is defined as the difference between the ten-year and three-year Treasury yields.

Additional controls are also provided in this study for the influence of model-free skewness (*SKEW*) and kurtosis (*KURT*), which are calculated based upon the method proposed by Bakshi, Kapadia and Madan (2003). Since our empirical analysis focuses on the S&P 500 index, we construct the VIX by the methodology of CBOE as the implied volatility index.[12]

In the following sub-sections, we propose three alternative approaches to the

---

[10]   Examples include Ang and Bekaert (2007), Bollerslev, Tauchen and Zhou (2009), Vilkov and Xiao (2013), Conrad, Dittmar and Ghysel (2013), Chang, Christoffersen and Jacobs (2013) and Kozhan, Neuberger and Schneider (2013)

[11]   We follow Bollerslev et al. (2009) to use intra-day data to construct the realized monthly variance. As discussed in the prior studies, including Andersen, Bollerslev, Diebold and Labys (2001) and Hansen and Lunde (2006), the selection of the sampling frequency is the trade-off between data continuity and market microstructure noises. Five minutes is the most frequently adopted frequency for the calculation of stock realized volatility.

[12] Details on the computation of the VIX and the construction of the VIX term structure are provided in Appendix A.

incorporation of the information implied in the VIX term structure for the prediction

of future excess returns.

3.2.1 Forward implied variance

Our first approach adheres closely to our theoretical model. $E_t(\sigma^2_{IV,t+1})$, by

definition, refers to the expectation at time t of the implied variance at time $t+1$

which gauges the level of volatility during the period from time $t+1$ to $t+2$.

Conceptually, this is the forward variance for the period from time $t+1$ to $t+2$;

therefore, the forward implied variance, $2VIX^2_{t,2} - VIX^2_{t,1}$, which is computed from the

squared VIX term structure, is used to replace $E_t(\sigma^2_{IV,t+1})$.

Given that one month (30 calendar days) is commonly used as the time horizon

in the VIX index, we also set the time unit as one month (30 days) in the present

study and then rewrite Equation (13) as:

$$E_t(r_{t+30} - r_f) = A'_1 * DVIX^2_{t,30} + A_2 * VIX^2_{t,30} + K, \tag{14}$$

where

$$DVIX^2_{t,30} = FVIX^2_{t,(30,60)} - VIX^2_{t,30} = 2VIX^2_{t,60} - VIX^2_{t,30} - VIX^2_{t,30} = 2(VIX^2_{t,60} - VIX^2_{t,30}).$$

We run the following regression in order to examine whether the information

implied in the VIX term structure can predict excess returns:

$$ER_{t,t+h} = \alpha + \beta_1 VIX^2_{t,30} + \beta_2 DVIX^2_{t,30} + \epsilon_t, \tag{15}$$

where $ER_{t,t+h}$ refers to the excess return for the period from time t to $t+h$, which is

defined as the return on the S&P 500 index minus the three-month T-bill rate. According to our theoretical model, β_1 and β_2, which respectively represent the predictive contribution of the implied variance level and the term structure with regard to excess returns, are both expected to be positive.

3.2.2 The first and second principal components

In the bond markets, Cochrane and Piazzesi (2005) indicate that amongst the first three components generated from 'principal component analysis' (PCA), which are identified as the level, slope and curvature of the interest rate term structure, the term structure of the interest rates is found to provide abundant information on excess bond returns. Furthermore, Feunou et al. (2014) apply the PCA to extract the systematic factors across the term structure of option-implied variance and empirically suggest that the first two components are sufficient to explain the changes on future S&P 500 index returns.

Since $DVIX^2_{t,30} = 2(VIX^2_{t,60} - VIX^2_{t,30})$ is a conceptual measure of the slope of the implied variance term structure for the period from 30 to 60 days, it would seem natural to question whether the slope of the implied variance term structure also contains useful information on the future excess returns of the underlying asset.

Following the approaches adopted in several studies within the extant literature, we apply PCA to the squared VIX across all of the available maturity periods and

take the first and second principal components, $PC1$ and $PC2$, as the information

proxy of the implied variance term structure. These two components essentially

represent the level and slope of the implied variance term structure, which are the

respective conceptual equivalents of $VIX_{t,30}^2$ and $DVIX_{t,30}^2$;[13] hence, the regression

model is amended to:

$$ER_{t,t+h} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \epsilon_t. \tag{16}$$

3.2.3 Two-factor stochastic volatility framework proxies

Egloff et al. (2010) extended the (SV) model of Heston (1993) to allow the central

tendency of the variance to be another stochastic process; thus, they revised the

decomposition of the squared VIX to:

$$VIX_{t,\tau}^2 = \omega_1 V_t + \omega_2 m_t + (1 - \omega_1 - \omega_2)\theta_m \tag{17}$$

where $\quad \omega_1 = \dfrac{1-e^{-\kappa_v\tau}}{\kappa_v\tau}, \quad \omega_2 = \dfrac{1+\dfrac{\kappa_m}{\kappa_v-\kappa_m}e^{-\kappa_v\tau}-\dfrac{\kappa_v}{\kappa_v-\kappa_m}e^{-\kappa_m\tau}}{\kappa_m\tau}.$

Consequently, the squared VIX indices for various maturity periods can be

transformed to the maturity-independent instantaneous variance, V_t, the stochastic

central tendency, m_t, and the long-run mean, θ_m.[14]

It is quite a straightforward matter for researchers to adopt instantaneous

variance as the proxy for the level of the VIX term structure, and given the

---

[13] From our analysis in the present study, $PC1$ and $PC2$ are found to be capable of explaining almost 95% of the variation in the VIX term structure.

[14] Luo and Zhang (2012a) also proposed a similar decomposition of the squared VIX under a framework with an independent stochastic long-run mean variance.

mean-reversion property of volatility, the relative relationship between the instantaneous variance and the stochastic central tendency, $m_t - V_t$, may well reveal the direction in which the VIX level is likely to move, which is the analogue of the slope of the VIX term structure.

We use the efficient iterative two-step procedure, suggested by Christoffersen, Heston and Jacob (2009), to estimate the parameters and to generate the time series of V_t and m_t. The two-stage procedure is implemented as follows:

Step 1: We solve the following optimization in order to estimate the time series of $(\hat{V}_t, \hat{m}_t)$, $t = 1, 2, \ldots, T$.

$$\left(\hat{V}_t, \hat{m}_t\right) = argmin \sum_{j=1}^{n_t} \left(VIX_{t,\tau_j}^{Mkt\,2} - VIX_{t,\tau_j}^2\right)^2, t = 1, 2, \cdots, T, \qquad (18)$$

where VIX_{t,τ_j}^{Mkt} and VIX_{t,τ_j} respectively denote the market and theoretical VIX levels for the maturity τ_j at time t.

Step 2: We collect the (V_t, m_t) time series in order to estimate the $\left(\hat{\theta}_m, \hat{\kappa}_\tau, \hat{\kappa}_m\right)$ parameter by implementing the following minimization:

$$\left(\hat{\theta}_m, \hat{\kappa}_\tau, \hat{\kappa}_m\right) = argmin \sum_{t=1}^{T} \sum_{j=1}^{n_t} \left(VIX_{t,\tau_j}^{Mkt\,2} - VIX_{t,\tau_j}^2\right)^2. \qquad (19)$$

The iteration procedure between Steps 1 and 2 is then carried out until the convergence criterion in the objective function of Step 2 is reached; the prediction regression is therefore revised to:

$$ER_{t,t+h} = \alpha + \beta_1 V_t + \beta_2(m_t - V_t) + \epsilon_t. \tag{20}$$

4. EMPIRICAL RESULTS

Our empirical analysis begins with an investigation into the information content of the VIX term structure for the subsequent-period excess returns, with the variables being compiled from the three different approaches described in the previous section. If predictive power is discernible, we then go on to explore how long such power may persist. We subsequently determine whether any profitable trading strategy can be formed based upon the predictive power of the VIX term structure. Finally, we carry out tests to verify the robustness of the results.

4.1 Forward Implied Variance Predictions

The $DVIX_{t,30}^2$ variable is compiled from the forward variance implied in the VIX term structure, and given that this variable is perfectly in line with our theoretical model, we first run the regression model specified in Equation (15). Based upon the control variables detailed at the end of Sub-section 3.2 (Methodology), the regression model is re-specified as:

$$ER_{t,t+30} = \alpha + \beta_1 VIX_{t,30}^2 + \beta_2 DVIX_{t,30}^2$$
$$+\beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \epsilon_t. \tag{21}$$

The results on the full model and the various restricted models are reported in Table 3. As shown in Models (1) and (2), when $VIX_{t,30}^2$ and $DVIX_{t,30}^2$ are individually

examined, both β_1 and β_2 are found to be insignificant, albeit positive; however, when the two variables are included in Model (3), β_2 becomes significantly positive at the 10% level, whilst β_1 remains insignificant.

When $VIX_{t,30}^2$ and $DVIX_{t,30}^2$ are respectively run with the control variables in Models (4) and (5), $DVIX_{t,30}^2$ is found to be more informative than $VIX_{t,30}^2$ on future excess returns. Finally, as shown in the full model, Model (6), both variables are found to be significantly positive at the 5% level.

<Table 3 is inserted about here>

Overall, the information implied in the squared VIX term structure is found to play an important role in the prediction of excess returns in the S&P 500 index, although the squared VIX level is also found to be informative to some extent. The important role of $DVIX_{t,30}^2$ also gains support from the higher incremental R^2 values that are found in those models in which $DVIX_{t,30}^2$ is included.

In summary, the positive signs of β_1 and β_2 are consistent with our theoretical predictions, although they are not always found to be significant. When comparing the relative contributions of $VIX_{t,30}^2$ and $DVIX_{t,30}^2$ to the prediction of future excess returns, we find that the latter is more informative than the former and that neither can completely replace the other. Therefore, in addition to the squared VIX level, the squared VIX term structure is also found to contain significant information of

relevance to the prediction of excess returns in the S&P 500 index.

4.2 First and Second Principal Component Predictions

Given that the first and second principal components of the squared VIX term structure, $PC1$ and $PC2$, represent its level and slope, we use the PCA approach for the squared VIX values, under various time horizons, as the means of generating $PC1$ and $PC2$, and then run the model specified in Equation (16). The regression model with control variables included is re-specified as:

$$ER_{t,t+30} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t$$

$$+\beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \epsilon_t.$$

(22)

The results on the full model and the various restricted models, which are shown in Table 4, reveal that the first and second principal component results are generally in line with those obtained from forward implied variance, with both $PC1$ and $PC2$ being found to be positively related to future excess returns, which is consistent with our predictions. When examining those models without the control variables, Models (1)-(3), $PC2$ appears to be more informative than $PC1$, since β_2 is found to have greater significance than β_1; however, when the control variables are included, β_1 in Model (4) is found to be significantly positive at the 1% level, and β_2 in Model (5) is found to be significantly positive at the 5% level, although β_2 becomes insignificant in Model (6), the full model.

<Table 4 is inserted about here>

Overall, in addition to the squared VIX level, the slope of the squared VIX term structure is also found to contribute to the prediction of excess returns in the S&P 500 index although $PC2$ can explain only 5.5% of the variation in the squared VIX term structure, whilst $PC1$ is capable of explaining 93.11%. Our empirical findings provide strong support for the informativeness of the slope factor of the VIX term structure on future excess returns in the S&P 500 index.

4.3 Two-factor Stochastic Volatility Framework Proxy Predictions

Since the two-factor stochastic volatility model provides an alternative instrument for the transformation of the maturity-dependent VIX values to maturity-independent instantaneous variance and its stochastic central tendency, we generate the time series of (V_t, m_t) and then run the regression model specified in Equation (20). The regression model with the control variables included is re-specified as:

$$ER_{t,t+30} = \alpha + \beta_1 V_t + \beta_2 (m_t - V_t)$$
$$+ \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \epsilon_t. \tag{23}$$

The results on the full model and the various restricted models, which are reported in Table 5, are generally found to be consistent with those of the previous two alternative approaches. In particular, both β_1 and β_2 are found to be significantly positive at the 1% level in the full model. Therefore, in addition to the instantaneous

20

variance level, its position relative to the stochastic central tendency is also found to play a significant role in terms of predicting excess returns.

As regards the incremental contribution to such prediction, β_2 is clearly more informative than β_1, since β_2 is found to be significantly positive in most of the models. In summary, the findings from our third alternative approach also confirm the importance of the information provided by the squared VIX term structure with regard to the prediction of excess returns in the S&P 500 index.

<Table 5 is inserted about here>

4.4 Predictions across Time Horizons

The focus so far in this study has been placed on a one-month prediction horizon in order to match the maturity of the options used to compute the VIX index, and since our alternative approaches have provided consistent findings on the important role of the squared VIX term structure in the prediction of excess returns, it would seem natural to enquire just how far forward the excess returns can be predicted. In order to investigate this intriguing question, we employ various horizons (2, 3, 6, 9 and 12 months) of the excess returns as the dependent variables in the regression models, with the results from the three alternative approaches being reported in Table 6.

<Table 6 is inserted about here>

As shown in Models (1) and (3) of Table 6, for all horizons, both β_1 and β_2 are

found to be significantly positive at the level of at least 5%, which thereby indicates that the squared VIX level and the term structure are not only informative for shorter excess return horizons, but also for longer horizons. Since this predictive power can be satisfactory for horizons of up to a full year, it is clearly of interest to explore the relative contributions of the squared VIX level and the term structure to the prediction of excess returns across such horizons.

In order to facilitate an exploration of the way in which the term structure factor incrementally contributes to such prediction, we compare the R^2 values for both the full model and the restricted model (where $\beta_2 = 0$). The R^2 values of the restricted model, for various prediction horizons, are shown in the penultimate row of each panel in Table 6, with the final row showing the percentage increase in R^2 as a result of the inclusion of the term structure variables. Regardless of the approach adopted for the compilation of the variables, the percentage contribution made by the term structure variable is found to be most notable for the one-month prediction horizon, followed by a general decline over longer prediction horizons. For example, in Panel C, as compared to the model with V_t and the control variables, the R^2 in the full model increases by a factor of about fourteen, from 0.0030 to 0.0410.

When using PCA as an alternative instrument for the compilation of the variables, the contribution of the squared VIX term structure is not as impressive as

that of the squared VIX level; for example, the β_2 coefficient is found to be significantly positive only in the six-month case. Nevertheless, the sign remains positive across all prediction horizons, and the conclusions drawn from the R^2 values are consistent with those drawn from the other two alternative approaches.

Overall, whilst both the VIX level and the VIX term structure are found to be equally informative with regard to future excess returns in the S&P 500 index across all of the horizons investigated, the VIX term structure is particularly informative for shorter-horizon excess returns.

4.5 Trading Strategy Tests

According to the regression results discussed above, we have found that the information derived from the squared VIX term structure, as well as the squared VIX level, is of use in determining future excess returns in the S&P 500 index; however, our regression analysis has been based upon an in-sample framework. Therefore, as a guide for trading strategies, we take our analysis a step further to investigate the out-of-the-sample performance of the level and term structure based upon predictions generated from regression models using historical data. In specific terms, we run the models with the variables representing the squared VIX level, the squared VIX term structure, or both factors, with all of the control variables included, to determine whether trading strategies using the information derived from the squared VIX term

structure can be more profitable than those using the squared VIX level only.

The trading strategies include: (i) setting up a long position with a single unit asset when the forecast exceeds a critical value; (ii) setting up a short position with a single unit asset when the forecast is below the negative critical value; and (iii) doing nothing when the forecast is within the positive and negative critical values. The critical values range from 0 to 0.1 ($=10\%$) and the prediction horizons for the excess returns are set at one, two and three months. The sample period used to generate the forecasts covers a total of six months, with the predictions being implemented based upon a rolling window procedure.

The results on the three alternative approaches used to compile the variables representing the information on the VIX term structure are illustrated in Figures 1-3, with Figure 1 illustrating the results for $VIX^2_{t,\,30}$ ($DVIX^2_{t,\,30}$) as the variable representing the squared VIX level (term structure).

<Figure 1 is inserted about here>

It is difficult to distinguish between the performances of the three models when the prediction horizon is set at one or two months; thus, as we can see from the first and second graphs in Figure 1 (*h=30* and *h=60*), the model with $DVIX^2_{t,30}$ outperforms the model with $VIX^2_{t,30}$ only for the critical value between 0.02 and 0.04 (0.03) in the one-month (two-month) prediction.

However, when examining the other prediction horizons, we find that the model with $DVIX_{t,30}^2$ outperforms the model with $VIX_{t,30}^2$ for most of the critical values. In particular, the former outperforms the latter in almost every case in the three-month prediction. Interestingly, the model with both $VIX_{t,30}^2$ and $DVIX_{t,30}^2$ is unable to outperform the model with only $DVIX_{t,30}^2$, which therefore suggests that the information quality of $DVIX_{t,30}^2$ may be better than that of $VIX_{t,30}^2$.

The results with $PC1$ ($PC2$) as the variable representing the squared VIX level (VIX term structure) are shown in Figure 2, whilst those with V_t ($m_t - V_t$) as the variable representing the squared VIX level (term structure) are shown in Figure 3. As we can see from these figures, the general patterns are similar to those for $VIX_{t,30}^2$ and $DVIX_{t,30}^2$; however, in contrast to the earlier results, for the two-month prediction, the model with the information derived from the squared VIX term structure is found to be only slightly better than that with the information derived from the squared VIX level. As regards the one- and three-month predictions, the term structure is found to overwhelmingly outperform the squared VIX level, whilst for most of the critical values, the model which includes both the squared VIX level information and the VIX term structure information is also found to make higher profits than that with only the VIX level information.

<Figures 2 and 3 are inserted about here>

In summary, regardless of which approach is adopted to extract the information variables, a trading strategy which follows the forecasts generated from the information on the squared VIX term structure clearly improves and outperforms a strategy which follows the forecasts generated from only the squared VIX level. Therefore, both the in-sample regression analysis and the out-of-the-sample trading strategy highlight the merits of the information implied in the squared VIX term structure with regard to the prediction of excess returns in the S&P 500 index, although the squared VIX level is also found to be informative to some extent.

4.6 Robustness Analysis

In order to ensure that we have a sufficiently large number of observations, we have run the regression models using overlapping data (i.e. daily one-month returns); however, despite having adopted robust Newey-West standard errors for our analyses, we may still wonder whether the findings would remain unchanged if we were to use non-overlapping data. Thus, in order to explore whether the findings are sensitive to data selection, we also use non-overlapping data to rerun the empirical analysis presented above. The results of the full models for one-, two- and three-month prediction horizons are reported in Table 7.

<Table 7 is inserted about here>

The general finding from Table 7 is that the variable representing the information

on the squared VIX term structure is more informative than that representing the squared VIX level information. When examining the one- and three- month prediction horizons, we find that the coefficient on the term structure variable, β_2, is always significantly positive, a factor which is not dependent on the approach adopted for the compilation of the information variables; however, the coefficient on the level variable, β_1, is found to be positive in all case but insignificant. As regards the two-month prediction, only one coefficient is found to be significantly positive; that is, β_2 with $m_t - V_t$ as the term structure variable.

Overall, our findings based upon both overlapping and non-overlapping data are generally consistent. Both analyses provide support for the important role of the squared VIX term structure in the determination of future excess returns in the S&P 500 index. The information content of the squared VIX term structure is clearly significant, and indeed, a strong complement to that of the squared VIX level.

5.   CONCLUSIONS

We construct a theoretical model based upon the stochastic volatility (SV) model proposed by Heston (1993) to associate the squared VIX level and VIX term structure with the excess returns of the S&P 500 index. Our findings reveal the existence of a positive relationship between excess returns and the squared VIX level, which is consistent with the traditional capital asset pricing model (CAPM).

Furthermore, a positive relationship is also found to exist between excess market returns and the squared VIX term structure, which is consistent with the empirical results of Bakshi et al. (2011).

We use three alternative empirical approaches to support our theoretical model, from which we find that the squared VIX term structure is more informative than the squared VIX level, although predictive power is found to exist for both factors across various time horizons. The incremental contribution of the VIX term structure with regard to the prediction of excess returns is particularly discernible for shorter horizons. Finally, we find that the adoption of trading strategies based upon the forecasts generated from the information on the squared VIX term structure are clearly superior to those based upon only the squared VIX level.

REFERENCES

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001), 'The Distribution of Realized Exchange Rate Volatility', *Journal of the American Statistical Association*, **96**: 42-55.

Ang, A. and G. Bekaert (2007), 'Stock Return Predictability: Is it There?', *Review of Financial Studies*, **20**: 651-707.

Bakshi, G.S., N. Kapadia and D. Madan (2003), 'Stock Return Characteristics, Skew Laws and the Differential Pricing of Individual Equity Options', *Review of Financial Studies*, **16**: 101-43.

Bakshi, G.S., G. Panayotov and G. Skoulakis (2011), 'Improving the Predictability of Real Economic Activity and Asset Returns with Forward Variances inferred from Option Portfolios', *Journal of Financial Economics*, **100**: 475-95.

Banerjee, P.S., J.S. Doran, and D.R. Peterson (2007), 'Implied Volatility and Future Portfolio Returns', *Journal of Banking and Finance*, **31**: 3183-3199.

Blair, J.B., S.H. Poon and S.J. Taylor (2001), 'Forecasting S&P 100 Volatility: The Incremental Information Content of Implied Volatilities and High-frequency Index Returns', *Journal of Econometrics*, **105**: 5-26.

Bollerslev, T., G. Tauchen and H. Zhou (2009), 'Expected Stock Returns and Variance Risk Premium', *Review of Financial Studies*, **22**: 4463-92.

Britten-Jones, M. and A. Neuberger (2000), 'Option Prices, Implied Price Process and Stochastic Volatility', *Journal of Finance*, **55**: 839-66.

Campbell, J. and R. Shiller (1991),' Yield Spreads and Interest Rate Movements: A Bird's Eye View', *Review of Economic Studies*, **58**: 495-514.

Carr, P. and D. Madan (1998), 'Towards a Theory of Volatility Trading', in Robert Jarrow ed., *Volatility Estimation Techniques for Pricing Derivatives*, London: Risk Books, pp. 417-27.

Carr, P. and L. Wu (2009), 'Variance Risk Premiums', *Review of Financial Studies*, **22**: 1311-41.

Chang, B.Y., P. Christoffersen, and K. Jacobs (2013), 'Market Skewness Risk and the Cross Section of Stock Returns', *Journal of Financial Economics*, **107**: 46-68.

Christensen, B.J. and N.R. Prabhala (1998), 'The Relation between Implied and Realized Volatility', *Journal of Financial Economics*, **50**: 125-50.

Christoffersen, P., S. Heston and K. Jacobs (2009), 'The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work so Well', *Management Science*, **55**: 1914-32.

Cochrane, J.H. and M. Piazzesi (2005), 'Bond Risk Premia', *American Economic Review*, **95**: 138-60.

Conrad, J., R.F. Dittmar and E. Ghysels (2013), 'Ex-ante Skewness and Expected Stock Returns', *Journal of Finance*, **68**: 85-124.

Cox, J., J. Ingersoll and S. Ross (1981), 'A Reexamination of the Traditional Hypotheses about the Term Structure of Interest Rates', *Journal of Finance*, **36**: 321-46.

Demeterfi, K., E. Derman, M. Kamal and J. Zou (1999), 'A Guide to Volatility and Variance Swaps', *Journal of Derivatives*, **6**: 9-32.

Duan, J.C. and C.Y., Yeh (2011), 'Price and Volatility Dynamics implied by the VIX Term Structure', Working Paper, National University of Singapore.

Egloff, D., M. Leippold and L. Wu (2010), 'The Term Structure of Variance Swap Rates and Optimal Variance Swap Investment', *Journal of Financial and Quantitative Analysis*, **45**: 1279-1310.

Fama, E. and R. Bliss (1987), 'The Information in Long-maturity Forward Rates', *American Economic Review*, **77**: 680-92.

Feuno, B., J.-E. Fontaine, A. Taamouti, and R. Tedongap (2014), 'Risk Premium, Variance Premium, and the Maturity Structure of Uncertainty', *Review of Finance*, **18**: 219-269.

Fleming, J., (1998), 'The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices', *Journal of Empirical Finance*, **5**: 317-345.

Giot, P. (2005), 'The Relationship between Implied Volatility Indexes and Stock Index Returns', *Journal of Portfolio Management*, **31**: 92-100.

Guo, H., and R.F. Whitelaw, 'Uncovering the Risk-return Relation in the Stock Market', *Journal of Finance*, **61**: 1433-1463.

Hansen, P.R. and A. Lunde (2006), 'Realized Variance and Market Microstructure Noise', *Journal of Business and Economic Statistics*, **24**: 127-61.

Heston, S. (1993), 'A Closed-form Solution for Options with Stochastic Volatility

with Applications to Bond and Currency Options', *Review of Financial Studies*, **6**: 327-43.

Jiang, G. and Y. Tian (2005), 'Model-free Implied Volatility and its Information Content', *Review of Financial Studies*, **18**: 1305-42.

Johnson, T.L. (2012), 'Equity Risk Premia and the VIX Term Structure', Working Paper, University of Texas at Austin.

Kozhan, R., A. Neuberger and P. Schneider (2013), 'The Skew Risk Premium in the Equity Index Market', *Review of Financial Studies*, **26**: 2174-203.

Luo, X.-G. and J.E. Zhang (2012a), 'The Term Structure of VIX', *Journal of Futures Markets*, **32**: 1092-23.

Luo, X.-G. and J.E. Zhang (2012b), 'Expected Stock Returns and Forward Variance', Working paper, University of Zhejiang.

Merton, R.C. (1973), 'An Inter-temporal capital asset pricing model', *Econometrica*, **41**: 867-87.

Mixon, S. (2007), 'The Implied Volatility Term Structure of Stock Index Options', *Journal of Empirical Finance*, **14**: 333-54.

Moench, E. (2012), 'Term Structure Surprises: The Predictive Content of Curvature, Level and Slope', *Journal of Applied Econometrics*, **27**: 574-602.

Newey, W. and K. West (1987), 'A Simple, Positive, Semi-definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica*, **55**: 703-8.

Vilkov, G. and Y. Xiao (2013), 'Option-implied Information and Predictability of

Extreme Returns', Working paper, Goethe University, Frankfurt.

Whaley, R.E. (2000), 'The Investor Fear Gauge', *Journal of Portfolio Management*, **26**: 12-17.

Zhang, J.E. and Y.-Z. Zhu (2006), 'VIX Futures', *Journal of Futures Markets*, **26**: 521-31.

Table 1 Correlation matrix of excess future monthly returns, VIX30 and DVIX30

$ER_{t,t+30}$ is the excess return for the period from time t to $t+30$, defined as the return on the S&P 500 index minus the three-month T-bill rate; $VIX^2_{t,30}$ is the squared VIX with a 30 calendar-day maturity; and $DVIX^2_{t,30}$ is the difference between the squared VIX with a 60 calendar-day maturity and the squared VIX with a 30 calendar-day maturity, scaled by 2, that is: $DVIX^2_{t,30} = 2(VIX^2_{t,60} - VIX^2_{t,30})$.

	$ER_{t,t+30}$	$VIX^2_{t,30}$	$DVIX^2_{t,30}$
$ER_{t,t+30}$	1.0000		
$VIX^2_{t,30}$	0.0047	1.0000	
$DVIX^2_{t,30}$	0.1132	−0.6167	1.0000

Table 2 Correlation matrix of squared VIX values under different horizons

$VIX^2_{t,h}$ is the squared VIX with an h calendar-day maturity, where $h = 30, 60, 90, 180, 270$ and 360.

	$VIX^2_{t,30}$	$VIX^2_{t,60}$	$VIX^2_{t,90}$	$VIX^2_{t,180}$	$VIX^2_{t,270}$	$VIX^2_{t,360}$
$VIX^2_{t,30}$	1.0000					
$VIX^2_{t,60}$	0.9852	1.0000				
$VIX^2_{t,90}$	0.9622	0.9899	1.0000			
$VIX^2_{t,180}$	0.9031	0.9522	0.9767	1.0000		
$VIX^2_{t,270}$	0.8519	0.9128	0.9460	0.9834	1.0000	
$VIX^2_{t,360}$	0.7541	0.8222	0.8583	0.9108	0.9362	1.0000

Table 3   Forward implied variance predictions

This table presents the results based upon the following regression model:

$$ER_{t,t+30} = \alpha + \beta_1 VIX_{t,30}^2 + \beta_2 DVIX_{t,30}^2 + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t$$

where $ER_{t,t+30}$, is the excess return for the period from time t to $t+30$, defined as the return on the S&P 500 index minus the three-month T-bill rate; $VIX_{t,30}^2$ is the squared VIX with a 30 calendar-day maturity; and $DVIX_{t,30}^2$ is the difference between the squared VIX with a 60 calendar-day maturity and the squared VIX with a 30 calendar-day maturity, scaled by 2, that is: $DVIX_{t,30}^2 = 2(VIX_{t,60}^2 - VIX_{t,30}^2)$; VRP_t is the variance risk premium, that is: $VRP_t = VIX_{t,30}^2 - RV_{t,30}$, where $RV_{t,30}$ is the realized variance for the period from $t-30$ to t; $SKEW_t$ ($KURT_t$) is the model-free skewness (kurtosis) with a 30 calendar-day maturity (Bakshi et al., 2003); TMS_t is the term spread defined as the difference between the ten-year and three-month Treasury yields; and DFS_t is the term spread defined as the difference between Moody's BAA and AAA bond yield indices (Bollerslev et al., 2009). The standard errors (S.E.) are calculated based upon the Newey-West method, with the lag being equal to the number of overlapping horizons. The sample period runs from 2 January 1998 to 31 August 2012.

One-month Horizon	Models											
	(1)		(2)		(3)		(4)		(5)		(6)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
α	0.0020	0.0042	0.0017	0.0033	−0.0050	0.0055	0.0003	0.0120	0.0023	0.0121	0.0039	0.0126
β_1	0.0047	0.0881	–	–	0.1190	0.0984	0.1532	0.1136	–	–	0.3598**	0.1556
β_2	–	–	0.2880	0.1850	0.4760*	0.2510	–	–	0.3820*	0.2110	0.6216**	0.2678
β_3	–	–	–	–	–	–	0.0068	0.0047	−0.0042	0.0045	0.0055	0.0051
β_4	–	–	–	–	–	–	−0.0015	0.0023	−0.0011	0.0021	−0.0005	0.0021
β_5	–	–	–	–	–	–	−0.0003	0.0004	−0.0002	0.0004	−0.0001	0.0004
β_6	–	–	–	–	–	–	−0.0009	0.0023	−0.0009	0.0023	−0.0011	0.0024
β_7	–	–	–	–	–	–	−0.0028	0.0136	−0.0046	0.0140	−0.0160	0.0149
No. of Obs.	3,644		3,664		3,664		3,644		3,644		3,644	
R^2	0.0001		0.0132		0.0220		0.0072		0.0171		0.0361	

Table 4 First and second principal component predictions

This table presents the results based upon the following regression model:

$$ER_{t,t+30} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t$$

where $ER_{t,t+30}$, is the excess return for the period from time t to $t+30$, defined as the return on the S&P 500 index minus the three-month T-bill rate; $PC1_t$ and $PC2_t$ are the first and second principal components; $VRP_{t,30}$ is the variance risk premium, that is: $VRP_t = VIX^2_{t,30} - RV_{t,30}$, where $RV_{t,30}$ is the realized variance for the period from $t-30$ to t; $SKEW_t$ ($KURT_t$) is the model-free skewness (kurtosis) with a 30 calendar-day maturity (Bakshi et al., 2003); TMS_t is the term spread defined as the difference between the ten-year and three-month Treasury yields; and DFS_t is the term spread defined as the difference between Moody's BAA and AAA bond yield indices (Bollerslev et al., 2009). The standard errors (S.E.) are calculated based upon the Newey-West method, with the lag being equal to the number of overlapping horizons. The sample period runs from 2 January 1998 to 31 August 2012.

One-month Horizon	Models											
	(1)		(2)		(3)		(4)		(5)		(6)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
α	0.0022	0.0033	0.0023	0.0032	0.0023	0.0032	0.0254*	0.0141	0.0089	0.0122	0.0263*	0.0143
β_1	0.0012	0.0018	–	–	0.0012	0.0017	0.0075***	0.0028	–	–	0.0061**	0.0027
β_2	–	–	0.0108*	0.0060	0.0108*	0.0063	–	–	0.0149**	0.0062	0.0097	0.0060
β_3	–	–	–	–	–	–	0.0102*	0.0056	-0.0040	0.0039	0.0051	0.0050
β_4	–	–	–	–	–	–	-0.0008	0.0022	-0.0011	0.0022	-0.0006	0.0022
β_5	–	–	–	–	–	–	-0.0001	0.0004	-0.0002	0.0004	–0.0001	0.0004
β_6	–	–	–	–	–	–	-0.0013	0.0025	-0.0014	0.0024	-0.0016	0.0025
β_7	–	–	–	–	–	–	-0.0141	0.0142	-0.0088	0.0136	-0.0175	0.0145
No. of Obs.	3,644		3,664		3,664		3,644		3,644		3,644	
R^2	0.0031		0.0132		0.0167		0.0283		0.0201		0.0342	

Table 5 ***Two-factor stochastic volatility framework proxy predictions***

This table presents the results based upon the following regression model:

$$ER_{t,t+30} = \alpha + \beta_1 V_t + \beta_2(m_t - V_t) + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t$$

where $ER_{t,t+30}$, is the excess return for the period from time t to $t+30$, defined as the return on the S&P 500 index minus the three-month T-bill rate; θ_t and m_t are the respective instantaneous variance and stochastic central tendency; $VRP_{t,30}$ is the variance risk premium, that is: $VRP_t = VIX^2_{t,30} - RV_{t,30}$, where $RV_{t,30}$ is the realized variance for the period from $t\text{-}30$ to t; $SKEW_t$ ($KURT_t$) is the model-free skewness (kurtosis) with a 30 calendar-day maturity (Bakshi et al., 2003); TMS_t is the term spread defined as the difference between the ten-year and three-month Treasury yields; and DFS_t is the term spread defined as the difference between Moody's BAA and AAA bond yield indices (Bollerslev et al., 2009). The standard errors (S.E.) are calculated based upon the Newey-West method, with the lag being equal to the number of overlapping horizons. The sample period runs from 2 January 1998 to 31 August 2012.

One-month Horizon	Models											
	(1)		(2)		(3)		(4)		(5)		(6)	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
α	0.0030	0.0037	0.0016	0.0034	-0.0086	0.0066	0.0002	0.0119	0.0010	0.0116	0.0024	0.0122
β_1	-0.0141	0.0760	–	–	0.1780	0.1140	0.0314	0.1090	–	–	0.4190***	0.1620
β_2	–	–	0.0981	0.0791	0.2730**	0.1330	–	–	0.1800*	0.0970	0.4270***	0.1480
β_3	–	–	–	–	–	–	0.0027	0.0049	-0.0062	0.0049	0.0026	0.0050
β_4	–	–	–	–	–	–	-0.0016	0.0023	-0.0010	0.0021	-0.0001	0.0021
β_5	–	–	–	–	–	–	-0.0002	0.0004	-0.0002	0.0004	-0.0005	0.0004
β_6	–	–	–	–	–	–	-0.0008	0.0023	-0.0011	0.0023	-0.0016	0.00250
β_7	–	–	–	–	–	–	-0.0001	0.0133	-0.0044	0.0132	-0.0184	0.0144
No. of Obs.	3,644		3,664		3,664		3,644		3,644		3,644	
R^2	0.0001		0.0108		0.0241		0.0036		0.0166		0.0411	

Table 6   Time horizon predictions

This table presents the results based upon the following regression models:

$$ER_{t,t+h} = \alpha + \beta_1 VIX^2_{t,30} + \beta_2 DVIX^2_{t,30} + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \qquad \text{Model (1)}$$

$$ER_{t,t+h} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \qquad \text{Model (2)}$$

$$ER_{t,t+h} = \alpha + \beta_1 V_t + \beta_2 (m_t - V_t) + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \qquad \text{Model (3)}$$

where h = 30, 60, 90, 180, 270 and 360, with the sample period running from 2 January 1998 to 31 August 2012. The table presents the coefficients on only the main explanatory variables, $VIX^2_{t,30}$, $DVIX^2_{t,30}$, $PC1_t$, $PC2_t$, V_t and $m_t - V_t$, along with the standard errors (S.E.) which are calculated based upon the Newey-West method, where the lags are equal to the number of overlapping horizons. We present the R^2 of the regression models with the two main factors and all of the control variables. The penultimate and final row of each panel are presented as the R^2 of the regression models only with the level factor and all control variables and the ratio of the bottom third row to the penultimate row minus one respectively.

Variables	Horizons											
	One-month		Two-month		Three-month		Six-month		Nine-month		One-year	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Panel A: Model (1)												
β_1	0.3600**	0.1560	0.6590***	0.2290	1.0510***	0.3520	1.8280***	0.5120	2.2700***	0.6510	2.2900***	0.7810
β_2	0.6220**	0.2680	0.8590**	0.3990	1.1910**	0.4740	2.2030***	0.6370	2.5400***	0.7050	2.3320***	0.7800
No. of Obs.	3,644		3,623		3,602		3,539		3,476		3,421	
R^2	0.0365		0.0502		0.0923		0.1788		0.2171		0.2390	
R^2 with $VIX^2_{t,30}$ & controls	0.0070		0.0221		0.0540		0.1192		0.1670		0.2074	
Increase in R^2	4.14		1.27		0.70		0.50		0.30		0.15	

38

Table 6 (*Contd.*)

Variables	Horizons											
	One-month		Two-month		Three-month		Six-month		Nine-month		One-year	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Panel B: Model (2)												
β_1	0.0061**	0.0027	0.0126***	0.0041	0.0202***	0.0063	0.0344***	0.0093	0.0427***	0.0124	0.0420***	0.0149
β_2	0.0097	0.0060	0.0100	0.0091	0.0123	0.0109	0.0333*	0.0183	0.0324	0.0238	0.0282	0.0296
No. of Obs.	3,644		3,623		3,602		3,539		3,476		3,421	
R^2	0.0340		0.0581		0.1053		0.2057		0.2394		0.2532	
R^2 with $PC1_t$ & controls	0.0281		0.0543		0.1022		0.1941		0.2324		0.2492	
Increase in R^2	0.21		0.07		0.03		0.06		0.03		0.02	
Panel C: Model (3)												
β_1	0.4190***	0.1620	0.8030***	0.2490	1.2630***	0.3690	2.2570***	0.5750	2.7240***	0.7440	2.6720***	0.9080
β_2	0.4270***	0.1480	0.7280**	0.2310	1.0280***	0.2990	1.8800***	0.4650	2.1270***	0.5860	1.9490***	0.7190
No. of Obs.	3,644		3,623		3,602		3,539		3,476		3,421	
R^2	0.0410		0.0686		0.1137		0.2153		0.2436		0.2538	
R^2 with V_t & controls	0.0030		0.0111		0.0362		0.0961		0.1450		0.1923	
Increase in R^2	12.67		5.18		2.14		1.24		0.68		0.32	

Table 7 Predictions with non-overlapping data

This table presents the results based upon the following regression models:

$$ER_{t,t+h} = \alpha + \beta_1 VIX^2_{t,30} + \beta_2 DVIX^2_{t,30} + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \quad \text{Model (1)}$$

$$ER_{t,t+h} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \quad \text{Model (2)}$$

$$ER_{t,t+h} = \alpha + \beta_1 V_t + \beta_2 (m_t - V_t) + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t \quad \text{Model (3)}$$

where h = 30, 60 and 90, and the sample is non-overlapping, with the sample period running from 2 January 1998 to 31 August 2012. The standard errors (S.E.) are obtained based upon the ordinary least square (OLS) method.

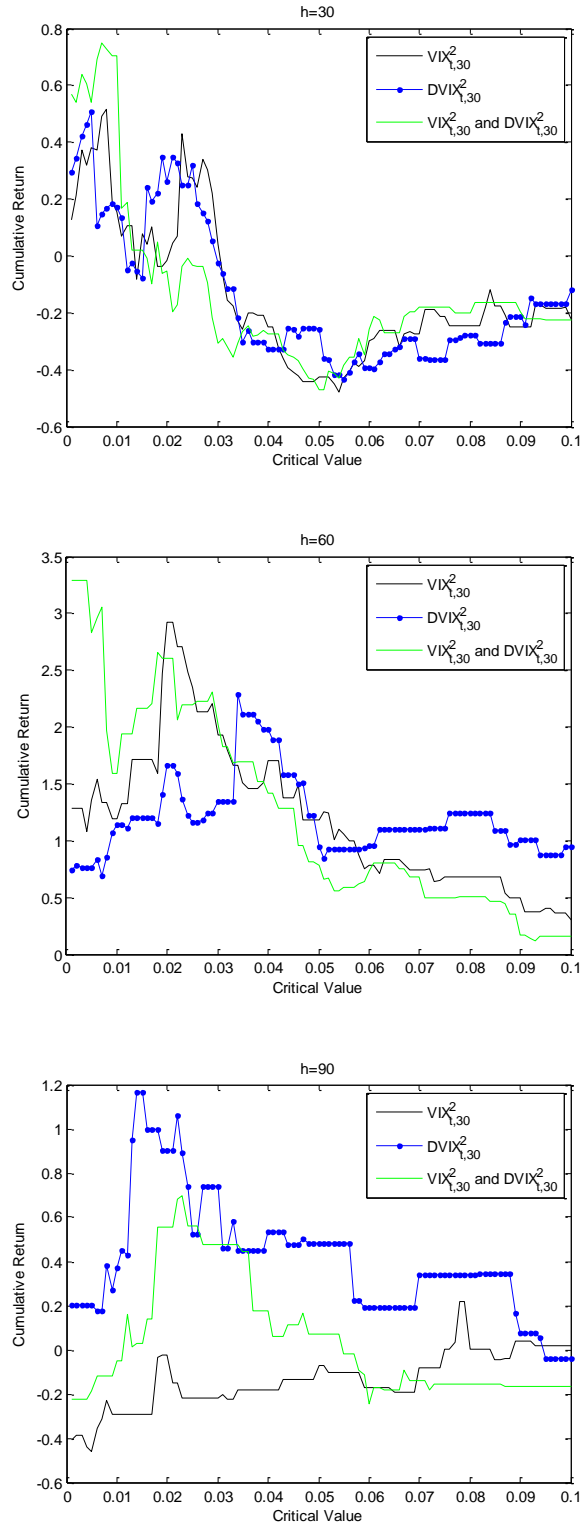| | Horizons | | | | | |
| Variables | One-month | | Two-month | | Three-month | |
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Panel A: Model (1)						
α	0.0061	0.0126	-0.0066	0.0274	0.0097	0.0378
β_1	0.0270	0.1870	0.4818	0.4631	0.3924	0.6235
β_2	0.6220**	0.2750	-0.6415	0.7971	2.5645***	0.7288
β_3	-0.0062	0.0071	0.0133	0.0166	-0.0656	0.0447
β_4	-0.0027	0.0031	-0.0074	0.0072	-0.0024	0.0107
β_5	-0.0006	0.0006	-0.0012	0.0013	0.0002	0.0023
β_6	-0.0004	0.0033	0.0017	0.0071	0.0058	0.0099
β_7	-0.0150	0.0135	-0.0223	0.0293	-0.0963**	0.0418
No. of Obs.	174		86		57	
R^2	0.049		0.041		0.218	
Panel B: Model (2)						
α	0.0145	0.0171	0.0381	0.0369	0.0771	0.0590
β_1	0.0010	0.0037	0.0086	0.0087	0.0160	0.0133
β_2	0.0180**	0.0073	0.0193	0.0216	0.0460*	0.0251
β_3	-0.0035	0.0070	-0.0009	0.0135	-0.0087	0.0426
β_4	-0.0028	0.0031	-0.0046	0.0070	-0.0014	0.0113
β_5	-0.0005	0.0005	-0.0007	0.0001	0.0001	0.0002
β_6	-0.0013	0.0033	0.0004	0.0071	-0.0018	0.0101
β_7	-0.0170	0.0135	-0.0442	0.0284	-0.0740*	0.0429
No. of Obs.	174		86		57	
R^2	0.049		0.059		0.134	
Panel C: Model (3)						
α	0.0038	0.0123	0.0021	0.0257	-0.0039	0.0389
β_1	0.1840	0.2030	0.5360	0.4460	1.1430	0.7270
β_2	0.3610**	0.1530	0.8950**	0.3900	1.3630***	0.4930
β_3	-0.0049	0.0067	-0.0109	0.0149	-0.0316	0.0422
β_4	-0.0019	0.0031	-0.0022	0.0071	-0.0009	0.0110
β_5	-0.0004	0.0005	-0.0004	0.0010	-0.0001	0.0023
β_6	-0.0010	0.0033	-0.0004	0.0070	-1.75e-05	0.0099
β_7	-0.0175	0.0133	-0.0426	0.0280	-0.0776*	0.0414
No. of Obs.	174		86		57	
R^2	0.0652		0.0821		0.1822	

Figure 1   Cumulative returns of forward implied variance trading strategies

Note: The strategy, based upon turnover rates of 30, 60 and 90 calendar days and an estimation period of 180 calendar days, is to take up long (short) positions in the S&P 500 index when estimated returns exceed (are below) the critical value x (-x), and do nothing when estimated returns are in the interval, [-x, x]. The strategy is based upon the following estimated regression model:

$$ER_{t,t+h} = \alpha + \beta_1 VIX_{t,30}^2 + \beta_2 DVIX_{t,30}^2 + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_{t.}$$
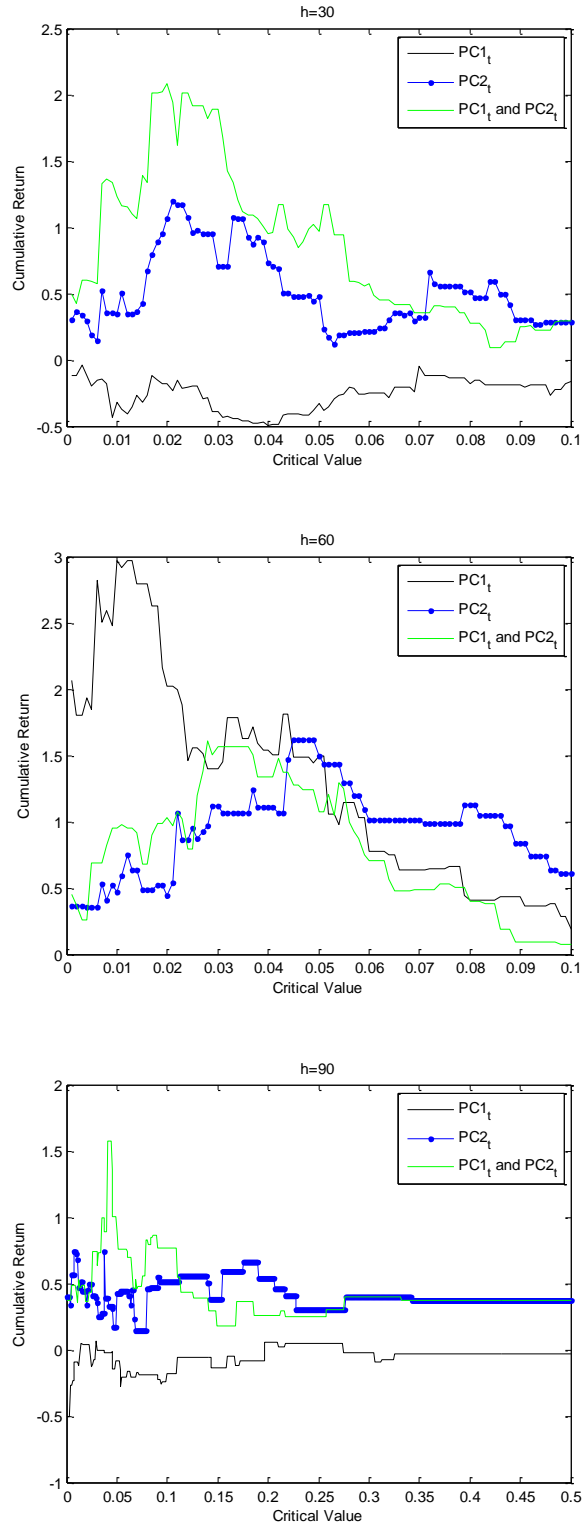
Figure 2   Cumulative returns of first and second principal component trading strategies

Note: The strategy, based upon turnover rates of 30, 60 and 90 calendar days and an estimation period of 180 calendar days, is to take up long (short) positions in the S&P 500 index when estimated returns exceed (are below) the critical value x (-x), and do nothing when estimated returns are in the interval, [-x, x]. The strategy is based upon the following estimated regression model:

$$ER_{t,t+h} = \alpha + \beta_1 PC1_t + \beta_2 PC2_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t$$
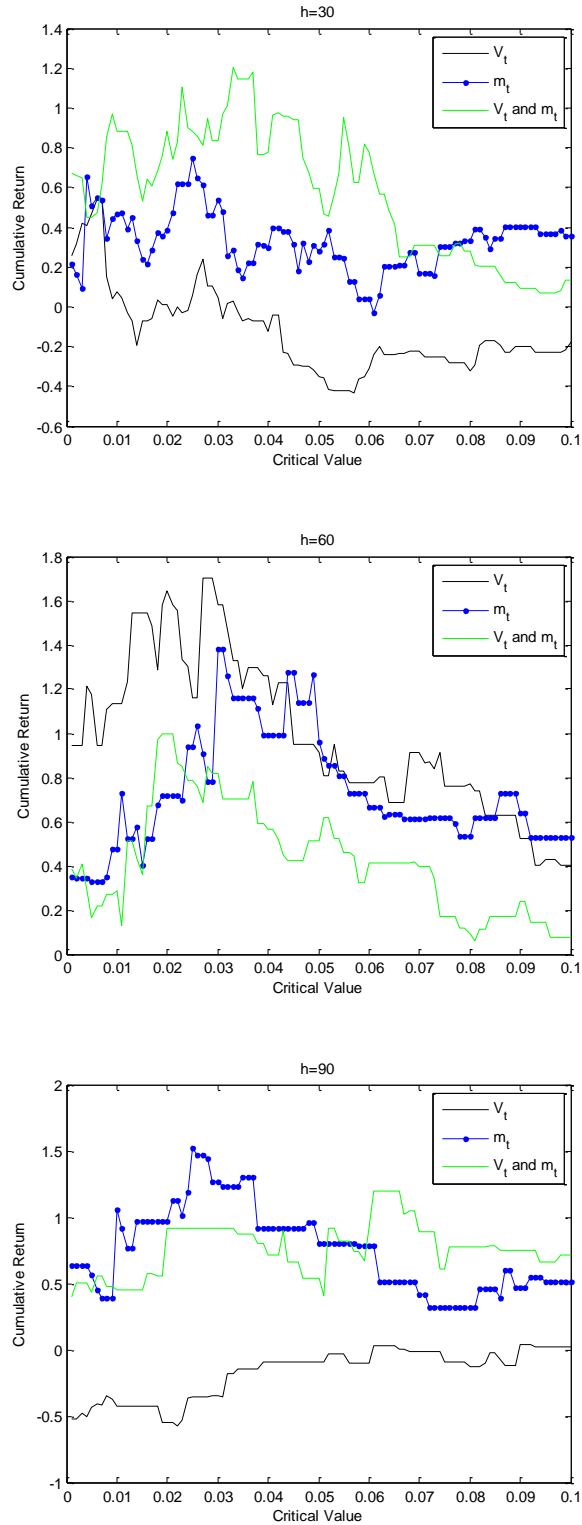
Figure 3   Cumulative returns of the two-factor SV framework proxy trading strategies

Note: The strategy, based upon turnover rates of 30, 60 and 90 calendar days and an estimation period of 180 calendar days, is to take up long (short) positions in the S&P 500 index when estimated returns exceed (are below) the critical value x (-x), and do nothing when estimated returns are in the interval, [-x, x]. The strategy is based upon the following estimated regression model:

$$ER_{t,t+h} = \alpha + \beta_1 V_t + \beta_2(m_t - V_t) + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 KURT_t + \beta_6 TMS_t + \beta_7 DFS_t + \varepsilon_t$$

APPENDIX A

We follow the CBOE methodology to calculate the squared VIX as:

$$VIX_{t,\tau}^2 \equiv 365 * \left(\frac{2}{\tau} \sum_i \frac{\Delta K_i}{K_i^2} e^{r_f * \frac{\tau}{365}} Q(K_i) - \frac{1}{\tau} \left(\frac{F}{K_0} - 1 \right)^2 \right).$$

where K_i is the strike price of the ith out-of-the-money (OTM) option; ΔK_i is the interval between two strike prices, defined as $\Delta K_i = (K_{i+1} - K_{i-1})/2$; in particular, ΔK_i for the lowest strike price is simply the difference between the lowest and the next higher strike price, i.e., $K_{i+1} - K_i$; similarly, the ΔK_i for the highest strike price is equal to $K_i - K_{i-1}$. r_f refers to the risk-free rate; τ is the time to expiration defined as the number of calendar days; $Q(K_i)$ is the midpoint of the bid-ask spread for each option with strike price K_i; F refers to the implied forward index level derived from the nearest-the-money index option prices based upon put-call parity; and K_0 is the first strike price below the forward index level.

We use the interpolation similar to that suggested by the CBOE to construct the VIX term structure with six maturities as:

$$VIX_{t,\tau}^2 = \left[T_1 * VIX_{t,T_1}^2 \left(\frac{T_2 - \tau}{T_2 - T_1} \right) + T_2 * VIX_{t,T_2}^2 \left(\frac{\tau - T_1}{T_2 - T_1} \right) \right] * \frac{1}{\tau}$$

where $\tau = 30$, 60, 90, 180, 270 or 360 days, and T_1 and T_2 are the two nearest maturities embracing τ. A similar methodology is employed to calculate the term structure of the forward squared VIX as:

$$FVIX_{t,T_1,T_2}^2 = VIX_{t,T_2}^2 \left(\frac{T_2}{T_2 - T_1} \right) - VIX_{t,T_1}^2 \left(\frac{T_1}{T_2 - T_1} \right), T_1 < T_2.$$