

A dimension-invariant cascade model for VIX futures

Zhiguang Wang¹ | Brice Dupoyet²

¹Ness School of Management and Economics, South Dakota State University, Brookings, South Dakota ²Department of Finance, Florida International University, Miami, Florida

Correspondence

Zhiguang Wang, Ness School of Management and Economics, South Dakota State University, Box 2220, Brookings, SD 57007. Email: zhiguang.wang@sdstate.edu

Abstract

We propose a new stochastic volatility model by allowing for a cascading structure of volatility components. The model, under a minor assumption, allows us to add as many components as desired with no additional parameters, effectively defeating the curse of dimensionality often encountered in traditional models. We derive a semi-closed-form solution to the VIX futures price, and find that our six-factor model with only six parameters can closely fit spot VIX and VIX futures prices from 2004 to 2015 and produce out-of-sample pricing errors of magnitudes similar to those of in-sample errors.

K E Y W O R D S

cascade, dimension-invariant, term structure, VIX futures

1 | INTRODUCTION

VIX futures and options contracts have now 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 stabilized at nine in 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 (Lin, 2007; Zhang & Huang, 2010; Zhang & Zhu, 2006; Zhu & Zhang, 2007) to two factors (Christoffersen, Jacobs, Ornthanalai, & Wang, 2008; Egloff, Leippold, & Wu, 2010; Luo & Zhang, 2012; Zhang, Shu, & Brenner, 2010), and more recently to three factors as in Lu and Zhu (2010). In particular, Lu and Zhu (2010) find that the third factor is statistically significant for explaining the variance term structure.

Lu and Zhu's (2010) three-factor model largely represents the state of the art, or at least the most comprehensive model, in VIX futures pricing. Their framework offers a rich structure able to accommodate five strips of VIX futures contracts in sample. The multifactor model improves significantly on the pricing of short-term contracts (30 and 60 days). However, their results reveal some weaknesses: (a) Their three-factor model generates large errors for 90- and 270-day contracts; (b) the empirical results are not based on original VIX futures data but on interpolated (smoothed) prices, with the interpolation potentially hiding the actual pricing errors of the model; (c) their out-of-sample test is limited to 8 days, making it difficult for one to judge the merit of the model.

Another shortcoming of current VIX futures factor-based term structure models is that they suffer from the curse of dimensionality. One generally needs three extra parameters for each additional factor. Lu and Zhu's (2010) two-factor model has between seven and eight parameters, while their three-factor model contains between 10 and 12 parameters. The likely overfitting of the five strips of VIX futures contracts is a distinctive concern.

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 of Calvet, Fisher, and Wu (2018). The cascading volatility model essentially has one governing factor with multiple layers or commonly referenced as factors. Each layer, identified by its speed of mean reversion, mean-reverts to the next layer until it converges to a constant long-run mean. A faster (slower) moving layer represents economic information occurring at a higher (lower) frequency, for example intraday (vs. quarterly) news arrival. Since such a structure allows one to add as many layers as desired without any additional

parameters, it is thus dimension-free. The additional layers do not require extra modeling or higher computational cost. As a result, we pick six layers to match the daily average number of VIX futures contracts for our modeling and empirical exercises. ¹ The flexibility in choosing the number of components enables rich dynamics in the term structure of both the 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 (UKF) method, we estimate a six-factor model on price data from March 2004 to December 2015.² We find that our model can generate low in-sample pricing errors and equally desirable pricing precision 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 UKF methodology in Section 3. We report the in-sample and out-of-sample performance of the model in Section 4, and conclude in Section 5.

2 | MODEL

-WILEY-

2.1 | \mathbb{P} -Measure dynamics

Let V_t denote the instantaneous variance. V_t is the ending point of an *n*-layer cascading volatility $\sigma_{j,t}^2$ process, where *j* stands for the *j*th component. Therefore, $V_t = \sigma_{n,t}^2$. The higher frequency (or faster moving) component $\sigma_{j,t}^2$ reverts to the lower frequency component $\sigma_{j-1,t}^2$, until it reaches a constant long-run mean θ_v . Each component *j* is identified by their speed of mean reversion κ_j . The structure is presented as follows:

$$d\sigma_{j,t}^{2} = \kappa_{j}(\sigma_{j-1,t}^{2} - \sigma_{j,t}^{2})dt + \omega_{j}dW_{j,t}, \quad j = 1, 2, ...n,$$
(1)

$$\sigma_{0,t}^2 = \theta_{\nu},\tag{2}$$

$$\sigma_{n,t}^2 = \sigma_t^2 \equiv V_t,\tag{3}$$

$$\kappa_j = \kappa_1 \beta^* j, \quad \beta > 1, \quad j > 1, \tag{4}$$

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

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{5}$$

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

$$\mu(X) = \begin{pmatrix} \kappa_1 \theta_v \\ 0 \\ \cdots \\ 0 \\ K_0 \end{pmatrix} + \begin{pmatrix} -\kappa_1 & 0 & 0 & 0 \\ \kappa_2 & -\kappa_2 & 0 & 0 \\ 0 & \cdots & \cdots & 0 \\ 0 & 0 & \kappa_n & -\kappa_n \end{pmatrix} X,$$
(6)
$$\Sigma(X) = \begin{pmatrix} \omega_1^2 & 0 & 0 & 0 \\ 0 & \omega_2^2 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \omega_n^2 \end{pmatrix} .$$
(7)

¹Calvet et al. (2018) select 15 factors for 15 daily interest rates.

²We also estimated a three-factor model that performed adequately but not as well as the six-factor model both in-sample and out-of-sample.

2

2.2 | Q-Measure dynamics

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:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \prod_{j=1}^{n} \exp\left(-\int_{0}^{t} \gamma_{j,s} \omega_{j} dW_{j,s} - \frac{1}{2} \int_{0}^{t} \gamma_{j,s}^{2} \omega_{j}^{2} ds\right).$$
(8)

WILEY 3

The instantaneous variance dynamics under the risk-neutral measure become

$$d\sigma_{j,t}^2 = -\gamma_{j,t}\omega_j^2 dt + \kappa_j (\sigma_{j-1,t}^2 - \sigma_{j,t}^2) dt + \omega_j dW_{j,t}^Q.$$
(9)

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

$$dX_t = \mu^Q(X_t)dt + \sigma^Q(X_t)dW_t, \tag{10}$$

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

$$\mu^{Q}(X) = \underbrace{\begin{pmatrix} \kappa_{1}\theta_{v} - \gamma_{1}\omega_{1}^{2} \\ -\gamma_{2}\omega_{2}^{2} \\ \cdots \\ -\gamma_{n}\omega_{n}^{2} \end{pmatrix}}_{K_{0}^{Q}} \qquad (11)$$

$$+ \underbrace{\begin{pmatrix} -\kappa_{1} - \lambda_{11}\omega_{1}^{2} & -\lambda_{12}\omega_{2}^{2} & \cdots & -\lambda_{1n}\omega_{n}^{2} \\ \kappa_{2} - \lambda_{21}\omega_{1}^{2} & -\kappa_{2} - \lambda_{22}\omega_{2}^{2} & \cdots & -\lambda_{2n}\omega_{n}^{2} \\ \cdots & \cdots & \cdots & \cdots \\ -\lambda_{n1}\omega_{1}^{2} & \cdots & \kappa_{n} - \lambda_{nn-1}\omega_{n-1}^{2} - \kappa_{n} - \lambda_{nn}\omega_{n}^{2} \end{pmatrix}}_{K_{1}^{Q}} X, \qquad (12)$$

$$\Sigma^{Q}(X) = \underbrace{\begin{pmatrix} \omega_{1}^{2} & 0 & 0 & 0 \\ 0 & \omega_{2}^{2} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \omega_{n}^{2} \end{pmatrix}}_{H_{0}^{Q}}.$$

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, that is, $\lambda_j = 0$. Constant risk premia are commonly assumed in VIX derivatives pricing as in Lu and Zhu (2010) and Zhang and Zhu (2006). Similarly, for estimation purposes we will later make the simplifying assumption of setting the ω_n diffusion parameters constant across the cascading layers, thereby allowing for as many factors as desired without any increase in the number of parameters whatsoever. The risk-neutral dynamics of $\mu^Q(X)$ and $\Sigma^Q(X)$ become

$$\mu^{Q}(X) = \underbrace{\begin{pmatrix} \kappa_{1}\theta_{\nu} - \gamma_{1}\omega_{1}^{2} \\ -\gamma_{2}\omega_{2}^{2} \\ \cdots \\ -\gamma_{n}\omega_{n}^{2} \end{pmatrix}}_{K_{0}^{Q}} + \underbrace{\begin{pmatrix} -\kappa_{1} & 0 & \cdots & 0 \\ \kappa_{2} & -\kappa_{2} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \kappa_{n} & -\kappa_{n} \end{pmatrix}}_{K_{1}^{Q}} X,$$
(13)

WANG AND DUPOYET

$$\Sigma^{Q}(X) = \underbrace{\begin{pmatrix} \omega_{1}^{2} & 0 & 0 & 0\\ 0 & \omega_{2}^{2} & 0 & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \omega_{n}^{2} \end{pmatrix}}_{H_{0}^{Q}}.$$
(14)

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

$$V_t \equiv \sigma_{n,t}^2 = \theta_v + \sum_{j=1}^n a_j(t)(\sigma_{j,0}^2 - \theta_v) + \sum_{j=1}^n b_j(t)\gamma_j\omega_j^2 + \sum_{j=1}^n \omega_j \int_0^t a_j(t-s)dW_{j,s},$$
(15)

$$a_j(t) = (\mathbb{K}_j^* \cdots * \mathbb{K}_n)(t) / \kappa_j$$
(16)

$$=\sum_{i=j}^{n}\frac{\kappa_{j+1}...\kappa_{n}}{\prod_{s=j,s\neq i}^{n}(\kappa_{s}-\kappa_{i})}e^{-\kappa_{i}t}$$
(17)

$$b_j(t) = \left(\sum_{i=j}^n a_i(t) - 1\right) / \kappa_j,$$
 (18)

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 $\mathbb{K}_j = \kappa_j e^{-\kappa_j t} \mathbf{1}_{[t \ge 0]}$.

Proof. See Appendix A

▲ WILEY-

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

$$VIX_t^2 = \frac{1}{\tau} E_t^{\mathbb{Q}} \left[\int_t^{t+\tau} \sigma_{n,s}^2 ds \right],\tag{19}$$

where τ is fixed at 30 days according to the CBOE.

By inserting Equation (15) into Equation (19), we obtain VIX_T^2 , the spot 30-day variance at expiration time *T*, as a linear combination of the *n* factors:

$$VIX_{T}^{2} = \frac{1}{\tau} E_{T}^{\mathbb{Q}} \Biggl[\int_{T}^{T+\tau} \Biggl[\theta_{v} + \sum_{j=1}^{n} a_{j}(s)(\sigma_{j,T}^{2} - \theta_{v}) + \sum_{j=1}^{n} b_{j}(s)\gamma_{j}\omega_{j}^{2} \Biggr] ds \Biggr] + \frac{1}{\tau} E_{T}^{\mathbb{Q}} \Biggl[\sum_{j=1}^{n} \omega_{j} \int_{T}^{T+\tau} a_{j}(s - u) dW_{j,u} ds \Biggr] = \frac{1}{\tau} \Biggl[\int_{T}^{T+\tau} \Biggl[\theta_{v} + \sum_{j=1}^{n} a_{j}(s)(\sigma_{j,T}^{2} - \theta_{v}) + \sum_{j=1}^{n} b_{j}(s)\gamma_{j}\omega_{j}^{2} \Biggr] ds \Biggr]$$
(20)
$$= \theta_{v} + \frac{1}{\tau} \Biggl[\Biggl[\int_{T}^{T+\tau} \sum_{j=1}^{n} a_{j}(s) ds \Biggr] (\sigma_{j,T}^{2} - \theta_{v}) \Biggr] + \frac{1}{\tau} \int_{T}^{T+\tau} \sum_{j=1}^{n} b_{j}(s)\gamma_{j}\omega_{j}^{2} ds = \theta_{v} + \sum_{j=1}^{n} A(j)(\sigma_{j,T}^{2} - \theta_{v}) + \frac{1}{\tau} \int_{T}^{T+\tau} \sum_{j=1}^{n} b_{j}(s)\gamma_{j}\omega_{j}^{2} ds = \tilde{A}' X_{T} + B,$$

where³

$$\begin{split} A(j) &= \frac{1}{\tau} \int_{T}^{T+\tau} a_{j}(s) ds \\ &= \frac{1}{\tau} \sum_{i=j}^{n} \frac{\kappa_{j+1} \dots \kappa_{n}}{\prod_{s=j,s\neq i}^{n} (\kappa_{s} - \kappa_{i})} \frac{e^{-\kappa_{i}T} - e^{-\kappa_{i}(T+\tau)}}{\kappa_{i}}, \\ B &= \theta_{v} \left(1 - \sum_{j=1}^{n} A(j) \right) + \frac{1}{\tau} \int_{T}^{T+\tau} \sum_{j=1}^{n} b_{j}(s) \gamma_{j} \omega_{j}^{2} ds \\ &= \theta_{v} \left(1 - \sum_{j=1}^{n} A(j) \right) + \frac{1}{\tau} \int_{T}^{T+\tau} \sum_{j=1}^{n} \frac{\sum_{i=j}^{n} a_{i}(s) - 1}{\kappa_{j}} \gamma_{j} \omega_{j}^{2} ds \\ &= \theta_{v} \left(1 - \sum_{j=1}^{n} A(j) \right) + \sum_{j=1}^{n} \left(\sum_{i=j}^{n} A(i) - 1 \right) \frac{\gamma_{j} \omega_{j}^{2}}{\kappa_{j}}, \\ \tilde{A} &= (A(1), A(2), \dots A(n))', \\ X_{T} &= \tilde{\sigma}_{T}^{2} = (\sigma_{1,T}^{2}, \sigma_{2,T}^{2}, \dots \sigma_{n,T}^{2})'. \end{split}$$

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

$$F(t, T) = E_t^{\mathbb{Q}} \left[\sqrt{VIX_T^2} \right].$$
(21)

WILEY - WILEY

Schürger (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:

$$E[\sqrt{z}] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sz}]}{s^{3/2}} ds.$$
 (22)

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:

$$\dot{\beta}(t;s) = -K_1^T \beta(t;s), \tag{23}$$

$$\dot{\alpha}(t;s) = -K_0 \cdot \beta(t;s) - \frac{1}{2} \beta(t;s)^T H_0 \beta(t;s),$$
(24)

with boundary conditions $\alpha(T; s) = 0$ and $\beta(T; s) = -s\tilde{A}$. H_0 and K_1 are from Equations (13) and (14), respectively.

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

$$F(t, T) = E_t^{\mathbb{Q}} \left[\sqrt{VIX_T^2} \right] = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - E[e^{-sz}]}{s^{3/2}} ds$$
(25)

$$=\frac{1}{2\sqrt{\pi}}\int_{0}^{\infty}\frac{1-e^{-sB+\alpha(t;s)+\beta(t;s)\cdot\sigma_{t}^{2}}}{s^{3/2}}ds,$$
(26)

³Note that $\sigma_{j,T}$ is used instead of the original notation $\sigma_{j,0}$ in Equation (15), because the initial time "0" for the time interval $[T, T + \tau]$ is T.

• WILEY-

where $\alpha(t; s)$ and $\beta(t; s)$ are determined by Equations (23) and (24). Given that $X_t = \sigma_t^2$ is defined in Equation (10), the VIX futures pricing function F(t, T) can be expressed as a nonlinear function of state variables X_t , that is, $F(t, T) = f(X_t)$.

3 | DATA AND METHODOLOGY

3.1 | Data

We obtain daily VIX futures closing prices along with spot VIX prices from the CBOE, with data ranging from 2004 to 2016. On any given day, the number of contracts listed varies from 4 to 11 during the sample period. To visualize the sample time series, we linearly interpolate the VIX futures prices to obtain the prices with 1-, 2-, 3-, 6-, 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, by which long-maturity futures contracts exhibit higher prices than short-maturity futures contracts.

Table 1 reports both spot VIX and VIX futures prices of 3,019 trading days ranging from March 2004 to August 2016. In contrast to the term structure of VIX futures prices, VIX futures volatility levels exhibit normal backwardation in their term structure, that is, higher volatility levels are associated with shorter maturities. Both the spot VIX and VIX futures display positive skewness and excess kurtosis.

3.2 | UKF implementation

The VIX futures pricing formula derived in Equation (26) is a nonlinear function of the unobservable state variables $\sigma_{j,t}^2$. We propose to estimate the model using UKF. The UKF 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 particularly well suited to nonlinear state-space models (see Wan & Van der Merwe, 2001). The application of UKF to model estimation has been employed in the derivatives literature in recent years. Christoffersen, Dorion, Jacobs, and Karoui (2014) provide a detailed examination of the nonlinear Kalman filtering, with a focus on UKF applications to affine term structure models of interest rates. Carr and Wu (2007) have applied the UKF methodology to currency options and Carr and Wu (2010) to equity options and credit default swaps.

To implement the UKF methodology, we discretize the state transition shown in Equation (10) in the following matrix form:



FIGURE 1 Spot VIX and VIX futures prices [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Summary statistics

Contract	Mean	SD	Median	Minimum	Maximum	Skewness	Kurtosis
Spot	19.4930	9.3293	16.5800	9.8900	80.8600	2.5027	11.2765
1 m	20.3563	8.0071	18.0375	11.2636	65.6763	1.9664	7.7926
2 m	20.9884	7.3580	19.0321	12.1952	60.1391	1.6835	6.5813
3 m	21.3855	6.8616	19.6643	12.7787	54.6700	1.4332	5.4088
6 m	22.1791	6.0889	20.7862	13.6385	45.0000	0.9864	3.6439
9 m	22.6469	5.6536	21.3757	9.2482	43.9900	0.7866	3.1008

 $dX_t = -K_1^Q (K_0^{Q*} - X_t) dt + \sqrt{\Sigma} dW_t,$

 $X_t = (I - \Phi) K_0^{Q*} + \Phi X_{t - \triangle t} + \sqrt{\Sigma \triangle t} u_t,$

with

$$\begin{split} \Phi &= e^{K_1^Q \bigtriangleup t}, \\ K_0^{Q*} &= -[K_1^Q]^{-1} K_0^Q, \\ [K_1^Q]^{-1} &= \begin{pmatrix} -\frac{1}{\kappa_1} & 0 & \cdots & 0\\ -\frac{1}{\kappa_1} & -\frac{1}{\kappa_2} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ -\frac{1}{\kappa_1} & -\frac{1}{\kappa_2} & \cdots & -\frac{1}{\kappa_n} \end{pmatrix}. \end{split}$$

Furthermore, the daily time interval $\Delta t = 1/365$, and *I* is the identity matrix. K_0^Q , K_1^Q , Σ are defined in Equations (13) and (14), while u_t follows the standard normal distribution. We further define the measurement equation for VIX futures prices as follows:

$$F_t = f(X_t) + e_t. \tag{28}$$

WILEY 7

(27)

where F_t is the observed VIX futures price, $f(X_t)$ is the pricing formula given in Equation (26), and the error term e_t follows a normal distribution with a mean of zero and a standard deviation σ_e . 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 σ_e across contracts.⁴

For a set of M contracts available on a given day, we have the following pricing equations:

. . .

$$F_{1,t} = f_1(X_t, \Theta) + e_{1,t},$$
(29)

$$F_{2,t} = f_2(X_t, \Theta) + e_{2,t},$$
(30)

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

$$e_{m,t} \sim N(0, \sigma_e) \quad \forall \ m = 1, 2, ..., M.$$
(31)

where Θ is the parameter vector in the model that includes κ_1 , b, ω , γ , θ_{ν} .

⁴We considered the alternative of removing the restriction of identical errors across contracts and found no significant improvement in the 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.

* WILEY-	TURËS RKFTS
----------	----------------

TABLE 2 Parameter estimates for the six-factor model

	κ1	β	ω	γ	$ heta_{ u}$	$\sigma_{\!e}$
Estimate	0.6301	71.2912	0.9548	-0.6193	0.0285	0.0002
SE	(0.0195)	(1.8359)	(0.0197)	(0.0241)	(0.0309)	(0.0000)

Given the observations and the state transition Equation (27), we employ the UKF for parameter and state estimation. Specifically, we set the objective function for the nonlinear least squares for *N* days of observations as $\sum_{t=1}^{N} \sum_{m=1}^{M} (F_{m,t} - f_m(X_t, \Theta))^2$. ⁵ By minimizing the objective function under the UKF scheme, we can estimate the parameters Θ and the pricing errors standard deviation parameter σ_e along with the state variables $X_t = \{x_{1t},...,x_{nt}\}$. 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, namely, the standard deviation of the error term σ_e . Our approach differs from the previous literature such as Lu and Zhu (2010) and Zhang et al. (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 six-factor model to the term structure of daily VIX futures from March 2004 to December 2015.⁶ Parameter estimates and their standard errors are reported in Table 2. We can see that all parameters are statistically significant at the 1% level, except the long-run mean θ_{v} . A positive and significant β parameter lends empirical support to the cascading feature of mean reversion. The large β value is indeed consistent with our proposed cascading feature, as the higher frequency volatility layers mean-revert back to the lower levels at a much higher speed than the lower frequency volatility layers. Additionally, γ is negative, confirming the well-known negative variance risk premium. Despite the assumption of identical errors across contracts, we still manage to obtain a statistically significant standard deviation parameter for the pricing errors *e*. The low value of 0.0002 indicates a good match between our model and the data.

4.2 Model's flexibility of generating term structure

To illustrate the flexibility of our model in its ability to generate the term structure of VIX futures, we reproduce various shapes of term structure using the parameter estimates from Table 2. Figure 2 displays contango, inverse hump, backwardation, and a combination of contango and backwardation, on May 30, June 12, June 26, and July 10, 2007, respectively. The various term structure shapes indicate that our model is capable of quickly adapting to changing market conditions.

4.3 | Pricing performance

We evaluate the in-sample and out-of-sample performance of our six-factor model. The in-sample pricing errors, defined as daily average dollar error = market price – model price across all contracts, are presented in Figure 3. The sample period runs from March 2004 to December 2015. Consistent with the small magnitude of the estimated error term σ_e , our six-factor model generates low pricing errors during both noncrisis and crisis periods.

⁵Another approach is maximum likelihood estimation, but we found the convergence of nonlinear least squares to be faster than that of maximum likelihood estimation.

⁶The choice of six factors, as opposed to more, is made to achieve a balance between flexibility and computational burden. Calvet et al. (2018) investigate both the in-sample fitting and the out-ofsample predictive power of more factors in the modeling of the interest rate term structure, in which they used a maximal 15-factor model to fit the 15 maturity strips of LIBOR and swap rates. However, the square root function in our VIX futures pricing formula (see Equation (26)) makes its computation significantly more challenging than that of their interest rate pricing formula.



FIGURE 2 Sample term structures of VIX futures [Color figure can be viewed at wileyonlinelibrary.com]



FIGURE 3 In-sample pricing errors [Color figure can be viewed at wileyonlinelibrary.com]

The out-of-sample pricing errors are shown in Figure 4. The out-of-sample period runs from January to August 2016. Figure 4 demonstrates that the six-factor model can produce out-of-sample errors of magnitudes similar to those of in-sample errors.⁷ In fact, the pricing errors are rather consistently ranging between -0.02 and 0.04. It is worth noting that our estimation method differs from those commonly used in the previous literature—We use the original price series as opposed to the interpolated price series as mentioned in Section 3.2. As a comparison, Lu and Zhu's (2010) three-factor model and Zhang et al.'s (2010) two-factor model produce in-sample dollar pricing errors (RMSE) mostly in the range of 1.3–2.6, or between 5% and 15% for the sample from 2004 to 2008. Lu and Zhu (2010) and Zhang et al. (2010) report about 5% out-of-sample errors for 2 months and 8 days, respectively.

⁷Following the referee's suggestion, we also estimated the model using the first half of the sample (2004 to early 2010), and tested its out-of-sample performance using the second half of the sample (late 2010-2016). We confirmed the pricing stability of our six-factor model with the balanced sample and test data.

9



FIGURE 4 Out-of-sample pricing errors [Color figure can be viewed at wileyonlinelibrary.com]

Therefore, our six-factor model appears able to handle the extra noise in the price series and produce low insample and out-of-sample pricing errors.

5 | CONCLUSION

We introduce a new stochastic volatility model that relies on a cascading structure of volatility components and increasing levels of speeds of mean reversion. We subsequently apply this new framework to the pricing of VIX futures contracts. Our term structure model of VIX futures prices is designed to overcome the curse of parameter dimensionality: While one can add a myriad of bells and whistles to a model, the number of parameters often grows at a rate of about three additional parameters per factor or feature, making the estimation a complex task yielding estimates in which one cannot necessarily have a high degree of confidence. The flexibility of the cascading structure, with the help of a small assumption, allows us to include as many factors as needed with no additional parameters required. In this paper, we apply a six-factor model with only six parameters to both VIX spot and futures prices from 2004 to 2016. The semi-closed-form solution further enables fast calibration. Without resorting to a jump process, our unique new stochastic cascading volatility model is able to produce low pricing errors both in-sample and out-of-sample. In addition to the term structure of VIX futures, the cascading structure can be broadly applied when the model calls for dynamic mean reversion and/or multifactor characteristics. Examples of those include historical volatility prediction in the time series literature, the risk-return trade-off under short-run and long-run volatility components in the asset pricing literature, and commodity futures pricing in agricultural and energy economics.

ACKNOWLEDGMENTS

We thank the Editor, Bob Webb, the two anonymous referees, Liuren Wu, Claudio Tebaldi (discussant), Chien-Ling Lo, and various participants at the 2017 Financial Management Association conference for their helpful comments. We also thank discussants Ray Chou and Gang Li and various participants at the 2018 Multinational Finance Society and 2018 Global Finance conferences for their helpful insights.

ORCID

Zhiguang Wang D http://orcid.org/0000-0002-8014-026X

REFERENCES

- Calvet, L., Fisher, A., & Wu, L. (2018). Staying on top of the curve: A cascade model of term structure dynamics. *Journal of Financial and Quantitative Analysis*, *53*(2), 937–963.
- Carr, P., & Wu, L. (2007). Stochastic skew in currency options. Journal of Financial Economics, 86, 213-247.
- Carr, P., & Wu, L. (2010). Stock options and credit default swaps: A joint framework for valuation and estimation. Journal of Financial Econometrics, 8, 409-449.
- Christoffersen, P., Dorion, C., Jacobs, K., & Karoui, L. (2014). Nonlinear kalman filtering in affine term structure models. *Management Science*, 60(9), 2248–2268.
- Christoffersen, P., Jacobs, K., Ornthanalai, C., & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. Journal of Financial Economics, 90(3), 272–297.
- Duffie, D., Pan, J., & Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-diffusions. Econometrica, 68, 1343–1376.
- Egloff, D., Leippold, M., & Wu, L. (2010). The term structure of variance swap rates and optimal variance swap investments. *Journal of Financial and Quantitative Analysis*, 45(5), 1279.
- Lin, Y. (2007). Pricing VIX futures: Evidence from integrated physical and risk-neutral probability measures. *Journal of Futures Markets*, 27, 1175–1217.
- Lu, Z., & Zhu, Y. (2010). Volatility components: The term structure dynamics of VIX futures. *Journal of Futures Markets*, *30*(3), 230–256. Luo, X., & Zhang, J. E. (2012). The term structure of VIX. *Journal of Futures Markets*, *32*(12), 1092–1123.
- Schürger, K. (2002). Laplace Transforms and Suprema of Stochastic Processes. In Sandmann, K., & Schönbucher, P.J. (Eds.), Advances in Finance and Stochastics (pp. 285–294). Berlin, Heidelberg: Springer.
- Wan, E., & VanderMerwe, R. (2001). The unscented Kalman filter. In Simon, H. (Ed.), *Kalman filtering and neural networks* (pp. 221–280). New York: Wiley.
- Zhang, J. E., & Zhu, Y. (2006). VIX futures. Journal of Futures Markets, 26, 521-531.
- Zhang, J. E., & Huang, Y. (2010). The CBOE sp 500 three-month variance futures. Journal of Futures Markets, 30, 48-70.
- Zhang, J. E., Shu, J., & Brenner, M. (2010). The new market for volatility trading. Journal of Futures Markets, 30(9), 809-833.
- Zhu, Y., & Zhang, J. (2007). Variance term structure and VIX futures pricing. International Journal of Theoretical and Applied Finances, 10, 111–127.
- Wang, Z., & Dupoyet, B. (2019). VIX futures. Retrieved from https://cfe.cboe.com/market-data/historical-data-archive#VX.

APPENDIX A: PROOF OF PROPOSITION 1

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

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

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

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

$$\begin{aligned} \sigma_{n,t}^2 &= e^{-\kappa_n t} \sigma_{n,0}^2 + \frac{e^{-\kappa_n t} - 1}{\kappa_n} \gamma_n \omega_n^2 + \int_0^t \kappa_n e^{-\kappa_n (t - t_{n-1})} \sigma_{n-1,t_{n-1}}^2 dt_{n-1} + \int_0^t \omega_n e^{-\kappa_n (t - t_{n-1})} dW_{n,t_{n-1}}, \\ \sigma_{n-1,t_{n-1}}^2 &= e^{-\kappa_{n-1} t} \sigma_{n-1,0}^2 + \frac{e^{-\kappa_{n-1} t} - 1}{\kappa_{n-1}} \gamma_{n-1} \omega_{n-1}^2 \\ &+ \int_0^{t_{n-1}} \kappa_{n-1} e^{-\kappa_{n-1} (t - t_{n-2})} \sigma_{n-2,t_{n-2}}^2 dt_{n-2} + \int_0^{t_{n-1}} \omega_{n-1} e^{-\kappa_{n-1} (t - t_{n-2})} dW_{n-1,t_{n-2}}. \end{aligned}$$

Substituting $\sigma_{n-1,t}^2$ back into $\sigma_{n,t}^2$ yields

$$\begin{aligned} \sigma_{n,t}^{2} &= e^{-\kappa_{n}t}\sigma_{n,0}^{2} + \frac{e^{-\kappa_{n}t} - 1}{\kappa_{n}}\gamma_{n}\omega_{n}^{2} + \int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}e^{-\kappa_{n-1}t_{n-1}}\sigma_{n-1,0}^{2}dt_{n-1} \\ &+ \left[\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\frac{e^{-\kappa_{n-1}t_{n-1}} - 1}{\kappa_{n-1}}dt_{n-1}\right]\gamma_{n-1}\omega_{n-1}^{2} \\ &+ \int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\int_{0}^{t_{n-1}} \kappa_{n-1}e^{-\kappa_{n-1}(t_{n-1}-t_{n-2})}\sigma_{n-2,t_{n-2}}^{2}dt_{n-2}dt_{n-1} \\ &+ \int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\int_{0}^{t_{n-1}} \omega_{n-1}e^{-\kappa_{n-1}(t_{n-1}-t_{n-2})}dW_{n-1,t_{n-2}}dt_{n-1} + \int_{0}^{t} \omega_{n}e^{-\kappa_{n}(t-t_{n-1})}dW_{n,t_{n-1}}. \end{aligned}$$

We can continue the iterative substitution in $\sigma_{n,t}^2$ with

$$\sigma_{n-2,t_{n-2}}^2 = e^{-\kappa_{n-2}t}\sigma_{n-2,0}^2 + \frac{e^{-\kappa_{n-2}t}-1}{\kappa_{n-2}}\gamma_{n-2}\omega_{n-2}^2 + \int_0^{t_{n-2}} \kappa_{n-2}e^{-\kappa_{n-2}(t-s)}\sigma_{n-3,t_{n-3}}^2 dt_{n-3} + \int_0^{t_{n-2}} \omega_{n-2}e^{-\kappa_{n-2}(t_{n-2}-t_{n-3})}dW_{n-2,t_{n-3}}$$

and obtain

$$\begin{aligned} \sigma_{n,t}^{2} &= e^{-\kappa_{n}t}\sigma_{n,0}^{2} + \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}e^{-\kappa_{n-1}t_{n-1}}dt_{n-1}\right)\sigma_{n-1,0}^{2} \\ &+ \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\int_{0}^{t_{n-1}} \kappa_{n-1}e^{-\kappa_{n-1}(t_{n-1}-t_{n-2})}e^{-\kappa_{n-2}t_{n-2}}dt_{n-2}dt_{n-1}\right)\sigma_{n-2,0}^{2} + \cdots \\ &+ \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\cdots\int_{0}^{t_{2}} \kappa_{2}e^{-\kappa_{2}(t_{2}-t_{1})}e^{-\kappa_{1}t_{1}}dt_{1}\dots dt_{n-1}\right)\sigma_{1,0}^{2} \\ &+ \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\cdots\int_{0}^{t_{1}} \kappa_{1}e^{-\kappa_{1}(t_{1}-t_{0})}dt_{0}dt_{1}\dots dt_{n-1}\right)\theta_{v} \\ &+ \frac{e^{-\kappa_{n}t}-1}{\kappa_{n}}\gamma_{n}\omega_{n}^{2} + \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-s)}\frac{e^{-\kappa_{n-1}t_{n-1}}-1}{\kappa_{n-1}}dt_{n-1}\right)\gamma_{n-1}\omega_{n-1}^{2} + \cdots \\ &+ \left(\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\cdots\int_{0}^{t_{2}} \kappa_{1}e^{-\kappa_{1}(t_{2}-t_{1})}\frac{e^{-\kappa_{1}t_{1}}-1}{\kappa_{1}}dt_{1}\dots dt_{n-1}\right)\gamma_{1}\omega_{1}^{2} \\ &+ \omega_{n}\int_{0}^{t} e^{-\kappa_{n}(t-t_{n-1})}dW_{n,t_{n-1}} + \omega_{n-1}\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\int_{0}^{t_{n-1}} e^{-\kappa_{n-1}(t_{n-1}-t_{n-2})}dW_{n-1,t_{n-2}}dt_{n-1} \\ &+ \cdots + \omega_{1}\int_{0}^{t} \kappa_{n}e^{-\kappa_{n}(t-t_{n-1})}\cdots\int_{0}^{t_{1}} e^{-\kappa_{1}(t_{1}-t_{0})}dW_{1,t_{0}}dt_{1}\dots dt_{n-1}. \end{aligned}$$

 We^8 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

$$a_j(t) = (\mathbb{K}_j * \dots * \mathbb{K}_n)(t) / \kappa_j,$$

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

$$\mathbb{K}_n = \kappa_n e^{-\kappa_n t} \mathbf{1}_{[t \ge 0]}.$$

APPENDIX B: SOLUTIONS TO RICATTI EQUATIONS (23) AND (24)

In this section, we report the solution to the cascade model of stochastic volatility for the case of n = 6. Specifically, the solutions to the β s are given as follows:

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

where the coefficients are defined as

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

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

$$\begin{aligned} \alpha'(t) &= a_{11}e^{-k_{1}t} + \frac{1}{2}b_{11}^{2}\omega_{1}^{2}e^{-2k_{1}t} + b_{11}b_{12}\omega_{1}^{2}e^{-(2b+1)k_{1}t} + b_{11}b_{13}\omega_{1}^{2}e^{-(3b+1)k_{1}t} \\ &+ b_{11}b_{14}\omega_{1}^{2}e^{-(4b+1)k_{1}t} + b_{11}b_{15}\omega_{1}^{2}e^{-(5b+1)k_{1}t} + b_{11}b_{16}\omega_{1}^{2}e^{-(6b+1)k_{1}t} \\ &+ A_{2}e^{-2bk_{1}t} + A_{3}e^{-3bk_{1}t} + A_{4}e^{-4bk_{1}t} + A_{5}e^{-5bk_{1}t} + A_{6}e^{-6bk_{1}t} \\ &+ A_{7}e^{-7bk_{1}t} + A_{8}e^{-8bk_{1}t} + A_{9}e^{-9bk_{1}t} + A_{10}e^{-10bk_{1}t} + A_{11}e^{-11bk_{1}t} + A_{12}e^{-12bk_{1}t} \end{aligned}$$

with the boundary condition $\alpha(0) = 0$ and the coefficients defined as

How to cite this article: Wang Z, Dupoyet B. A dimension-invariant cascade model for VIX futures. *J Futures Markets*. 2019;1–14. <u>https://doi.org/10.1002/fut.22042</u>