Term Structure of VIX Futures: A Cascade Model*

Zhiguang Wang†

Abstract

The current factor-based term structure models of VIX futures suffer from over parameterization and the curse of dimensionality: three additional parameters for an extra factor. The state-of-the-art 3-factor model proposed by Lu and Zhu (2010) has twelve parameters and tends to over-fit an average of five strips of daily VIX futures. To defeat the curse of dimensionality, we propose a new model of volatility by allowing for a cascading structure of volatility components, which allows us to add as many components as desired with additional cost of one more parameter. The flexibility in choosing the number of components enables a rich dynamics in the term structure of the cash VIX and VIX futures. We derive a semi-closed form solution to the VIX futures price. We compare the in-sample and out-of-sample performance of our model to the previous models. The preliminary results show that our model outperforms the existing factor-based models.

1 Introduction

VIX futures along with its options have become the second most actively traded contracts by the Chicago Board of Exchange (CBOE). The number of VIX futures contract months has

---

*The theoretical part of the paper has been finished and the data have been collected and processed. The empirical estimation is expected to be done by late March.
†Assistant Professor and Stahly Scholar in Financial Economics, Department of Economics at South Dakota State University, Brookings, SD 57007, USA
increased from two in early 2004 to nine in 2013. VIX futures pricing is always a focal point of academic research. Along with the expansion of VIX futures contracts, the literature of stochastic volatility model has evolved from single factor (Zhang and Zhu (2006); Lin (2007); Zhu and Zhang (2007); Zhang and Huang (2010) to two factors (Christoffersen et al. 2008; Egloff et al. 2009; Zhang et al. (2010), more recently to three factors by Lu and Zhu (2010). In particular, Lu and Zhu find that the third factor is statistically significant for variance term structure.

Lu and Zhu’s three-factor model largely represents the state of art, at least the most sophisticated model, for VIX futures pricing. Their model offers a rich structure to accommodate five strips of VIX futures in sample. The multi-factor model improves significantly on the short-term contract (30-day and 60-day). However, the results still show weakness: (1) their 3-factor model still generates large errors for 90-day and 270-day contracts; (2) the empirical results are not based on the original VIX futures data, but the interpolated (smoothed) data. The interpolation potentially hide the actual pricing errors of their model; (3) their out-of-sample test is only limited to eight days, which makes one hard to judge its merit.

Another shortcoming of the current factor-based term structure models of VIX futures suffer from over-parameterization and the curse of dimensionality. One needs three extra parameters for an extra factor. Lu and Zhu’s 2-factor model has seven or eight parameters and tends to over-fit an average of five strips of daily VIX futures, let alone their 3-factor model with ten or twelve parameters.

To address the above issues, we propose a new model of volatility by allowing for a cascading structure of volatility components similar to the interest rate model by Calvet et al. (2012). The cascading volatility model essentially has one governing factor with multiple layers. Such a structure allows us to add as many layers as desired with at most one new parameter for
each additional component. The flexibility in choosing the number of components enables a rich dynamics in the term structure of the cash VIX and VIX futures, which helps improve the in-sample and out-of-sample empirical performance of the model.

The rest of the paper is organized as follows. We first derive a semi-closed form solution to the VIX futures price in Section 2. We describe daily cash VIX and VIX futures data from 2004 to 2012 and further discuss the estimation method of the Unscented Kalman Filtering method in Section 3. We compare the in-sample and out-of-sample performance of our model to the previous models in Section 4. Section 5 concludes.

2 Model

2.1 $\mathbb{P}$-Measure Dynamics

Denote $V_t$ instantaneous variance. $V_t$ is an ending point of the cascading volatility $\sigma^2_{j,t}$, where $j$ stands for the $j$th component. Therefore, $V_t = \sigma^2_{n,t}$. The higher-frequency (or faster moving) component $\sigma^2_{j,t}$ reverts to the lower-frequency component $\sigma^2_{j-1,t}$, until it reaches a constant long-run mean $\theta_v$. The structure is presented as follows:

$$
\begin{align*}
  d\sigma^2_{j,t} &= \kappa_j(\sigma^2_{j-1,t} - \sigma^2_{j,t})dt + \omega_j dW_{j,t}, \quad j = 1, 2, \ldots, n \\
  \sigma^2_{0,t} &= \theta_v \\
  \sigma^2_{n,t} &= \sigma^2_t \equiv V_t \\
  \kappa_j &= \kappa_1 \beta^{j-1}, \quad \beta > 1
\end{align*}
$$

Denote $X_t = (\sigma^2_{1,t}, \sigma^2_{2,t}, \ldots, \sigma^2_{n,t})'$. The drift and diffusion terms of $X_t$ are denoted as $\mu(X)$ and $\Sigma(x)$. The dynamics of $n$-dimensional volatility cascade can be rewritten in matrix form as follows:

$$
dX_t = \mu(X_t)dt + \Sigma(X_t)dW_t
$$
where both $\mu(X)$ and $\Sigma(X)$ have an affine structure:

$$
\mu(X) = 
\begin{pmatrix}
\kappa_1 \theta_v \\
0 \\
\vdots \\
0
\end{pmatrix}
+ 
\begin{pmatrix}
-k_1 & 0 & 0 \\
\kappa_2 & -k_2 & 0 \\
0 & \ldots & \ldots & 0
\end{pmatrix}
X
$$

(6)

$$
\Sigma(X) = 
\begin{pmatrix}
\omega_1^2 & 0 & 0 \\
0 & \omega_2^2 & 0 \\
\vdots & \vdots & \ddots \\
0 & 0 & 0 & \omega_n^2
\end{pmatrix}
$$

(7)

2.2 Q-Measure Dynamics

In order to price VIX futures, we specify a measure change from P-measure (physical) to Q-measure (risk neutral) through the Radon-Nikodym derivative:

$$
\frac{dQ}{dP} = \prod_{j=1}^{n} \exp \left( - \int_0^t \gamma_{j,s} \omega_j dW_{j,s} - \frac{1}{2} \int_0^t \gamma_{j,s}^2 \omega_j^2 ds \right)
$$

(9)

The instantaneous variance dynamics under the risk-neutral measure becomes:

$$
d\sigma_{j,t}^2 = -\gamma_{j,t} \omega_j^2 dt + \kappa_j (\sigma_{j-1,t}^2 - \sigma_{j,t}^2) dt + \omega_j dW_{j,t}^Q
$$

(10)

We assume affine risk premia $\gamma_{j,t} = \gamma_j + \lambda_j^T X_t$ by denoting $\lambda_j = (\lambda_{j,1}, \ldots, \lambda_{j,n})$ and obtain the risk-neutral dynamics in the following matrix form:

$$
dX_t = \mu^Q(X_t) dt + \Sigma^Q(X_t) dW_t
$$

(11)
where both $\mu^Q(X)$ and $\Sigma^Q(X)$ have an affine structure:

\[
\begin{align*}
\mu^Q(X) &= \begin{pmatrix}
\kappa_1 \theta_v - \gamma_1 \omega_1^2 \\
-\gamma_2 \omega_2^2 \\
\vdots \\
-\gamma_n \omega_n^2
\end{pmatrix}
\end{align*}
\]

\[
\begin{pmatrix}
(1) \\
(2) \\
\vdots \\
(2n)
\end{pmatrix}
\]

\[
\begin{pmatrix}
\kappa_0^Q \\
\kappa_1^Q \\
\kappa_2^Q \\
\vdots \\
\kappa_n^Q
\end{pmatrix}
\]

\[
\begin{align*}
\Sigma^Q(X) &= \begin{pmatrix}
\omega_1^2 & 0 & 0 & 0 \\
0 & \omega_2^2 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \omega_n^2
\end{pmatrix}
\end{align*}
\]

\[
\begin{pmatrix}
(1) \\
(2) \\
\vdots \\
(2n)
\end{pmatrix}
\]

\[
\begin{pmatrix}
H_0^Q \\
K_0^Q \\
K_1^Q \\
\vdots \\
K_n^Q
\end{pmatrix}
\]

\[
\begin{align*}
X &= \begin{pmatrix}
-\kappa_1 - \lambda_{11} \omega_1^2 & -\lambda_{12} \omega_2^2 & \ldots & -\lambda_{1n} \omega_n^2 \\
\kappa_2 - \lambda_{21} \omega_1^2 & -\kappa_2 - \lambda_{22} \omega_2^2 & \ldots & -\lambda_{2n} \omega_n^2 \\
\vdots & \vdots & \ddots & \vdots \\
-\lambda_{n1} \omega_1^2 & \ldots & \kappa_n - \lambda_{nn-1} \omega_{n-1}^2 & -\kappa_n - \lambda_{nn} \omega_n^2
\end{pmatrix}
\end{align*}
\]

However, the complex structure in $K_1^Q$ renders no analytically tractable solution to VIX futures prices. We make a simplifying assumption of constant risk premia, i.e. $\lambda_j = 0$. Constant risk premia are commonly assumed in VIX derivatives pricing as in Lu and Zhu
The risk-neutral dynamics of $\mu^Q(X)$ and $\Sigma^Q(X)$ become

$$
\begin{align*}
\mu^Q(X) &= 
\begin{pmatrix}
\kappa_1 \theta_v - \gamma_1 \omega^2_1 \\
-\gamma_2 \omega^2_2 \\
... \\
-\gamma_n \omega^2_n \\
\end{pmatrix}
\begin{pmatrix}
\begin{bmatrix}
-K_1 & 0 & ... & 0 \\
-K_2 & -K_2 & ... & 0 \\
& & & \vdots \\
0 & ... & -K_n & -K_n \\
\end{bmatrix}
\end{pmatrix} X
\end{align*}
$$

(14)

$$
\Sigma^Q(X) = 
\begin{pmatrix}
\omega^2_1 & 0 & 0 & 0 \\
0 & \omega^2_2 & 0 & 0 \\
... & ... & ... & ... \\
0 & 0 & 0 & \omega^2_n \\
\end{pmatrix}
\begin{pmatrix}
\begin{bmatrix}
H_0^Q \\
H_1^Q \\
\end{bmatrix}
\end{pmatrix}
$$

(15)

Based on the risk-neutral dynamics, we derive the instantaneous variance $V_t \equiv \sigma^2_{n,t}$ in the following proposition.

**Proposition 1.** The instantaneous variance $V_t$ can be represented by

$$
V_t \equiv \sigma^2_{n,t} = \theta_v + \sum_{j=1}^{n} a_j(t)(\sigma^2_{j,0} - \theta_v) + \sum_{j=1}^{n} b_j(t) \gamma_j \omega^2_j + \sum_{j=1}^{n} \omega_j \int_{0}^{t} a_j(t - s)dW_{j,s}
$$

(16)

$$
\begin{align*}
a_j(t) &= (K_j \ast ... \ast K_n)(t)/\kappa_j \\
&= \sum_{i=j}^{n} \prod_{s=j,s\neq i}^{n} (\kappa_s - \kappa_i) e^{-\kappa_i t}
\end{align*}
$$

(17)

(18)

$$
\begin{align*}
b_j(t) &= \prod_{j=1}^{n} (1 - e^{-\kappa_j t})
\end{align*}
$$

(19)

where response function $a_j(t)$ is the convolution of exponential functions $K_j = \kappa_j e^{-\kappa_j t} \mathbb{1}_{[t \geq 0]}$.

Proof. See Appendix □
2.3 Derivation of VIX Futures Pricing Formula

Given the instantaneous variance under the risk-neutral measure, we can write the square of the spot VIX as

$$VIX_t^2 = \frac{1}{\tau} E_t^Q \left[ \int_t^{t+\tau} \sigma_{n,s}^2 ds \right]$$

(20)

where $\tau$ is fixed as 30 days according to the CBOE.

By inserting Equation (16) into Equation (20), we obtain $VIX_T^2$ as a linear combination of the n factors:

$$VIX_T^2 = \frac{1}{\tau} E_T^Q \left[ \int_T^{T+\tau} \left[ \theta_v + \sum_{j=1}^{n} a_j(s)(\sigma_{j,T}^2 - \theta_v) + \sum_{j=1}^{n} b_j(s)\gamma_j\omega_j^2 \right] ds \right]$$

$$+ \frac{1}{\tau} E_T^Q \left[ \sum_{j=1}^{n} \omega \int_T^{T+s} a_j(s - u)dW_{j,u}ds \right]$$

$$= \frac{1}{\tau} \left[ \int_T^{T+\tau} \left[ \theta_v + \sum_{j=1}^{n} a_j(s)(\sigma_{j,T}^2 - \theta_v) + \sum_{j=1}^{n} b_j(s)\gamma_j\omega_j^2 \right] ds \right]$$

$$= \theta_v + \frac{1}{\tau} \left[ \left( \int_T^{T+\tau} \sum_{j=1}^{n} a_j(s)ds \right)(\sigma_{j,T}^2 - \theta_v) \right] + \frac{1}{\tau} \int_T^{T+\tau} \sum_{j=1}^{n} b_j(s)\gamma_j\omega_j^2 ds$$

$$+ \frac{1}{\tau} \int_T^{T+\tau} \sum_{j=1}^{n} b_j(s)\gamma_j\omega_j^2 ds$$

$$= \theta_v + \frac{1}{\tau} \sum_{j=1}^{n} A(j)(\sigma_{j,T}^2 - \theta_v) + \frac{1}{\tau} \int_T^{T+\tau} \sum_{j=1}^{n} b_j(s)\gamma_j\omega_j^2 ds$$

$$= \tilde{A}'\tilde{\sigma}_T + B$$

(21)
where

\[ A(j) = \frac{1}{\tau} \sum_{i=j}^{n} \prod_{s=j, s \neq i}^{n} (\kappa_s - \kappa_i) \left( e^{-\kappa_i T} - e^{-\kappa_i (T+\tau)} \right) \]

\[ B = \theta^Q_t (1 - \frac{1}{\tau} \sum_{j=1}^{n} A(j) + \frac{1}{\tau} \int_{T}^{T+\tau} \sum_{j=1}^{n} b_j(s) \gamma_j \omega^2_j ds) \]

\[ \tilde{\mathbf{A}} = (A(1), A(2), ... A(n))' \]

\[ \tilde{\sigma}^2_T = (\sigma^2_{1,T}, \sigma^2_{2,T}, ... \sigma^2_{n,T})' \]

Let \( F(t, T) \) be the futures price at time \( t \) expiring at time \( T \), we have

\[ F(t, T) = E^Q_t \left[ \sqrt{VIX^2_T} \right]. \tag{22} \]

Schurger (2002) shows that the expectation of the square root of a variable \( Z = VIX^2_T \) can be expressed in terms of moment generating functions as follows

\[ E[\sqrt{z}] = \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} \frac{1 - E[e^{-sz}]}{s^2} ds \tag{23} \]

Thanks to the affine structure of \( \sigma^2_{j,T} \), the moment generating function \( \Psi(s) = E[e^{-sz}] = E[e^{-s(\tilde{\mathbf{A}}' \tilde{\sigma}^2_T + B)}] \) admits an exponential affine form according to Duffie, Pan and Singleton (2000). The affine solution takes the form \( \Psi(s) = \exp(\alpha(t) + \beta(t) \cdot \tilde{\sigma}_t) \) with \( \alpha(t) \) and \( \beta(t) \) satisfying the following Riccati equations:

\[ \dot{\beta}(t) = \rho_1 - K_1^T \beta(t) - \frac{1}{2} \beta(t)^T H_1 \beta(t) - l_1(\theta(\beta(t)) - 1) \tag{24} \]

\[ \dot{\alpha}(t) = \rho_0 - K_0 \cdot \beta(t) - \frac{1}{2} \beta(t)^T H_0 \beta(t) - l_0(\theta(\beta(t)) - 1) \tag{25} \]

The above equations can be solved symbolically with Matlab. With \( \alpha \) and \( \beta \), VIX futures
price can be expressed as follows:

$$F(t, T) = E_t^Q \left[ \sqrt{VIX_t^2} \right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sz}]}{s^2} ds \quad (26)$$

$$= \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-s[B + \alpha(t) + \beta(t)\sigma_t]}}{s^2} ds \quad (27)$$

where $\alpha(t)$ and $\beta(t)$ are determined by Equations (25).

3 Data and Methodology

3.1 Data

We obtain daily VIX futures close prices along with cash VIX prices from the Chicago Board of Exchange (CBOE), ranging from 2004 to 2012. The daily interest rates are obtained from Federal Reserve St. Louis website. Daily interest rates are linearly interpolated to match the expiries of VIX futures. On any given day, there are 5 to 9 contracts traded.

3.2 Methodology

VIX futures pricing formula given in Equation (27) is a non-linear function of the state variables $\sigma^2_{j,t}$. We propose to estimate the model using maximum likelihood estimation via Unscented Kalman Filtering (UKF). UKF is especially suited for non-linear state-space models (see Wan and van der Merve (2001)). The application of UKF to model estimation has been employed in the derivatives literature in recent years (Carr and Wu (2007) for currency options and Carr and Wu (2010) for equity options and credit default swap).

We discretize the state transition given by Equation (10) in the following matrix form:

$$X_{t+1} = K_0^Q \Delta t + K_1^Q X_t + \sqrt{\Sigma} u_{t+1}$$  \quad (28)
where $\Delta t = 1/252$, $u_{t+1}$ follows standard normal distribution, and $K_0, K_1, \Sigma$ are defined in Equations (14-15). We further define the measurement equation for VIX futures prices as follows:

$$F_t = f(X_t) + e_t.$$  \hspace{1cm} (29)

where $F_t$ is the observed VIX futures price, $f(X_t)$ is the pricing formula given in Equation (27), and $e_t$ follows normal distribution with mean 0 and standard deviation $\sigma_e$.

The unscented Kalman filter approximates the posterior state density using a set of sample points. These sample points produce the true mean and covariance of the normally distributed state variables. The posterior mean and variance/covariance of futures prices, which is nonlinear function of state variables, can be approximated based on the propagated sample points. The technical details are referred to Wan and van der Merve (2001).

4Results

We evaluate the in-sample and out-of-sample performance of our model and compare it to Lu and Zhu’s (2010) models. [The results are expected to be available by late March.]

5Conclusion

Our term structure model of VIX futures defeats the curse of dimension problem. The flexibility of our cascading structure allows us to include as many factors as we prefer. The semi-closed form solution further enables fast calibration. We expect our model to outperform the previous models both in- and out-of-sample.

APPENDIX
We apply Ito’s lemma to $e^{\kappa_n t} \sigma_{n,t}^2$ and obtain

$$d(e^{\kappa_n t} \sigma_{n,t}^2) = \kappa_n e^{\kappa_n t} \sigma_{n,t}^2 dt + e^{\kappa_n t} d\sigma_{n,t}^2$$

$$= \kappa_n e^{\kappa_n t} \sigma_{n,t}^2 dt + e^{\kappa_n t} (\gamma_n \omega_n^2 dt + \kappa_n (\sigma_{n-1,t}^2 - \sigma_{n,t}^2) dt + \omega_n dW_{n,t})$$

$$= e^{\kappa_n t} (\kappa_n \sigma_{n-1,t}^2 dt - \gamma_n \omega_n^2 dt + \omega_n \sigma_{n,t} dW_{n,t})$$

Integrating the above equality and dividing it by $e^{\kappa_n t}$ yield

$$\sigma_{n,t}^2 = e^{-\kappa_n t} \sigma_{n,0}^2 + (1 - e^{-\kappa_n t}) \gamma_n \omega_n^2 + \int_0^t \kappa_n e^{-\kappa_n (t-s)} \sigma_{n-1,s}^2 ds + \int_0^t \omega_n e^{-\kappa_n (t-s)} dW_{n,s}$$

$$\sigma_{n-1,t}^2 = e^{-\kappa_n (t-s)} \sigma_{n-1,0}^2 + (1 - e^{-\kappa_n (t-s)}) \gamma_{n-1} \omega_{n-1}^2 + \int_0^t \kappa_{n-1} e^{-\kappa_n (t-s)} \sigma_{n-2,s}^2 ds + \int_0^t \omega_{n-1} e^{-\kappa_n (t-s)} dW_{n-1,s}$$

$$\sigma_{n,t}^2 = e^{-\kappa_n t} \sigma_{n,0}^2 + (1 - e^{-\kappa_n t}) \gamma_n \omega_n^2 + \int_0^t \kappa_n e^{-\kappa_n (t-s)} e^{-\kappa_n-1 t} \sigma_{n-1,0}^2 ds$$

$$+ (1 - e^{-\kappa_n t})(1 - e^{-\kappa_n-1 t}) \gamma_{n-1} \omega_{n-1}^2 + \int_0^t \kappa_{n-1} e^{-\kappa_n (t-s)} \sigma_{n-1,0}^2 ds + \int_0^t \omega_{n-1} e^{-\kappa_n (t-s)} dW_{n-1,s}$$

$$+ \int_0^t \kappa_{n-1} e^{-\kappa_n (t-s)} \sigma_{n-2,s}^2 ds + \int_0^t \omega_{n-1} e^{-\kappa_n (t-s)} dW_{n,s}$$

$$= e^{-\kappa_n t} \sigma_{n,0}^2 + \left( \int_0^t \kappa_n e^{-\kappa_n (t-s)} e^{-\kappa_n-1 t} ds \right) \sigma_{n-1,0}^2$$

$$+ \left( \int_0^t \kappa_{n-1} e^{-\kappa_n (t-s)} \sigma_{n-1,0}^2 ds \right) \sigma_{n-2,0}^2$$

$$+ \sum_{j=1}^n a_j(t) \sigma_{j,0}^2 + \theta(t) (1 - \sum_{j=1}^n a_j(t)) + \sum_{j=1}^n b_j(t) \gamma_j \omega_j^2 + \sum_{j=1}^n \omega_j \int_0^t a_j(t-s) dW_{j,s}$$
where

\[ a_j(t) = (K_j \ast \ldots \ast K_n)(t)/\kappa_n \]

\[ K_j = \kappa_n e^{-\kappa_n t} \mathbb{1}_{[t \geq 0]} \]

\[ b_j(t) = \prod_{j=1}^{n} (1 - e^{-\kappa_j t}) \]

References


