Unifying Variance Swap Term Structures, SPX and VIX Derivatives

Bo Zhao[a]

[a] *Cass Business School, London EC1Y 8TZ, UK*

Abstract

We propose an instantaneous variance process and a return process to jointly price SPX and VIX derivatives. The instantaneous variance process is described by a two-factor model. In the interest of analytical tractability, these processes are structured by affine processes. The main distinct feature of the model is that the factor coefficients are time-varying and they are designed to be bonded with the term structure of variance swaps. The model incorporates additional characteristics that the jump of the return process and the jump of the instantaneous variance processes are more recognizable in the short term than in the long term, the term structure of variance swaps is versatilely rich to be able to accommodate many desired features, and the correlation between shocks to returns and shocks to variances is stochastic.

Introduction

The concept of variance swap appears first in Neuberger (1990), Neuberger (1994), and Carr and Madan (1998). Since then, many derivative instruments written on volatility are actively traded in financial market. Among them, VIX futures were launched in March 2004, VIX options were introduced February 2006. These volatility products allow investors to take views on implied/realized/forward volatilities, and hedge volatility risks of equity market positions. The unprecedented trading, investing and hedging in volatilities have led to a parallel systematic development of dynamic volatility modelings. The aim of this paper is to bring together a volatility model and a return process in which they match stylized facts of financial markets and they are consistently reconciled to each other.

As far as parametric models are concerned, there are two main approaches in getting variance/volatility swaps. Buehler (2006), Bergomi (2004), Bergomi (2005), Bergomi (2008), Cont and Kokholm (2011) set up first the forward variance, which is described by an exponential martingale, to explore variance swaps, instantaneous variance and vanilla options. However, these models are short of analytical tractability. To value the underlying options, Monte Carlo simulations are needed in Bergomi (2004), Bergomi (2005), Bergomi (2008), Cont and Kokholm (2011). For the second vein, the straightforward approach is to specify the return dynamics. The quadratic variation of the log return process yields variance swaps. Under the Black-Scholes framework, the underlying asset is modeled by a geometric

Email address: `bo.zhao.1@city.ac.uk` (Bo Zhao)

First Draft: November 21, 2011, Current Draft: January 13, 2012.

Brownian motion with a constant volatility. The implied smile function and the variance swap are both trivially constant, which are empirically biased. To highlight features of smiles and smirks, the Black-Scholes model is soon surrendered by stochastic volatility processes or/and jump-diffusion models, such as the CGMY dynamics (Carr et al. (2005)), Heston, Merton and Bates dynamics (Broadie and Jain (2008)), the Sato Process (Madan and Yor (2010)). However, a single-factor variance process is still criticized on the ground that it is too rigid to manage versatile natures of volatility smiles across different strikes and maturities. These models are short of empirical justification, especially they are not flexible enough to capture the time variation and cross-sectional variation property of the term structure. Coming into play is a much more improved and strengthened model which assumes there are several factors spontaneously steering volatility movements(Lee and Engle (1999), Christoffersen et al. (2008), Gatheral (2008), Fonseca et al. (2008), Christoffersen et al. (2009), Egloff et al. (2010)). The existence of multiple factors is empirically tested and supported in Skiadopoulos et al. (1999), Fengler et al. (2003). Christoffersen et al. (2009) shows that two-factor variance processes not only provide much more variabilities in engineering the level and slope of the implied smile function of the Index, but also they yield more adaptable volatility term structure.

A multi-factor stochastic model, coupled with jump components both in the return process and volatility process, is what we believe a realistic and robust model. The construction of our model is based on the intuition that there is a line joining the initial instantaneous variance and the long run mean variance due to the mean-reverting property of variance. The variance behaviour at the initial stage and the behaviour at the long run stage are conceivably driven by different dynamical factors. Indeed, the current economic condition determines the initial variance dynamics, but not the long term dynamics. The latter is considered to be stable irrespective of any initial market conditions. The interpolating of these two dynamical factors produces the evolution dynamics of the variance process. The distinctive characteristic that separates our model from the other multi-factor models (Gatheral (2008), Christoffersen et al. (2009), Egloff et al. (2010)) is that the factor coefficients are time-varying and they play a role in interpolating the initial and long run variance process. The exact functional form of coefficients is chosen to agree with the versatile nature of the variance term structure. In the interest of analytical tractability, the two dynamic factors are structured from affine processes. Up to solutions of Riccati ODEs, valuations of variance (volatility) derivatives and underlying instruments have closed analytical forms.

Besides the requirement of the multi-factor feature, another key element to fortify model robustness is attributed to jumps. Numerous studies have suggested that the presence of jumps forms an integral feature of financial markets. In addition, Todorov and Tauchen (2011) show that simultaneous jumps from the return and the variance with opposite signs are statistically significant. Das and Sundaram (1998) find that jump diffusions don't work at long maturities, but only for short maturities. While stochastic volatility models can't effectively capture high levels of skewness and kurtosis at short maturities. In a study of comparing differential information of the near term versus long term options, Bakshi et al. (2000) find that for the near term options, models with stochastic volatility and jumps in the return process perform better among other alternative models, while for the long term options(LEAPS), models with stochastic volatility alone produce a noticeably better result than models with both stochastic volatility and jumps in the return process. Adding jumps to the return process only helps to deteriorate the overall performance for the long term options. These studies confirm that both the return process and the instantaneous variance process consist of jump components, and jumps or spikes are more identifiable in the short term than in the long term. To meet these empirically justified features, in our variance model, the short term factor is driven by an affine jump-diffusion process, and the long term factor is driven by an affine diffusion process. For the specification of the simultaneous jumps from the return and the variance, we follow Duffie et al. (2000).

In addition to the multi-factor and jump features, our model is naturally endowed with stochastic leverage effect. In currency options, stochastic skew effect is empirically realized (Carr and Wu (2007)). Stochastic skewness can be introduced by randomizing the correlation between the return and increments in volatility. In our model, the stochastic leverage effect is elicited from different dynamical behaviours of the long and the short term factors.

Our time-varying two-factor model is also generic in that it can be easily and naturally transplanted to modelling interest rate dynamics and default rate dynamics. Apart from studying dynamics of stochastic volatility, our model construction thus sheds some new light on the study of other asset families, such as interest rates and credit default rates.

The paper is organized as follows. In section 1, we give a brief coverage of multi-dimensional affine processes. Section 2 presents the time-varying two factor variance model. In section 3, we present results on term structure of variance swaps based on several existing models as well as different damping functions. In section 4, we introduce the CBOE volatility index (VIX), and we price volatility derivatives from the time-varying two factor model. In section 5, we proceed to pricing the equity option. In section 6, we examine the impact to variance/volatility derivatives in the presence of return jumps. Section 7 briefly introduces the data of our analysis. Section 8 deals with the implementation. Finally, section 9 presents the main results covering the term structure of variance swaps, convexity adjustment, SPX option pricing and VIX option pricing.

1. Affine Jump Diffusion Process

Let $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, Q)$ be a filtered probability space, where Q is an equivalent martingale measure under the risk neutral world and the filtration satisfies the usual conditions. Let $X = (x_1, \dots, x_n)^\intercal$ be an n-dimensional $\mathscr{F}$-adapted stochastic process, solving the stochastic differential equation (SDE)

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dJ_t,$$

where W is a standard n-dimensional Wiener process, J is a pure jump process whose jump intensity rate is $\iota(X_t)$ and whose jump size is characterized by its bilateral Laplace transform $\mathscr{L}_J(\cdot)$.

For reasons of computational tractability, we assume the state variable X is an affine jump-diffusion process (AJD, Duffie et al. (2000)). The affine dependence is structured by

$$
\begin{aligned}
&\mu(X_t) = K_0 + K_1 X_t, \quad K_0 \in \mathbb{R}^n, K_1 \in \mathbb{R}^{n\times n},\\
&(\sigma(X_t)\sigma^\intercal(X_t))_{j,k} = (H_0)_{j,k} + (H_1)^\intercal_{j,k} X_t, \quad H_0 \in \mathbb{R}^{n\times n}, H_1 \in \mathbb{R}^{n\times n\times n},\\
&\iota(X_t) = L_0 + L_1^\intercal X_t, \quad L_0 \in \mathbb{R}, L_1 \in \mathbb{R}^n.
\end{aligned}
$$

The main advantage in choosing AJD lies in the fact that the Laplace transform of X_t will be exponentially affine. Under technical regularity conditions (see Duffie et al. (2000)), we have

$$\mathbb{E}(e^{-\omega^\intercal X_t}|\mathscr{F}_s) = e^{-\alpha^\intercal(t-s,\omega)X_s - \beta(t-s,\omega)},$$

where the coefficients $\alpha(t,\omega) \in \mathbb{R}^n$, $\beta(t,\omega) \in \mathbb{R}$ satisfy a Riccati ODE system

$$\frac{\partial \alpha}{\partial t} = K_1^\intercal \alpha - \frac{1}{2}\alpha^\intercal H_1 \alpha - L_1(\mathscr{L}_J(\alpha) - 1)$$

$$\frac{\partial \beta}{\partial t} = K_0^\intercal \alpha - \frac{1}{2}\alpha^\intercal H_0 \alpha - L_0(\mathscr{L}_J(\alpha) - 1)$$

with boundary conditions

$$\alpha(0,\omega) = \omega, \quad \beta(0,\omega) = 0.$$

1.1. Variance Swaps

Let us assume the instantaneous variance process of the underlying asset is a linear combination of state variable X, that is,

$$\sigma_t^2 = \Theta_t^\intercal X_t,$$

where $\Theta_t = (\theta_1(t), \ldots, \theta_n(t))^\intercal$ is an n-dimensional (column) vector whose entries are non random and possibly time varying. In literature, the multi-factor affine model, mostly in the situation where the coefficients are constant, is extensively employed in modeling interest rate (Duffie and Kan (1996), Dai and Singleton (2000)), volatility (Christoffersen et al. (2009), Egloff et al. (2010)), and default rate (Duffie and Singleton (1999)). Whereas the multi-factor model with (non-random) time-varying coefficients is rarely administrated. There are two main reasons. First, the model with constant coefficients is simple and easy to implement. Second, in order to specify these time-varying coefficients, it requires a justifiable explanation for what is the interpretation behind the construction. The rationale underpinning the proposed time-varying two-factor affine model comes from the term structure of forward variance in which the initial forward variance level will always progressively march towards the equilibrium level, the details of which will be presented in section 2.

The concept of variance swaps is similar to the concept of interest rate swaps. *A variance swap* is a forward contract in which the holder pays a predetermined value called *the variance delivery price* or *the variance swap rate* in order to receive at the maturity the annualized realized variance accrued over the lifespan of the contract. Without arbitrage, the variance delivery price is equal to the risk neutral expectation on the annualized future realized variance of the underlying. The future realized variance is a summation of variances induced from the quadratic variation of the diffusion part and the quadratic variation of the jump component on the return process. After taking the conditional risk neutral expectation, we have

$$\mathcal{VS}(0,T) = \mathcal{VS}^C(0,T) + \mathcal{VS}^J(0,T), \tag{1}$$

where $\mathcal{VS}^C(0,T) = \frac{1}{T}\mathbb{E}\left(\int_0^T \sigma_t^2 dt\right)$ is the result of conditional risk neutral expectation on the realized variance caused from the diffusion part of the return, $\mathcal{VS}^J(0,T)$ is caused from the jump component of the return and it vanishes if the underlying dynamics are purely continuous. A standard practice from pricing a log contract (Neuberger (1990), Neuberger (1994), Carr and Wu (2006)) yields that variance swaps can be replicated by vanilla options, that is,

$$\mathcal{VS}(0,T) = \frac{2e^{rT}}{T}\int_0^\infty \frac{\mathcal{O}(K,T)}{K^2}dK + \epsilon(0,T), \tag{2}$$

where $\mathcal{O}(K,T)$ denotes the value of an out-of-the-money option with strike K and maturity T, and $\epsilon(0,T)$ is a small number induced from jumps on the return process. If the return process is immune from jumps, then variance swaps can be perfectly replicated by vanilla options, i.e.,

$$\mathcal{VS}(0,T) = \mathcal{VS}^C(0,T) = \frac{2e^{rT}}{T}\int_0^\infty \frac{\mathcal{O}(K,T)}{K^2}dK.$$

When the return process consists of jumps, both $\mathcal{VS}^J(0,T)$ and $\epsilon(0,T)$ are non-varnishing and small. Furthermore, $\mathcal{VS}^J(0,T) \neq \epsilon$. Objects of $\mathcal{VS}^J(0,T)$ and $\epsilon(0,T)$ will be examined in section 6 after the dynamic model is specified. With the abuse of terminology, in the sequel, $\mathcal{VS}(0,T), \mathcal{VS}^C(0,T)$ and $\frac{2e^{rT}}{T}\int_0^\infty \frac{\mathcal{O}(K,T)}{K^2}dK$ are all called variance swap rates.

In order to get variance swap related products, we need to find the (conditional) expectation of X_t.

$$\mathbb{E}(X_t|\mathscr{F}_s) = -\frac{\partial}{\partial\omega}\mathbb{E}(e^{-\omega X_t}|\mathscr{F}_s)\Big|_{\omega=0} = \left(\frac{\partial}{\partial\omega}\alpha^\intercal(t-s,0)\right)X_s + \frac{\partial}{\partial\omega}\beta(t-s,0).$$

Without arbitrage, the delivery price of a time-t_1 forward variance swap over period $[t_1,t_2]$ is given by

$$\begin{aligned}
\mathcal{VS}^C(t_1,t_2) :&= \mathbb{E}\left(\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\sigma_t^2 dt\Big|\mathscr{F}_{t_1}\right)\\
&= \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\Theta_t^\intercal\mathbb{E}(X_t|\mathscr{F}_{t_1})dt\\
&= A(t_1,t_2)^\intercal X_{t_1} + B(t_1,t_2),
\end{aligned}$$

where

$$A(t_1,t_2) = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} \frac{\partial}{\partial\omega}\alpha(t-t_1,0)\Theta_t dt,$$

$$B(t_1,t_2) = \frac{1}{t_2-t_1}\int_{t_1}^{t_2} \Theta_t^{\intercal}\frac{\partial}{\partial\omega}\beta(t-t_1,0)dt.$$

Again, it is seen that the time-t_1 forward variance swaps are affine in state variable X_t. However, the affine coefficients are not constant, but time-varying. Nevertheless, the Laplace transform of forward variance swaps is still exponentially affine. In penalizing the time-varying affine coefficients, the functions φ, ϕ being determined below will depend on t_1, t separately rather than on the difference $t_1 - t$.

$$\mathbb{E}(e^{-\omega\mathcal{VS}^C(t_1,t_2)}|\mathscr{F}_t) = e^{-\omega B(t_1,t_2)}\mathbb{E}(e^{-\omega A^{\intercal}(t_1,t_2)X_{t_1}}|\mathscr{F}_t)$$
$$= e^{-\psi^{\intercal}(t,t_1,t_2)X_t-\varphi(t,t_1,t_2)}.$$

Under technical regularity conditions, φ and ϕ satisfy a Riccati system of ODEs

$$\frac{\partial\psi}{\partial t} = -K_1^{\intercal}\psi + \frac{1}{2}\psi^{\intercal}H_1\psi + L_1(\mathscr{L}_J(\psi)-1),$$

$$\frac{\partial\varphi}{\partial t} = -K_0^{\intercal}\psi + \frac{1}{2}\psi^{\intercal}H_0\psi + L_0(\mathscr{L}_J(\psi)-1),$$

with boundary conditions

$$\psi(t_1,t_1,t_2) = A(t_1,t_2)\omega, \varphi(t_1,t_1,t_2) = \omega B(t_1,t_2).$$

1.2. Instantaneous Forward Variances and Forward Volatilities

In the interest rate market, forward interest rates are the rates of interest implied by current zero rates of time in the future. They are the future interest rates which can be locked in by trading current bond portfolio. Analogously, forward volatilities are the market price of volatilities that an investor can lock in today to obtain volatility exposure over specific some range of future times.

Definition 1.1. *Given the instantaneous variance process $\{\sigma_t^2\}_{t\geq 0}$, the instantaneous T-forward variance observed at t is defined by*

$$\zeta_t(T) := \mathbb{E}(\sigma_T^2|\mathscr{F}_t).$$

The instantaneous T-forward volatility observed at t is defined by

$$\eta_t(T) := \sqrt{\zeta_t(T)}.$$

Immediately, the T-forward variance observed at t and the time-t forward variance swaps over $[t, T]$ entail the relation below to convert one to the other,

$$\zeta_t(T) = \frac{\partial\big((T-t)\mathcal{VS}^C(t,T)\big)}{\partial T} = \mathcal{VS}^C(t,T) + (T-t)\frac{\partial \mathcal{VS}^C(t,T)}{\partial T},$$
$$\mathcal{VS}^C(t,T) = \frac{1}{T-t}\int_t^T \zeta_t(s)ds.$$

The forward volatility is also termed by *the local volatility*, which lies at the heart of volatility development by Dupire (1994). Notice, the forward volatility is not the conditional expectation of volatility, but the square root of the forward variance. From Definition 1.1, it is seen that the forward variance is a martingale, whereas the forward volatility is not. From the martingale representation theory, the forward variance can be represented by an Ito integral with respect to some Brownian motion processes. Starting from the forward variance curve, which is represented by an exponential martingale, to investigate variance swaps is exactly the modeling approach followed by Buehler (2006), Bergomi (2005), Bergomi (2008) and Cont and Kokholm (2011).

2. The Two-Factor Variance Model

Having disposed of the main structural layout of the variance model, we proceed to build a two-factor variance model and specify relevant functions in such a way that the Riccati ODE system admits closed solutions. The benefit in doing that is it allows us to find tractable solutions to both variance and volatility derivative products.

The time-varying two-factor variance model we propose is described by

$$\sigma_t^2 = \sigma_\infty^2(1-D(t))x_t + \sigma_0^2 D(t) y_t,$$

where

- σ_0^2 and σ_∞^2 are respectively called *the initial instantaneous variance* and *the steady state mean variance.*

- $D(t)$ is a damping function. It satisfies $D(0) = 1$ and $D(\infty) = 0$. The damping function plays a role in interpolating points of the initial instantaneous variance and the steady state mean variance.

- $\{x_t\}_{t\geq 0}$ is a positive affine diffusion process. It satisfies the normalization and initial conditions:

  $$x_0 = 1, \lim_{t\to\infty}\mathbb{E}(x_t) = 1.$$

  $\{y_t\}_{t\geq 0}$ is a positive affine jump diffusion process, whose initial condition satisfies

  $$y_0 = 1.$$

  These conditions are specified so that consistences in the sense that $\lim_{t\to\infty}\mathbb{E}(\sigma_t^2) = \sigma_\infty^2$ and $\lim_{t\to 0}\sigma_t^2 = \sigma_0^2$ are satisfied. Together with properties of D, it follows that in the short term, the variance process is mainly dominated by $\{y_t\}_{t\geq 0}$. In the long term, the process $\{x_t\}_{t\geq 0}$ dominates. In regard to these features, the processes $\{x_t\}_{t\geq 0}$ and $\{y_t\}_{t\geq 0}$ are hereafter called respectively *the permanent factor* (or *the long term factor*) and *the transitory factor* (or *the short term factor*).

The model specification indicates the jump component resides in the transitory factor term. As a result, the jump of variance dynamics will be more pronounced in the short term. The model also indicates that the variance process has the mean reverting property and the steady state mean is σ_∞^2. In addition, inherited from the steady state of the permanent factor $\{x_t\}_{t\geq 0}$, the variance process has a steady state. All these characteristics are well justified from stylized facts of variance dynamics.

To fully examine variance dynamics, we are left with the task of describing two key components: the damping function, the long term and short term factors.

2.1. Damping

Example 2.1 (Single damping)**.**

$$D(t) = e^{-kt}.$$

Example 2.2 (Overdamping)**.**

$$D(t) = ae^{-k_1 t} + (1-a)e^{-k_2 t}.$$

Example 2.3 (Critical damping)**.**

$$D(t) = (1+at)e^{-kt} \quad \textit{where } a, k > 0.$$

These damping models are solutions to a first and second order homogeneous linear ODEs with constant coefficients. Analogously, they can be easily extended to solutions to an n-th order linear ODEs to accommodate more sophisticated damping shapes. Examples (2.2) and (2.3) are called *harmonic oscillators* in classical mechanics (Kreyszig (2005)). Example (2.3) is called the Nelson and Siegel (1987) damping model.

All of these damping models have the characteristic that they approach zero as time goes to infinity. The damping is monotonic for the single damping. In contrast, the curve of the overdamping or critical damping may exhibit humps. Figure 1 illustrates three different shapes of critical damping model.

The key role played by the damping function is that it connects the initial variance with the steady state mean variance. The shape of the term structure of variance swaps is largely influenced by the shape feature of the damping. Damping models are thus naturally and intrinsically associated with the term structure of variance swaps or volatility swaps. For the damping function, without otherwise stated, we use the single-damping model, i.e., $D(t) = e^{-kt}$, where k is called *the damping constant.* Empirically, the term structure usually displays humps. Though the single damping is monotonic, later in Section 3, we show the hump can also be elicited from jumps of the short term factor.

2.2. Long Term and Short Term Factors

It is known that the volatility process is usually negatively correlated with the return process of the underlying asset. The negative correlation between the return and the volatility is called *the leverage effect*. Taking account of the leverage effect, the innovations driving the long term and short term factors can't be arbitrary. If we assume the innovation part of the return process is driven by $\{b_t\}_{t\geq 0}$, innovations of the long term and short term factors are driven respectively by $\{w_{1,t}\}_{t\geq 0}$ and $\{w_{2,t}\}_{t\geq 0}$. Let ρ be the negative correlation of the leverage effect, then there exists Brownian motions b_1, b_2, both independent to b, such that $w_{1,t} = \rho b_t + \sqrt{1-\rho^2} b_{1,t}$ and $w_{2,t} = \rho b_t + \sqrt{1-\rho^2} b_{2,t}$. It leads to $dw_{1,t} \cdot dw_{2,t} = \rho^2 dt + (1-\rho^2) db_{1,t} \cdot db_{2,t}$. Since $-dt \leq db_{1,t} \cdot db_{2,t} \leq dt$, therefore the correlation between the long and short term factors should be no less than $2\rho^2 - 1$. In particular, it suggests that, for an expensive leverage effect, $w_{1,t}$ and $w_{2,t}$ must not be independent to each other. In this paper, we simple assume the long and short term factors are driven by the same innovation, namely, $w_{1,t} = w_{2,t}$.

To explicitly specify the two factors, we provide the following bivariate process. The bivariate process is engineered in such a way that the pricing of VIX derivatives is analytically available.

$$\begin{cases} dx_t = \lambda_1(1-x_t)dt + \xi_1\sqrt{x_t}dw_t, & x_0 = 1, \\ dy_t = \lambda_2(1-x_t)dt + \xi_2\sqrt{x_t}dw_t + dj_t, & y_0 = 1. \end{cases}$$

For the bivariate process, the long term factor $\{x_t\}_{t\geq 0}$ is described by a CIR process. with the Feller condition $\frac{2\lambda_1}{\xi_1^2} \geq 1$ such that the process $\{x_t\}_{t\geq 0}$ is strictly positive. The process $\{x_t\}_{t\geq 0}$ is stationary. Its steady state has a Gamma distribution. Since the steady state of the variance is derived from the steady state of the permanent factor, in the long term, the variance process has a Gamma distribution. Component j is a compound Poisson process whose jump size is exponentially distributed with mean γ and whose intensity rate is a constant ι. The short factor $\{y_t\}_{t\geq 0}$is not a stationary process due to the positive sized jump component. If more flexibilities of variance dynamics are desired, the long run mean level of the short term factor can be different from 1. From the specification of long and short term

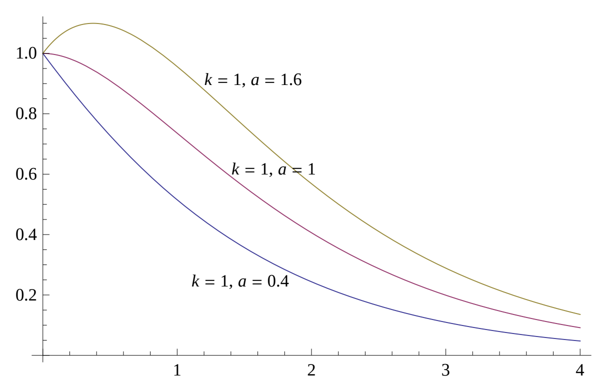

Figure 1: the critical damping model

factors, their conditional means are given by

$$\mathbb{E}(x_t|\mathscr{F}_s) = (x_s - 1)e^{-\lambda_1(t-s)} + 1,$$
$$\mathbb{E}(y_t|\mathscr{F}_s) = (t-s)\iota\gamma + y_s + \lambda_2 \frac{e^{-\lambda_1(t-s)} - 1}{\lambda_1}(x_s - 1).$$

The result of conditional means confirms that $\{x_t\}_{t\geq 0}$ is mean reverting, whereas $\{y_t\}_{t\geq 0}$ is not due to its jump component. From the conditional means, we compute A_1, A_2 and B. These gadgets will be frequently needed in order to having variance swap related products.

$$A_1(t_1,t_2) = \Big(\mathcal{E}(\lambda_1(t_2-t_1))(1-e^{-kt_2}) - \frac{ke^{-kt_1}}{k+\lambda_1}\Phi(k,\lambda_1,t_2-t_1)\Big)\sigma_\infty^2 - \frac{\lambda_2 e^{-kt_1}}{k+\lambda_1}\Phi(k,\lambda_1,t_2-t_1)\sigma_0^2,$$
$$A_2(t_1,t_2) = e^{-kt_1}\mathcal{E}(k(t_2-t_1))\sigma_0^2,$$
$$B(t_1,t_2) = \left(1 - \mathcal{E}(\lambda_1(t_2-t_1)) - \frac{\lambda_1 e^{-kt_1}}{k+\lambda_1}\Phi(k,\lambda_1,t_2-t_1)\right)\sigma_\infty^2 + e^{-kt_1}\left(\frac{\gamma\iota}{k}\Phi(k,0,t_2-t_1) + \frac{\lambda_2}{k+\lambda_1}\Phi(k,\lambda_1,t_2-t_1)\right)\sigma_0^2,$$

where

$$\mathcal{E}(x) = \frac{1-e^{-x}}{x}, \quad \Phi(x,y,t) = \mathcal{E}(xt) - e^{-xt}\mathcal{E}(yt).$$

3. Term Structure of Variance Swaps

The relation between the variance swap delivery price and its time to maturity is often called *the term structure of variance swaps*. The relation conveyed in a mathematical way is a function sending maturity T to the variance swap delivery price $\mathcal{VS}^C(0,T)$.

In the yield curve literature, the term structure is generally decomposed into a handful of latent factors, labelled by the level, the slope and the curvature to explain the entire yield curve. Intuitively, *the level term* is the one which is persistent across all maturities. Pertaining to those factors used in Nelson and Siegel (1987), the function

$$\mathcal{E}(kt) = \frac{1-e^{-kt}}{kt},$$

which starts at 1 and diminishes monotonically to 0 at infinity, is called *the slope term*. The parameter k controls the speed of decay. The function

$$\Phi(k,0,t) = \frac{1-e^{-kt}}{kt} - e^{-kt},$$

which starts initially at 0, increases and eventually decays to 0 at infinity, is called *the curvature term*. The curvature term genuinely produces humps. An important insight of there factors is that they may also be interpreted as the long term, short term and medium term.

In additional to these three factors, we introduce a fourth factor called *the aberrancy term*[1] by the function

$$\Psi(kt) = \frac{1 - e^{-kt}}{kt} - e^{-kt} - \frac{1}{2}kte^{-kt}.$$

Aberrancy is the term used to depict the third derivative of the curve. Its geometric interpretation is that it measures locally the deviation of symmetry at a given point on the curve. The aberrancy term provides an additional degree of freedom to control locally the asymmetric behavior of the curve. Figure 2 plots the shapes of the four factors.

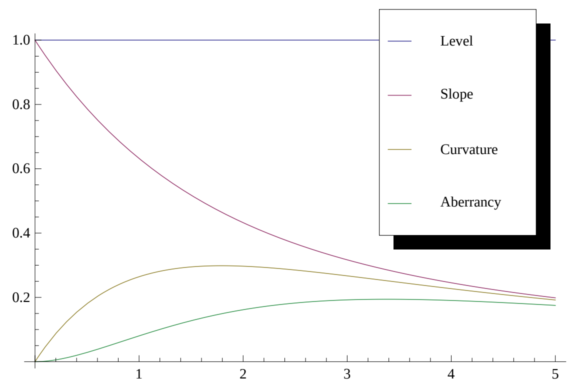

Figure 2: The Four Factors

To explore the role played by the damping and jump more rigorously, we examine the term structure of variance swaps obtained from different damping functions. But before we do, we first inspect the term structure attained from several existing dynamic models as an attempt to display their resemblances and dissimilarities.

3.1. The Black-Scholes Model

The pioneering Black-Scholes model is very basic in that the volatility is a constant. As a result, the term structure of variance swaps is a constant. It only produces the level term.

3.2. The Basic Affine Process(Duffie and Gârleanu (2001))

The SDE

$$dv_t = \lambda(\theta - v_t)dt + \sigma\sqrt{v_t}dw_t + dj_t,$$

where $\{j_t\}_{t\geq 0}$ is a jump process whose intensity rate is ι and whose jump size is exponentially distributed with mean γ, is called *the basic affine process*. Let us assume the instantaneous variance process is governed by the basic affine process, then the term structure of variance swaps becomes

$$\mathcal{VS}^C(0,T) = (\frac{\gamma\iota}{\lambda} + \theta - v_0)\mathcal{E}(\lambda T) + \theta + \frac{\gamma\iota}{\lambda}. \tag{3}$$

[1]The original notion is *déviation*, introduced by Transon (1841). *Aberrancy* is the modern translated term. For the geometric interpretation, see Wilczynski (1916).

It demonstrates that the term structure consists of only the level factor and the slope factor. As a result, the curve is monotonic.

3.3. The Two-Factor Process(Balduzzi et al. (1998))

The two-factor variance process, governed by the following SDEs

$$\begin{cases} dv_t = \lambda(v_t' - v_t)dt + \xi\sqrt{v_t}dw_{1t}, \\ dv_t' = \kappa(\theta - v_t')dt + \eta\sqrt{v_t'}dw_{2t}. \end{cases}$$

is generally employed in modeling variance dynamics (Gatheral (2008), Christoffersen et al. (2009), Egloff et al. (2010)). The process $\{v_t\}_{t\geq 0}$ depicts the instantaneous variance, and $\{v_t'\}_{t\geq 0}$ is the stochastic central tendency towards the long run mean level θ.

The resulting term structure of variance swaps (Egloff et al. (2010)) from the two-factor process turns into

$$\mathcal{VS}^C(0,T) = \theta + \frac{k(\theta - v_0) + \lambda(v_0 - v_0')}{\lambda - \kappa}\mathcal{E}(\lambda T) + \frac{\lambda(v_0' - \theta)}{\lambda - \kappa}\mathcal{E}(\kappa T). \tag{4}$$

It indicates that the term structure consists of three terms, one level term followed by two different slope terms. Each of the slope terms is controlled by a different decay rate. It improves the model (3) by adding one more slope term. However, the hump is not evident.

3.4. The Single Damping Model

For the single damping model, we have

$$\begin{aligned} \mathcal{VS}^C(0,T) &= A_1(0,T) + A_2(0,T) + B(0,T) \\ &= \sigma_\infty^2 + (\sigma_0^2 - \sigma_\infty^2)\mathcal{E}(kT) + \sigma_0^2\frac{\gamma\iota}{k}\Phi(k,0,T). \end{aligned} \tag{5}$$

The corresponding instantaneous forward variance curve is

$$\zeta_0(T) := \mathbb{E}\sigma_T^2 = \sigma_\infty^2 + \left(\sigma_0^2 - \sigma_\infty^2\right)e^{-kT} + \sigma_0^2\gamma\iota T e^{-kT}.$$

The term structure model (5) is divided into three components: the level, the slope and the curvature. In comparing to the interest rate, it is exactly the yield curve model of Nelson and Siegel (1987). The main difference is that the curvature coefficient of model (5) is always positive and the model (5) is deduced naturally from the time-varying two factor variance model. By specifying functional forms of the Nelson-Sigel term structure model with time-varying factors, Christensen et al. (2009), and Christensen et al. (2011) also derive multi-variate affine processes which maintain the Nelson-Sigel factor loading structure.

It is worth emphasizing that λ_1 and λ_2, the speed of reversion of the long and short term factors, as well as ξ_1 and ξ_2, don't enter into equation (5). In contrast, the term structure of previous models relies on its speed of reversion to control the slope. The advantage of our model is that λ_1, λ_2, ξ_1 and ξ_2 provide unique freedoms to control implied smiles. The term structure model (5) is solely determined by $(\sigma_0, \sigma_\infty, k, \gamma\iota)$. Conceptually, σ_0^2 is the initial level of variance displacement from the equilibrium level σ_∞^2. The parameter k, the damping constant, plays a key role in determining panoramic shape of the term structure. The parameter $\gamma\iota$, stemming from the jump part of the short term factor, locally adds additional flexibility of humps. If there is no jump, the term structure model (5) degenerates to the long and the short term components. If the damping is not included, i.e. $k = 0$, it further degenerates to the long term component. Viewing from the upgraded direction, the term structure of variance swaps implied from the Black-Scholes model only produces the long term component. If a stochastic Heston model (with jump) is assumed, the slope term is added. If in addition, the damping and the jump are added, the term structure displays additional humps.

3.5. Critical Damping Model

For the critical damping model, the forward variance curve has the form

$$\zeta_0(T) = \sigma_\infty^2 + \left(a\gamma\iota\sigma_0^2 T^2 + \left(\sigma_0^2(a+\gamma\iota) - a\sigma_\infty^2\right)T + \sigma_0^2 - \sigma_\infty^2\right)e^{-kT}.$$

The curve is a solution to a third order homogeneous linear ODE whose characteristic function has three identical root k. The corresponding term structure of variance swaps is

$$\mathcal{VS}^C(0,T) = \sigma_\infty^2 + (\sigma_0^2 - \sigma_\infty^2)\mathcal{E}(kT) + \frac{a\left(\sigma_0^2 - \sigma_1^2\right) + \gamma\iota\sigma_0^2}{k}\Phi(k,0,T) + \frac{2a\gamma\iota\sigma_0^2}{k^2}\Psi(kT),$$

Interestingly, it decomposes into the level, slope, curvature and aberrancy term. The parameter a manages the behavior of the aberrancy term as well as possibly the curvature term. The jump component also reins the aberrancy term.

3.6. Overdamping Model

For the overdamping model, the instantaneous forward variance rate becomes

$$\zeta_0(T) = \sigma_\infty^2 + (1-a)\left(\sigma_0^2 - \sigma_\infty^2 + \gamma\iota\sigma_0^2 T\right)e^{-k_2 T} + a\left(\sigma_0^2 - \sigma_\infty^2 + \gamma\iota\sigma_0^2 T\right)e^{-k_1 T},$$

which is a solution to a fourth order homogeneous linear ODE whose characteristic function has two distinctive roots, k_1 and k_2, each of which has multiplicity 2. The corresponding term structure of variance swaps turns into

$$\begin{aligned}\mathcal{VS}^C(0,T) = \sigma_\infty^2 &+ \left(\sigma_0^2 - \sigma_\infty^2\right)a\mathcal{E}(k_1 T) + \left(\sigma_0^2 - \sigma_\infty^2\right)(1-a)\mathcal{E}(k_2 T) \\ &+ \frac{a\gamma\iota\sigma_0^2}{k_1}\Phi(k_1,0,T) + \frac{(1-a)\gamma\iota\sigma_0^2}{k_2}\Phi(k_2,0,T).\end{aligned}$$

The term structure consists of 5 terms, among which there are two slope terms and two curvature terms. Each of the two slope terms is controlled by a different damping constant, so are the two curvature terms. The parameter a assigns a weight to each of the slope terms as well as the curvature terms. The term assigned with a larger weight will be called *the* *primary term*, the less weighted one will be called *the secondary term*. The overall effect of the two curvature terms is that the curve exhibits possibly two humps. The impact of the two slope terms together is that it develops into one monotonic slope. The term structure model is virtually a Nelson-Siegel-Svensson type(Svensson (1994)).

From the study of these existing dynamic models, we see the resulting term structures only consist of the level term and the slope term. Lacking the curvature term, the hump, an empirically justified feature, is missing. For our time-varying two-factor model, a combination of the damping and the jump effectively creates humps [2]. Multiple humps can be produced by properly adjusting the damping model. Without the jump (i.e., $\kappa\iota = 0$), the term structure curve is purely determined by the damping model except the situation when $\sigma_0 = \sigma_\infty$, in which case, the term structure curve is flat. These examinations suggest that the jump and damping form essential parts in modeling the term structure of variance swaps.

4. Volatility Index

Variance swaps are much more widely traded in the market than volatility swaps due to the fact they can be replicated by vanilla options. While volatility swaps have not generally thought to be replicated possibly. However, in financial markets, the volatility index is conventionally referenced rather than the variance index.

In 1993, the Chicago Board of Options Exchange(CBOE) introduced the CBOE volatility index called VIX. VIX_t measures the implied volatility of S&P 500 index options of the next 30 days after time t. Mathematically speaking, VIX_t^2 approximates the conditional risk neutral expectation of the annualized realized variance over the next 30 days. The original calculation of VIX was revamped to accommodate a more robust method in September 2003. The principal idea of the new calculation is based on the fact that variance swaps can be replicated by the corresponding option prices of the underlying. To calculate VIX^2 price, the CBOE first uses available OTM SPX options to approximate $\frac{2e^{rT_1}}{T_1}\int_0^\infty \frac{\mathcal{O}(K,T_1)}{K^2}dK$ and $\frac{2e^{rT_2}}{T_2}\int_0^\infty \frac{\mathcal{O}(K,T_2)}{K^2}dK$ from the two nearest maturities T_1 and T_2, the linear interpolated value of these two quantities yields the price of VIX^2. When the nearest time to maturity is less than 8 days, the CBOE switches to the next-nearest maturity in order to avoid microstructure effects. For detailed computation and specification of VIX, we refer to Carr and Wu (2006).

In this paper, we simply assume

$$\text{VIX}_t^2 = \mathcal{VS}^C(t, t+\hbar) + \mathcal{VS}^J(t, t+\hbar) - \epsilon(t, t+\hbar), \tag{6}$$

where $\hbar = \frac{1}{12}$ represents an annualized 30 calendar days. The calculation (6) is exact up to approximation errors stemming from two major sources: option interpolations/extrapolations from discrete strikes, a linear interpolation/extrapolation from two maturities.

In equation (6), the proportional value of $\mathcal{VS}^J(t, t+\hbar) - \epsilon(t, t+\hbar)$ to $\mathcal{VS}^C(t, t+\hbar)$ is

[2]In the time-varying two-factor model, bumps can also be elicited if the long run mean level of the short term factor is different from the long run mean level of the long term factor.

small, we define

$$\widetilde{\mathrm{VIX}}_t^2 := \mathcal{VS}^C(t, t+\hbar) = \frac{1}{h}\mathbb{E}\left(\int_t^{t+\hbar} \sigma_s^2 ds|\mathscr{F}_t\right). \tag{7}$$

If the turn process is immune from jumps, then $\widetilde{\mathrm{VIX}}_t = \mathrm{VIX}_t$.

4.1. $\widetilde{\mathrm{VIX}}$ Dynamics

From the time-varying two-factor variance model specification and definition (6), after a rearrangement of terms, we have

$$\begin{aligned}
\widetilde{\mathrm{VIX}}_t^2 =& A_1(t, t+\hbar)x_t + A_2(t, t+\hbar)y_t + B(t, t+\hbar) \\
=& \mathcal{E}(\lambda_1 h)(x_t - 1)\sigma_\infty^2 + \sigma_\infty^2 && (8)\\
&+ e^{-kt}\left(\mathcal{E}(kh)\sigma_0^2 y_t - e^{-kh}\mathcal{E}(\lambda_1 h)\sigma_\infty^2 x_t - \frac{\Phi(k,\lambda_1,h)}{k+\lambda_1}(\lambda_2\sigma_0^2 + k\sigma_\infty^2)x_t\right) && (9)\\
&+ e^{-kt}\left(\frac{\gamma\iota}{k}\Phi(k,0,h)\sigma_0^2 + \frac{\Phi(k,\lambda_1,h)}{k+\lambda_1}(\lambda_2\sigma_0^2 - \lambda_1\sigma_\infty^2)\right). && (10)
\end{aligned}$$

It indicates that $\widetilde{\mathrm{VIX}}_t^2$ can be decomposed into three components. Component (8) simply describes the long term behavior of implied variances. It is primarily controlled by the permanent factor $\{x_t\}_{t\geq 0}$. The long term component consists of a level factor and a slope factor. Component (9) is the stochastic transitory term, which consists of a mixture of stochastic slope and curvature factors. The stochastic slope and curvature factors are managed by both the long term factor $\{x_t\}_{t\geq 0}$ and the short term factor $\{y_t\}_{t\geq 0}$. Both of stochastic slope and curvature factors fade away by the damping constant k. Component (10) is the deterministic transitory term. It slowly becomes weaker due to the damping constant k.

The dynamic movement of $\widetilde{\mathrm{VIX}}^2$ goes as follows. At $t = 0$, $\widetilde{\mathrm{VIX}}_0^2$ is decomposed into a level, a slope and a curvature term. The slope and the curvature factor are controlled by the damping k. Initially, the process $\{y_t\}_{t\geq 0}$ dominates the process $\{x_t\}_{t\geq 0}$. As a result, jumps are more identifiable in the short term. $\widetilde{\mathrm{VIX}}_t^2$ consists of the same level and a mixture of stochastic slope and curvature terms. With the lapse of time, the process $\{x_t\}_{t\geq 0}$ dominates the process $\{y_t\}_{t\geq 0}$. $\widetilde{\mathrm{VIX}}_t^2$ becomes nearly a continuous diffusion process. The stochastic curvature component fades away. As $t \to \infty$, $\{y_t\}_{t\geq 0}$ only consists of the level term and a stochastic slope term. The slope term is shifted from the original damping k to λ_1. The steady state distribution of VIX_t^2 is generated from the steady state distribution of the permanent factor $\{x_t\}_{t\geq 0}$. The resulting different jump behaviours between the near term and long term options are consistent with the empirical study by Bakshi et al. (2000).

4.2. VIX Instruments

On March 26, 2004, the CBOE launched a new exchange, the Chicago Futures Exchanges (CFE), and started to trade futures on VIX. VIX futures (VX) are standard futures contract. VIX futures price is defined by

$$F(T) := \mathbb{E}(\mathrm{VIX}_T).$$

VIX futures is simply a contract on a forward 30-day implied volatilities.

VIX options were launched on Feb 24, 2006. They are no different than standard options written on stocks. The payoff at the expiry date T of a VIX vanilla call option with strike K is given by

$$(\mathrm{VIX}_T - K)^+.$$

The payoff at the expiry date T of a VIX vanilla put option with strike K is given by

$$(K - \mathrm{VIX}_T)^+.$$

We are thus led to the valuation of VIX derivatives. Before we do that, we first price derivatives of $\widetilde{\mathrm{VIX}}$. The valuation of VIX derivatives will be provided after the return process is specified. Since the Laplace transform of random variable $\widetilde{\mathrm{VIX}}_t^2$ is known in closed form. To price $\widetilde{\mathrm{VIX}}$ derivatives, we resort to the Laplace transform approach. Valuations of variance derivatives as an inverse Laplace problem are also examined in Friz and Gatheral (2005).

4.3. Laplace Transform of $\mathcal{VS}^C(t_1, t_2)$

The conditional Laplace transform of $\mathcal{VS}^C(t_1, t_2)$ is given by

$$\mathbb{E}(e^{-\omega \mathcal{VS}^C(t_1,t_2)} | \mathscr{F}_t) = e^{-\psi_1(t)x_t - \psi_2(t)y_t - \varphi(t)}, \quad \omega \in \mathbb{C}.$$

In general, ω is an element taken from the region bounded left by a vertical line segment.

Following the material investigated in Section 1.1, the coefficients $\psi_1(t), \psi_2(t), \varphi(t)$ satisfy a Riccati ODE system

$$\begin{cases} \dfrac{\partial \psi_1}{\partial t} = \lambda_1 \psi_1 + \lambda_2 \psi_2 + \dfrac{1}{2}\xi_1^2 \psi_1^2 + \dfrac{1}{2}\xi_2^2 \psi_2^2 + \xi_1 \xi_2 \psi_1 \psi_2, \\ \dfrac{\partial \psi_2}{\partial t} = 0, \\ \dfrac{\partial \varphi}{\partial t} = -\lambda_1 \psi_1 - \lambda_2 \psi_2 + \iota(\dfrac{1}{1+\psi_2 \gamma} - 1), \end{cases}$$

with boundary conditions

$$\psi_1(t_1, \omega) = A_1(t_1, t_2)\omega, \quad \psi_2(t_1, \omega) = A_2(t_1, t_2)\omega, \quad \varphi(t_1, \omega) = \omega B(t_1, t_2).$$

The solution to the Riccati ODE system is

$$\begin{cases} \psi_1(t) = \omega \left(A_1 + \dfrac{(1 - e^{-\vartheta(t_1 - t)})a}{(1 - e^{-\vartheta(t_1 - t)})(\vartheta - b) - 2\vartheta} \right) \\ \psi_2(t) = \omega A_2, \\ \varphi(t) = \omega B - \dfrac{\lambda_1(\lambda_1 - \vartheta)}{\xi_1^2}(t_1 - t) + \dfrac{2\lambda_1}{\xi_1^2} \ln \left(1 - \dfrac{\vartheta - b}{2\vartheta}(1 - e^{-\vartheta(t_1 - t)}) \right) \\ \qquad - \dfrac{\omega A_2}{\xi_1}(\lambda_1 \xi_2 - \lambda_2 \xi_1)(t_1 - t) - \iota(\dfrac{1}{1 + \omega A_2 \gamma} - 1)(t_1 - t), \\ \end{cases}$$

where

$$\vartheta = \sqrt{\lambda_1^2 - 2\xi_1 A_2(\xi_1\lambda_2 - \xi_2\lambda_1)\omega},$$

$$a = 2A_1\lambda_1 + 2A_2\lambda_2 + \omega\left(A_2\xi_2 + A_1\xi_1\right)^2,$$

$$b = \lambda_1 + \xi_1\omega(\xi_1 A_1 + \xi_2 A_2).$$

4.4. Pricing $\widetilde{\mathrm{VIX}}$ Derivatives

From the integral representation of square root function,

$$\sqrt{x} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-\omega x}}{\omega^{\frac{3}{2}}} d\omega,$$

immediately, the $\widetilde{\mathrm{VIX}}$ futures can be evaluated by

$$\mathbb{E}\widetilde{\mathrm{VIX}}_T = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \mathbb{E}(e^{-\omega \widetilde{\mathrm{VIX}}_T^2})}{\omega^{\frac{3}{2}}} d\omega.$$

For a $\widetilde{\mathrm{VIX}}$ call option with strike K and maturity T, let $f(x) = (\sqrt{x} - K)^+$ be its intrinsic value. Let $\mathrm{Re}(\omega) > 0$, the Laplace transform of the intrinsic value is given by

$$\begin{aligned} \mathcal{L}(f)(\omega) &= \int_0^\infty (\sqrt{x} - K)^+ e^{-\omega x} dx \\ &= \frac{\sqrt{\pi}}{2\omega^{\frac{3}{2}}} \operatorname{erfc}(K\sqrt{\omega}) \end{aligned}$$

where $\operatorname{erfc}(\cdot)$ is the complementary error function. The inverse Laplace transform is given by

$$f(x) = \frac{1}{2\pi i} \lim_{\eta \to \infty} \int_{\zeta - i\eta}^{\zeta + i\eta} \mathcal{L}(f)(\omega) e^{\omega x} d\omega,$$

which is (absolutely) convergent when $\zeta > 0$. For a suitable ζ, the expectation

$$\mathbb{E}(f(\widetilde{\mathrm{VIX}}_T^2)) = \frac{1}{2\pi i} \lim_{\eta \to \infty} \int_{\zeta - i\eta}^{\zeta + i\eta} \mathcal{L}(f)(\omega) \mathbb{E}(e^{\omega \widetilde{\mathrm{VIX}}_T^2}) d\omega.$$

is convergent. In contrast, a mollification technique in Friz and Gatheral (2005) is needed in order to avoiding oscillation and ensuring the convergence of the integration.

Consequently, the $\widetilde{\mathrm{VIX}}$ call option price is given by

$$C(K,T) = \frac{e^{-rT}}{2\pi i} \int_{\zeta - i\infty}^{\zeta + i\infty} \mathcal{L}(f)(\omega) \mathbb{E}(e^{\omega \widetilde{\mathrm{VIX}}_T^2}) d\omega.$$

In the same manner, the corresponding put option price at the same strike and maturity is given by

$$P(K,T) = \frac{e^{-rT}}{2\pi i} \int_{\zeta - i\infty}^{\zeta + i\infty} \mathcal{L}(g)(\omega) \mathbb{E}(e^{\omega \widetilde{\mathrm{VIX}}_T^2}) d\omega,$$

where

$$\mathcal{L}(g)(\omega) = \int_0^\infty (K - \sqrt{x})^+ e^{-\omega x} dx = \frac{K}{\omega} - \frac{\sqrt{\pi}}{2\omega^{\frac{3}{2}}} \operatorname{erf}(K\sqrt{\omega})$$

and $\operatorname{erf}(\cdot)$ is the error function. Put option prices can also be obtained from the put-call parity.

5. SPX Option Valuations

To price SPX options, the important point remaining concerns the leverage effect. Leverage effect is the correlation between shocks to returns and shocks to variances. In light of analytical tractability of affine process, aggregating the return process of the underlying with the bivariate process, we assume under the risk neutral world, they form a three dimensional affine process described below:

$$\begin{cases} dx_t = \lambda_1(1 - x_t)dt + \xi_1\sqrt{x_t}dw_t, \quad x_0 = 1, \\ dy_t = \lambda_2(1 - x_t)dt + \xi_2\sqrt{x_t}dw_t + dj_t, \quad y_0 = 1, \\ dz_t = \left(r - \delta - \frac{1}{2}\left((1 - e^{-kt})\sigma_\infty^2 x_t + e^{-kt}\sigma_0^2 y_t\right) - \iota\mu\right) dt + \rho\sqrt{((1 - e^{-kt})\sigma_\infty^2 + e^{-kt}\sigma_0^2)x_t}dw_t \\ \qquad + \sqrt{(1 - \rho^2)(1 - e^{-kt})\sigma_\infty^2 x_t + e^{-kt}\sigma_0^2(y_t - \rho^2 x_t)}db_t + e^{-kt}dq_t, \quad z_0 = \ln S_0. \end{cases}$$

Here, $\{w_t\}_{t\geq 0}$ and $\{b_t\}_{t\geq 0}$ are two independent Brownian motions, r is the constant riskless interest rate, δ is the constant proportional dividend rate, S_0 is the initial price of the underlying. The diffusion correlation is given by

$$\rho\sqrt{\frac{(1-e^{-kt})\sigma_\infty^2 x_t + e^{-kt}\sigma_0^2 x_t}{(1-e^{-kt})\sigma_\infty^2 x_t + e^{-kt}\sigma_0^2 y_t}}, \tag{11}$$

where ρ is subject to

$$-1 \leq \rho \leq 1, \quad (1-\rho^2)(1-e^{-kt})\sigma_\infty^2 x_t + e^{-kt}\sigma_0^2(y_t - \rho^2 x_t) \geq 0.$$

The above diffusion correlation suggests the leverage effect is both time-varying and stochastic. The diffusion correlation initially starts from ρ. A perturbation, caused by different behavior movements of x_t and y_t prompts the course deviation of the instantaneous diffusion correlation away from ρ. The deviation gets weakened as a result of the damping k. Eventually, the diffusion correlation progressively moves towards ρ. The time-varying and stochastic nature of leverage effect is empirically supported from studies of currency options by Carr and Wu (2007). The diffusion correlation is not limited to being of the form (11). Different forms of diffusion correlations can be configured to accommodate more flexible features as needed.

Following the return and variance jump specifications from Duffie et al. (2000), we assume the jump component of return is depicted by a compound Poisson process $\{q_t\}_{t\geq 0}$, whose intensity rate is ι and whose size of jump is normally distributed and correlated to the size of jump from process $\{j_t\}_{t\geq 0}$. More precisely, we assume the joint Laplace transform function of jumps of j_t and q_t is given by

$$\mathscr{L}_{j_t,q_t}(\omega_1,\omega_2) = \frac{e^{-\theta\omega_2+\frac{1}{2}v\omega_2^2}}{1+\varrho\gamma\omega_2+\gamma\omega_1}.$$

Translating into the probability density, it means the marginal distribution of jump size in j_t is exponentially distributed with mean γ. Conditional on the jump size l of j_t, the jump size of q_t is normally distributed with mean $\theta+\varrho l$ and variance v. Here, ϱ is interpreted to be the jump correlation. Simultaneous jumps with opposite signs in return and in volatility are well justified empirically (Todorov and Tauchen (2011)). The empirical testing from Bakshi et al. (2000) confirm the jump diffusion process for the near term options and diffusion process for the long term options are more desirable than other alternatives. Therefore, we add an exponential damping function in front of q_t. As a result, the jump from the log return process is much more evident in the short term, the log return process becomes nearly a continuous diffusion process in the long term. The joint Laplace transform of j_t and $e^{-kt}q_t$ becomes

$$\mathscr{L}_{j_t,e^{-kt}q_t}(\omega_1,\omega_2) = \frac{\exp(-\theta\omega_2 e^{-kt}+\frac{1}{2}v\omega_2^2 e^{-2kt})}{1+\gamma\omega_1+\varrho\gamma\omega_2 e^{-kt}}.$$

The corresponding joint density function is given by

$$p_{j_t, e^{-kt}q_t}(x,y) = \frac{1}{\sqrt{2\pi v}\gamma} \exp\left(kt - \frac{x}{\gamma} - \frac{\left(e^{kt}y - (\theta + \varrho x)\right)^2}{2v} \right).$$

Finally, $\mu = \mathscr{L}_{j_t, e^{-kt}q_t}(0,-1) - 1$ so that the process $\{e^{z_t-(r-\delta)t}\}_{t\geq 0}$ forms a martingale under the risk neutral world.

We proceed to computer the (conditional) Laplace transform of log return z_t. Let

$$\mathbb{E}(e^{-\omega z_{t_1}} | \mathscr{F}_t) = e^{-\psi_1(t)x_t - \psi_2(t)y_t - \psi_3(t)z_t - \varphi(t)}, \quad \mathrm{Re}(\omega) > 0,$$

then the coefficients ψ_1, ψ_2, ψ_3 and φ are determined by a Riccati ODE system

$$\begin{cases} \dfrac{\partial \psi_1}{\partial t} = \lambda_1\psi_1 + \lambda_2\psi_2 + \dfrac{1}{2}(1-e^{-kt})\sigma_\infty^2\psi_3 + \dfrac{1}{2}(\xi_1\psi_1 + \xi_2\psi_2 + \rho\sqrt{(1-e^{-kt})\sigma_\infty^2 + e^{-kt}\sigma_0^2}\psi_3)^2 \\ \qquad + \dfrac{1}{2}\left((1-e^{-kt})\sigma_\infty^2 - \rho^2((1-e^{-kt})\sigma_\infty^2 + e^{-kt}\sigma_0^2)\right)\psi_3^2, \\ \dfrac{\partial \psi_2}{\partial t} = \dfrac{1}{2}e^{-kt}\sigma_0^2\psi_3 + \dfrac{1}{2}e^{-kt}\sigma_0^2\psi_3^2, \\ \dfrac{\partial \psi_3}{\partial t} = 0 \\ \dfrac{\partial \varphi}{\partial t} = -(r-\delta-\iota\mu)\psi_3 - \lambda_1\psi_1 - \lambda_2\psi_2 + \iota\left(\dfrac{\exp(-\theta\psi_3 e^{-kt} + \frac{1}{2}v\psi_3^2 e^{-2kt})}{1+\gamma\psi_2 + \varrho\gamma\psi_3 e^{-kt}} - 1\right), \end{cases}$$

with boundary conditions

$$\psi_1(t_1,\omega) = 0, \quad \psi_2(t_1,\omega) = 0, \quad \psi_3(t_1,\omega) = \omega, \quad \varphi(t_1,\omega) = 0.$$

Immediately, it implies $\psi_3(t) = \omega$, $\psi_2(t) = \frac{1}{2k}(\omega+\omega^2)(e^{-kt_1} - e^{-kt})\sigma_0^2$ and ψ_1 is the solution to the Riccati ODE

$$\begin{aligned} \frac{\partial \psi_1}{\partial t} &= \frac{1}{2}\xi_1^2\psi_1^2 + \left(\lambda_1 + \xi_1\xi_2\psi_2 + \omega\xi_1\rho\sqrt{(1-e^{-kt})\sigma_\infty^2 + e^{-kt}\sigma_0^2}\right)\psi_1 \\ &\quad + \lambda_2\psi_2 + \frac{1}{2}\omega(1+\omega)(1-e^{-kt})\sigma_\infty^2 + \frac{1}{2}\xi_2^2\psi_2^2 + \omega\xi_2\rho\sqrt{(1-e^{-kt})\sigma_\infty^2 + e^{-kt}\sigma_0^2}\psi_2 \end{aligned}$$

with boundary condition $\psi_1(t_1) = 0$. ψ_1 can be solved numerically, such as by the Runge-Kutta method. Once the numerical solution ψ_1 is attained, the solution to φ is given by

$$\begin{aligned} \varphi(t) = &\frac{\lambda_2}{2k^2}(\omega+\omega^2)\sigma_0^2(e^{-kt_1}(1+k(t_1-t)) - e^{-kt}) + \lambda_1\int_t^{t_1}\psi_1(s)ds + (\iota + (r-\delta+\iota)\omega)(t_1-t) \\ &- \omega\iota\int_t^{t_1}\frac{\exp(\theta e^{-ks} + \frac{1}{2}ve^{-2ks})}{1-\varrho\gamma e^{-ks}}ds - \iota\int_t^{t_1}\frac{\exp(-\theta\omega e^{-ks} + \frac{1}{2}v\omega^2 e^{-2ks})}{1+\gamma\psi_2(s) + \varrho\gamma\omega e^{-ks}}ds, \end{aligned}$$

Having obtained numerically the Laplace transform of log return

$$\mathbb{E}(e^{-\omega z_t}) = e^{-\psi_1(0)-\psi_2(0)-\phi(0)-\omega \ln S_0},$$

we use the fast Fourier transform method presented in Carr and Madan (1999) to find the pricings of SPX options.

6. Roles of Jumps from the Return Process

If the return process is immune from jumps, variance swap rates are equivalent to the quadratic variation of the diffusion part of the return process and variance swaps can be completely replicated by vanilla options. The presence of jumps in the return process is confirmed and justified in empirical work (Bates (1996)). Consequently, to derive quadratic variations of the return process, some correction terms are needed in the presence of jumps. The aim of this section is to examine the impact of return jumps to variance swap rates and volatility derivatives.

Let M be the Poisson random measure of jumps from the return process and let ν be its corresponding intensity measure. From equation (6), two random variables gaining further attentions are

$$\mathcal{VS}^J(t_1,t_2) = \frac{1}{t_2-t_1}\int_{t_1}^{t_2}\int_{\mathbb{R}} x^2 dM(s,x),$$

$$\epsilon(t_1,t_2) = -\frac{2}{t_2-t_1}\int_{t_1}^{t_2}\int_{\mathbb{R}} (e^x - 1 - x - \frac{x^2}{2})dM(s,x). \tag{12}$$

The establishment of equation (12) can be found in Carr and Wu (2006). To adjust valuations of volatility derivatives, we compute the Laplace transform of $\mathcal{VS}^J(t_1,t_2) - \epsilon(t_1,t_2)$. To correct terms to variance swap rates as a result of return jumps, we quantify $\mathcal{VS}^J(0,T)$ and $\epsilon(0,T)$.

Proposition 6.1. *For* $\mathrm{Re}(\omega) > 0$

$$\begin{aligned}&\mathbb{E}\exp\left(-\omega(\mathcal{VS}^J(t_1,t_2)-\epsilon(t_1,t_2))\right)\\ \approx &\exp\left(-\iota(t_2-t_1) - \frac{\iota}{2\gamma\varrho}\sqrt{\frac{\pi(t_2-t_1)}{\omega}}e^{\frac{\theta}{\gamma\varrho}}\int_{t_1}^{t_2}\left(1+\operatorname{erf}\left(\frac{c}{2\gamma\varrho}+\frac{\theta}{c}\right)\right)e^{kt+\frac{c^2}{4\gamma^2\varrho^2}}dt\right),\end{aligned}$$

where

$$c = \sqrt{2v + \frac{t_2-t_1}{\omega}e^{2kt}}.$$

Proof.

$$
\begin{aligned}
&\mathbb{E}\exp\left(-\omega(\mathcal{VS}^J(t_1,t_2)-\epsilon(t_1,t_2))\right)=\mathbb{E}\exp\left(-\frac{2\omega}{t_2-t_1}\int_{t_1}^{t_2}\int_{\mathbb{R}}(e^x-1-x)dM(s,x)\right)\\
&=\exp\left(\int_{t_1}^{t_2}\int_{\mathbb{R}}(e^{-\frac{2\omega}{t_2-t_1}(e^x-1-x)}-1)\nu(dx,dt)\right) \qquad (13)\\
&\approx\exp\left(\int_{t_1}^{t_2}\int_{\mathbb{R}}(e^{-\frac{\omega x^2}{t_2-t_1}}-1)\nu(dx,dt)\right)\\
&=\exp\left(-\iota(t_2-t_1)-\frac{\iota}{2\gamma\varrho}\sqrt{\frac{\pi(t_2-t_1)}{\omega}}e^{\frac{\theta}{\gamma\varrho}}\int_{t_1}^{t_2}\left(1+\operatorname{erf}\left(\frac{c}{2\gamma\varrho}+\frac{\theta}{c}\right)\right)e^{kt+\frac{c^2}{4\gamma^2\varrho^2}}dt\right).
\end{aligned}
$$

Equation (13) is obtained from exponential formula for Poisson random measure (see Cont and Tankov (2004)).

Results from Section 4.3 and Proposition 6.1 allow one to get the Laplace transform of VIX_T^2, i.e.,

$$\mathbb{E}(e^{-\omega\,\mathrm{VIX}_T^2})\approx\mathbb{E}(e^{-\omega\widetilde{\mathrm{VIX}}_T^2})\mathbb{E}(e^{-\omega(\mathcal{VS}^J(T,t+\hbar)-\epsilon(T,t+\hbar))})$$

Notice, it is an approximation because $\mathcal{VS}^C(t_1,t_2)$ and $\mathcal{VS}^J(t_1,t_2)-\epsilon(t_1,t_2)$ are not generally independent. Simply replacing $\widetilde{\mathrm{VIX}}_T^2$ with VIX_T^2, materials from Section 4.4 straightforwardly yield the valuations of VIX derivatives.

Adjustments to variance swap rates as a result of return jumps concern $\mathcal{VS}^J(0,T)$ and $\epsilon(T)$. To compute them, we apply the Laplace transform of return jumps.

$$
\begin{aligned}
\mathcal{VS}^J(0,T)=&\frac{1}{T}\int_0^T\int_{\mathbb{R}}x^2\nu(dx,dt)=\frac{\iota}{T}\int_0^T\frac{\partial^2}{\partial\omega_2^2}\mathscr{L}_{j_t,e^{-kt}q_t}(0,\omega_2)|_{\omega_2=0}dt\\
=&\iota(2\gamma\theta\varrho+2\gamma^2\varrho^2+\theta^2+v)\mathcal{E}(2kT).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathcal{VS}(0,T)&=\mathcal{VS}^C(0,T)+\mathcal{VS}^J(0,T)\\
&=\sigma_\infty^2+(\sigma_0^2-\sigma_\infty^2)\mathcal{E}(kT)+\iota(2\gamma\theta\varrho+2\gamma^2\varrho^2+\theta^2+v)\mathcal{E}(2kT)+\sigma_0^2\frac{\gamma\iota}{k}\Phi(k,0,T).
\end{aligned}
$$

Adding jumps to the return process, the adjustment to variance swap rates is that a second slope term is created.

If option replicated variance swaps are used to delegate variance swaps, in the event of

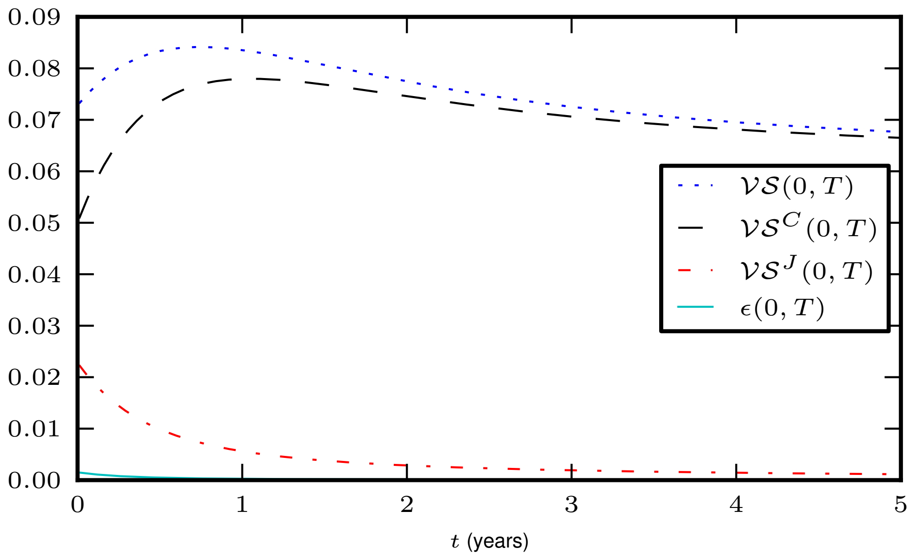

Figure 3: adjustment terms caused by return jumps where $\sigma_0^2 = 0.05$, $\sigma_\infty^2 = 0.06$, $k = 2$, $\gamma = 3$, $\iota = 1$, $\varrho = -0.01$,$\theta = -0.1$, $v = 0.005$.

return jumps, the adjustment term is given by (Carr and Wu (2006))

$$\begin{aligned}
&\epsilon(0,T) = -\frac{2}{T}\int_0^T\int_{\mathbb{R}}(e^x - 1 - x - \frac{x^2}{2})\nu(dx,dt)\\
=& -\frac{2\iota}{T}\int_0^T \mathscr{L}_{j_t,e^{-kt}q_t}(0,-1) - 1 + \left(\frac{\partial}{\partial\omega_2}\mathscr{L}_{j_t,e^{-kt}q_t}(0,\omega_2) - \frac{1}{2}\frac{\partial^2}{\partial\omega_2^2}\mathscr{L}_{j_t,e^{-kt}q_t}(0,\omega_2)\right)|_{\omega_2=0}dt\\
=&2\iota - \frac{2\iota}{T}\int_0^T \frac{\exp(\theta e^{-ks} + \frac{1}{2}ve^{-2ks})}{1-\varrho\gamma e^{-ks}}ds + 2\iota(\theta+\gamma\varrho)\mathcal{E}(kT) + \iota(2\gamma\theta\varrho + 2\gamma^2\varrho^2 + \theta^2 + v)\mathcal{E}(2kT).
\end{aligned}$$

The adjustment term can also be approximated by

$$\begin{aligned}
\epsilon(0,T) =& -\frac{2}{T}\int_0^T\int_{\mathbb{R}}(e^x - 1 - x - \frac{x^2}{2})\nu(dx,dt)\\
\approx& -\frac{1}{3T}\int_0^T\int_{\mathbb{R}} x^3 d\nu(x,t) = \frac{\iota}{3T}\int_0^T \frac{\partial^3}{\partial\omega_2^3}\mathscr{L}_{j_t,e^{-kt}q_t}(0,\omega_2)|_{\omega_2=0}dt\\
=& -\iota(\frac{\theta^3}{3} + \gamma\theta^2\varrho + 2\gamma^2\theta\varrho^2 + 2\gamma^3\varrho^3 + \theta v + \gamma\varrho v)\mathcal{E}(3kT).
\end{aligned}$$

To rectify variance swap rates from option replicated variance swap rates, an extra slope term needs to be compensated. In the case of SPX, usually $\theta < 0$ and $\varrho < 0$ hold, which implies $\epsilon(0,T) > 0$. Under this circumstance, the exponential skewness in the terminology of Carr et al. (2011) is negative and the multiplier which measures the relative price of a variance swap rate to a log contract is larger than 2.

Under the negative exponential skewness situation, these correction terms satisfy $0 < \epsilon(0,T) < \mathcal{VS}^J(0,T)$. Figure 3 reports the adjustment terms $\epsilon(0,T)$ and $\mathcal{VS}^J(0,T)$ along with the terms $\mathcal{VS}^C(0,T)$ and $\mathcal{VS}(0,T)$. In general, diffusion variations contributing to variance swap rates far outweigh jump variations.

7. Data

The Standard & Poor's 500 Index (SPX) is a capitalization-weighted index of 500 large-cap common stocks across a broad range of industries. In the security option market, SPX option is a European type option which is offered on the CBOE(Chicago Board Options Exchange). It trades with expiries with three near-term months followed by further additional months from the March quarterly cycle(March, June, September and December). In addition, the exchange may list Long-term Equity Anticipation Securities(LEAPS) contracts that expires from 12 to 60 months from the date of issue. The expiration date is the Saturday following the third Friday of each expiring month. The CBOE Volatility Index (VIX) option traded on the CBOE is also a European style with expiration months up to six contract months. The maturity date is the Wednesday that is thirty days prior to the third Friday of the calendar month immediately following the expiring month.

In our study, option prices of SPX and VIX are collected from Market Data Express (MDE). In producing option prices, we follow the common convention by taking the average of the bid and ask price for each strike. In processing SPX/VIX option prices, two types of options are deleted, the option whose spread is twice larger than the bid price and the corresponding trading volume is 0, the SPX (resp. VIX) option whose price is less than 0.3 (resp. 0.15) and whose trading volume is 0.

Theoretically, option prices should satisfy certain no-arbitrage shape restriction conditions. Assume $\{(K_i, C_t(K_i), P_t(K_i)\}_{i=1}^n$ is the option price set on trading date t with maturity date T, where K_i, C_t, P_t are the strike, call option price and put option price. Let the futures and the discount factor respectively be F_t and $e^{-r_t(T-t)}$, then no arbitrage option price should satisfy the monotonicity condition

$$-e^{-r_t(T-t)} \leq C_t^{'}(K) \leq 0, 0 \leq C_t^{'}(K) \leq e^{-r_t(T-t)},$$

the convexity condition

$$C^{''}(K) \geq 0, P^{''}(K) \geq 0$$

and the option pricing constraint

$$e^{-r_t(T-t)}(F_t - K)^+ \leq C_t(K) \leq e^{-r_t(T-t)}F_t, e^{-r_t(T-t)}(K - F_t)^+ \leq P_t(K) \leq e^{-r_t(T-t)}K,$$

where $(K - F_t)^+ = \max(0, K - F_t)$.

However, the three no arbitrage constraints of cross-section market option prices are generally not satisfied. We apply the shape restriction to SPX/VIX options following the methodology presented in Aït-Sahalia and Duarte (2003). In performing the shape restriction procedure, the futures and the discount factor are needed. the SPX/VIX futures contract is traded on the Chicago Futures Exchangess (CFE). However, SPX/VIX futures and SPX/VIX options are traded in two different markets, suggesting that they may not be recorded at the same time. For the reason of avoiding non-synchronous microstructure noise, we don't collect SPX/VIX futures and the discount factor or the riskless interest rate from the market. Instead, the value of the futures F_t and the discount factor are computed from option prices according to the put-call parity,

$$C_t - P_t = e^{-r_t(T-t)}F_t - Ke^{-r_t(T-t)},$$

the futures F_t and the discount factor $e^{-r_t(T-t)}$ are obtained by minimizing the weighted sum of squared residual (SSR)

$$SSR = \sum_{i=1}^{n} \omega_i (C_t(K_i) - P_t(K_i) - F_t e^{-r_t(T-t)} + K_i e^{-r_t(T-t)})^2,$$

where $\omega_i, i = 1, \ldots, n$ are the weights to reflect relative liquidity of options.

Date	Asset	Level	Maturity	No. of Strikes
Aug 28, 2008	VIX	19.43	Sep 17, 2008	12
			Oct 22, 2008	10
			Nov 19, 2008	15
			Dec 17, 2008	9
			Jan 21, 2009	11
			Feb 18, 2009	11
	SPX	1300.68	Sep 20, 2008	69
			Oct 18, 2008	75
			Nov 22, 2008	69
			Dec 20, 2008	51
			Mar 21, 2009	36
			Jun 20, 2009	34
			Dec 19, 2009*	31
			Jun 19, 2010*	24
			Dec 18, 2010*	29

* LEAPS

Table 1: Listed are data information of SPX and VIX options traded on August 28, 2008.

Having obtained the futures and the discount factor, we apply the shape restriction approach presented in Aït-Sahalia and Duarte (2003) to convert market option prices to these satisfying no arbitrage shape restrictions.

In a final step, because out-of-the-money (OTM) calls and puts are relatively more liquidly traded than in-the-money(ITM) options, we collect the OTM calls and puts. For SPX/VIX options, we further use the Black option model to convert each OTM option price to a unique value of total implied variance. The option data information, with which we investigate our proposed time-varying two factor model, is listed in Table 1.

8. Calibrations

The parameter calibration step consists of establishing three objective functions, arising from term structure of variance swaps, VIX options and SPX options. Assume maturity dates of SPX options are $\{T_1, T_2, \ldots, T_n\}$, we compute first variance swaps $\{\mathcal{VS}(0, T_j)\}_{j=1}^n$ based on the fact that they are completely replicated from vanilla options up to a correction term caused by jumps. In theory (Carr and Wu (2006)), for a given maturity T, with deterministic interest rate,

$$\begin{aligned}\frac{2e^{rT}}{T}\int_0^\infty \frac{\mathcal{O}(K,T)}{K^2}dK &= \mathcal{VS}^C(0,T) + \mathcal{VS}^J(0,T) - \varepsilon(0,T) \\ &\approx \mathcal{VS}^C(0,T),\end{aligned}$$

where $\mathcal{O}(K,T)$ denotes the value of an out-of-the-money option with strike K and maturity T. However, available option prices in the traded financial market are not continuous. Since there is a one-to-one mapping between option prices and implied volatilities, a smoothed implied volatility curve uniquely produces a smoothed option price function.

In addition to those n synthetic variance swap rates from maturities $\{T_i\}_{i=1}^n$, there is one natural variance swap rate expiring in the next 30 days, that is, the square of the initial VIX price. The model implied initial VIX price is

$$\mathrm{VIX}_0^2 = \mathcal{VS}^C(0,h) + \mathcal{VS}^J(0,h) - \epsilon(0,h) \approx \mathcal{VS}^C(0,h).$$

Since VIX_0^2 is derived geometrically from the linear interpolation/extrapolation of the first two of those n synthetic variance swap rates, including it in them may be unnecessary to do parameter calibrations from a statistic point of view. However, from a financial point of perspective, it can be viewed as assigning a higher weight to the near term variance swaps to reflect relative liquidity and to put more emphasis on VIX price. This consideration reconciles together SPX and VIX derivatives.

In our study, there are totally 9 maturities of SPX options on August 28, 2008. Together with the VIX price, it is sufficient to estimate the parameter set $\Theta_1 = \{\sigma_\infty, \sigma_0, k, \gamma\iota\}$. The parameter set Θ_1 is obtained by minimizing squared distance between option replicated and model implied variance swaps, between market and model implied initial VIX price, i.e., Θ_1 minimizes the root mean squared error(RMSE) given by

$$RMSE_1 = \sqrt{\frac{1}{n+1}\left(\sum_{i=1}^n\left(\frac{2e^{rT}}{T}\int_0^\infty \frac{\mathcal{O}_S^{\mathcal{M}}(K,T)}{K^2}dK - \mathcal{VS}^C(0,T_i)\right)^2 + (\mathrm{VIX}_0^2 - \mathcal{VS}^C(0,h))^2\right)}.$$

Other than Θ_1, the set of parameters controlling the valuations of VIX derivatives is $\Theta_2 = \{\lambda_1, \lambda_2, \xi_1, \xi_2, \gamma\}$. To calibrate them, we computer the squared distance between market implied normalized[3] OTM VIX option prices and the corresponding model implied ones across all available strikes and maturities, i.e., Θ_2 minimizes

$$\mathrm{RMSE}_2 = \sqrt{\frac{\sum_{i,j}(\widetilde{\mathcal{O}}_V^{\mathcal{M}}(K_{i,j}, T_i) - \widetilde{\mathcal{O}}_V(K_{i,j}, T_i))^2}{\sum_{i,j} 1}}.$$

The remaining set of parameters is $\Theta_3 = \{\rho, \varrho, \theta, v\}$. They are calibrated from minimizing the squared distance between market implied and model implied OTM SPX option prices, i.e., Θ_3 minimizes

$$\mathrm{RMSE}_3 = \sqrt{\frac{\sum_{i,j}(\widetilde{\mathcal{O}}_S^{\mathcal{M}}(K_{i,j}, T_i) - \widetilde{\mathcal{O}}_S(K_{i,j}, T_i))^2}{\sum_{i,j} 1}}.$$

Ideally, the three groups of parameters Θ_1, Θ_2 and Θ_3 can be calibrated progressively, known as the method of dimensionality reduction. In consideration of the market liquidity, SPX options are among the most highly liquid options on the market as opposed to VIX options. Furthermore, in the event that the return process of SPX contains jumps, the objection functions RMSE_1 and RMSE_2 should incorporate adjustment terms relying on Θ_3. In light of these considerations, calibrations are performed by minimizing a weighed sum of

$$w_1 \,\mathrm{RMSE}_1 + w_2 \,\mathrm{RMSE}_2 + w_3 \,\mathrm{RMSE}_3, \quad \text{where} \quad w_1 + w_2 + w_3 = 1.$$

The minimization is a nonlinear least square optimization problem. Numerical methods available to solve a minimization problem include Gauss-Newton algorithm(GNA), Leverberg-Marguardt algorithm(LMA), genetic programming. The solution from implementing the LMA or the GNA primarily depends on the initial point, that means in many cases, a local minimum value is achieved rather than the global minimum. A global searching heuristic of differential evolution (DE) method (Price et al. (2005)) is implemented first, which is then followed by the LMA to solve the minimization problem.

Table 2 reports the calibrated parameters and the three RMSEs. Both the speed of reversion and the volatility of volatility of the short term are more severe than those of the long term. It demonstrates the uncertainty in the short term is higher than in the long term. From the specification of the leverage effect, the high level of uncertainty in the short term also increases the level of a second risk factor which is not explained by the risk factor of

[3]Given a European option price $\mathcal{O}(K, T)$ with strike K and maturity T, its normalized option price is given by $\widetilde{\mathcal{O}}(K, T) = e^{rT} \frac{\mathcal{O}(K,T)}{F_T}$ where F_T is the corresponding futures price. The normalized option price simply measures the relative option price to the present value of futures.

Parameters					
		Descriptions		weights	RMSEs
Θ_1	k	damping	3.51	0.225	1.3×10^{-3}
	σ_0^2	initial variance	0.026		
	σ_∞^2	long run variance	0.065		
	$\gamma\iota$	jump mean(VIX) $\times$ jump intensity	1.57		
Θ_2	λ_1	long speed of reversion	1.57	0.05	9.0×10^{-3}
	λ_2	short speed of version	9.81		
	ξ_1	long vol of vol	1.69		
	ξ_2	short vol of vol	2.08		
	γ	jump mean(VIX)	11.4		
Θ_3	ρ	diffusion correlation	-0.99	0.725	3.1×10^{-4}
	ϱ	jump correlation	-0.0032		
	θ	jump mean(SPX)	-0.156		
	v	jump variance(SPX)	0.023		

Table 2: The groups $\Theta_1, \Theta_2, \Theta_3$ are calibrated from SPX/VIX options on August 28, 2008.

the return. The RMSE of SPX options is 3.1×10^{-4}. Its equivalent value in dollar term is around $1300 \times 3.1 \times 10^{-4} \approx \$0.40$. The RMSE of VIX options is 9.0×10^{-3}. Its equivalent dollar value is approximately $22 \times 9.0 \times 10^{-3} \approx \$0.198$.

9. Results

9.1. Term Structure of Variance Swaps

Figure 4 displays the term structure of variance swaps together with the current VIX level. The current VIX price practically passes through a straight line determined by option replicated variance swaps from the two nearest maturities. The calibrated long run variance level $\sigma_\infty^2 = 0.065$, which suggests that the steady state instantaneous volatility is close to 0.255. The initial variance level is $\sigma_0^2 = 0.026$, which is about 0.161 in terms of volatility measurement. The damping is valued at 3.51. The half-life associated with the exponential damping is approximately 0.197 years or 72 days, suggesting the variance level will be half the distance towards its steady level after approximately 72 days.

9.2. Convexity Adjustment

A standard result from Jensen's inequality implies the following inequality relationship between variance futures $\mathbb{E}\,\mathrm{VIX}_T^2$ and volatility futures $\mathbb{E}\,\mathrm{VIX}_T$:

$$\mathbb{E}\,\mathrm{VIX}_T = \mathbb{E}\sqrt{\mathrm{VIX}_T^2} \leq \sqrt{\mathbb{E}\,\mathrm{VIX}_T^2}.$$

Hence, volatility futures is bounded above by the square root of variance futures. The difference between the square root of variance futures and volatility futures is called the *convexity correction*.

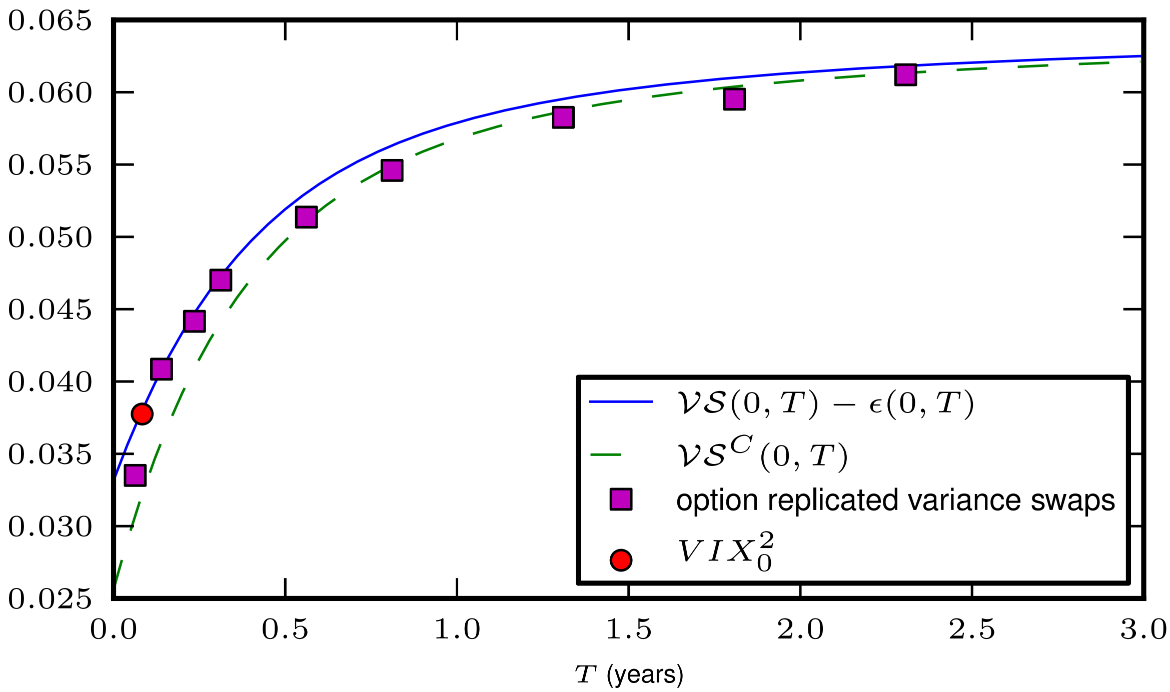

Figure 4: term structure of variance swaps

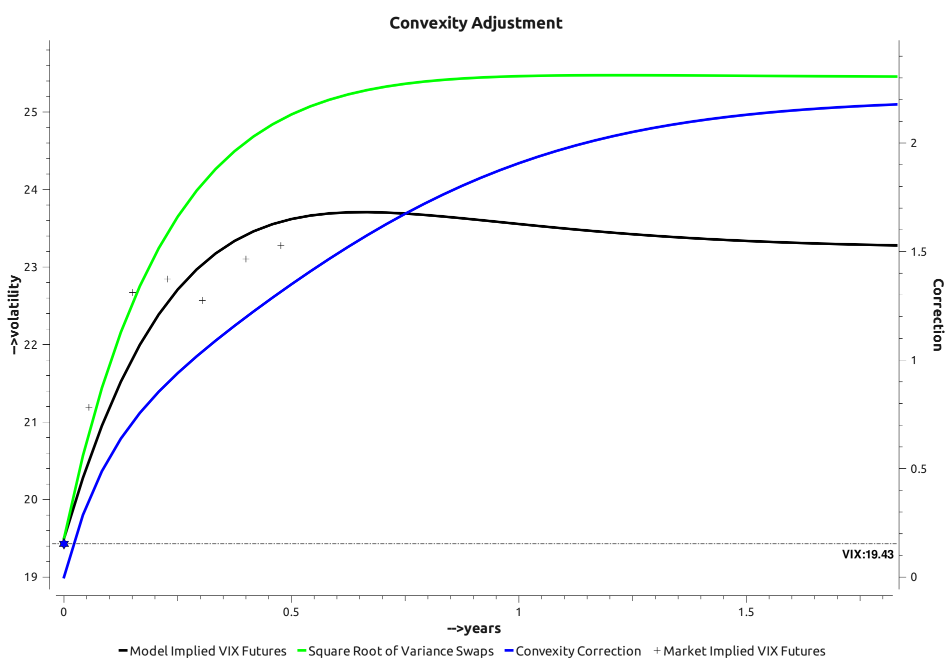

Figure 5: Convexity adjustment

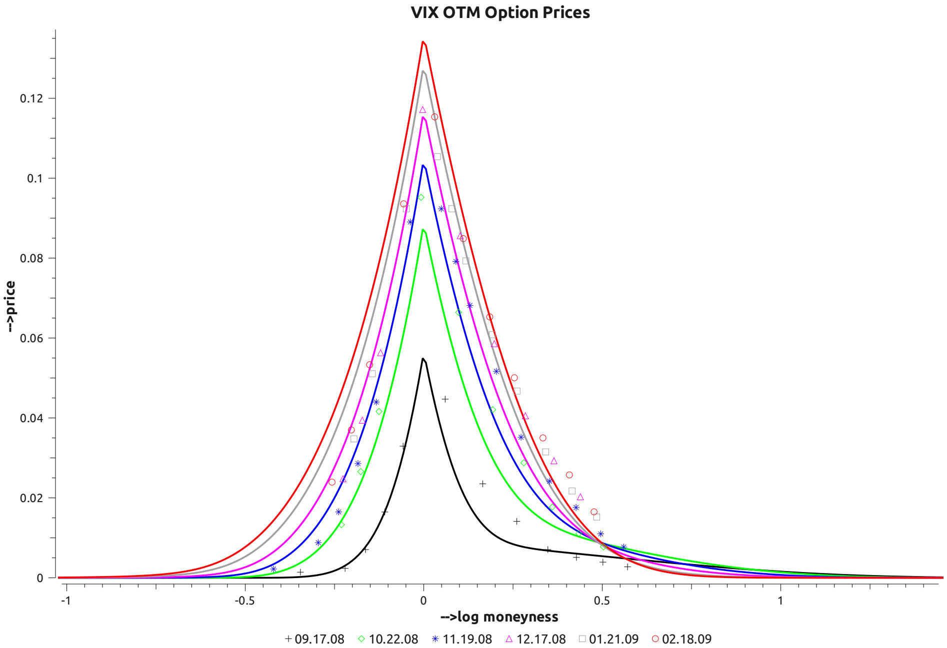

Figure 6: OTM VIX option prices

Figure 5 reports VIX futures and the corresponding square root of 30-days forward variance swaps. The convexity correction is plotted by the right Y-axis. The correction starts initially at 0, then it gets wider as time goes but becomes stable eventually. From the model specification, we have $\lim_{T\to\infty}\sqrt{\mathbb{E}\,\mathrm{VIX}_T^2} = \lim_{T\to\infty}\sqrt{\mathbb{E}\widetilde{\mathrm{VIX}}_T^2} = \sigma_\infty$. From equations (8)(9)(10), the equilibrium VIX futures $\lim_{T\to\infty}\mathbb{E}\,\mathrm{VIX}_T$ depends on both Θ_1 and Θ_2.

9.3. OTM VIX Options

Figure 6 plots the OTM VIX option prices on August 28, 2008. For each long term maturity, the model implied call prices are consistently undervalued at the high extreme strikes and the model implied put prices are consistently overvalued at the low extreme strikes. The reason is that the Heston dynamics are mis-specified. For an intuitive explanation, Figure 7 draws the total implied variance curves if the instantaneous variance process is described by the Heston dynamics. The implied variance curves apparently have downward sloping shapes in both tails. However, demonstrated in Figure 8, empirically, the implied variance curves are upward sloping. Transforming into option prices, at the high extreme strikes, it is not surprising that the OTM call prices derived from the Heston model are systematically underestimated, and at the low extreme strikes, the put prices are overestimated. The mean-reverting processe, such as the Ahn-Gao process (Ahn and Gao (1999)) and the GARCH diffusion process (Bollerslev et al. (1994)), might be a better choice to rectify this tail inconsistency. Empirical supportings for these two processes over the Heston are advocated in Aït-Sahalia and Kimmel (2007), and Bakshi et al. (2006). However, in the spirit of model

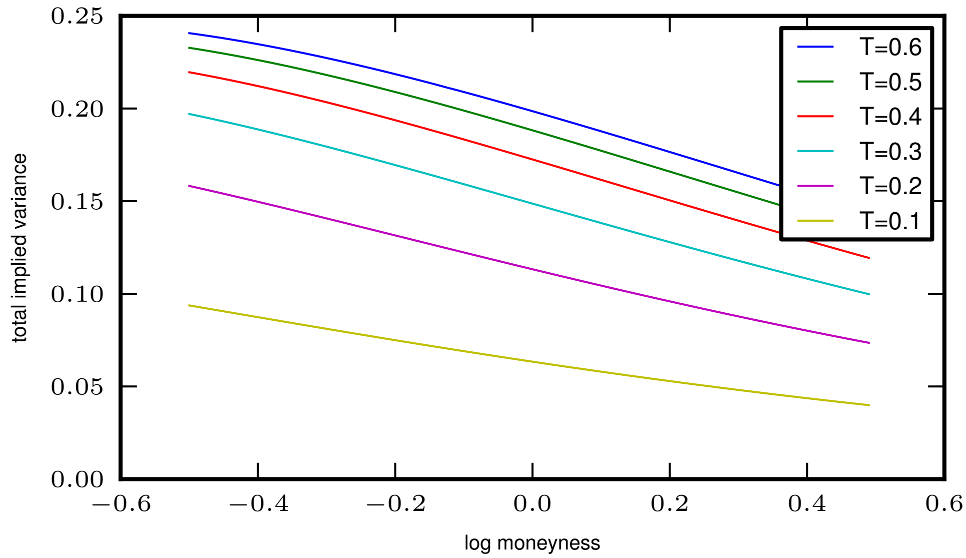

Figure 7: total implied variances of VIX options if the instantaneous variance process follows the Heston dynamics: $dx_t = \lambda(\mu - x_t) + \xi\sqrt{x_t}dw_t$ where $\lambda = 1.6, \mu = 0.065, \xi = 0.45$.

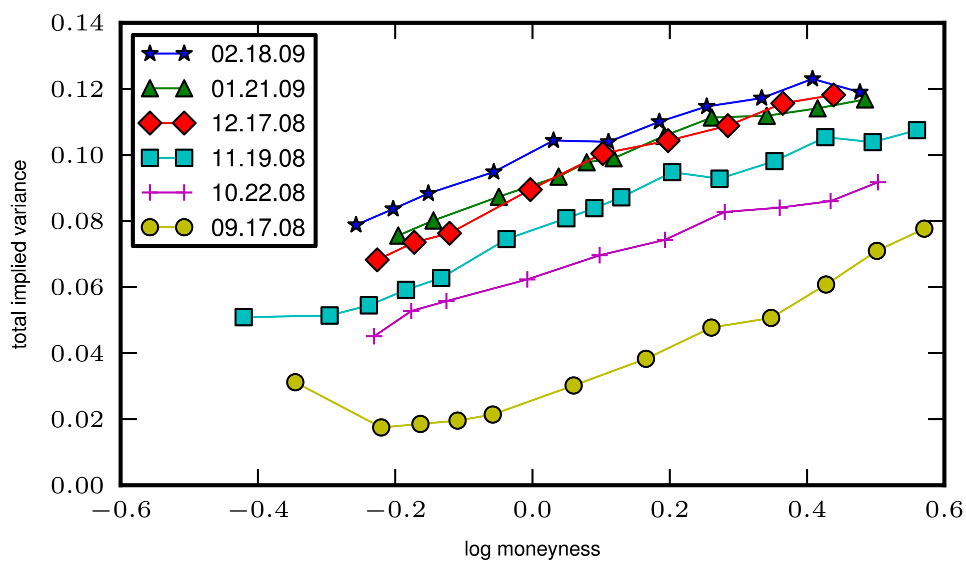

Figure 8: total implied variances of VIX options on August 28, 2008

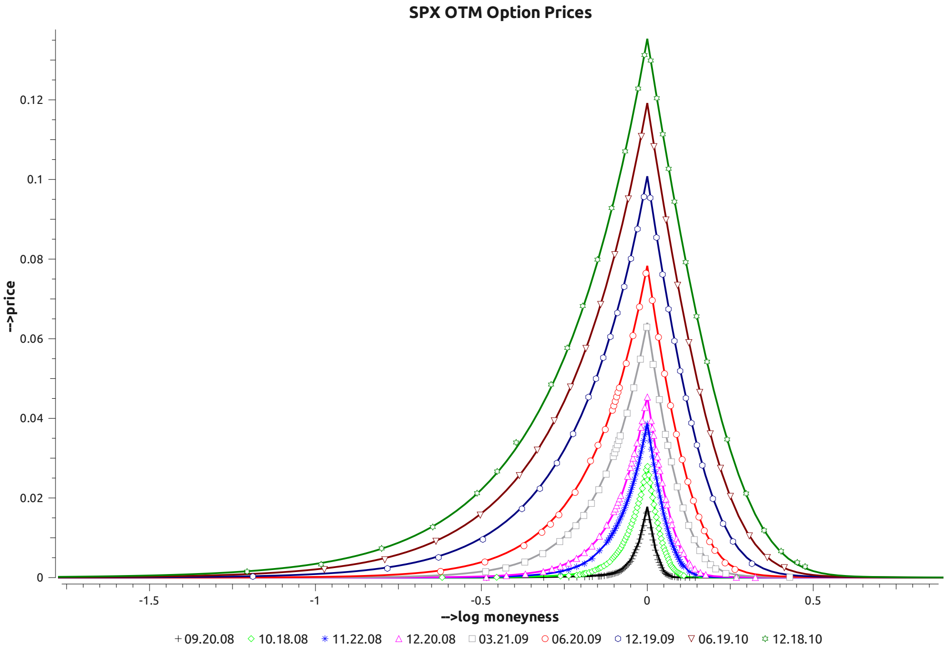

Figure 9: OTM SPX option prices

tractability, they may short of closed analytical solutions to derivative valuations.

9.4. OTM SPX Options

Figure 9 presents the OTM SPX option prices across all maturities and available strikes on August 28, 2008. Figure 10 reports the corresponding total implied variance curves against the forward log moneyness. The model fits generally well to SPX derivatives.

Conclusion

The main feature of the two-factor model is that its factor coefficients are time-varying. The time-varying coefficients are coherently bonded with the term structure of variance swaps. In the interest of model tractability, the factors of the model are built upon affine processes. We find the model fits well to SPX options. However, the VIX dynamics are misspecified with the engagement of affine Heston dynamics. Examination of other mean-reverting processes, at the sacrifice of analytical tractability, is left for the future research. In addition to the Heston misspecification, the jump direction of instantaneous variance process of the two factor model is refrained from granting negative values. To fortify model robustness, an extension could also include permitting both negative and positive jumps.

References

Ahn, A., Gao, B., 1999. A Parametric Nonlinear Model of Term Structure Dynamics. The Review of Financial Studies 12 (4), 721–762.

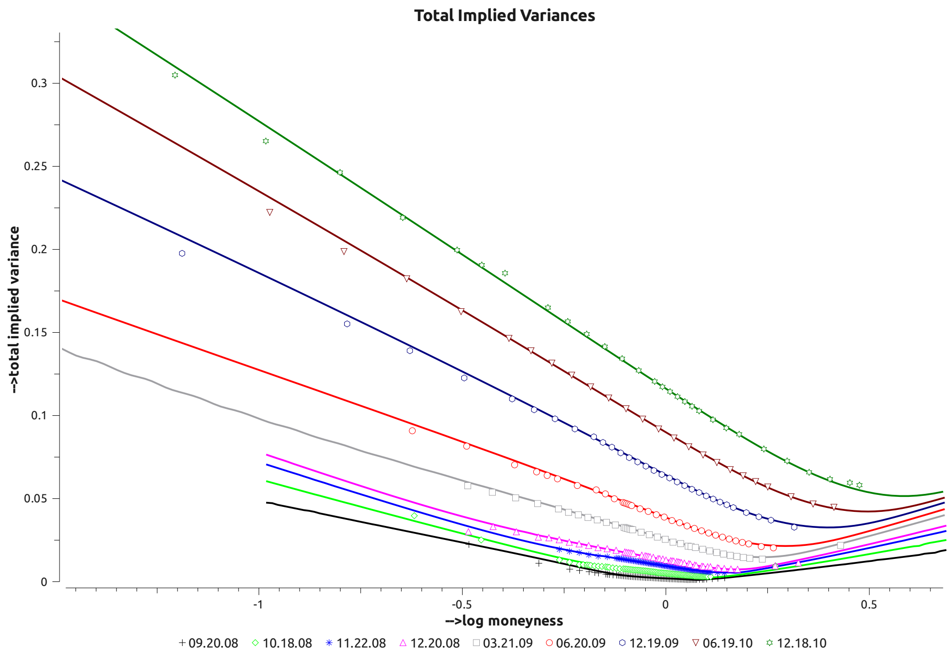

Figure 10: Total implied variances of SPX options

Aït-Sahalia, Y., Duarte, J., 2003. Nonparametric option pricing under shape restrictions. Journal of Econometrics 116, 9–47.

Aït-Sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics 83, 413–452.

Bakshi, G., Cao, C., Chen, Z., 2000. Pricing and hedging long-term options. Journal of Econometrics 94 (1-2), 277–318.

Bakshi, G., Ju, N., Ou-Yang, H., 2006. Estimation of continuous-time models with an application to equity volatility dynamics. Journal of Financial Economics 82 (1), 227 – 249.

Balduzzi, P., Das, S. R., Foresi, S., February 1998. The central tendency: A second factor in bond yields. The Review of Economics and Statistics 80 (1), 62–72.

Bates, D. S., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. Review of Financial Studies 9, 69–107.

Bergomi, L., 2004. Smile Dynamics. Risk, 117–123.

Bergomi, L., 2005. Smile Dynamics II. Risk, 67–73.

Bergomi, L., 2008. Smile Dynamics III. Risk, 90–96.

Bollerslev, T., Engle, R. F., Nelson, D. B., 1994. Arch models. In: Engle, R. F., McFadden, D. L. (Eds.), Handbook of Econometrics. Vol. IV. Amsterdam: North Holland, Ch. 49, pp. 2959–3038.

Broadie, M., Jain, A., 2008. The effect of jumps and discrete sampling on volatility and variance swaps. International Journal of Theoretical and Applied Finance (IJTAF) 11 (08), 761–797.

Buehler, H., 2006. Consistent variance curve models. Finance and Stochastic 10, 178–203.

Carr, P., Geman, H., Madan, D. B., Yor, M., 2005. Pricing options on realized variance. Finance and Stochastics 9, 453–475.

Carr, P., Lee, R., Wu, L., 2011. Variance swaps on time-changed levy processes. Finance and Stochastics, 1–21.

Carr, P., Madan, D., 1998. Towards a theory of volatility trading. In: jarrow, R. (Ed.), Volatility: New Estimation Techniques for Pricing Derivatives. Risk Publications, pp. 417–427.

Carr, P., Madan, D., 1999. Option valuation using the fast fourier transform. Journal of Computational Finance 2, 61–73.

Carr, P., Wu, L., 2006. A Tale of Two Indices. Journal of Derivatives 12, 13–29.

Carr, P., Wu, L., 2007. Stochastic skew in currency options. Journal of Financial Economics 86, 213–247.

Christensen, J. H., Diebold, F. X., Rudebusch, G. D., 2011. The affine arbitrage-free class of nelson-siegel term structure models. Journal of Econometrics 164 (1), 4 – 20, annals Issue on Forecasting.

Christensen, J. H. E., Diebold, F. X., Rudebusch, G. D., November 2009. An arbitrage-free generalized nelson–siegel term structure model. Econometrics Journal 12 (3), C33–C64.

Christoffersen, P., Jacob, K., Ornthanalai, C., Wang, Y., 2008. Option Valuation with Long-run and Short-run Volatility Components. Journal of Financial Economics 90, 272–297.

Christoffersen, P. F., Heston, S., Jacobs, K., 2009. The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. Management Science 55, 1914–1932.

Cont, R., Kokholm, T., 2011. A consistent pricing model for index options and volatility derivatives. Mathematical Finance.

Cont, R., Tankov, P., 2004. Financial modelling with jump processes. Chapman & Hall/CRC financial mathematics series. Chapman & Hall/CRC.

Dai, Q., Singleton, K. J., October 2000. Specification analysis of affine term structure models. Journal of Finance 55 (5), 1943–1978.

Das, S. R., Sundaram, R. K., 1998. Of smiles and smirks: A term-structure perspective. JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS 34, 211–240.

Duffie, D., Gârleanu, N., 2001. Risk and Valuation of Collaterzlied Debt Obligations. Financial Analysis Journal 57, 41–59.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6 (4), 379–406.

Duffie, D., Pan, J., Singleton, K., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.

Duffie, D., Singleton, K., 1999. Modeling Term Structures of Defaultable Bonds. The Review of Financial Studies 12, 687–720.

Dupire, B., 1994. Pricing with a smile. RISK 7 (1), 18–20.

Egloff, D., Leippold, M., Wu, L., October 2010. The term structure of variance swap rates and optimal variance swap investments. Journal of Financial and Quantitative Analysis 45 (05), 1279–1310.

Fengler, M. R., Härdle, W., Villa, C., 2003. The dynamics of implied volatilies: A common principal components approach. Review of Derivatives Research 6, 179–202.

Fonseca, J. D., Grasselli, M., Tebaldi, C., 2008. A Multifactor Volatility Heston Model. Quantitative Finance 8 (6), 591–604.

Friz, P., Gatheral, J., December 2005. Valuation of volatility derivatives as an inverse problem. Quantitative Finance 5 (6), 531–542.

Gatheral, J., 2008. Consistent modeling of spx and vix options. Tech. rep., Merrill Lynch.

Kreyszig, E., 2005. Advanced Engineering Mathematics, 9th Edition. Wiley.

Lee, G. G. J., Engle, R. F., 1999. A permanent and transitory component model of stock return volatility. In: Engle, R. F., White, H. (Eds.), Cointegration Causality and Forecasting A Festschrift in Honor of Clive WJ Granger. Oxford University Press, pp. 475–497.

Madan, D., Yor, M., 2010. The s&p 500 index as a sato process travelling at the speed of the vix. Applied Mathematical Finance, 1–18.

Nelson, C. R., Siegel, A. F., October 1987. Parsimonious modeling of yield curves. Journal of Business 60 (4), 473–89.

Neuberger, A., 1990. Volatility trading. London Business School Working Paper.

Neuberger, A., 1994. The log contract. Journal of Portfolio Management 20, 74–80.

Price, K., Storn, R. M., Lampinen, J. A., 2005. Differential Evolution: A Practical Approach to Global Minimization. Natural Computing Series. Springer.

Skiadopoulos, G., Hodges, S., Clewlow, L., 1999. The Dynamics of the S&P 500 Implied Volatility Surface. Journal of Economic Literature 3, 262–282.

Svensson, L., 1994. Estimating and interpreting foreward interest rates: Sweden 1992-1994. Papers 579, Stockholm - International Economic Studies.

Todorov, V., Tauchen, G., 2011. Volatility jumps. Journal of Business and Economic Statistics, Forthcoming.

Transon, A., 1841. Recherches sur la courbure des lignes et des surfaces. Journal de mathmatiques pures et appliques, 191–208.

Wilczynski, E. J., 1916. Some remarks on the historical development and the future prospects of the differential geometry of plane curves. Bulletin of the American Mathematical Society 22 (7), 317–329.