A Dimension-invariant Cascade Model for VIX Futures

Zhiguang Wang* Brice Dupoyet†

*Associate Professor, Department of Economics at South Dakota State University, Brookings, SD 57007, USA. zhiguang.wang@sdstate.edu
†Associate Professor, Department of Finance at Florida International University, Miami, FL 33199, USA. dupoyetb@fiu.edu
A Dimension-invariant Cascade Model for VIX Futures

Abstract

We propose a new model of volatility by allowing for a cascading structure of volatility components. The cascading feature is achieved by introducing an increasing structure to the speed of mean reversion. It allows us to add as many components as desired with no additional parameter, effectively defeating the curse of dimensionality often seen in traditional models. The flexibility in choosing the number of components enables rich dynamics in the term structure of both spot VIX and VIX futures, without the need to introduce price jumps. We derive a semi-closed form solution to the VIX futures price, and find that our 6-factor model with only 6 parameters can closely fit spot VIX and VIX futures data from 2004 to 2015 and produce out of sample pricing errors of magnitudes similar to those of in-sample errors.

1 Introduction

VIX futures and options contracts have become the second most actively traded contracts on the Chicago Board of Exchange (CBOE). The number of VIX futures contracts’ months has increased from four in March 2004 to ten in 2009, and eventually stablized at nine until 2016. VIX futures pricing has always been a focal point of academic research. Along with the expansion of VIX futures contracts, the literature on stochastic volatility models has evolved from a 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. (2010); Zhang et al. (2010); Luo and Zhang (2012)), and more recently to three factors by Lu and Zhu (2013). 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, or at least the most sophisticated model, for VIX futures pricing. Their model offers a rich structure able to accommodate five strips of VIX futures contracts in sample. The multi-factor model improves significantly on the short-term contract (30-day and 60-day). However, their results still show some weaknesses: (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 on interpolated (smoothed) data. The interpolation potentially hides the actual pricing errors of the model; (3) their out-of-sample test is only limited to eight days, which makes it difficult for one to judge its merit.

Another shortcoming of VIX futures current factor-based term structure models is that they suffer from the curse of dimensionality. One generally needs three extra parameters for each extra factor being added. Lu and Zhu’s 2-factor model has seven or eight parameters, while their 3-factor model contains between 10 and 12 parameters. The likely overfitting of the five strips of VIX futures contracts is a distinctive issue.

To address the above issues, we propose a new model of volatility by allowing for a cascading structure of volatility components, motivated by the interest rate model by Calvet et al.
The cascading volatility model essentially has one governing factor with multiple layers. Such a structure allows one to add as many layers as desired without any additional parameters, hence it is dimension-free. The flexibility in choosing the number of components enables rich dynamics in the term structure of both spot VIX and VIX futures, which in turn helps improve the in-sample and out-of-sample empirical performance of the model. Using the unscented Kalman filtering method, we estimate a 6-factor model on price data from March 2004 to December 2015. We find that our model can generate low in-sample pricing errors and equally desirable pricing performance out of sample (from January to August 2016).

The rest of the paper is organized as follows. We first derive a semi-closed form solution to the VIX futures price for the generic n-factor cascading model of volatility in Section 2. We describe daily spot VIX and VIX futures data from 2004 to 2016 and further discuss the Unscented Kalman Filtering method in Section 3. We report the in-sample and out-of-sample performance of our model in Section 4, and Section 5 concludes.

2 Model

2.1 \(\mathbb{P}\)-Measure Dynamics

Denote \(V_t\) instantaneous variance. \(V_t\) is the ending point of a cascading volatility \(\sigma^2_{j,t}\) process, 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:

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

Let \(X_t = (\sigma^2_{1,t}, \sigma^2_{2,t}, \ldots, \sigma^2_{n,t})'\). 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 the following affine structure:

\[
\mu(X) = \begin{pmatrix} \kappa_1 \theta_v \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 & 0 & 0 \\ \kappa_2 & -\kappa_2 & 0 & 0 \\ 0 & \ldots & \ldots & 0 \\ 0 & 0 & \kappa_n & -\kappa_n \end{pmatrix} X
\]

\[
\Sigma(X) = \begin{pmatrix} K_0 \\ K_1 \end{pmatrix}
\]
\[ \Sigma(X) = \begin{pmatrix} 
\omega_1^2 & 0 & 0 & 0 \\
0 & \omega_2^2 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \omega_n^2 
\end{pmatrix} \]  \hspace{1cm} (7)

\[ \Sigma^Q(X) = \begin{pmatrix} 
\omega_1^2 & 0 & 0 & 0 \\
0 & \omega_2^2 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \omega_n^2 
\end{pmatrix} \]  \hspace{1cm} (8)

### 2.2 \( Q \)-Measure Dynamics

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

\[ \frac{dQ}{dP} = \prod_{j=1}^{n} \exp \left( - \int_{0}^{t} \gamma_{j,s} \omega_j dW_{j,s} - \frac{1}{2} \int_{0}^{t} \gamma_{j,s}^2 \omega_j^2 ds \right) \]  \hspace{1cm} (9)

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

\[ d\sigma_{j,t}^2 = -\gamma_{j,t} \omega_j^2 dt + \kappa_j (\sigma_{j-1,t}^2 - \sigma_{j,t}^2) dt + \omega_j dW_{j,t}^Q \]  \hspace{1cm} (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:

\[ dX_t = \mu^Q(X_t) dt + \sigma^Q(X_t) dW_t \]  \hspace{1cm} (11)

where both \( \mu^Q(X) \) and \( \Sigma^Q(X) = [\sigma^Q(X_t)] [\sigma^Q(X_t)]' \) have the following affine structure:

\[ \mu^Q(X) = \begin{pmatrix} 
\kappa_1 \theta_v - \gamma_1 \omega_1^2 \\
-\gamma_2 \omega_2^2 \\
\vdots \\
-\gamma_n \omega_n^2 
\end{pmatrix} + \begin{pmatrix} 
-\kappa_1 - \lambda_{11} \omega_1^2 \\
\kappa_2 - \lambda_{21} \omega_1^2 \\
\vdots \\
-\lambda_{n1} \omega_1^2 
\end{pmatrix} X \]  \hspace{1cm} (12)

\[ \Sigma^Q(X) = \begin{pmatrix} 
\omega_1^2 & 0 & 0 & 0 \\
0 & \omega_2^2 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \omega_n^2 
\end{pmatrix} \]  \hspace{1cm} (13)
However, the complex structure in $K_1^Q$ yields no analytically tractable solution to VIX futures prices. We make the simplifying assumption of a constant risk premium, i.e. $\lambda_j = 0$. Constant risk premia are commonly assumed in VIX derivatives pricing as in Lu and Zhu (2010) and Zhang and Zhu (2006). The risk-neutral dynamics of $\mu^Q(X)$ and $\Sigma^Q(X)$ become

$$
\mu^Q(X) = \begin{pmatrix}
\kappa_1 \theta_v - \gamma_1 \omega_1^2 \\
-\gamma_2 \omega_2^2 \\
\ddots \\
-\gamma_n \omega_n^2 \\
\end{pmatrix} + \begin{pmatrix}
-\kappa_1 & 0 & \cdots & 0 \\
0 & -\kappa_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & -\kappa_n \\
\end{pmatrix} X 
$$

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

From these risk-neutral dynamics, we derive the instantaneous variance $V_t \equiv \sigma_{n,t}^2$ in the following proposition.

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

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

$$
a_j(t) = (K_j \ast \cdots \ast K_n)(t)/\kappa_j 
$$

$$
a_j(t) = \sum_{i=j}^n \frac{K_j+\cdots+K_n}{\prod_{s=j, s \neq i}(\kappa_s - \kappa_i)} e^{-\kappa_i t} 
$$

$$
b_j(t) = \sum_{i=j}^n a_i(t) - 1)/\kappa_j 
$$

where $\sigma_{j,0}^2$ is the initial instantaneous variance of component $j$ and the 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 A

2.3 Derivation of the VIX Futures Pricing Formula

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

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

where $\tau$ is fixed at 30 days according to the CBOE.
By inserting Equation (16) into Equation (20), we obtain $VIX_T^2$, the spot 30-day variance at expiration time $T$, as a linear combination of the $n$ factors:

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

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

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

\[
= \tilde{A}' X_T + B
\]

where

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

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

\[
= \theta_v \left( 1 - \sum_{j=1}^{n} A(j) \right) + \frac{1}{T} \int_T^{T+\tau} \sum_{j=1}^{n} \sum_{i=j}^{n} a_i(s) - \frac{1}{\kappa_j} \gamma_j \omega_j^2 ds
\]

\[
= \theta_v \left( 1 - \sum_{j=1}^{n} A(j) \right) + \sum_{j=1}^{n} \left( \sum_{i=j}^{n} A(i) - 1 \right) \frac{\gamma_j \omega_j^2}{\kappa_j}
\]

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

\[
X_T = \sigma_T^2 = (\sigma_{1,T}^2, \sigma_{2,T}^2, ... \sigma_{n,T}^2)
\]

\[\text{Note that } \sigma_{j,T} \text{ is used instead of the original notation } \sigma_{j,0} \text{ in Equation (16), because the initial time "0" for the time interval } [T, T + \tau] \text{ is } T.\]
Letting \( F(t, T) \) be the futures price at time \( t \) expiring at time \( T \), we have

\[
F(t, T) = E_t^Q \left[ \sqrt{VIX_T^2} \right].
\]  

(22)

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

\[
E[\sqrt{z}] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sz}]}{s^{3/2}} ds
\]

(23)

Thanks to the affine structure of \( VIX_T^2 \) in terms of \( X_T \), the moment generating function \( \Psi(s) = E[e^{-sz}] = E[e^{-s(\tilde{A} X_T + B)}] \) admits an exponential affine form according to Duffie, Pan and Singleton (2000). The affine solution takes the form \( \Psi(s) = \exp(-sB + \alpha(t; s) + \beta(t; s) \cdot X_t) \) with \( \alpha(t; s) \) and \( \beta(t; s) \) satisfying the following Ricatti equations:

\[
\dot{\beta}(t; s) = -K_T^T \beta(t; s)
\]

(24)

\[
\dot{\alpha}(t; s) = -K_0 \cdot \beta(t; s) - \frac{1}{2} \beta(t; s)^T H_0 \beta(t; s)
\]

(25)

with boundary conditions \( \alpha(T; s) = 0 \) and \( \beta(T; s) = -s\tilde{A} \). \( H_0 \) and \( K_1 \) are from Equation (14) and (15), respectively.

The above equations can be solved in closed-form (see Appendix B for a 3-component case). With \( \alpha \) and \( \beta \) derived, the VIX futures price can be computed as follows:

\[
F(t, T) = E_t^Q \left[ \sqrt{VIX_T^2} \right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sz}]}{s^{3/2}} ds
\]

(26)

\[
= \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-s[B+\alpha(t; s)+\beta(t; s)\tilde{A}^2]}}{s^{3/2}} ds
\]

(27)

where \( \alpha(t; s) \) and \( \beta(t; s) \) are determined by Equations (24) and (25). Given that \( X_t = \tilde{\sigma}_t^2 \) is defined in Equation (14), the VIX futures pricing function \( F(t, T) \) can be expressed as a non-linear function of state variables \( X_t \), i.e. \( F(t, T) = f(X_t) \).

3 Data and Methodology

3.1 Data

We obtain daily VIX futures closing prices along with spot VIX prices from the Chicago Board of Exchange (CBOE), ranging from 2004 to 2016. On any given day, the number of contracts listed varies from 4 to 11 during the sample period. In order to visualize the sample time series, we linearly interpolate the VIX futures prices to obtain the prices with 1-month, 2-month, 3-month, 6-month and 9-month maturities. VIX futures prices are shown in Figure (1). We can see that, on average, VIX futures prices exhibit a contango shape, in
which futures price with a longer maturity is higher than that with a shorter maturity.

Table I reports spot VIX and VIX futures prices of 3,019 trading days from March 2004 to August 2016. In contrast to the term structure of the price of VIX futures, the volatility of VIX futures prices exhibits a term structure of normal backwardation, i.e. higher volatility levels over shorter maturities. Both spot VIX and VIX futures show positive skewness and excess kurtosis.

![Spot VIX and VIX Futures Prices (2004-2016)](image)

**Figure 1: Spot VIX and VIX Futures Prices**

<table>
<thead>
<tr>
<th>Contract</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1m</td>
<td>20.3563</td>
<td>8.0071</td>
<td>18.0375</td>
<td>11.2636</td>
<td>65.6763</td>
<td>1.9664</td>
<td>7.7926</td>
</tr>
<tr>
<td>2m</td>
<td>20.9884</td>
<td>7.3580</td>
<td>19.0321</td>
<td>12.1952</td>
<td>60.1391</td>
<td>1.6835</td>
<td>6.5813</td>
</tr>
<tr>
<td>3m</td>
<td>21.3855</td>
<td>6.8616</td>
<td>19.6643</td>
<td>12.7787</td>
<td>54.6700</td>
<td>1.4332</td>
<td>5.4088</td>
</tr>
<tr>
<td>6m</td>
<td>22.1791</td>
<td>6.0889</td>
<td>20.7862</td>
<td>13.6385</td>
<td>45.0000</td>
<td>0.9864</td>
<td>3.6439</td>
</tr>
<tr>
<td>9m</td>
<td>22.6469</td>
<td>5.6536</td>
<td>21.3757</td>
<td>9.2482</td>
<td>43.9900</td>
<td>0.7866</td>
<td>3.1008</td>
</tr>
</tbody>
</table>
3.2 Methodology

The VIX futures pricing formula derived in Equation (27) is a non-linear function of the unobservable state variables $\sigma_{j,t}^2$. We propose to estimate the model using Unscented Kalman Filtering (UKF). 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, nonlinear functions of state variables, can be approximated based on the propagated sample points. 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. Christoffersen et al. (2014) provide a detailed examination of the nonlinear Kalman Filtering, especially the UKF to affine term structure model of interest rates. Carr and Wu (2007) have applied the UKF method to currency options and Carr and Wu (2010) to equity options and credit default swaps.

To implement the UKF, we discretize the state transition shown in Equation (11) in the following matrix form:

$$
\begin{align*}
\text{d}X_t &= -K_Q^Q(K_0^Q - X_t)dt + \sqrt{\Sigma} dW_t \\
X_t &= (I - \Phi)K_0^Q + \Phi X_{t-\Delta t} + \sqrt{\Sigma \Delta t} u_t
\end{align*}
$$

with

$$
\Phi = e^{K_Q^Q \Delta t},
$$

$$
K_0^Q = -[K_1^Q]^{-1} K_0^Q,
$$

$$
[K_1^Q]^{-1} = \begin{pmatrix}
-\frac{1}{\kappa_1} & 0 & \cdots & 0 \\
-\frac{1}{\kappa_2} & -\frac{1}{\kappa_2} & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
-\frac{1}{\kappa_n} & -\frac{1}{\kappa_n} & \cdots & -\frac{1}{\kappa_n}
\end{pmatrix}
$$

Furthermore, the daily time interval $\Delta t = 1/365$, and $I$ is the identity matrix. $K_0^Q, K_1^Q, \Sigma$ are defined in Equations (14-15), while $u_t$ follows the standard normal distribution. We further define the measurement equation for VIX futures prices as follows:

$$
F_t = f(X_t) + \epsilon_t.
$$

where $F_t$ is the observed VIX futures price, $f(X_t)$ is the pricing formula given in Equation (27), and the error term $\epsilon_t$ follows a normal distribution with a mean of zero and a standard deviation $\sigma_\epsilon$. Since we have a varying number of VIX futures contracts (between 4 and 11 depending on the day), we simplify the error structure by assuming a constant $\sigma_\epsilon$ across contracts.\(^2\)

\(^2\)We considered the alternative of removing the restriction of identical errors across contracts and found no significant improvement in empirical results. However, the gain in computational speed with a simpler structure outweighs the gains in accuracy, especially when the number of contracts increases dramatically.
For $M$ number of contracts on a given day, we have the following pricing equations:

$$
F_{1,t} = f_1(X_t, \Theta) + e_{1,t} \tag{30}
$$

$$
F_{2,t} = f_2(X_t, \Theta) + e_{2,t} \tag{31}
$$

$$
\ldots
$$

$$
F_{M,t} = f_M(X_t, \Theta) + e_{M,t} \tag{32}
$$

where $\Theta$ is the parameter vector in the model, including $\kappa_1, b, \omega, \gamma, \theta_v, e$.

Given the observations and the state transition Equation (28), we employ the unscented Kalman filtering for parameter and state estimation. Specifically, we set the objective function for the nonlinear least squares for $N$ days of observations as:

$$
\sum_{t=1}^{N} \sum_{m=1}^{M} (F_{m,t} - f_m(X_t, \Theta))^2. \tag{28}
$$

By minimizing the objective function under the UKF scheme, we can estimate the parameters $\Theta$ along with the state variables $X_t = \{x_{1t}, \ldots, x_{nt}\}$. A practical challenge to implementing the UKF methodology in our setting is the estimation of the error variance matrix for $e_{m,t}$ for a varying $M$ number of contracts. By assuming an identical error distribution across contracts, we only need to update one single value, the standard deviation of the error term $e$. Our approach differs from Lu and Zhu (2010) in the sense that we use the original data directly, instead of resorting to smoothing out the data series by interpolation.

4 Results

4.1 Parameter Estimation

As previously mentioned, there are between 4 and 11 contract months traded on any given day. Adding the spot VIX to the mix, the median number of observations per day is 9. We choose to apply a 6-factor model to the term structure of daily VIX futures from March 2004 to December 2015. Parameter estimates and their standard errors are reported in Table II. We can see that all parameters are statistically significant at the 1% level, except the long-run mean $\theta_v$. A positive and significant $b$ parameter lends empirical support to the cascading feature of mean reversion. Additionally, $\gamma$ is negative, confirming the well-known negative variance risk premium. Despite the shortcomings of the assumption of identical errors across contracts, we can still identify a statistically significant estimate for pricing errors $e$. The low value of 0.0002 indicates a good match between our model and the data.

4.2 Pricing Performance

We evaluate the in-sample and out-of-sample performance of our 6-factor model. The in-sample pricing errors are presented in Figure 2. The sample period runs from March 2004 to December 2015. Consistent with the low value of estimated error term $e$, our 6-factor

---

3Another approach is maximum likelihood estimation, but we found the convergence of nonlinear least squares is faster than that of maximum likelihood estimation.
model generates low pricing errors during both non-crisis and crisis periods. The out-of-sample pricing errors are shown in Figure 3. The out-of-sample period runs from January to August 2016. As shown in Figure 3, the 6-factor model can produce out-of-sample errors of magnitudes similar to those of in-sample errors. In fact, the pricing errors are rather consistently between -0.02 to 0.04.

![Daily Average Pricing Errors](image)

Figure 2: In-Sample Pricing Errors

5 Conclusion

Our term structure model of VIX futures prices is designed to address the curse of parameter dimensionality. The flexibility of the cascading structure allows us to include as many factors as needed with no additional parameter required. In particular, we apply a 6-factor model with only 6 parameters to VIX spot and futures prices from 2004 to 2016. The semi-closed

---

*In fact, we only need 5 parameters to achieve the same performance, because the long-run mean parameter $\theta_v$ is not statistically significant.*
form solution further enables fast calibration. Without resorting to jump process, our multi-factor model can produce low pricing errors both in- and out-of-sample.
APPENDIX

A  Proof of Proposition

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

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

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

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

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

Substituting $\sigma_{n-1,t}$ back into $\sigma_{n,t}$ yields:

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

We can continue the iterative substitution in $\sigma_{n,t}$ with

$$\sigma_{n-2,t_n-2}^2 = e^{-\kappa_2 t} \sigma_{n-2,0}^2 + e^{-\kappa_2 t} \gamma_n \omega_n^2$$

and obtain
\[ \sigma_{n,t}^2 = e^{-\kappa_n t} \sigma_{n,0}^2 + \left( \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} e^{-\kappa_{n-1} t_{n-1}} dt_{n-1} \right) \sigma_{n-1,0}^2 \]
\[ + \left( \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-1}} \kappa_{n-1} e^{-\kappa_{n-1} (t_n-t_{n-2})} e^{-\kappa_{n-2} t_{n-2}} dt_{n-2} dt_{n-1} \right) \sigma_{n-2,0}^2 + \ldots \]
\[ + \left( \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-2}} \kappa_2 e^{-\kappa_2 (t_2-t_1)} e^{-\kappa_1 t_1} dt_1 \ldots dt_{n-1} \right) \sigma_{1,0}^2 \]
\[ + \left( \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-1}} \kappa_1 e^{-\kappa_1 (t_1-t_0)} dt_0 \ldots dt_{n-1} \right) \theta_v \]
\[ + \frac{e^{-\kappa_n t} - 1}{\kappa_n} \gamma_n \omega_n^2 + \left( \int_0^t \kappa_n e^{-\kappa_n (t-s)} e^{-\kappa_{n-1} t_{n-1}} - 1 \right) \frac{1}{\kappa_{n-1}} \gamma_{n-1} \omega_{n-1}^2 + \ldots \]
\[ + \left( \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-1}} \kappa_1 e^{-\kappa_1 (t_1-t_0)} e^{-\kappa_1 t_1} - 1 \right) \frac{1}{\kappa_1} \gamma_1 \omega_1^2 \]
\[ + \omega_n \int_0^t e^{-\kappa_n (t-t_n-1)} dW_{n,t_{n-1}} + \omega_{n-1} \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-1}} e^{-\kappa_{n-1} (t_{n-1}-t_{n-2})} dW_{n-1,t_{n-2}} dt_{n-1} \]
\[ + \ldots + \omega_1 \int_0^t \kappa_n e^{-\kappa_n (t-t_n-1)} \int_0^{t_{n-1}} e^{-\kappa_1 (t_1-t_0)} dW_{1,t_0} dt_{1} \ldots dt_{n-1} \]

We further apply Fubini’s theorem to the \(dW\) items and obtain

\[ \sigma_{n,t}^2 = \sum_{j=1}^n a_j(t) \sigma_{j,0}^2 + \theta_v \left( 1 - \sum_{j=1}^n a_j(t) \right) + \sum_{j=1}^n b_j(t) \gamma_j \omega_j^2 + \sum_{j=1}^n \omega_j \int_0^t a_j(t-s) dW_{j,s} \]

where

\[ a_j(t) = (K_j \ast ... \ast K_n)(t) / \kappa_j \]
\[ b_j(t) = (\sum_{i=j}^n a_i(t) - 1) / \kappa_j \]
\[ K_n = \kappa_n e^{-\kappa_n t} \mathbb{1}_{[t \geq 0]} \].

\(^5\)Note that \(\theta_v = \sigma_{0,t}^2\) applies \(\forall t\), i.e. \(\theta_v = \sigma_{0,0}^2\).
B Solutions to Ricatti Equations 24 and 25

The solutions to $\beta$’s are given as follows:

$$\begin{align*}
\beta_1 &= b_{12}e^{-2bk_1t} + b_{13}e^{-3bk_1t} + b_{14}e^{-4bk_1t} + b_{15}e^{-5bk_1t} + b_{16}e^{-6bk_1t} + b_{11}e^{-k_1t} \\
\beta_2 &= b_{22}e^{-3bk_1t} + b_{23}e^{-3bk_1t} + b_{24}e^{-4bk_1t} + b_{25}e^{-5bk_1t} + b_{26}e^{-6bk_1t} \\
\beta_3 &= b_{33}e^{-3bk_1t} + b_{34}e^{-4bk_1t} + b_{35}e^{-5bk_1t} + b_{36}e^{-6bk_1t} \\
\beta_4 &= b_{44}e^{-6bk_1t} + b_{45}e^{-5bk_1t} + b_{46}e^{-6bk_1t} \\
\beta_5 &= b_{55}e^{-5bk_1t} + b_{56}e^{-6bk_1t} \\
\beta_6 &= b_{66}e^{-6bk_1t}
\end{align*}$$

where the coefficients are defined as:

$$\begin{align*}
&b_{11} = \frac{720b^5u_6}{(2b - 1)(3b - 1)(4b - 1)(5b - 1)(6b - 1)} + \frac{120b^4u_5}{(2b - 1)(3b - 1)(4b - 1)(5b - 1)} \\
&+ \frac{24b^3u_4}{(2b - 1)(3b - 1)(4b - 1)} + \frac{6b^2u_3}{(2b - 1)(3b - 1)} + \frac{2bu_2}{2b - 1} + u_1 \\
&b_{12} = -\frac{2b(u_2 + 3u_3 + 6u_4 + 10u_5 + 15u_6)}{2b - 1} \\
&b_{13} = \frac{6b(u_3 + 4u_4 + 10u_5 + 20u_6)}{3b - 1} \\
&b_{14} = -\frac{12b(u_4 + 5u_5 + 15u_6)}{4b - 1} \\
&b_{15} = \frac{20b(u_5 + 6u_6)}{5b - 1} \\
&b_{16} = -\frac{30bu_6}{6b - 1}.
\end{align*}$$

In the interest of space, we list only the ODE for $\alpha$ in lieu of its lengthy formula

$$\begin{align*}
\alpha'(t) &= a_{11}e^{-k_1t} + \frac{1}{2}b_{11}^2\omega_1^2e^{-2k_1t} + b_{11}b_{12}\omega_1^2e^{-(2b+1)k_1t} + b_{11}b_{13}\omega_1^2e^{-(3b+1)k_1t} \\
&+ b_{11}b_{14}\omega_1^2e^{-(4b+1)k_1t} + b_{11}b_{15}\omega_1^2e^{-(5b+1)k_1t} + b_{11}b_{16}\omega_1^2e^{-(6b+1)k_1t} \\
&+ A_2e^{-2bk_1t} + A_3e^{-3bk_1t} + A_4e^{-4bk_1t} + A_5e^{-5bk_1t} + A_6e^{-6bk_1t} \\
&+ A_7e^{-7bk_1t} + A_8e^{-8bk_1t} + A_9e^{-9bk_1t} + A_{10}e^{-10bk_1t} + A_{11}e^{-11bk_1t} + A_{12}e^{-12bk_1t}
\end{align*}$$
with the boundary condition \( \alpha(0) = 0 \) and the coefficients defined as:

\[
\begin{align*}
a_{11} &= \kappa_1 \theta_v b_{11} \\
a_{12} &= \kappa_1 \theta_v b_{12} - \gamma_1 \omega_1^2 (b_{12} + b_{22}) \\
a_{13} &= \kappa_1 \theta_v b_{13} - \gamma_1 \omega_1^2 (b_{13} + b_{23} + b_{33}) \\
a_{14} &= \kappa_1 \theta_v b_{14} - \gamma_1 \omega_1^2 (b_{14} + b_{24} + b_{34} + b_{44}) \\
a_{15} &= \kappa_1 \theta_v b_{15} - \gamma_1 \omega_1^2 (b_{15} + b_{25} + b_{35} + b_{45} + b_{55}) \\
a_{16} &= \kappa_1 \theta_v b_{16} - \gamma_1 \omega_1^2 (b_{16} + b_{26} + b_{36} + b_{46} + b_{56} + b_{66}) \\
A_2 &= a_{12} \\
A_3 &= a_{13} \\
A_4 &= a_{14} + \frac{b_{12}^2 \omega^2}{2} + \frac{b_{22}^2 \omega^2}{2} \\
A_5 &= a_{15} + b_{12} b_{13} \omega^2 + b_{22} b_{23} \omega^2 \\
A_6 &= a_{16} + b_{12} b_{14} \omega^2 + b_{22} b_{24} \omega^2 + \frac{b_{13}^2 \omega^2}{2} + \frac{b_{23}^2 \omega^2}{2} + \frac{b_{33}^2 \omega^2}{2} \\
A_7 &= b_{12} b_{15} \omega^2 + b_{13} b_{14} \omega^2 + b_{22} b_{25} \omega^2 + b_{23} b_{24} \omega^2 + b_{33} b_{34} \omega^2 \\
A_8 &= b_{12} b_{16} \omega^2 + b_{13} b_{15} \omega^2 + b_{22} b_{26} \omega^2 + b_{23} b_{25} \omega^2 + b_{33} b_{35} \omega^2 \\
&\quad+ \frac{b_{14}^2 \omega^2}{2} + \frac{b_{24}^2 \omega^2}{2} + \frac{b_{34}^2 \omega^2}{2} + \frac{b_{44}^2 \omega^2}{2} \\
A_9 &= b_{13} b_{16} \omega^2 + b_{14} b_{15} \omega^2 + b_{23} b_{26} \omega^2 \\
&\quad+ b_{24} b_{25} \omega^2 + b_{33} b_{36} \omega^2 + b_{34} b_{35} \omega^2 + b_{44} b_{45} \omega^2 \\
A_{10} &= b_{14} b_{16} \omega^2 + b_{24} b_{26} \omega^2 + b_{34} b_{36} \omega^2 + b_{44} b_{46} \omega^2 \\
&\quad+ \frac{b_{15}^2 \omega^2}{2} + \frac{b_{25}^2 \omega^2}{2} + \frac{b_{35}^2 \omega^2}{2} + \frac{b_{45}^2 \omega^2}{2} + \frac{b_{55}^2 \omega^2}{2} \\
A_{11} &= b_{15} b_{16} \omega^2 + b_{25} b_{26} \omega^2 + b_{35} b_{36} \omega^2 + b_{45} b_{46} \omega^2 + b_{55} b_{56} \omega^2 \\
A_{12} &= \frac{b_{16}^2 \omega^2}{2} + \frac{b_{26}^2 \omega^2}{2} + \frac{b_{36}^2 \omega^2}{2} + \frac{b_{46}^2 \omega^2}{2} + \frac{b_{56}^2 \omega^2}{2} + \frac{b_{66}^2 \omega^2}{2}.
\end{align*}
\]
References


